2018

Asymmetric Biotic Interactions and Abiotic Niche Differences Revealed by a Dynamic Joint Species Distribution Model

Nina K. Lany
Michigan State University

Phoebe Zarnetske
Michigan State University

See next page for additional authors

Follow this and additional works at: https://digitalcommons.uri.edu/bio_facpubs

The University of Rhode Island Faculty have made this article openly available. Please let us know how Open Access to this research benefits you.

This is a pre-publication author manuscript of the final, published article.

Terms of Use
This article is made available under the terms and conditions applicable towards Open Access Policy Articles, as set forth in our Terms of Use.

Citation/Publisher Attribution

This Article is brought to you for free and open access by the Biological Sciences at DigitalCommons@URI. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.
Running title: Spatio-temporal joint distribution model

Asymmetric biotic interactions and abiotic niche differences revealed by a dynamic joint species distribution model

N.K. Lany$^{1,2}$, P.L. Zarnetske$^{1,2}$, E.M. Schliep$^3$, R.N. Schaeffer$^4$, C.M. Orians$^4$, D.A. Orwig$^5$, E.L. Preisser$^6$

$^1$Department of Forestry, Michigan State University, East Lansing, MI 48824 USA
$^2$Ecology, Evolutionary Biology, and Behavior Program, Michigan State University, East Lansing, MI 48824 USA
$^3$Department of Statistics, University of Missouri, Columbia, MO 65211 USA
$^4$Department of Biology, Tufts University, Medford, MA 02155 USA
$^5$Harvard Forest, Harvard University, Petersham, MA 01366 USA
$^6$Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881 USA

Corresponding author: Nina K. Lany

480 Wilson Road, MSU Department of Forestry
East Lansing, MI 48862
Phone: (517) 355-7671, Fax: (517) 432-1143
e-mail: lanynina@msu.edu
Abstract: A species’ distribution and abundance are determined by abiotic conditions and biotic interactions with other species in the community. Most species distribution models correlate the occurrence of a single species with environmental variables only, and leave out biotic interactions. To test the importance of biotic interactions on occurrence and abundance, we compared a multivariate spatio-temporal model of the joint abundance of two invasive insects that share a host plant - hemlock woolly adelgid (HWA; *Adelges tsugae*) and elongate hemlock scale (EHS; *Fiorinia externa*) - to independent models that do not account for dependence among co-occurring species. The joint model revealed that HWA responded more strongly to abiotic conditions than EHS. Additionally, HWA appeared to predispose stands to subsequent increase of EHS, but HWA abundance was not strongly dependent on EHS abundance. This study demonstrates how incorporating spatial and temporal dependence into a species distribution model can reveal the dependence of a species’ abundance on other species in the community. Accounting for dependence among co-occurring species with a joint distribution model can also improve estimation of the abiotic niche for species affected by interspecific interactions.

**Keywords:** *Adelges tsugae*, *Fiorinia externa*, invasive species, spatio-temporal species distribution model, species interactions, *Tsuga canadensis*

Introduction

Ecologists have long sought to understand how abiotic conditions and biotic interactions combine to determine a species’ distribution and abundance (Grinnell 1917, Andrewartha and Birch 1954, MacArthur 1972). The niche concept is often employed to conceptualize this balance (Chase and Leibold 2003). The effect of the environment on a species, with an emphasis on broad-scale abiotic conditions, has historically been associated with the Grinnellian niche (Grinnell 1917), while the impact of a species on the environment and local interactions with
other species have been associated with the Eltonian niche (Elton 1927). Subsequent ecological theory has integrated these paradigms to define a species’ niche as the range of biotic interactions and abiotic conditions under which a species has a positive population growth rate (Hutchinson 1957, Chase and Leibold 2003). Hutchinson (1957) distinguished the “fundamental niche” that encompasses the range of conditions under which a species could potentially exist from the “realized niche” that encompasses the typically smaller range of conditions under which a species can exist when competing with other species. The current definition additionally acknowledges predation and mutualism, as well as dispersal limitation (Peterson et al. 2011).

The distribution of a species can be interpreted as a projection of the realized niche onto geographic space (Pulliam 2000, Peterson et al. 2011). Despite the connection between both the biotic and abiotic components of a species niche and its geographic distribution, most species distribution modeling approaches correlate the occurrence of a single species with broad-scale environmental variables but omit biotic interactions. Because distribution and abundance often depend on other species in the community, explicitly incorporating biotic interactions into species distribution models is a research priority (Godsoe et al. 2015).

One way to accommodate biotic interactions is to model the joint distribution or abundance of species in a community with a multivariate generalized linear model that estimates the response of each co-occurring species to the abiotic environment. This approach explicitly accounts for residual dependence among species that can arise from either shared responses to an unmeasured covariate, or interactions among species (e.g. Ovaskainen et al. 2010, Pollock et al. 2014, Warton et al. 2015). Whereas the vast majority of species distribution models use static binary occurrence data, a time-series of abundance data provides more information on dynamic and density-dependent ecological processes (Pagel and Schurr 2012, Ehrlén and Morris 2015). In
addition, accounting for spatial autocorrelation can reflect underlying interactions among species and improve the precision of parameter estimates (Dormann et al. 2007, Ovaskainen et al. 2016). Here, we utilize a dynamic, spatially explicit joint species distribution model and long-term, spatially explicit data on the abundance of two invasive insect herbivores that share a common host plant – hemlock woolly adelgid (HWA; *Adelges tsugae*) and elongate hemlock scale (EHS; *Fiorinia externa*) – to test the hypotheses that: 1) the abiotic niches of these co-occurring species are different, and 2) the abundance of each of these species is dependent on biotic interactions with the other. We explicitly compare joint vs. independent models.

**Methods**

In the eastern USA, eastern hemlock (*Tsuga canadensis*) is host plant to HWA and EHS. HWA is a sessile xylem-feeding insect introduced to eastern North America from Japan and first documented in 1951 that has severely impacted eastern hemlocks and threatens to extirpate the species across its range (Orwig et al. 2012). EHS is also a sessile xylem-feeding insect introduced from Japan in 1908 that preferentially feeds on eastern hemlock needles and rarely kills its host tree (McClure 1980a). Fine-scale experiments have revealed exploitative competition between HWA and EHS at the scale on individual branches (Preisser and Elkinton 2008) and large-scale observations suggest HWA may facilitate EHS (Preisser et al. 2008).

We assessed the abundance of HWA and EHS on five occasions over 14 years at 142 forest stands across a latitudinal transect encompassing 7,500 km² in Connecticut (CT) (Orwig et al. 2002) and Massachusetts (MA) (Orwig et al. 2012). Stands were initially visited in 1997-1998 (CT) or 2002-2004 (MA), and each one of these stands were subsequently re-visited in 2005, 2007, 2009 and 2011. In the initial year of sampling, each stand was given an ordinal score representing the average infestation level of the stand (0 = 0 insects per meter of branch; 1 = 1-
10 insects/m; 2 = 11 – 100 insects/m; 3 = >100 insects/m). In subsequent years, 50 trees were haphazardly selected in each stand for observation. Fewer than 50 trees were sampled per stand in some highly-damaged stands, and stands impacted by logging or development during the study period were not sampled post-disturbance, resulting in a total of 27,050 observations. The median distance between pairs of stands was 56.7 km, and ranged from 0.2 to 165.2 km.

Daily temperature and precipitation data were obtained for each stand from 1996 to 2011 by interpolating 4 km² resolution climate data at the centroid of each eastern hemlock stand (PRISM Climate Group). For each stand-year, we calculated three weather variables known to affect HWA and EHS abundance: minimum temperature during the winter preceding the growing season, maximum summer temperature during the growing season, and total precipitation during the interval April 1 – September 30. We expect a positive relationship between winter temperature and insect abundance due to winter mortality (Cheah 2017) and between summer temperature and abundance due to the effects of temperature on development rate (Salom et al. 2002). Extremely warm summer temperatures, however, cause mortality for EHS (McClure 1989) and HWA during diapause (Sussky and Elkinton 2015). Heavy rains dislodge adelgid and scale insects (McClure 1989) and insects also benefit from feeding on drought-stressed trees (Koricheva et al. 1998), resulting in a negative relationship with summer precipitation. Minimum winter temperatures ranged from -12.4 to -28.4 °C, and were negatively correlated with latitude (r = -0.78, Appendix S1: Figure S1). Summer precipitation ranged from 422.7 to 1187.3 mm, and maximum summer temperature ranged from 30.0 to 38.5 °C. Neither summer precipitation (r = -0.07) nor summer temperature (r = -0.11) was strongly correlated with latitude, but both showed high inter-annual variation (Appendix S1: Figure S1). The
greatest correlation between covariates occurred between summer temperature and precipitation (r = 0.54). Data are archived at the Environmental Data Initiative (Orwig et al. 2017).

We modeled the joint abundance of the two insects with a multivariate generalized linear model with probit link function following the methods we developed in Schliep et al. (2018). We extended the probit link function to accommodate ordinal abundance categories by assuming that for each species s on tree j in stand i and time t, the observed ordinal abundance $Y^{(s)}_{i,j,t}$ resulted from a thresholding process on a latent (or unobserved true) multivariate Gaussian abundance $Z^{(s)}_{i,j,t}$. Here, $s = 1$ for HWA and $s = 2$ for EHS. Because the same trees were not sampled between years, we used a hierarchical structure to infer the stand-level mean ($K_{i,t}$) of the latent bivariate abundance $Z_{i,j,t}$ for each insect species in each year, such that $Z_{i,j,t} \sim \text{Multivariate Normal} (K_{i,t}, \Omega_{i})$. Larger values of $K_{i,t}$ indicate higher abundance of a species in a particular stand and year, while lower values indicate lower abundance. Tree-level dependence between species, the scale at which these species interact (Preisser and Elkinton 2008), was modeled with a 2x2 covariance matrix ($\Omega_{i}$) for each stand. The diagonals $\Omega_{1,1}$ and $\Omega_{2,2}$ describe the variance in abundance of each species on individual trees within a stand across all years, and the off-diagonal $\Omega_{1,2} = \Omega_{2,1}$ describes the within-stand covariance in abundance between the two species.

We defined the mean latent abundance of each species as $K_{i,t} = \alpha_t + \beta X_{i,t} + \rho K_{i,t-1} + \eta_{i,t}$, using the species- and year-specific random intercept $\alpha_t$, to capture variability across years and account for northward range expansion over the study period (see Schliep et al. 2018 for discussion of why a temporal random effect is necessary for these data), the term $\beta X_{i,t}$ to incorporate abiotic conditions specific to each stand-year, a lag-1 vector auto-regressive process $\rho K_{i,t-1}$ to capture temporal dependence, and a spatially correlated error term $\eta_{i,t}$ to capture spatial dependence. $X_{i,t}$ included weather-related covariates specific to each stand-year as both linear
and quadratic terms: minimum winter temperature, maximum summer temperature, and summer precipitation. All covariates were mean centered and standardized. $\beta$ was the 2 x 7 (linear and quadratic forms of each of the three predictor variables, plus the intercept) matrix of coefficients that described the response to abiotic conditions unique to each species and allowed comparison of the abiotic niche for each species. Inter- and intra-specific temporal dependence was modeled with the 2x2 lag-1 autoregressive matrix $\rho$. The off-diagonal elements of the parameter matrix $\rho$ ($\rho_{1,2}$ and $\rho_{2,1}$) described temporal dependence between species. For example, positive estimates of the off-diagonal parameter $\rho_{1,2}$ would indicate that average stand-level EHS latent abundance at time $t-1$ made a stand more susceptible to infestation by HWA at time $t$. Importantly, temporal dependence between species can be directional because the $\rho$ matrix is not necessarily symmetric. Spatially-correlated dependence within and among species not accounted for by model terms was captured with a linear model of coregionalization for the error term $\eta_{l,t}$. This permitted estimation of the effective range (the distance at which residual spatial correlation dropped below 0.05) for each species (Schliep et al. 2018). A large estimated effective range would indicate that important predictor variable(s) may be missing from the model.

We obtained inference in a Bayesian framework with non-informative and conjugate priors, and calculated marginal rank probability scores (RPS) to assess model fit (Schliep et al. 2018). We used the function 'Multivariate.Ordinal.Spatial.ModelX' available in the online supplement for Schliep et al. (2018). We evaluated evidence for the hypothesis that there is a difference in the abiotic niches of two species by comparing posterior estimates of the $\beta$ coefficients. To evaluate whether biotic interactions between the two insects mediate distribution and abundance (hypothesis 2), we evaluated the posterior estimates of $\rho_{1,2}$ and $\rho_{2,1}$. In addition, we specified independent models that did not include biotic interactions by setting the
parameters that describe temporal ($\rho_{1,2}$ and $\rho_{2,1}$), spatial (in the error term $\eta_{i,t}$), and tree-level ($\Omega_{i,2} = \Omega_{2,i}$ for each stand) dependence between species to zero. We compared the effective range of residual spatial correlation for each species from the joint model that accounts for dependence among species vs. independent models of the abundance of each species that do not account for dependence. Narrower credible intervals for the $\beta$ coefficients and smaller effective ranges in the dependent vs. independent model would indicate a better-specified, more robust model (Barry and Elith 2006). Markov chain Monte Carlo was run for 10,000 iterations and the first 2,000 were discarded as burn-in. No issues of convergence were detected in any of the models. An R script that runs the joint and independent models is provided in Appendix S2.

Results

The posterior mean of latent abundance of each species varied from year-to-year and also with latitude (Figure 1). In the joint model, HWA abundance was positively associated with minimum winter temperature as both linear and quadratic terms (Figure 2). HWA abundance was negatively associated with summer precipitation and positively associated with the square of summer precipitation (Figure 2). HWA abundance increased linearly according to maximum summer temperature (Figure 2). EHS abundance was positively and linearly associated with minimum winter temperature, but none of the other posterior coefficient estimates describing the abiotic niche for EHS were significantly different than zero according to the 95% credible intervals (Appendix S1: Table S1).

We found evidence for dependence between HWA and EHS. Both parameters that describe temporal dependence between the species ($\rho_{1,2}$ and $\rho_{2,1}$) had positive posterior means (Figure 3), indicating that higher EHS abundance at time $t-1$ was associated with higher HWA abundance at the subsequent time step, and vice versa. Zero was in the posterior credible interval
for $\rho_{2,t}$ (Appendix S1: Table S1), and the probability that $\rho_{2,t} > 0$ was 9.966. After accounting for all other model parameters, tree-level covariance across all years between the latent abundance of the two species ($\Omega_{t,2}$) was largely not significant from zero for the majority of eastern hemlock stands (118 of 142, Appendix S1: Figure S2). For the rest of the stands however, we did detect positive tree-level covariance in 19 stands, while five were negative. There was greater variability in abundance of both species among trees in southern stands ($\Omega_{t,1}$ and $\Omega_{2,2}$), especially for HWA (Appendix S1: Figure S2). There was positive spatial dependence between the two species at the stand level, and the effective range of residual spatial correlation was larger for EHS than for HWA (29.3 vs. 2.9 km, Figure 3).

Modeling the abundance of the two species jointly had a larger effect on EHS-specific parameters than on HWA-specific parameters. Posterior estimates for EHS tended to have narrower credible intervals in the joint distribution model (Figure 2), and the effective range of EHS residual spatial correlation was smaller in the joint model than in the independent model (26.7 km vs. 87.6 km, Figure 2). For HWA, however, the posterior coefficient estimates and the width of the credible intervals (Figure 2), as well as the effective range (Figure 2), were very similar in the independent vs. joint models. Marginal RPS did not indicate problems with lack of model fit, and were similar between the joint and independent models (Appendix S1: Figure S3).

**Discussion**

This study provides some of the first evidence that simultaneously modeling the abundance of multiple species in a community with a spatio-temporal joint species distribution model can indicate the degree to which a species’ distribution and abundance are dependent on biotic interactions with other species (but see Schliep et al. 2018). Our study also illustrates how this approach can improve estimation of the abiotic niche of species whose abundance is
dependent on other species. Analyses revealed differences in the abiotic niches of EHS and HWA. The positive relationship between minimum winter temperature and abundance was quadratic for HWA and linear for EHS. Therefore, we expect directional increases in winter temperature to benefit HWA more than EHS. Recent studies align with this expectation, showing that colder winter temperatures reduce HWA populations (Cheah 2017). HWA abundance was sensitive to abiotic conditions during the growing season, but EHS abundance was not. Higher HWA abundance was associated with extremely dry summers, perhaps because sap-sucking insects perform well when trees are water-stressed (Koricheva et al. 1998). HWA abundance was also positively associated with maximum summer temperatures – a pattern consistent with the ways temperature regulates development rate, an important life history characteristic for HWA (Salom et al. 2002). Taken together, these findings indicated that HWA was sensitive to extremes in abiotic conditions that may become more common as climate changes.

Hemlock woolly adelgid appeared to predispose stands to subsequent increase of EHS, but HWA abundance was not strongly dependent on EHS abundance. Evidence for dependence of EHS on HWA was found in the positive stand-level temporal dependence between the species ($\rho_{2,1}$, although the posterior credible interval for this parameter contained zero), and in the increased effective range of residual spatial autocorrelation combined with lower precision of parameter ($\beta$) estimates in the independent model, which does not account for dependence between species. Temporal dependence of HWA on EHS ($\rho_{1,2}$) was also positive, but the effective range and precision of the posterior distribution of the $\beta$ parameters were very similar in the independent vs. joint models for HWA. This asymmetric interaction is consistent with patterns observed after a single time step of sampling these eastern hemlock stands (initial year vs. 2005, Preisser et al. 2008) but differs from a fine-scale experiment in which HWA showed
reduced colonization on branches that were previously colonized by EHS, while EHS settlement was unaffected by previous HWA colonization (Miller-Pierce and Preisser 2012).

One interpretation of the result that HWA appeared to predispose stands to subsequent increase of EHS is that commensalism expanded the realized niche of EHS. The commensalism could have resulted from indirect interactions mediated by herbivore-induced changes in eastern hemlock primary and secondary metabolism. For instance, high HWA abundance could have facilitated EHS establishment and reproduction, as HWA infestation can increase foliar nitrogen levels (Soltis et al. 2015), an important factor determining EHS survival and fecundity (McClure 1980b). Another possibility is that HWA herbivory activates the salicylic acid (SA) defense pathway (Schaeffer et al. In Press), and thus compromises the ability of the host to activate the jasmonic acid (JA) defense pathway in response to subsequent EHS herbivory. Negative ‘cross talk’ in plant signaling pathways can inhibit plants from activating the JA pathway following induction of the SA pathway (Thaler et al. 2012), with downstream changes in metabolites and within-plant resource allocation that affect herbivores (Schweiger et al. 2014). Further research by Pezet et al. (2013) supports this interpretation – while HWA feeding (but not EHS) led to elevated methyl salicylate, EHS feeding more strongly increased green leaf volatiles. Green leaf volatiles can prime defenses and coordinate with the JA pathway to confer herbivore resistance (Christensen et al. 2013).

Commensalism could explain the long time period between EHS arrival and range expansion if EHS was unable to establish in new areas until HWA invasion made stands suitable for EHS infestation. An additional explanation is that EHS expanded northward more slowly because Allee effects had a stronger effect on EHS than on HWA (Taylor and Hastings 2005).

The sexual reproduction strategy of EHS likely required a greater number of individuals to
disperse to a site in order to overcome negative density-dependence at very small population size, slowing expansion. EHS may also be a poorer disperser than HWA. EHS and HWA have similar dispersal kernels in the absence of wind, but HWA crawlers are active earlier in spring when winds are strong and frequent (McClure 1989). Also, HWA produces 15 times more eggs per female than EHS (McClure 1989). These alternative explanations, however, cannot fully account for higher EHS abundance following a time step in which HWA abundance was higher.

It is important to highlight that although the joint species distribution model better described the ecology of this system, RPS indicated that the joint and independent models fit the data equally well. This result was expected because both models split the residual error into spatial and non-spatial correlation structures. The joint model captured dependence among species with model parameters, while the independent model captured that dependence as unexplained error that exhibited spatial correlation structure. The joint model better attributed variation in the abundance of each species to specific elements that were hypothesized to affect abundance a priori. Specifying a model that directly mapped to hypotheses about how the ecological system works was more informative than capturing those ecological processes with spatially-correlated errors that do not identify a specific process. However, the similarity of RPS between the two models adds to the evidence that when data are not available to fully specify a model containing all of the components hypothesized to strongly affect a system (which is often the case in ecological studies), accounting for spatial correlation of residual error can improve the robustness, fit, and predictive ability of species distribution models (Record et al. 2013).

This study demonstrates the benefits of accounting for biotic interactions with spatio-temporal joint species distribution models implemented in a multivariate generalized linear modeling framework. Accounting for spatial and temporal dependence among species improved
the precision of parameter estimates describing the abiotic niche for a species whose abundance
was highly dependent on interactions with another species in the community. Correctly
estimating the parameters that describe the abiotic niche of a species, and discovering whether
the distribution and abundance of a species is highly dependent on other species in the
community, are essential for tackling fundamental ecological questions, for making predictions
under climate change scenarios, and for conservation aims. Dynamic joint distribution models
such as the one presented here can help infer the underlying ecological processes that lead to
pattern and guide the design of future research.

Acknowledgements

NKL was supported by the Arnold and Mabel Beckman Foundation and Michigan State
University, and PLZ was supported by the USDA National Institute of Food and Agriculture,
Hatch project 1010055, and Michigan State University. We thank Sara Gómez and the many
graduate students, undergraduate researchers, and field technicians who have contributed to this
project over the years. This project was funded by the following grants: NSF DEB-0715504,
NSF DEB-1256769, NSF DEB-1256826, NIFA 2011-67013-30142 and is a contribution of the
Harvard Forest Long-Term Ecological Research Program (NSF DEB 06-20443).

Literature Cited

University of Chicago Press, Chicago.

43:413-423.


Orwig, D., N. Lany, and E. Preisser. 2017. Hemlock Woolly Adelgid and Elongate Hemlock Scale Surveys in Connecticut and Massachusetts 1997-2011. Environmental Data Initiative, [http://dx.doi.org/10.6073/pasta/55236414e515e94f5866d0b1e91475e0](http://dx.doi.org/10.6073/pasta/55236414e515e94f5866d0b1e91475e0).


Figure 1. Posterior mean of hemlock woolly adelgid (upper) and elongate hemlock scale (lower) latent abundance over time at 142 eastern hemlock stands located along a 165 km transect in Connecticut (CT) and Massachusetts (MA), USA.

Figure 2. Posterior distributions of model coefficients from joint vs. independent models of hemlock woolly adelgid (HWA) and elongate hemlock scale (EHS) abundance in Connecticut and Massachusetts, USA (1997-2011). Parameters describing the abiotic niche of each species ($\beta$) are shown in A). Although HWA abundance appeared independent of EHS abundance (the red and blue distributions were similar), including information on HWA abundance improved the precision of model parameters for EHS (red distributions were wider than blue distributions). Parameters describing temporal dependence are shown in B). Independent distribution models were specified by setting all parameters that describe dependence between species to zero. In C), the spatial extent of EHS effective range ($\varphi_{EHS}$) shrank considerably in the joint model that included HWA abundance. However, the effective range of HWA ($\varphi_{HWA}$) was similar in the independent vs. joint models.
Figure 1.