EXOTIC INVASIVE SPECIES DYNAMICS AND IMPACTS IN FOREST ECOSYSTEMS

by

Rachel T. Cook

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Forestry and Natural Resources

West Lafayette, Indiana

May 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Songlin Fei, Chair

Department of Forestry and Natural Resources

Dr. Jacob Hosen

Department of Forestry and Natural Resources

Dr. Andrew Liebhold

U.S. Forest Service, Northern Research Station

Approved by:

Dr. Robert Wagner

I would like to dedicate this Master's thesis to my wonderful support system and those who inspired me to pursue this research. First, my dad, for instilling in me a love and respect for forests and all living things. My mom, for endlessly guiding me in how to be self-sufficient. My partner Glenn, who provided constant emotional support as I grappled with data problems, analyses gone wrong, and the general stress of graduate school. My sister Marjorie for taking my video calls, night or day, and looking over my code while I talked out my confusion until we found a solution. My sister Dawn for always being available to provide advice, whether it be about life choices or her favorite new recipe. Thank you.

ACKNOWLEDGMENTS

We thank staff of USDA APHIS and Pennsylvania Department Agriculture for providing survey data and for providing feedback on this manuscript. The spotted lanternfly visual survey data used in or part of this publication was made possible, in part, by APHIS. This publication may not necessarily express the views or opinions of the APHIS. This research was supported by National Science Foundation Macrosystems Biology grant 1638702 to S.F. and A.L., and the USDA McIntire–Stennis program and USDA Forest Service support to S.F.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTARY TABLES

LIST OF SUPPLEMENTARY FIGURES

ABSTRACT

There are myriad of ecological questions to consider when assessing the spread and impact of non-native invasive forest pests. In my thesis, I investigated invasion dynamics in two ways. First, I used visual survey data collected from USDA Animal Plant and Health Inspection Service to analyze the historical changes in the range of spotted lanternfly *Lycorma delicatula* (White) (SLF) in the US to understand rates of spread and identify factors that influence risk of invasion. Second, I used historical gypsy moth outbreak data and water quality data to quantify the impact of gypsy moth (*Lymantria dispar* L.) defoliation on measures of water quantity and quality across the invaded range of the gypsy moth for the period 1999-2018. In the first study, I found there has been an upward trend in the number of counties newly invaded by SLF every year since initial discovery, with a consistent number of long-distance dispersals throughout the study period (2014-2019). Radial rates of spread, estimated using three methods, varied from 15.19 to 46.23 km/year. A Cox proportional hazards model identified two proxies for human-aided dispersal as significant drivers of SLF invasion. We anticipate that SLF will continue to spread, especially along human pathways of travel and transport. Efforts to manage SLF populations potentially could target these pathways to reduce rates of spread. In the second project, I found that overall, gypsy moth defoliation had a negative relationship with discharge, a slight positive relationship with water temperature, and a stream size-specific impact on dissolved oxygen. Expected differences in discharge with defoliation ranged from 16-25% reduction in discharge with varying degrees of defoliation. I conclude that observed changes in temperature and dissolved oxygen of between 1% and 6% with defoliation can have biologically meaningful impacts. The effects of defoliation observed across large regions are somewhat different from those observed in single watersheds, and offer important insights into the ecosystem impacts of gypsy moths. Taken together, these research projects provide increased understanding about the key driving factors contributing to forest pest spread as well as the impacts of forest pest dynamics on ecosystem services.

CHAPTER 1. INTRODUCTION

The number of established non-native invasive species in the US, including forest pests, is on the rise largely due to globalization and international trade (US Congress 1993, Hulme 2009), accumulating new exotic pests every year that are very unlikely to be eradicated. We know forest pests affect ecosystems in complex ways, and also have enormous associated economic costs. Invasions by non-native pests can pose significant ecological threats to forest ecosystems by reducing biodiversity and biomass, shifting forest composition, and destabilizing the community (Liebhold et al. 1995, Lovett et al. 2006, Fei et al. 2019) and have economic implications, costing over a billion dollars annually (Aukema et al. 2011). Despite implementation of biosecurity programs by national governments as well as by international conventions, an approximately constant number of damaging pests establish in the US every year, leading to a steady accumulation of non-native pests (Aukema et al. 2010). Given this growing problem and the limited number of options for its mitigation, information is needed on the spread and ecosystem impacts of these species.

Three stages are recognized that occur during the invasion of any non-native species: arrival in the new location (often referred to as introduction), establishment of reproducing populations, and spread into additional new areas (Sakai et al. 2001, Blackburn et al. 2011). Understanding how pests spread following their introduction and establishment is critical to the development of strategies for limiting further spread. By understanding what drives a species' spread, how it spreads, and the speed at which it spreads, we can prepare forest and land managers for new invasions. After a pest is established, quantifying the environmental impacts allows for prioritization of management efforts and allocation of resources. In this thesis, I assess spread and ecosystem impacts separately in two projects focused on different high-profile and damaging forest pests.

Spotted lanternfly, *Lycorma delicatula* (White) (Hemiptera: Fulgoridate), is a relatively new invasive pest in North America, and thus has a small but growing body of literature investigating its biology, behaviors, and spread. Spotted lanternfly (SLF) is the subject of a large amount of recent research as scientists quickly try to understand how far SLF will spread, where it is likely to spread, and how it will impact local ecosystems and economies. In contrast, gypsy

moth (*Lymantria dispar* L.) is a well-known pest that has been present in North America since 1869 with a large body of literature on a wide variety of topics.

Current knowledge about SLF biology includes information about its ability to feed on a wide range of host plant species and deposit eggs indiscriminately on a variety of surfaces (Parra et al. 2017, Urban 2019), allowing SLF to survive and reproduce in both forested and non-forested areas. There are no known sex or aggregation pheromones, and thus a species-specific trap is currently not available for use in SLF detection. Within the growing body of literature, there is still a knowledge gap in understanding how SLF has spread in North America and what factors influence its spread since first detection.

Though the biology and historical spread of gypsy moth are well understood, gypsy moth impacts on water quality is less well-studied. There have been small-scale studies on the effects of gypsy moth defoliation on water quality (Webb et al. 1995, Eshleman et al. 1998, Addy et al. 2018), many of which indicate negative impacts to water quantity, chemistry, and metabolism. However, there are few broad-scale studies that investigate the effect of gypsy moth defoliation over a wide time span and at the landscape level. In these projects, I aim to address these research gaps mentioned above in the field of invasive pest ecology using study systems of two well-known generalist pests, with a focus on spread dynamics and ecosystem impacts. The main objectives of the studies are to characterize and determine drivers of SLF spread and determine what impacts gypsy moth defoliation has on water quality.

CHAPTER 2. SPATIAL DYNAMICS OF SPOTTED LANTERNFLY, *LYCORMA DELICATULA*, INVASION OF THE NORTHEASTERN UNITED STATES

2.1 Introduction

Though most non-native pests fail to establish after arrival, those that successfully found reproducing populations can subsequently spread via a coupling of population growth with dispersal. The dispersal of many invading species is characterized by the simultaneous occurrence of local diffusion and occasional long-distance dispersal (Hastings 2005, Liebhold and Tobin 2008). Information on what factors drive spread of a non-native pest can guide management to contain its populations and reduce their impacts to ecosystems and economic costs (Sharov and Liebhold 1998, Liebhold and Kean 2019). Understanding the factors that drive spread is particularly important for newly established species, for which dispersal behaviors and population growth characteristics are often unknown.

Spotted lanternfly, *Lycorma delicatula* (White) (Hemiptera: Fulgoridae), is a non-native planthopper recently established in the United States. The species is native to southeast Asia, but recently invaded the USA in Berks County, Pennsylvania in 2014 (Barringer et al. 2015). Spotted lanternfly (SLF) is univoltine and lays egg masses on a variety of surfaces, including tree bark, stone, motor vehicles, and trains (Urban 2019). In addition to indiscriminate egg deposition, SLF also has a wide breadth of host use. This pest feeds on over 70 species of herbaceous and woody plants belonging to over 20 families, though it prefers tree of heaven (*Ailanthus altissima*), especially as a late instar (Dara et al. 2015, Parra et al. 2017). Notably, SLF feeds on apple (*Malus* spp.) and grape (*Vinus* spp.), both important agricultural plants in the Northeastern USA. Feeding on grape has reportedly resulted in lower fruit quality, less fruit production, and elevated mortality, though minimal impacts to fruit tree health have been reported (Urban 2019). The most conspicuous impact of SLF in forests is the accumulation of honeydew in the understory, which results in sooty mold growth that limits photosynthesis and growth of understory plants (Ding et al. 2006, Parra et al. 2017). There is also evidence that aggregation of SLF can cause weeping wounds on trees, resulting in crown dieback (Dara et al. 2015). While detrimental impacts on tree of heaven might be beneficial due to its status as an invasive plant, SLF is considered a serious pest due to its negative impacts to agricultural crops and native trees.

Despite regulations by the state of Pennsylvania that prohibit movement of any SLF living stage (e.g. egg masses, nymphs, adults) or material potentially harboring the pest (e.g. firewood, nursery stock, etc.) outside of a quarantine area, SLF has spread from Pennsylvania to seven surrounding states as of 2019 (**Figure 2-1**). Because SLF was detected in the USA recently, there is little information on how this species spreads or what drives its invasion. Though the body of knowledge on this insect is growing, many aspects of SLF spread, especially the role of environmental drivers, are unknown. Elucidating how this pest spreads can inform future management and survey efforts.

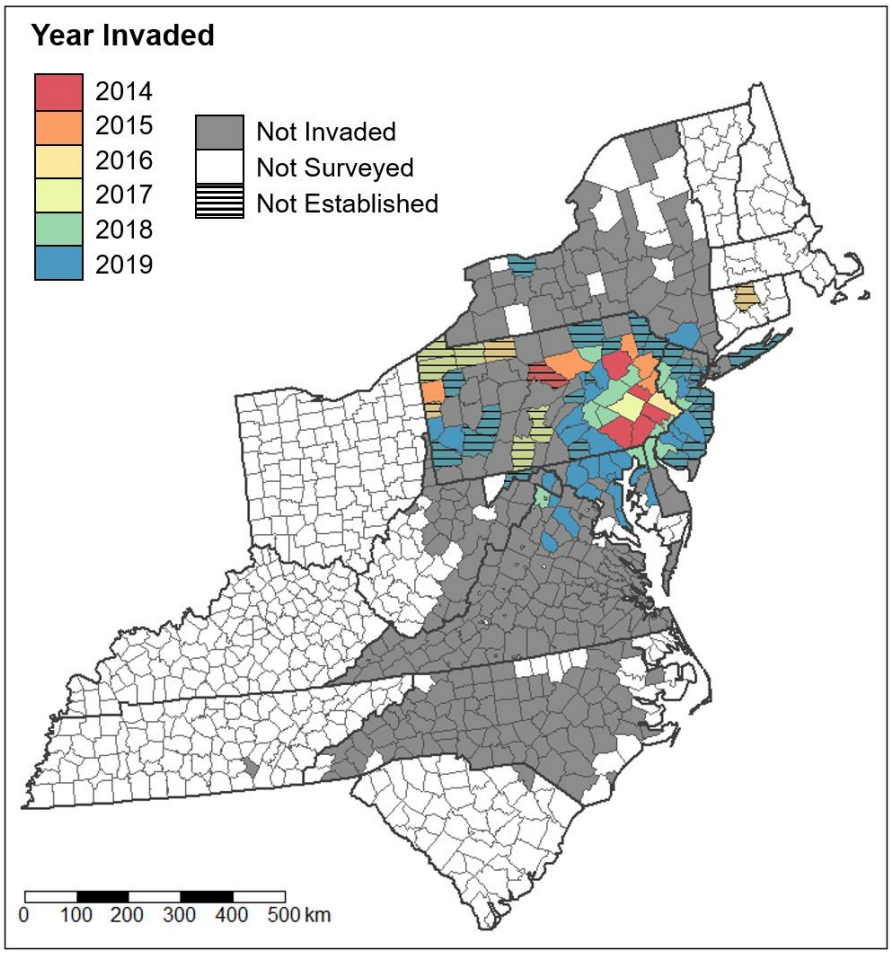

Figure 2-1. County-level distributions of spotted lanternfly (SLF) in the eastern USA. Distribution of SLF detections and establishments by year based on USDA Animal and Plant Health Inspection Service and Pennsylvania Department of Agriculture visual survey data. Counties with hash marks had SLF detections that failed to establish. We define a county as invaded when the county experiences at least two consecutive years of SLF detection. Counties with white color were not surveyed.

The ranges of introduced species are influenced by a multitude of anthropogenic factors and habitat features. For SLF, climatic niche models indicate that half of the USA, including most of the New England, Mid-Atlantic, and Pacific Coast states, is at risk of invasion (Wakie et al. 2020). While these climatic niche models provide valuable information on where SLF can potentially establish, analyses of spread can provide insight into how quickly SLF will arrive and what habitat and/or anthropogenic factors affect the dynamics of SLF spread. SLF has undergone several long-distance dispersal events that likely resulted from human-mediated transportation (Eddy 2018, Scheid 2020). Tree of heaven is also more abundant in urban areas, and thus human activities may increase both propagule pressure and habitat suitability. However, the rate of spread, including the frequency and distance of long-distance dispersal events, and drivers of spread have not been quantified. Therefore, we investigated how anthropogenic and habitat factors are related to SLF spread.

We analyzed the known geographical distribution of SLF (2014-2020) in the USA to quantify its rates of spread and identify factors that influence its invasion risk. Our goals were to: 1) describe the patterns of SLF spread following the initial detection in 2014, and 2) identify key drivers that are associated with SLF spread. For our second goal, we used known occurrences of SLF in conjunction with habitat and anthropogenic variables to determine the most important factors driving county-level invasion risk across the study area, defined below. We hypothesized that anthropogenic factors are important drivers of SLF spread, given the ability of this insect to lay inconspicuous eggs on a variety of materials, including motor vehicles and trains (Urban 2019).

2.2 Methods

The SLF distribution data analyzed in this study were derived from visual surveys conducted from 2014-2019 by the US Department of Agriculture, Animal and Plant Health Inspection Service (APHIS) and the Pennsylvania Department of Agriculture (PDA). We also used SLF county-level presence data for 2020 from the New York State Integrated Pest Management Program (Cornell 2021). Survey data include geospatial coordinates for survey locations as well as the number of SLF observed (if any). A total of 241,366 survey locations were obtained for this study (**Figure S2-1**). Given irregularity of survey locations and potential biases (e.g. surveys at expected SLF locations) and to render data at an equivalent scale as the 2020 presence data, we converted counts to county-level presence/absence records and used county as our unit of analysis.

The survey data contained many points that we identified as failed establishments in which SLF were observed in a county in a given year but were absent in surveys of the same county in subsequent years. These detections were likely either populations that failed to establish or regulatory incidents, such as dead SLF adults found in transported materials, and thus we did not treat them as invasions. Hereafter, we refer to detections as establishments plus failed establishments and establishments as only populations that persisted for more than one survey year in consecutive years within a county. Moreover, we categorized each invaded county in year *n* as contiguous or non-noncontiguous based on the presence or absence, respectively, of an invaded neighboring county in year *n*-1.

Described below are methods we used to 1) determine aspects of spread dynamics, such as jump distances and spread events into contiguous vs. non-contiguous counties, 2) compare three methods of estimating spread rates, and 3) fit a Cox proportional hazards model estimating time-to-invasion as a function of variables representing spatial proximity to existing SLF populations (henceforth referred to as spatial proximity), habitat suitability, and anthropogenic influences. Our study area was defined as the area of the eastern USA invaded in 2019 plus a buffer distance of 355 km, equal to the maximum observed jump distance (see "Characterization of spread events" in Methods). This study area was used for all subsequent analyses. Counties, which are the level at which quarantines and other management decisions are set, served as the unit of analysis for all analyses. All analyses were conducted using R version 4.0.2 (R Core Team 2020).

2.2.1 Characterization of spread events

To characterize spread, we quantified the number of yearly spread events into contiguous and non-contiguous counties, as well as the distribution of jump distances. Jump distance is defined as the distance between establishments or detections in non-contiguous counties in year *n* and the nearest previously invaded county in year *n*-1. We estimated jump distances for every newly invaded county by calculating the distance to the closest previously invaded county, as assuming new SLF establishments originate from the closest previously invaded county provides a conservative estimate. Distances were measured using county centroids. We repeated this process for each year, and summarized the distribution of jump distances (e.g. median, minimum, maximum). To determine if spatial proximity is related to whether or not a detection became an establishment (i.e. an invasion persisted), we separated jump distances by establishments and

failed establishments and used a Mann-Whitney U test to compare the distribution of jump distances between these two groups.

2.2.2 Spread rates

Because little is currently known about SLF spread patterns and different approaches can provide variable estimates of annual spread (Tobin et al. 2006), we compared three methods to calculate spread rate described by Gilbert and Liebhold (2010). The purpose of our comparison of these methods is to provide a range of possible spread rates as well as to determine robustness of each when applied to an insect at early stages of invasion.

The first method is to apply regression of the distance (centroid to centroid) of every county with positive establishment from the point of initial detection (Berks County, PA) as a function of years since initial detection (2014). The resulting slope of the estimated regression equation estimates the radial rate of spread measured in distance/year. The second method is to regress the square root of the invaded area (estimated by summing the area of invaded counties in each year) divided by π on time. The resulting slope of the estimated regression line estimates the radial spread rate in distance/year (e.g. effective range radius; Shigesada et al. 1995). Last, we calculated the average distance between invasion boundaries in consecutive years along radii emanating every 0.5 degrees from the centroid of Berks County, PA. We used radii at a frequent degree interval to obtain a high-resolution estimate of yearly distance between boundaries. We found invasion boundaries by fitting a convex hull polygon to the area of invasion in each year and subsequently converting the polygon edges to lines. The convex hull polygon in each year was stretched to the edges of non-contiguous invaded counties. The resulting average distance between boundaries on each radius between consecutive years can be used to estimate the annual radial spread rate (e.g. boundary displacement rate). Due to the nature of fitting a convex hull polygon around invaded counties, we used county boundaries as opposed to county centroids to calculate distances in boundary displacement estimations. In summary, distance regression is based on distance and year of sampling points from the origin where the species was first detected, while effective range radius considers area invaded over time. Boundary displacement estimates distance between invasion boundaries in consecutive years.

2.2.3 Dispersal kernel estimation

Dispersal kernels estimating risk of invasion as a function of distance have been developed for other invading forest insects (Orlova-Bienkowskaja and Bienkowski 2018). Given interspecific variation in spread rates (Fahrner and Aukema 2018), however, we estimated a SLF specific dispersal kernel, which, in turn, should enable more reliable estimates of the effects of SLF spatial proximity on invasion risk. Our analysis used 2015-2019 county-level SLF survey data from USDA APHIS and PDA and follows methods from Kovacs et al. (2010). A negative exponential function was used to model the probability, *p*, of each non-invaded county in the study area becoming invaded on an annual basis from 2015-2019:

$$p_{i,j} = e^{-\alpha d} \qquad (1)$$

where α is the parameter we sought to estimate and *d* is the distance in kilometers to a previously invaded county. To estimate α, we simulated county-level spread starting from the five initially invaded counties in 2014 using values of α between 0.01 and 0.10 in 0.001 intervals.

To simulate spread for a given α value, we calculated the distance from each non-invaded county *i* in year *n* to each invaded county *j* as of year *n*-1, as each county *j* invaded as of year *n*-1 could serve as a source for invasion into county *i* in year *n*. The distances from county *i* to each invaded county *j* were input into Equation 1, producing an estimate, *p*, for the probability of SLF invading from each county *j*. This probability value was then used to parameterize a Bernoulli distribution such that the probability of an event was equal to *p*. We then took a random draw from that Bernoulli distribution in which a draw of 1 or 0 would indicate invasion or non-invasion, respectively. This meant that there were *x* draws for each non-invaded county *i*, where *x*=number of invaded counties in year *n*-1. If any draw produced a 1, the county was categorized as invaded for the rest of the simulation (i.e. counties could not become uninvaded).

A single iteration of this process produced a simulated, county-level invasion at annual time steps (2015-2019) that may or may not have reflected the realized invasion. For each α value, we conducted 500 iterative simulations, starting with the initially invaded counties in 2014 and forecasting spread to 2019. Results were summarized with accuracy values - false negatives and positives, and true negatives and positives - compared with the actual invasion data from 2015-

2019. We selected the value of α that simultaneously resulted in the lowest number of false negatives and false positives when comparing actual spread to predicted spread.

2.2.4 Invasion drivers

Cox proportional hazards models can be used to estimate survival time based on predictor variables, including both static and time-varying predictors (Thomas and Reyes 2014). If we equate survival to a county persisting without invasion, we can use Cox proportional hazards models to evaluate which factors explain variation in time-to-invasion. Therefore, we used a Cox proportional hazards time-to-invasion model to evaluate potential drivers of SLF invasion at the county level, in a manner analogous to the implementation by Jules et al. (2002) and Ward et al. (2020).

The Cox proportional hazards model quantifies the probability of invasion at each one-year time step. Time steps ranged from 2014-2015 to 2018-2019. Predictor variables included static habitat variables (**Table S2-1**) and one time-varying predictor, spatial proximity. To quantify spatial proximity, we first used Equation 1, setting $\alpha = 0.045$ (i.e. determined from the dispersal kernel estimation process described above; see "Invasion drivers" in Results) and d as the distance in kilometers between each uninvaded county to all previously invaded counties. Spatial proximity, denoted *SpatialProx*, was then calculated for each county:

$$SpatialProx_i = 1 - \Pi(1 - p_{i,j}). \quad (2)$$

The other predictors included two anthropogenic variables and six habitat variables. The anthropogenic variables were human population from the U.S. Census and road density calculated by Liebhold et al. 2013 from the ArcGIS World Transportation reference layer (**Figure S2-2**) and each was considered a proxy for human-aided dispersal. The six habitat variables included forested area and five host availability terms expressed as basal area, host trees per acre, number of host trees per county, tree of heaven occurrence, and canopy cover (**Figure S2-3**). Forested land was obtained from the US Forest Service FIA MapMaker online data query system (https://www.nrs.fs.fed.us/fia/data-tools/mapping-tools). Percent forest canopy cover was obtained from the Forest Service's cartographic tree canopy cover product (USDA Forest Service 2016).

Host basal area and numbers of host trees per acre and county were obtained from the Forest Service's Forest Inventory and Analysis (FIA) program, using a published list of known SLF hosts from Barringer and Ciafre (2020). The FIA program is a long-term forest inventory program with one 0.40 hectare sample every 2,428 hectares, with most counties partially assessed annually since 2000. FIA assesses forest areas defined as at least 37 meters wide and 0.40 hectares in size, covered by at least 10% trees (Bechtold and Patterson 2005). We obtained plot-level basal area and stem density per acre from FIA records from 2015-2017. To estimate these variables at the county-level, we aggregated each by the summed county plot area for every known host with available FIA data by species code, obtained from the National Core Field Guild (USDA Forest Service 2019). We then estimated the number of each host species in a county by multiplying the estimated number of trees per acre by the total acres of forested land in each county. Because FIA only surveys forested areas, and tree of heaven is often found in developed or urbanized areas, number of tree of heaven observations were downloaded separately as point data from EDDMapS (EDDMapS 2021) and aggregated to the county level by summing the number of observations per county.

Prior to model development, we quantified pairwise correlations between our predictors to check for collinearity (defined as Pearson's product moment correlation coefficient ≥ 0.70). Based on this step, we removed road density and number of host trees per county due to collinearity with human population and forested area, respectively. We removed these two variables as opposed to human population and forested area because in preliminary models, they were more strongly associated (i.e. occurred in models with lower Akaike Information Criterion values) with SLF time-to-invasion than their co-varying counterparts. We then refined the model by applying a backward selection procedure that iteratively removed the variable associated with the highest *p*-value and refitting the model until only statistically significant predictors remained.

2.3 Results

2.3.1 Characterization of spread events

There was overall an upward trend in the number of newly invaded counties every year since initial discovery, although some counties contained failed establishments. There was drop in counties with establishment in 2016 and 2017, while the highest number of establishments was

observed in 2019 (**Table 2-1**). Similar to the number of newly invaded counties per year, number of counties with failed establishments generally increased across the study period and peaked in 2019. The highest percent of counties with failed establishment occurred in 2017, with 83% of detections failing to establish. The median yearly jump length into counties with detection and establishment ranged from 46 to 73 km and 50 to 92 km, respectively.

We did not find a significant difference between distributions of jump distances in established populations vs. failed establishments. Median jump distances across all years in failed establishments and established populations were 55 km and 71 km, respectively. A Mann-Whitney U test showed the distributions in the two groups did not significantly differ (W = 706, $p = 0.46$).

Table 2-1. Spread events summary. Number of observed contiguous (having at least one previously invaded neighboring county at time of invasion) and non-contiguous (having no previously invaded neighboring counties at time of invasion) newly invaded counties per year and median jump distances between invaded and uninvaded counties between consecutive years for both the non-persistent and the persistent counties.

	2014	**2015**	**2016**	**2017**	**2018**	**2019**	**Total**
Counties with detections (n)	5	6	4	6	18	47	86
Counties with establishment (n)	5	5	1	1	15	27	54
Counties with failed establishments (n)	0	1	3	5	3	20	32
% of Counties with failed establishments	-	16.7	75.0	83.3	16.7	42.6	-
Median jump length (km) into counties with detection	-	137.4	100.5	79.6	104.5	46.5	-
Median jump length (km) into counties with establishment	-	54.5	49.9	69.5	91.7	57.8	-
Counties with contiguous invasion	5	3	1	1	12	13	35
Counties with non-contiguous invasion	-	2	0	0	3	14	19

The SLF invasion began in eastern Pennsylvania, and many of the counties invaded in the surrounding area of eastern and central Pennsylvania were contiguous with previously invaded counties (**Figure 2-2).** In contrast, several populations in western Pennsylvania and northern Virginia resulted from invasion into non-contiguous counties, indicating long-distance jumps. There were no newly invaded non-contiguous counties in 2016 or 2017. Trends in the number of both contiguous and non-contiguous counties tracked the overall number of counties invaded, starting out low and increasing in 2018 and again in 2019 (**Table 2-1**). However, there were overall fewer non-contiguous counties invaded than contiguous counties across all years.

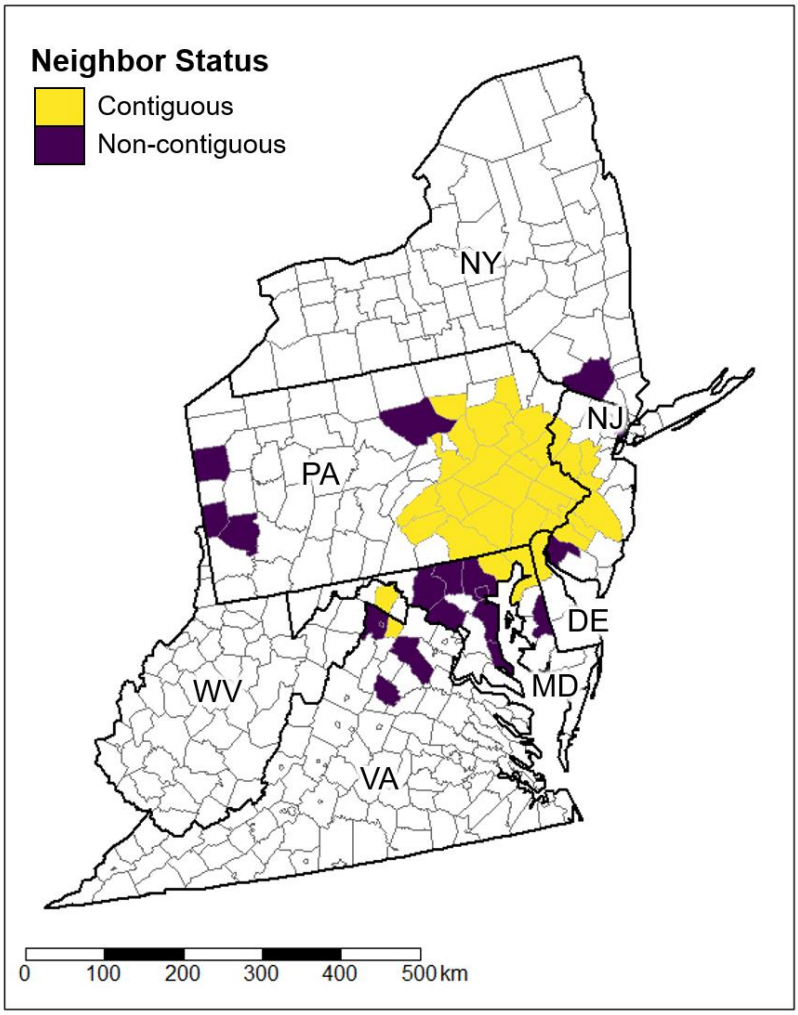

Figure 2-2. Contiguous and non-contiguous establishments of spotted lanternfly. Spatial distribution of contiguous (having at least one previously invaded neighboring county at time of invasion) and non-contiguous (having no previously invaded neighboring counties at time of invasion) counties across the study area.

Establishments showed similar patterns in numbers of new counties invaded and jump distances by year (**Table 2-1**), with lower values from 2014 to 2017 and an increase in 2018 and 2019. Median jump distances were greatest in 2017-2018 in the established counties, and were greatest in 2014-2015 in counties with detections. The year with the highest number of counties with newly established populations was 2019, whereas the year with the largest median jump distance (92 km) was 2018. Median jump distances were generally higher into counties with detections than counties with establishments. The overall maximum jump distance was 355 km into Mercer County in northwest Pennsylvania (**Figure 2-3**), while the median jump distance was 55 km for detections and 71 km for establishments.

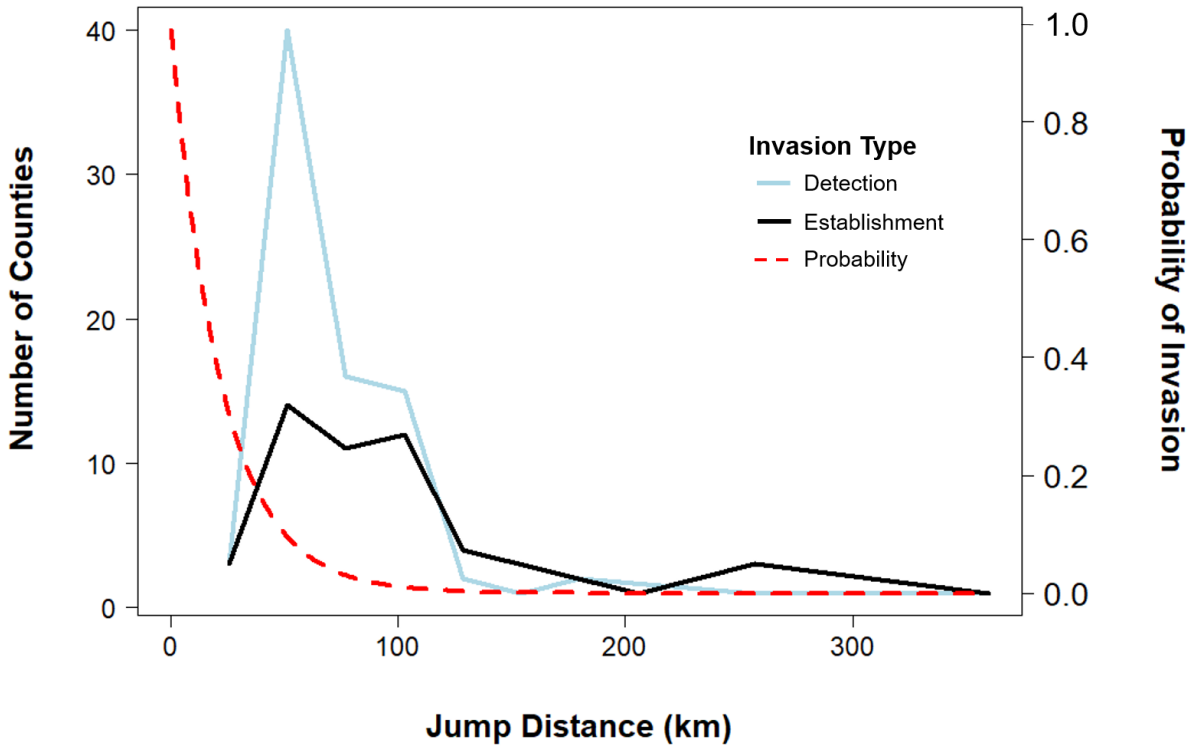

Figure 2-3. Jump distance distributions and probability of invasion by spotted lanternfly (SLF). Line graph of observed jump distances (the distance between new establishments in year *n* and the nearest previously invaded county in year *n*-1) for every newly invaded county for both establishments (black) and detections (blue). The red line indicates the probability of invasion by distance, based on the estimated SLF-specific negative exponential kernel function $p_{ij} = e^{-0.045d}$.

2.3.2 Spread rates

Estimated spread rates varied from 15-46 km per year among our three methods. Spread rate estimated by effective range radius was 46.2 km/year (SE = 7.19 km, 95% CI 26.26-66.20; **Figure 2-4A)**. Spread rate was estimated at 15.2 km/year (SE = 6.40 km, 95% CI 2.35-28.03) using distance regression (**Figure 2-4B**). Spread rate estimated by average boundary displacement (averaged over all years) was 38.6 km/year (range 0 to 75 km; **Figure 2-4C**), which was approximately 10 km less than estimated by effective range radius. The median boundary displacement across all years was 20.8 km/year. There was no difference in invaded area boundaries between 2016 and 2017, because the only newly invaded county was within the existing invasion boundary.

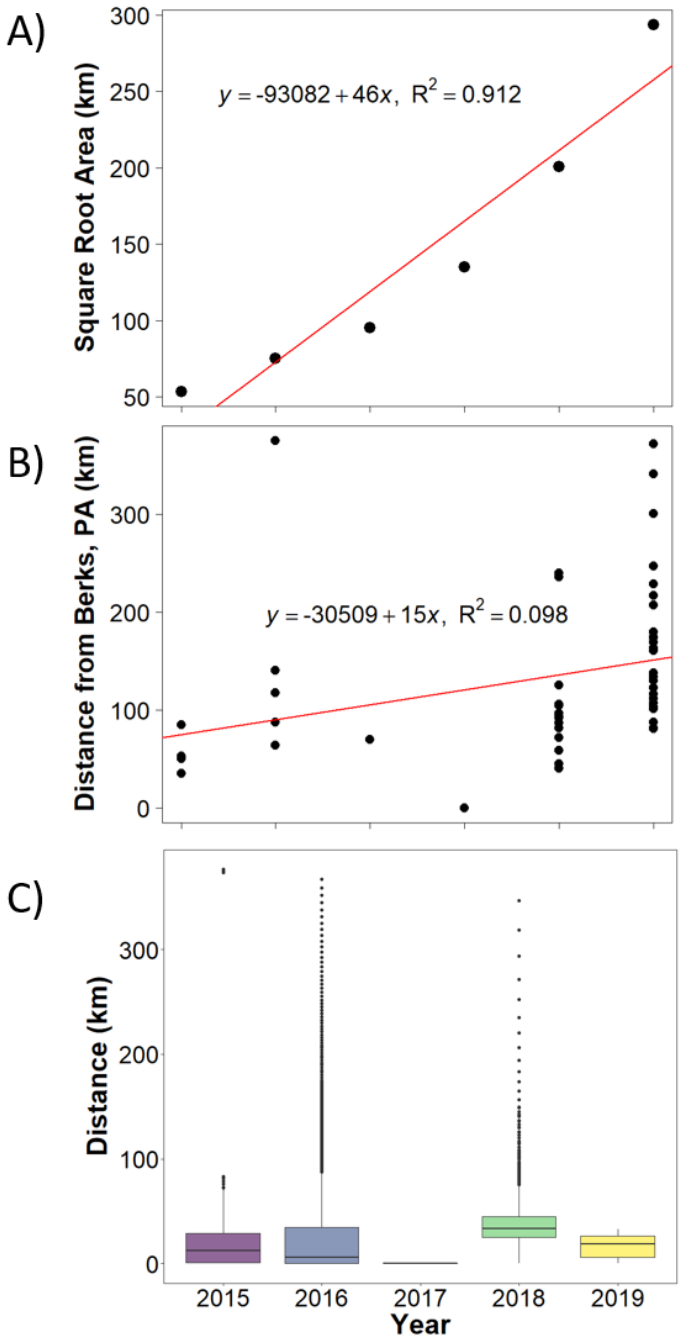

Figure 2-4. Estimated radial spread rates of spotted lanternfly (SLF). A. Plot of the square root cumulative county area containing SLF establishments divided by π by year of establishment. The slope of the regression is estimated at 46 km, providing an estimate of radial spread. **B.** Plot of distance from the centroid of the county with the first SLF detection point (Berks County, PA) by year of establishment. The slope of the regression is estimated at 15 km, providing an estimate of radial spread rate. **C.** Boxplots of boundary displacement distances between years of establishment, with average across all years of 38 km and median across all years of 21 km.

2.3.3 Invasion drivers

The best fitting value of α in the exponential dispersal kernel (*Equation* 1) was 0.045, which simultaneously resulted in the lowest number of false negatives and false positives. We used this value to estimate spatial proximity in the Cox proportional hazards model.

In the final Cox proportional hazards model, the hazards ratios for both spatial proximity and human population were greater than 1, indicating a positive relationship with increased risk of invasion (**Table 2-2**). Spatial proximity was identified as the strongest predictor (i.e. highest *Z*-value) with a notably high hazards ratio of ~40, followed by human population. No other covariates were statistically significant.

Table 2-2. Final Cox proportional hazards model summary. Summary statistics from final Cox proportional hazards model predicting time-to-invasion of SLF at the county level in the study area.

Predictor	Estimate (coefficient)	SE	Z	*p*-value	Hazards ratio (95% CI)
SLF spatial proximity	3.70	0.286	12.94	<0.0001	40.29 (23.01-70.54)
Human population	0.28	0.126	2.22	0.0265	1.32 (1.03-1.69)

2.4 Discussion

Spread of invasive species is often characterized by both short- and long-distance dispersal. In many systems, short-distance dispersal is caused by the natural movement of organisms (e.g. flight behavior) while long-distance dispersal is caused by accidental human movement (Hastings et al. 2005). Even small amounts of long-distance dispersal can result in greatly elevated rates of spread (Shigesada et al. 1995). So far in the SLF invasion, movement appears to consist of both short- and long-distance dispersal. Little is known about natural dispersal in this species. Our results indicate, however, that risk of long-distance movement increases with human population density, likely reflecting the propensity of SLF to become associated with objects transported by humans, such as when SLFs oviposit onto train cars and motor vehicles (Urban 2019).

A higher number of new establishments occurred in contiguous than in non-contiguous counties, but several long-distance jumps were observed and the frequency of jumps appears to be

increasing (**Table 2-1**). Human-mediated long-distance dispersal events are responsible for spread outside of the center of invasion, allowing for invasion of a larger geographic area than would be possible via insect movement alone. For example, the established population in northern Virginia (Frederick County) is believed to have originated from shipments from a stone yard in Pennsylvania (Eddy 2018). As SLF spreads, there may be increases in both long- and short-distance movement due to increases in numbers of source populations or increases in population size. A Mann Whitney U test showed no significant difference between jump distance distributions in detected but non-established vs. established populations, indicating that jump dispersal events are not necessarily more likely to persist if they are closer to the point of establishment (i.e. have closer spatial proximity). Ranges of jump distances were visually similar in range for both detected and established populations (**Figure 2-3**), signifying that established jumps went at least as far as jumps that failed to establish. Shigesada et al. (1995) demonstrated that such long-distance dispersal events typically result in faster rates of spread as well as accelerating patterns of radial spread. SLF spread rates could increase in this way, and we observed the largest increases in radial spread in the last two years of the study period (2018 and 2019), potentially indicating accelerating spread.

Our estimates of spread rate varied between methods, with the effective range radius method estimating the highest spread rate. The large differences observed between these methods may reflect the discontinuous nature of SLF spread. Measurement of the radial rate of spread of invading organisms was originally envisioned for continuous range expansion (e.g. Skellam 1951) and may not fully capture discontinuous spread such as observed here, which is also reflected in the low variance explained by distance regression spread estimation (r^2 =0.098) (**Figure 2-4B**). The effective range radius approach may provide a more representative measure of spread in this situation as it accounts for both the frequent long-distance dispersals and subsequent spread into the counties between contiguous and non-contiguous counties. For example, a long-distance dispersal event established a SLF population in northern Virginia in 2018, and in 2019, SLF spread to several counties between the eastern Pennsylvania invasion area and the new area in northern Virginia (**Figure 2-1**). The effective range radius approach accounts for the cumulative invaded area as these counties are occupied in subsequent years, whereas boundary displacement does not include those counties in estimates of radial spread. That is, counties closer to the previously

invaded area following a long-distance jump are enclosed by the convex hull polygon and do not influence future boundary displacements as they become invaded.

Based on the findings presented here, we estimate radial spread rate at around 40 km/year based on the average of the two more reliable methods (i.e. effective range radius and boundary displacement). If SLF were allowed to spread without any intervention, spread might be much higher given considerable management efforts are currently targeted to suppress SLF populations and limit movement. For example, active management programs conducted by USDA APHIS include egg scraping, sanitation (i.e. host tree removal) around SLF detections, and insecticide application to tree of heaven (USDA-APHIS 2018). Additionally, the State of Pennsylvania's quarantine on movement of goods out of the invaded area is implemented to limit spread of SLF. It is also important to note SLF is in the early stages of invasion, and the spread rate may increase as this pest continues to colonize new locations in the USA.

Results of the Cox proportional hazards model indicated that anthropogenic factors, specifically human density, are stronger drivers of SLF spread than forested area or availability of host trees. The role of humans in facilitating spread of invading organisms is a common phenomenon. Known international and domestic pathways of human-mediated spread of tree pests include transportation of pests on live plants (Liebhold 2012) and wood products (e.g. packing materials or movement of firewood) (Yemshanov et al. 2012), though pests can also be transported on non-host materials, such as on stone imports as with SLF. Domestic pathways of human-mediated spread include movement of firewood, transportation via vehicles (e.g. trains, motor vehicles), and "hitchhiking" on travel gear (e.g. hiking gear) and/or pets. Given that SLF lays eggs indiscriminately, human-mediated spread is not limited to host materials. Humans could facilitate the spread of this pest via travel (e.g. automobiles, trains) and movement of both host and non-host materials from an invaded area. Gilbert et al. (2004) came to similar conclusions in their analyses of the horse chestnut leafminer *Cameraria ohridella* Deschka & Dimic (Lepidoptera, Gracillariidae), finding that geographical variation in human population density explained most of the variation in historical spread of this species. Similarly, in an analysis of 79 damaging forest pests, Liebhold et al. (2013) found human population density associated with both spatial proximity and number of invasive forest pests per county across the USA. However, with all such analyses of historical spread, there is always some possibility that statistical associations may be caused in part by more intensive survey and reporting in more populated areas.

The invasion of tree of heaven in the eastern USA more than 200 years prior to the arrival of SLF may have facilitated the insect's initial establishment, causing an "invasional meltdown" (Feret 1985, Simberloff and Von Holle 1999) in which invasion of one species facilitates the invasion of another. Tree of heaven is the preferred host for SLF and SLF fitness (survival and fecundity) is maximized when feeding on tree of heaven, but this pest can survive and reproduce without access to tree of heaven (Uyi et al. 2020). In addition, SLF's ability to feed on a wide breadth of plant species (more than 70 species) increases likelihood of the insect encountering a suitable host following dispersal (Dara et al. 2015). The final Cox proportional hazards model did not include a significant effect of tree of heaven abundance on SLF spread, and therefore we found no evidence that this tree species has influenced SLF spread. Surveys for tree of heaven were conducted by many different people including volunteers and residents, and so it is possible the data are biased or incomplete despite verification by EDDMapS reviewers. In addition, as SLF invasion progresses, additional relationships to host trees or other environmental variables may become apparent or the importance of such variables may vary geographically.

Spatial proximity will remain an important predictor in the future spread of this pest, rendering estimation of SLF populations an important step in assessing spread. Current challenges in estimating SLF populations are primarily lack of long-term, systematic population assessment data and difficulties detecting small populations. The SLF-specific dispersal kernel we estimated here provided the best estimates of spatial proximity based on available distribution data but it was limited by the coarse spatial scale of county-level data and the limited temporal replication. We anticipate that as more data are collected on SLF populations, the estimated dispersal kernel could be refined and thus enhance model predictions.

There are a few limitations involved in our study. First, the data used in these analyses consisted of visual surveys that were located based on perceived risk of SLF establishment. These data were not collected in a systematic fashion, and thus there is potential for sampling bias. Though work is underway on developing traps to efficiently survey for SLF (Francese et al. 2020), a sensitive SLF-specific trapping system has not yet been widely implemented. The lack of a pest-specific trap increases risk of missed detections in visual assessments, especially for low population densities. Given these potential biases, we used counties as the unit of analysis, and the estimates of spread rate as well as drivers of local spread at a finer resolution may be different. We also assumed counties with only a single year of SLF detection indicated populations that failed to

establish and thus were not detected in future surveys. Failure to establish could be the result of stochastic dynamics or Allee effects, both of which can drive low-density, newly invaded populations to extinction (Liebhold and Tobin 2008). For example, Liebhold and Bascompte (2003) concluded that low density gypsy moth *Lymantria dispar* (L.) populations are likely to reach extinction without intervention, and in their analysis, most of the populations that did go extinct without treatment did so within a year of detection. Where management efforts are in place, failure to establish could also be the result of local eradication efforts. However, there is also a possibility that low-density populations did indeed persist, but due to difficulties in detecting this pest without specific lures or traps, small populations went undetected.

Focusing efforts on assessing populations and on estimating spatial proximity is important in describing and predicting spread of non-native pests. Our findings indicate that SLF has spread from 2014-2020 primarily through local diffusion with less frequent but consistent long-distance dispersal from previously established populations with influence from human populations. Based on the results presented here, we anticipate that SLF will continue to spread in the USA, though management and eradication efforts may effectively reduce population densities, reproductive potential, and ultimately rate of spread. Additional monitoring efforts to prevent and detect long-distance dispersals may prove useful, especially regarding transports of materials from areas with existing SLF populations.

2.5 Supplementary Materials

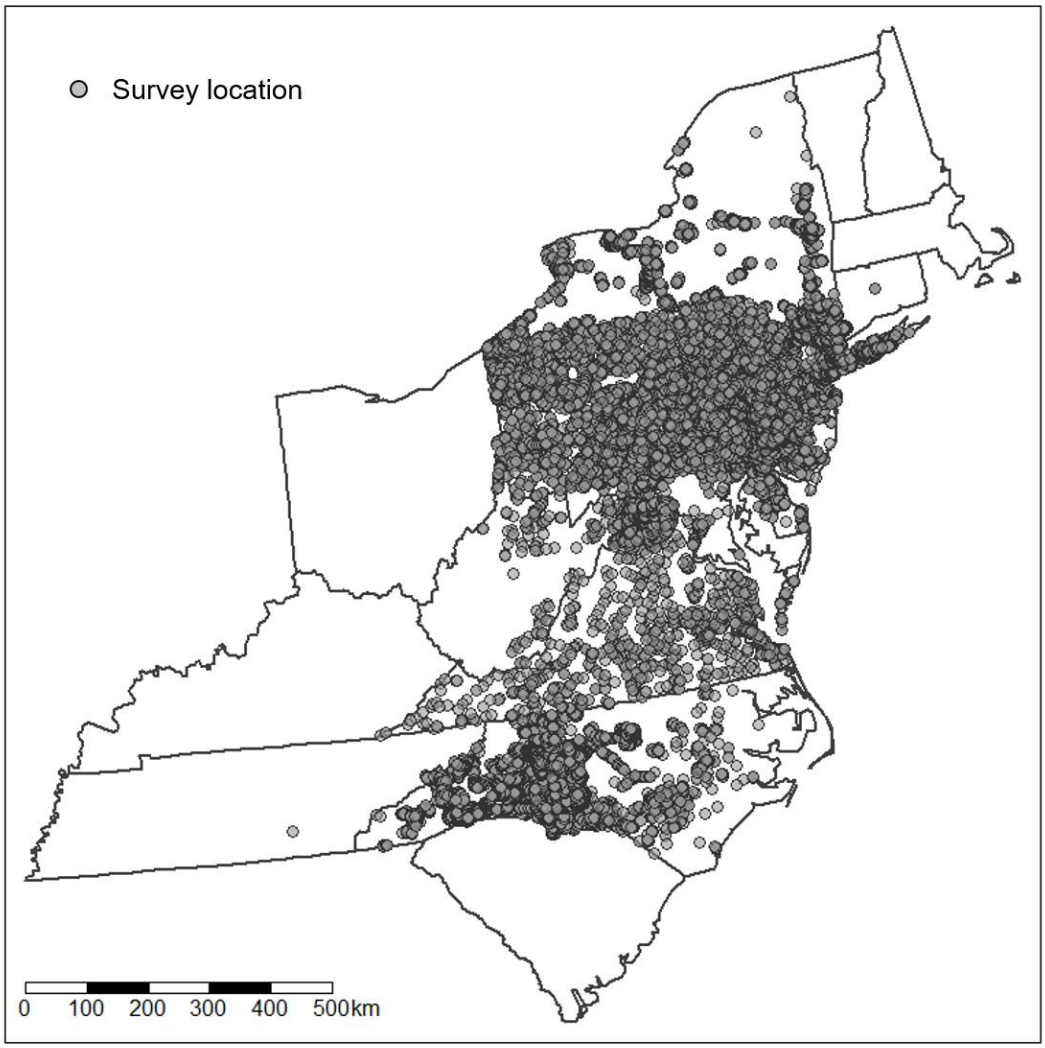

Figure S2-1. Distribution of survey locations. Locations of SLF visual surveys conducted by the US Animal and Plant Health Inspection Service and Pennsylvania Department of Agriculture.

Table S2-1. Predictor variable summary. Summary statistics of predictor variables used in Cox proportional hazards model development.

Variable	**Mean**	**Standard Deviation**
a) Habitat Variables		
Host Basal Area/Acre	14.50	4.92
Host Trees per Acre	31.84	10.07
Estimated Number of Host Trees	614.84	598.03
Tree of Heaven Observation Count	44.61	300.74
Forested Land (Acres)	17.57	16.61
Canopy Cover (%)	45.38	19.74
b) Anthropogenic Variables		
Human Population (2000)	151207	279291
Road Density	29.97	37.20

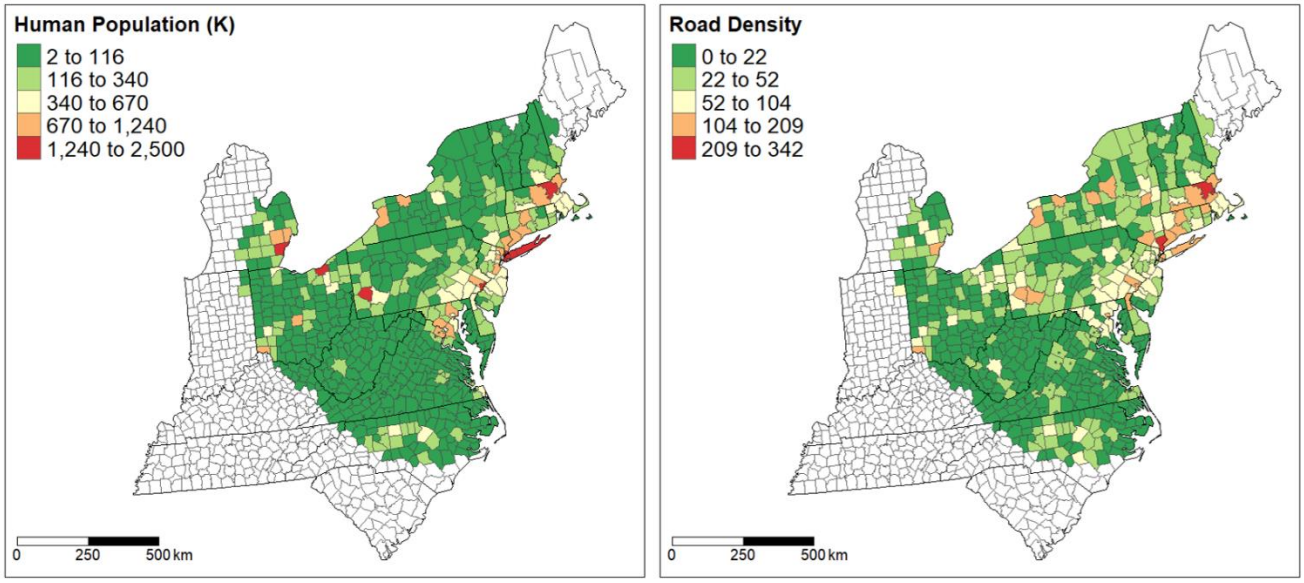

Figure S2-2. Anthropogenic predictor variable distributions. Distributions of anthropogenic predictor variables used in Cox proportional hazards model development.

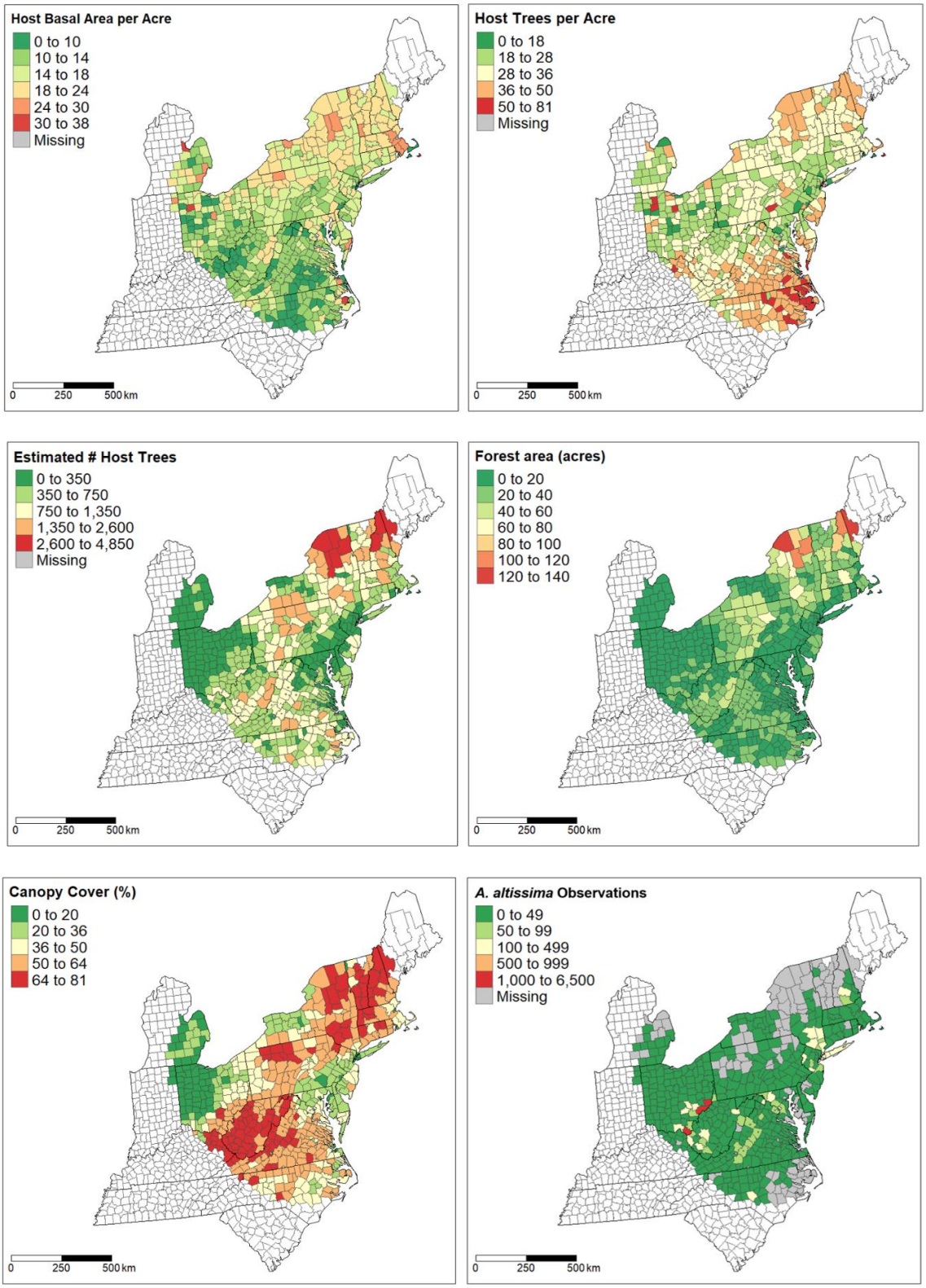

Figure S2-3. Habitat predictor variable distributions. Distributions of habitat predictor variables used in Cox proportional hazards model development.

CHAPTER 3. REGIONAL IMPACTS OF GYSPY MOTH DEFOLIATION ON WATER QUANTITY AND QUALITY

3.1 Introduction

Non-native pests can significantly alter ecosystem functions by reducing biodiversity and biomass, shifting forest composition, and destabilizing communities (Liebhold et al. 1995, Lovett et al. 2006, Fei et al. 2019). Several of the most important ecosystem services provided by forests are collection, filtration, and delivery of water into surface waters. These ecosystem services will become increasingly important with climate change, as predicted changes to air temperature, precipitation, and extreme weather events alter water availability and water quality (Arnell 1999, Danladi Bello et al. 2017). Despite the importance of both non-native pests and water resources, no broad-scale quantification of the impact of non-native species on water resources has been conducted.

Defoliation by non-native forest pests may have impacts on surface water similar to those caused by other landscape disturbances, such as clearcutting (Scoles et al. 1996), which can result in water redistribution, changing availability of water, and altered water chemistry (Schwarze and Beudert 2009, Schafer et al. 2014). Gypsy moth defoliation is known to change the amount and timing of water that is released from forested watersheds, as well as the concentrations of nutrients and other water quality indicators. Redistribution and changing availability and condition of water resources not only impact the forest in which forest pests reside, but downstream waters as well. The amount, timing, and condition of surface water resources play important roles in supporting ecosystems of both aquatic and terrestrial life (Ruegg et al. 2020), as well as providing important services to humans, such as drinking water, irrigation, disposal of wastewater, etc. This is particularly important because both invasions by non-native species and water availability are likely to be impacted by environmental changes resulting from climate change (Hellmann et al. 2008, Kreuzwieser and Gessler 2010).

Gypsy moth (*Lymantria dispar* L.) is a particularly well-known non-native pest with a long history in North America. Gypsy moth was accidentally introduced to the US in 1869 in Massachusetts by Etienne Leopold Trouvelot, when moth larvae he was cultivating for study of silk production escaped, and has since spread and caused considerable damage (Liebhold et al. 1989). Gypsy moth outbreaks can cause mass defoliation and, in some instances, mortality to a

wide breadth of tree species. Tree mortality is most common in forest stands with repeated defoliation over two or more consecutive years, even for healthy trees (Campbell and Sloan 1977). Gypsy moth larvae have a preference for oaks (*Quercus* spp.) and aspens (*Populus* spp.) in general, but when populations are high and food is limited, larvae will consume a large variety of foliage of other species (Lovett et al. 2006).

There are studies on the effects of gypsy moth defoliation on water quality and quantity at the watershed or ecosystem scale, typically over one or several years. Observed impacts include increased stream temperature (Collins 1961, Addy et al. 2018), lower dissolved oxygen (DO) but wider diel cycles (Addy et al. 2018), higher nitrate concentrations (Webb et al. 1995, Eshleman et al. 1998), increased stream metabolism (Addy et al. 2018), and higher streamflow (Corbett and Lynch 1987, Smith-Tripp et al. in review). While the impacts on water quality have been demonstrated at a small scale, they have infrequently been quantified and generalized at the landscape level.

The goal of this research is to detect and quantify patterns over decades at a broad spatial scale to synthesize major impacts of gypsy moth defoliation on water resources. The questions driving this research are: 1) Are there differences in water quality, temperature, and oxygenation between watersheds with and without gypsy moth defoliation? 2) If so, what are the magnitudes of these effects? 3) At what defoliation intensity are these differences evident? In this study, we use historical data to investigate the effect of gypsy moth defoliation on water quality and quantity over the range of gypsy moth invasion in the US from 1999 to 2018.

We hypothesize that gypsy moth defoliation is related to higher stream discharge given the reduction of evapotranspiration flux, particularly in the form of transpiration (Bearup et al. 2014), in trees by gypsy moth defoliation during the summer growing season (Dingman 2015, Smith-Tripp et al. in review). In addition, we hypothesize that in watersheds with defoliation, loss of foliar area and thus overhead shading from defoliation (and occasionally mortality) (Lance et al. 1987)leads to higher stream temperatures. In turn, increased stream temperatures will boost algal productivity and photosynthesis, and therefore increase DO levels. We investigated these three variables (stream discharge, water temperature, and DO) due to availability of large amounts of data over broad temporal and spatial scales. This project provides a benchmark for how streams and rivers in the US are affected by defoliation from gypsy moths during the past two decades.

3.2 Materials and Methods

3.2.1 Gypsy moth defoliation

We used the annual US Forest Service (USFS) Insect and Disease Detection Surveys (IDS), which are geospatial aerial defoliation and ground survey products identifying forest areas impacted by forest pests, diseases, and other disturbances. These surveys are part of a forest health protection initiative by the USFS to appraise forest conditions in terms of detriments and improvements over time and are stored in an online IDS database (Johnson and Wittwer 2006). The online database contains data dated back to 1999, and includes information about survey type, disturbance type and intensity, and descriptions of damage observed. Using these data from 1999 to 2018, we generated a spatial product containing years of survey and observation of gypsy moth defoliation presence.

In order to most accurately identify areas affected by gypsy moth defoliation, we removed records with less than 30% defoliation intensity reported, the lower threshold for aerial detection (Ciesla 2000). Due to inconsistency in defoliation intensity reports between earlier and more recent surveys, we converted records to presence or absence of gypsy moth defoliation. Thus, we rasterized defoliation polygons into yearly binary 4 km by 4 km rasters covering the gypsy moth expansion area from 1999 to 2018, where defoliated areas are assigned 1 and undefoliated areas are assigned 0. We overlaid these rasters with watersheds for sites of interest (described below) and summarized gypsy moth defoliation by watershed in terms of percent cover, i.e., how many cells out of the total in the watershed have gypsy moth defoliation. Henceforth this measure is referred to as defoliation or defoliation intensity.

3.2.2 Hydrology and water quality data

We queried the US Geological Survey National Water Information System (USGS NWIS) for public stream gage data using the dataRetrieval package in R (De Cicco et al. 2018) for the study area affected by gypsy moth as determined by USFS IDS data. Queries utilized the USGS parameter codes for discharge, temperature, and dissolved oxygen (00060, 00010, and 00300, respectively) for all available data in the period 1998 to 2018. USGS NWIS is a long-term nationwide water data repository that combines water quality and quantity data from both water sensors and historical field collections to report water information for almost 2 million sites across

the US and Territories. Stream gages vary in location along water bodies and have associated site information within the database, such as drainage area (watershed size), date of collection, station name, error codes, and other identification information.

For all sites, we performed a series of quality checks, removing records with error codes and erroneous records (e.g., temperatures over 40 degrees C, negative DO values). In addition, we removed any site downstream from a dam or reservoir, in heavily developed areas, or in primarily agricultural areas based on visual inspection. We removed these sites under the assumption that gypsy moth defoliation is unlikely to substantially impact water quality at a site without adequate forest cover. Our final site list contains only sites for which the watershed experienced gypsy moth defoliation at some point during the study period. We analyzed 244 total sites for discharge patterns, 186 for temperature, and 67 for DO (**Figure 3-1**). See **Table S3-1** for a full list of sites. We found the watershed for each site using the nhdplusTools R package (Blodgett 2019), querying NHDPlus using identifiers associated with each site. For statistical analysis, hydrology and water quality data were converted to monthly averages for each watershed. To conduct analyses on DO data, we converted DO concentration in mg/L to DO percent saturation using the rMR package (Moulton 2018) assuming barometric pressure at sea level and using temperature values for the same samples. This function converts DO concentration to DO percent saturation using solubility equations for oxygen in water. Henceforth we refer to DO percent saturation as DO.

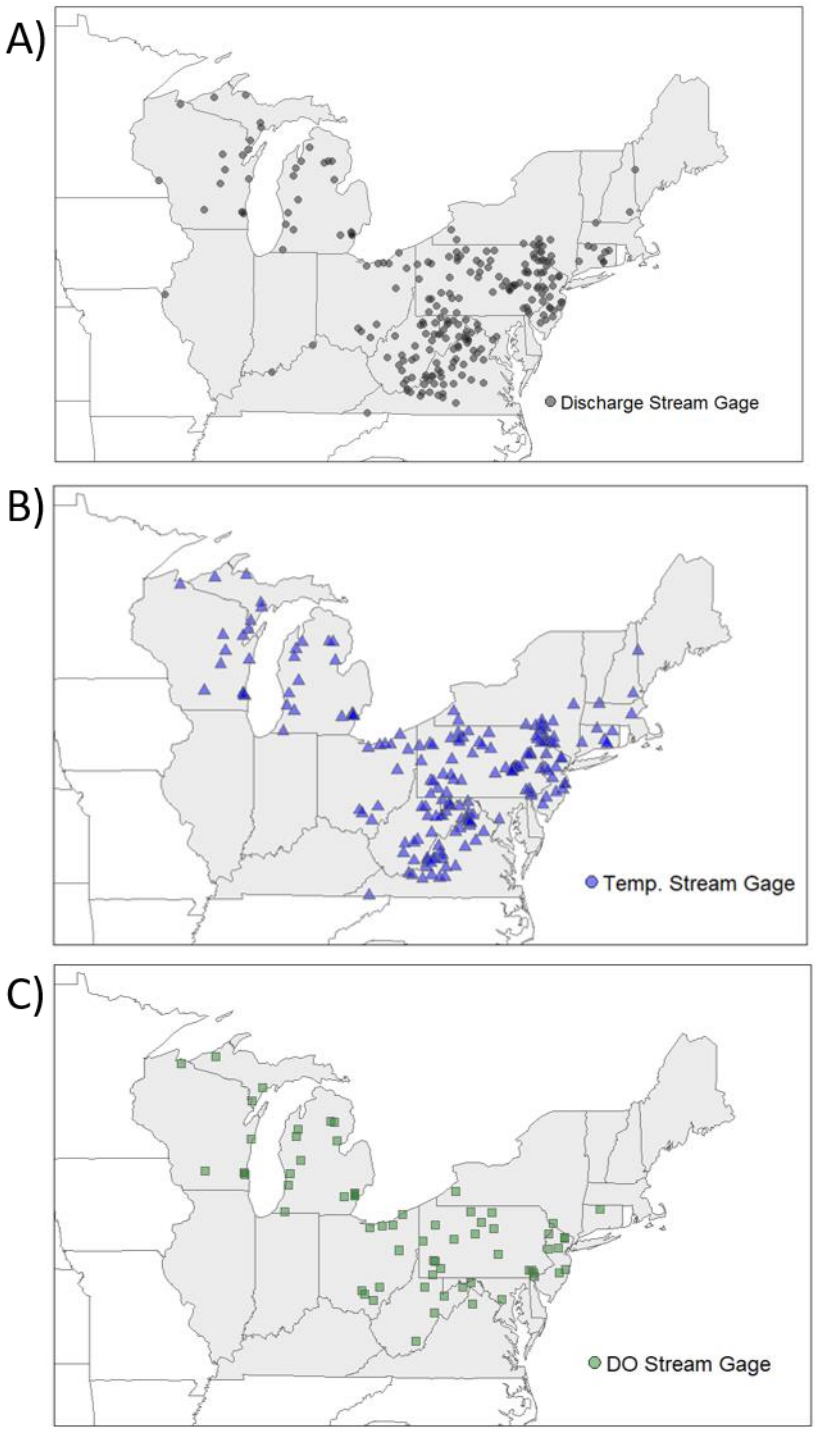

Figure 3-1. Stream gage map. Map of locations of USGS stream gages used in analyses for A) discharge, B) temperature, and C) dissolved oxygen.

3.2.3 Water balance

For each watershed within the dataset, we obtained monthly precipitation and evapotranspiration values. Monthly precipitation data come from PRISM at Oregon State University for the period 1998 to 2018 (PRISM Climate Group 2004). The PRISM Climate Group develops spatial climate data products that use over 13,000 national weather stations to interpolate precipitation values across a smoothed surface. Precipitation is interpolated based on input climate data, elevation, location, proximity to coastal areas, and other factors (Daly et al. 2008). We downloaded monthly averages directly from PRISM through R, using the prism package (Hart and Bell 2015) as gridded data with 4 km cell size.

Monthly evapotranspiration data come from USGS Center for Integrated Data Analytics (CIDA) for the period 2000 to 2017 (USGS 2019). These data were generated using a Simplified Surface Energy Balance model with ET inputs estimated from MODIS thermal imagery (Senay et al. 2013). The result is actual estimated ET, summarized in monthly and yearly data products. We obtained total monthly ET directly from USGS's GeoData Portal as gridded data with 1 km cell size.

Both precipitation and evapotranspiration data are available at the national level, which we spatially subset with the extent of the gypsy moth defoliation rasters we generated (described above). We found the total precipitation and evapotranspiration volume per watershed for each month between July and October using zonal statistics, then calculated water balance by subtracting water lost via evapotranspiration from water entering the watershed via precipitation. For statistical analysis, resulting precipitation, evapotranspiration, and water balance values were converted to monthly averages for each watershed.

3.2.4 Analyses

We used linear mixed models in combination with Satterthwaite-approximated Tukey tests to compare mean discharge, temperature, and DO in watersheds with defoliation versus without defoliation. For each fit mixed model, we used an ANOVA to identify significant interaction terms within the fitted models. Therefore, fitted mixed models were used as input into the ANOVAs. We subsequently extracted slopes of relationships from the fitted models to identify trends. In addition, the fitted mixed models were used as input to the Tukey test to compute and compare least-squares

means using the emmeans package in R (Lenth 2021). Mixed models allow for the analysis of interactions while accounting for repeated measures. Tukey tests conduct post-hoc mean comparison while pooling variance across the entire dataset used, instead of for individual groups being compared. We used a Satterhwaite adjustment to find "effective degrees of freedom", or to similarly pool degrees of freedom to correspond with the pooled variance in the Tukey test. In addition, Tukey tests protect against inflated risk of Type I error from conducting many significance tests among groups, and is thus a more conservative measure of significant differences between means (Abdi and Williams 2010).

For several analyses, we use runoff ratio (discharge divided by precipitation) to account for the amount of water entering the watershed versus exiting the watershed in the form of discharge. We split the data by major water basin (2-digit Hydrologic Unit Code, or HUC) and watershed area to identify trends among different regions and watershed sizes. Both 2-digit HUC and watershed area were provided with a site information table exported from NWIS. We binned samples by respective major water basin (2-digit HUCs 01-07 are represented in our study area, with varying amounts of representation across variables) and by watershed sizes of 0.1 - 10 km^2, 10 - 1,000 km^2, 1,000 - 10,000 km^2, or greater than 10,000 km^2. These subsets allowed us to determine patterns between defoliation and measures of water quality/quantity among various characteristics of the surface waters in our study area while maintaining approximately equivalent numbers of sites in each group. We chose these respective watershed sizes to capture comparable numbers of samples and sites in each grouping while investigating trends at a range of watershed areas. Approximately even splits among data are most appropriate based on the assumptions of linear mixed models and also help visual assessment of trends. While we aim to generalize trends across the entire study area, identifying if patterns emerge only under certain conditions provides additional information about the impacts of gypsy moth on water resources.

To model observational data (i.e. sample data compiled for water quality, climate, and gypsy moth defoliation), we conducted random forest (RF) using the randomForest package in R (Liaw and Wiener 2002) to quantify the impact of defoliation on discharge, temperature, and DO, and to predict response variables with varying defoliation conditions. RF is a non-parametric machine-learning algorithm that relies on regression trees using bootstrapped samples of the original data based on iteratively randomly chosen variables (Breiman 2001). The algorithm makes predictions for observations that were not used in model development, and calculates mean

squared error (MSE) from the difference in actual observations versus predicted observations. Given the large amount of non-normally distributed data in our study, the use of a non-parametric test that utilizes bootstrapping allows for better model performance and predictions.

We ran 500 trees using the default number of variables per bootstrapped sample (number of included predictor variables divided by 3) and used the resulting model to assess variable importance and predict response values based on ranges of input parameters. General model structure included a response variable as a function of all included predictor variables, the specified number of trees to generate (500), and an argument specifying the function to return importance values. The variables included in the RF algorithm for each response variable are shown in Table 3-1. Precipitation, watershed area, and defoliation were included as predictors in every RF based on their potential impact on each response variable, with additional specific predictors for each response variable. For all subsequent predictions, we created new datasets to generate predicted response variables values. For every set of predictions, we used input defoliation values of 0%, 20%, and 40%, expressed as proportion of watershed area, and watershed areas of 100 km^2, 1,000 km^2, 10,000 km^2, and 100,000 km^2. We used these defoliation intensities as they captured the range of defoliation intensities seen in the observational data while representing low, intermediate, and high defoliation events. Additional input predictor values varied among response variables modeled: we included evapotranspiration in the discharge RF, month and discharge in the temperature RF, and temperature and discharge in the DO RF.

Table 3-1. Random forest predictor variables. Predictor variables included in random forest algorithms by response variable with relative percent increase in MSE.

Variable	% Increase in MSE
a) Discharge	
Precipitation	130.51
Evapotranspiration	61.45
Watershed Area	51.77
Gypsy Moth Defoliation	38.22
b) Temperature	
Month	465.59
Watershed Area	132.33
Discharge	98.63
Precipitation	54.97
Gypsy Moth Defoliation	47.7
c) DO	
Watershed area	94.54
Temperature	24.67
Gypsy Moth Defoliation	21.84

3.3 Results

3.3.1 Discharge

We observed a trend of higher discharge and runoff ratios in sites without defoliation across our dataset. The slopes of the relationship between defoliation intensity and runoff ratio were generally negative, regardless of watershed, month, or major water basin grouping. There was an overall significant difference in monthly runoff ratios related to intensity of defoliation ($F_{1, 15804}$ = 47.90, p<0.001) using a linear mixed model. There was also a significant interaction between defoliation and watershed area ($F_{1, 1144.3}$ = 2.01, p<0.001), thus we found individual slopes of the

relationship between defoliation and runoff ratio for each watershed group (**Figure 3-2**). The slope of the relationship between defoliation and runoff ratio was negative regardless of watershed area. We also found a significant interaction between defoliation and month ($F_{3, 15601} = 3.087$, $p = 0.026$), and again found individual slopes of the relationship between defoliation and runoff ratio for each month. Slopes across all months were negative (**Figure S3-1**). There were no significant interactions between defoliation and major water basins. However, for every major water basin with enough data to plot, the trend between defoliation intensity and runoff ratio was negative with the exception of HUC 04, in which slope was slightly positive (**Figure S3-2**).

When split by watershed area, Tukey tests indicated the mean monthly runoff ratios at sites with defoliation were statistically significantly lower than in sites without defoliation in intermediate sized watersheds (100 - 1000 km^2 and 1000 - 10,000 km^2) (both $p < 0.001$, see **Table S3-2.** for group estimates and SEs). These tests indicated that in defoliated sites, runoff ratio was 0.028 (+- SE 0.006) lower for watersheds in the 100 - 1000 km^2 group and 0.029 (+- SE 0.006) lower for watersheds in the 1000 - 10,000 km^2 group. In other watershed sizes, the differences were insignificant, though we saw the same trend of higher runoff ratio in undefoliated sites regardless of watershed size (**Figure S3-3, Table S3-2).**

We also found significantly lower monthly runoff ratios in defoliated sites in July ($p = 0.007$), September ($p < 0.001$), and October ($p < 0.001$).

Tukey tests among major water basins indicated that only in HUC 02 the mean monthly runoff ratio at sites with defoliation was statistically significantly lower ($p < 0.001$) than in sites without defoliation (**Figure S4**, **Table S2**). The results of this test indicated that in HUC 02, the runoff ratio was 0.031 (+- SE 0.005) lower in defoliated sites than in undefoliated sites. Although differences in all other major water basins were not statistically different, we generally observed lower measures of discharge (both discharge and runoff ratio) in sites with defoliation than without across most major water basins (**Figure S2**). See Table **S2** for the results of all Tukey tests.

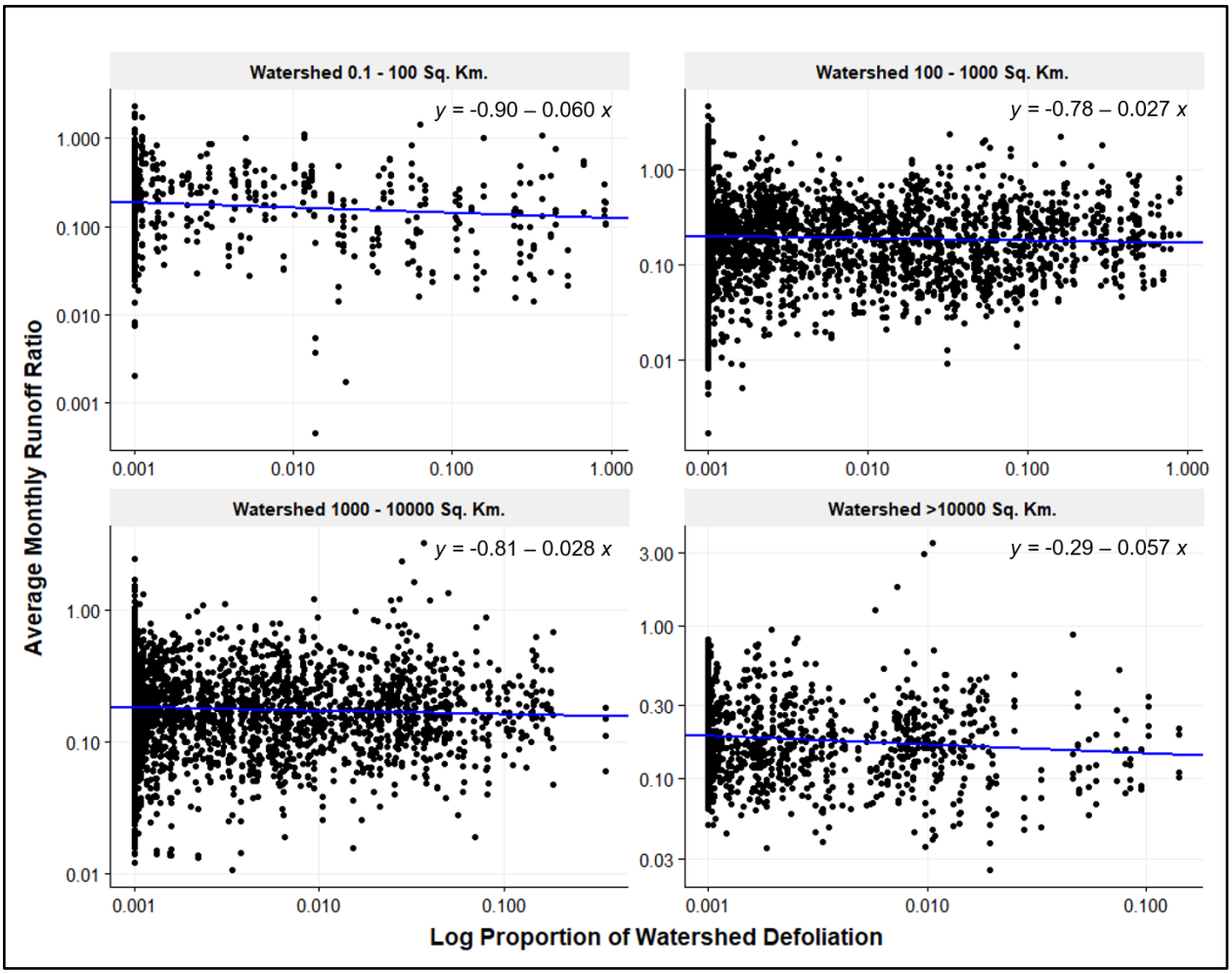

Figure 3-2. Runoff ratios by watershed area. Distribution of average monthly runoff ratios across observed defoliation intensity by watershed area. Reported equation extracted from mixed effect model.

The RF algorithm explained 60% of variance in monthly discharge. Defoliation had a lower importance value in terms of impact on MSE than all other predictor variables. Other variables more directly tied to hydrology (precipitation, evapotranspiration, and watershed area) had a stronger impact on MSE. Yet, gypsy moth defoliation explained a substantial amount of model variance (38% increase in MSE) (Table 3-1).

Predicted discharge was generally lower in watersheds with defoliation regardless of watershed size, aligning with the trends seen in sample data (**Figure 3-3**). RF discharge predictions were higher on average at 0% defoliation than 20% (22% higher) or 40% defoliation (24% higher), though the degree of difference varied between watershed areas and average monthly rainfall

(**Table 3-2**). In sites most heavily defoliated (greater than 60% of the watershed defoliated by gypsy moth), we found negative slopes between defoliation and runoff ratio (**Figure S5**), indicating that even at high defoliation, discharge decreases slightly with increasing defoliation at the same precipitation level.

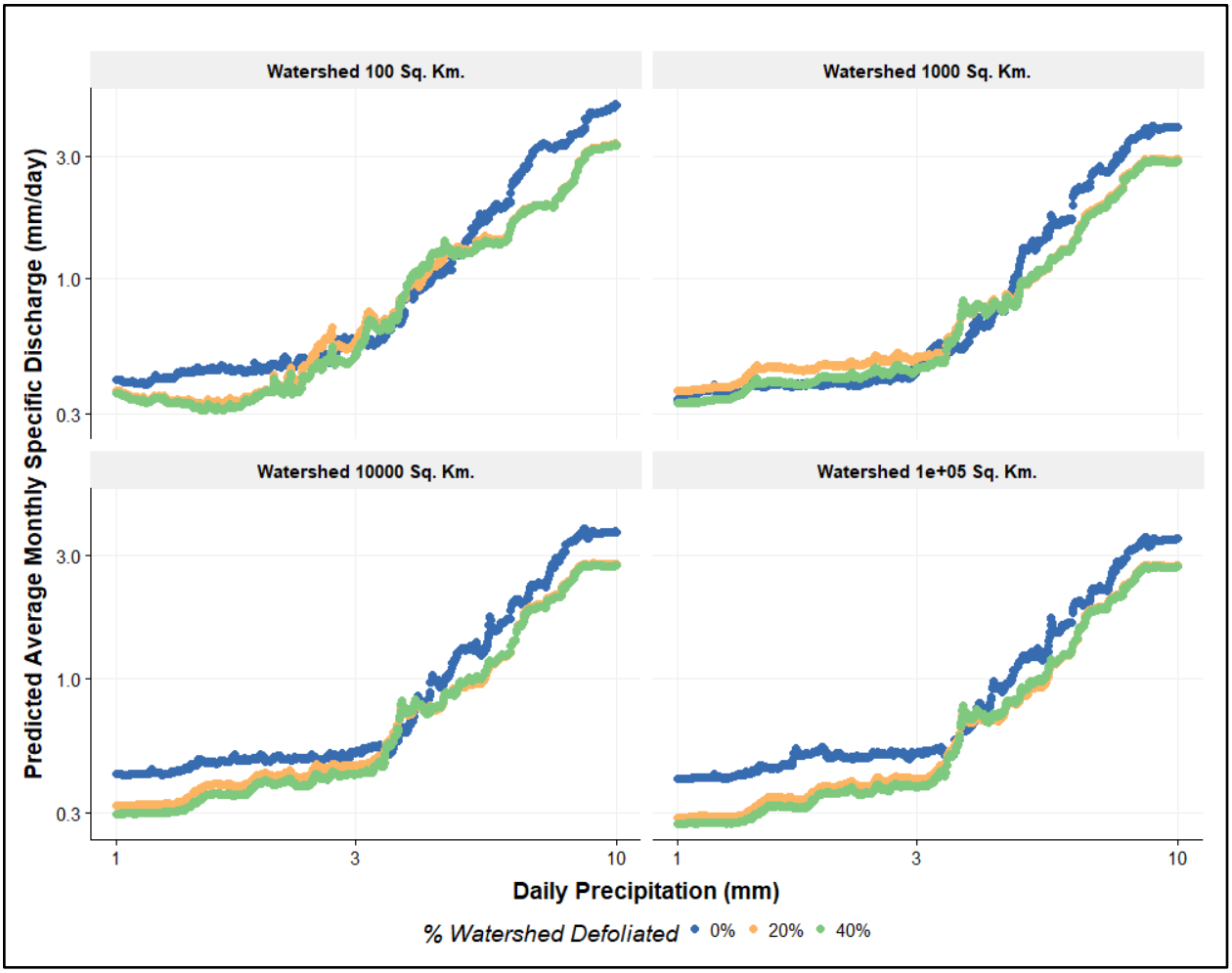

Figure 3-3. Specific discharge predictions. Specific discharge predicted from the random forest algorithm, using a set range of average daily precipitation, gypsy moth defoliation intensity, and watershed area values.

Table 3-2. Predicted percentage of differences in discharge from zero defoliation. Percent difference in predicted discharge and temperature values between 0% defoliation and 20% or 40% defoliation. Asterisks (*) indicate that the average predicted value was less than the average of predicted value at 0% defoliation.

	Mean			Percent Difference from 0%	
Watershed Size	**0% defoliation**	**20% defoliation**	**40% defoliation**	**20% defoliation**	**40% defoliation**
Discharge					
<100	0.0013	0.0010	0.0010	23.70*	25.31*
100-1000	0.0011	0.0010	0.0009	16.34*	20.12*
1000-10,000	0.0012	0.0009	0.0009	23.07*	25.21*
>10,000	0.0011	0.0009	0.0009	23.53*	24.95*
Temperature					
<100	16.87	16.49	16.58	2.25*	1.75*
100-1000	17.23	16.92	16.97	1.77*	1.51*
1000-10,000	19.04	19.28	19.28	1.26	1.28
>10,000	19.64	19.44	19.44	1.02*	1.01*
DO					
<100	96.47	94.08	93.31	2.51*	3.33*
100-1000	96.59	91.77	91.19	5.11*	5.75*
1000-10,000	93.42	96.40	95.94	3.14	2.67
>10,000	99.93	99.07	98.57	0.86*	1.37*

3.3.2 Temperature

We did not find an overall significant difference in monthly temperature directly related to intensity of defoliation. However, we observed slightly positive trends between defoliation and temperature across the dataset. In the mixed models with defoliation and watershed size, HUC, or month as predictors and temperature as the response variable, there were no significant interactions between defoliation and watershed size, HUC, or month. An examination of the slopes of these

relationships showed that in all but the largest watersheds, there was a positive slope between defoliation intensity and temperature (**Figure 3-4**), indicating an overall small positive relationship. Though Tukey tests indicated no statistically significant differences among months, watershed size, or major water basin, across all months of the study period, we observed a trend of higher mean temperature at sites with defoliation (mean = 18.74, SD = 4.79) than those without (mean = 18.15, SD = 4.55) (**Figure S3-6**, **Table S3-2**).

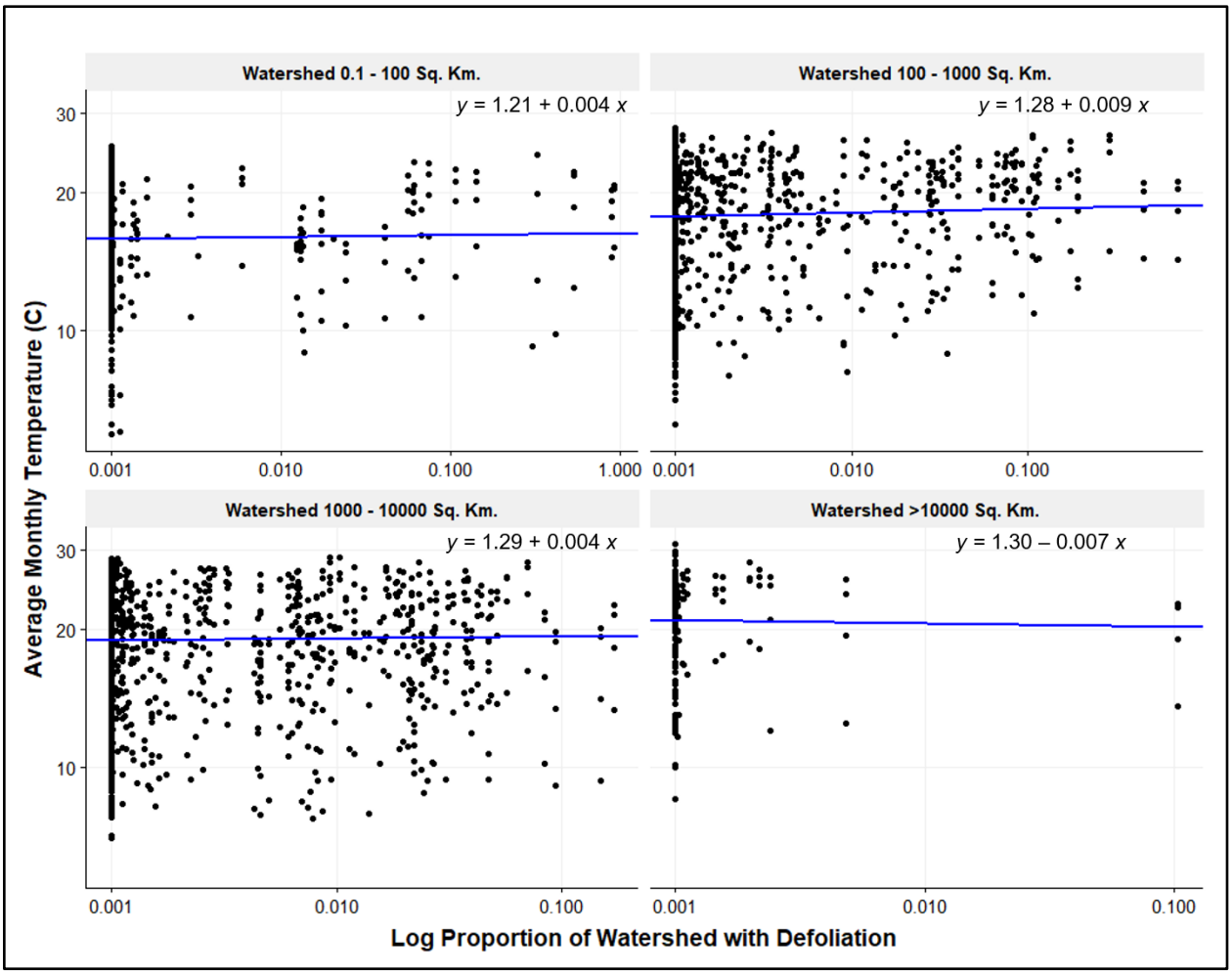

Figure 3-4. Temperature by watershed area. Distribution of average monthly water temperature across observed defoliation intensity by watershed area. Reported equation extracted from mixed effect model.

The RF algorithm explained 84% of variance in monthly temperature. Similar to the results for discharge RF, gypsy moth defoliation explained less model variance than other predictor variables, but still had substantial impact (**Table 3-1**). Month intuitively had the strongest influence on temperature given monthly variations in air temperature. Gypsy moth defoliation contributed to the percent increase in MSE within 10 percentage points of precipitation (48% and 55%, respectively).

Predicted temperature values were slightly lower for 20% and 40% defoliation than 0% defoliation, though this trend varied marginally by watershed size and by month (**Figure 3-5**). For example, predicted temperature at 0% defoliation was always higher than at 20% or 40% defoliation in September, but was generally lower than at 20% or 40% defoliation in October. However, regardless of watershed size, there were minimal differences in predicted temperature between defoliation percentages (1.01%-2.25% difference) (**Table 3-2**). We also used the RF algorithm to predict temperature values for a wider range of watershed sizes to identify if there are watershed sizes at which defoliation substantially impacts water temperature. In watersheds smaller than approximately 30 km^2, temperature predictions were higher at 20% and 40% defoliation than 0%. At watersheds between 30 km^2 and 200 km^2, the relationship switched and average predicted temperature was higher with no defoliation than 20% or 40% defoliation. Above 200 km^2, the difference became negligible (**Figure S3-7**).

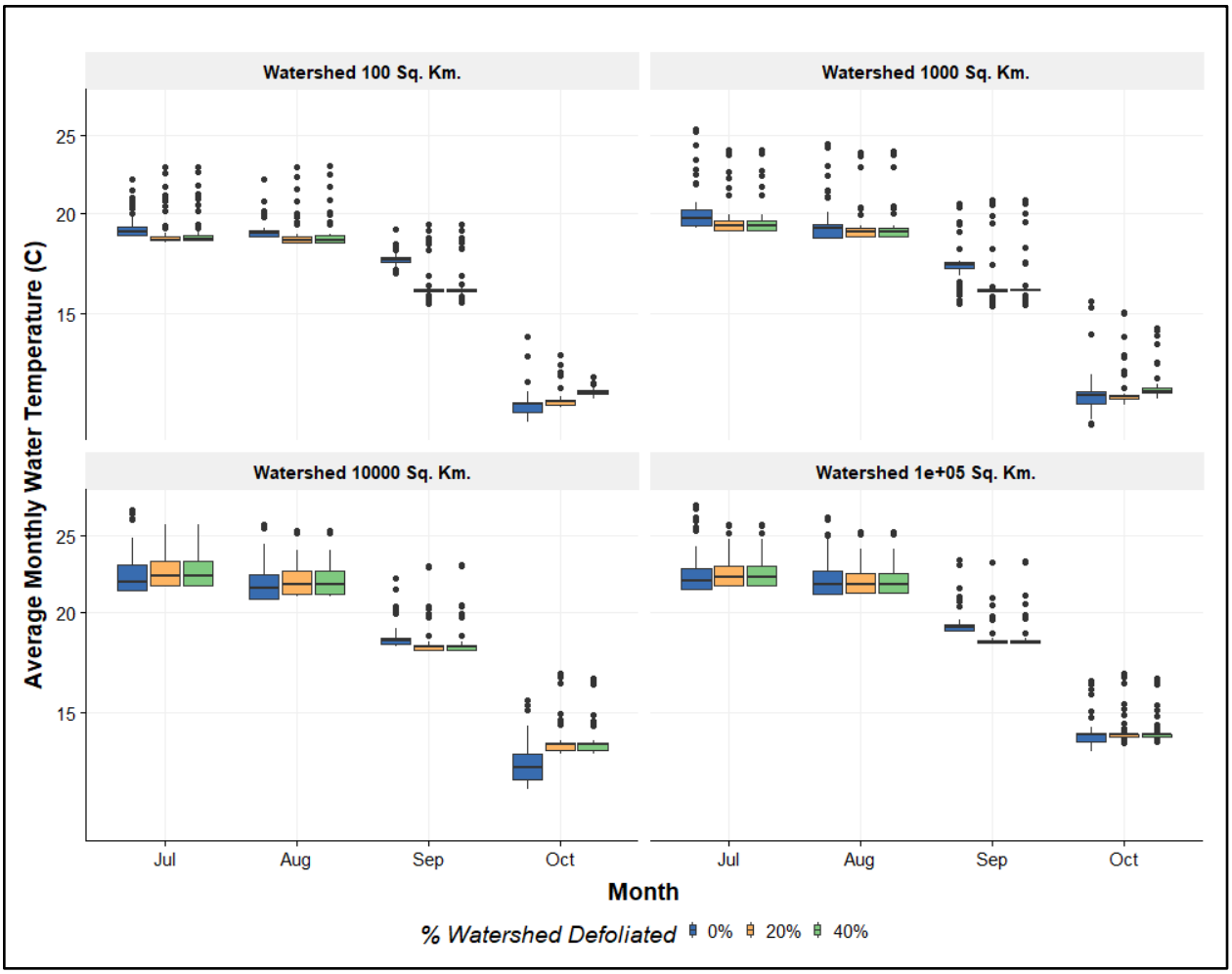

Figure 3-5. Predicted temperature by watershed area. Temperature predicted from the random forest algorithm across summer months, using a set range of precipitation, discharge, watershed area, and gypsy moth defoliation values.

3.3.3 Dissolved Oxygen

We found a significant negative relationship between defoliation intensity and monthly percent DO saturations related to intensity of defoliation based on the results of the linear mixed model. Trends between defoliation intensity and DO were generally slightly negative across the dataset. There was also a significant interaction between defoliation intensity and watershed area ($F_{3,\ 2319.82} = 3.25$, $p = 0.021$), thus we found individual slopes of the relationship between defoliation and DO for each watershed group (**Figure 3-6**). Slopes were negative in every watershed group except the 1,000 - 10,000 km^2 group. We also found a significant interaction between defoliation intensity and major water basin ($F_{2,\ 2319.47} = 3.62$, $p = 0.027$), and again found

individual slopes of the relationship between defoliation and DO for each major water basin with enough DO data to run mixed models independently (**Figure S3-8**). We therefore omitted HUCs 01, 03, 06, and 07 from these analyses and plots due to lack of adequate DO data. Slopes extracted from the mixed models were negative in HUCs 02 and 04, but positive in HUC 05. There was not a significant interaction between defoliation intensity and month. All Tukey tests comparing mean DO between defoliated and undefoliated sites by month, watershed area, and HUC were not significant (**Table S3-2**). Though the results of all Tukey tests indicated there were no statistically significant differences among defoliated and undefoliated sites, we observed a generally negative trend between defoliation intensity and DO.

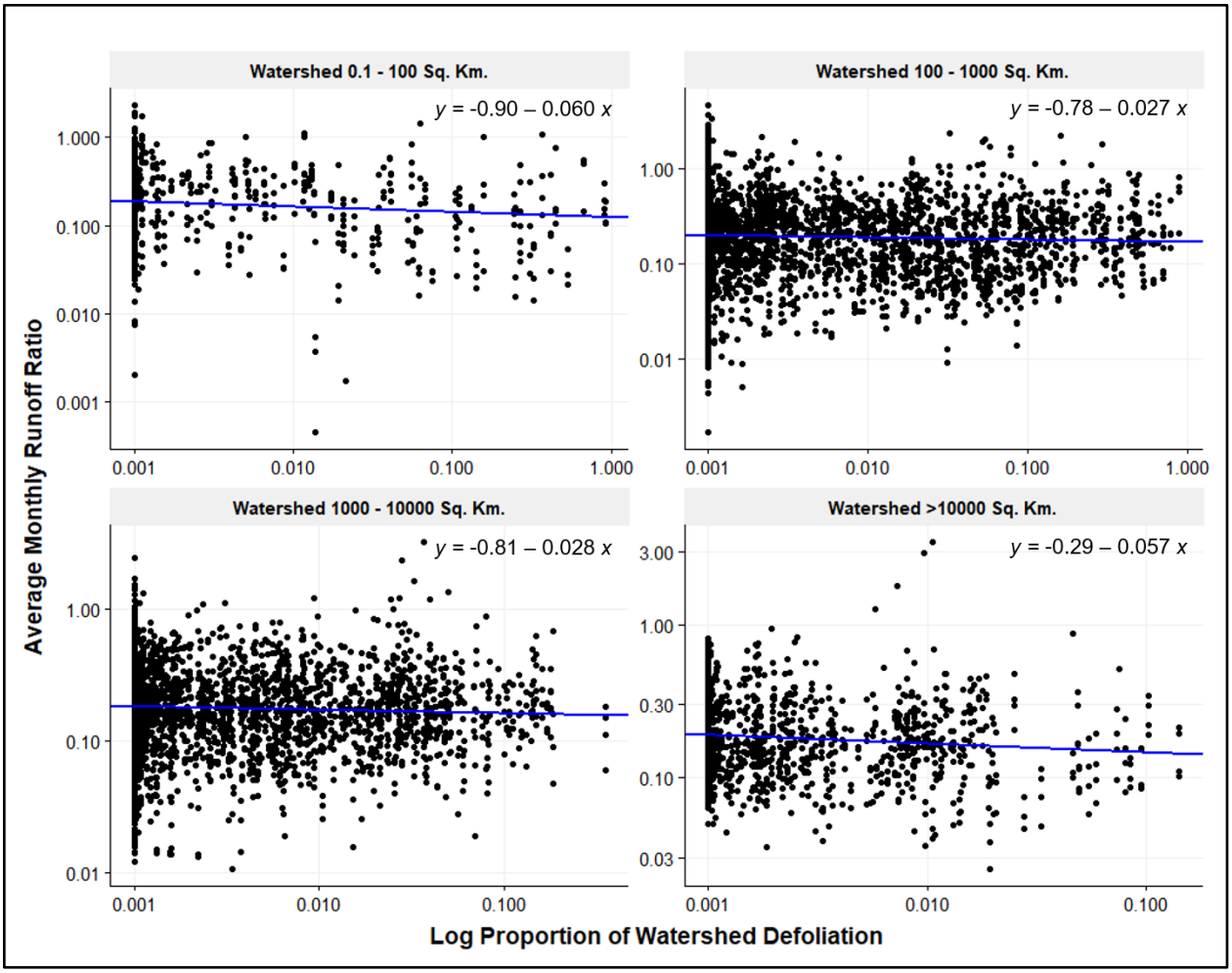

Figure 3-6. Dissolved oxygen by watershed area. Observed dissolved oxygen saturation (%) by gypsy moth defoliation intensity, split by watershed area.

The RF algorithm explained 56% of variance in monthly DO. Watershed area had the highest importance value in terms of predicting DO. Defoliation had a higher importance value than discharge or precipitation, and contributed to the percent increase in MSE within 5 percentage points of temperature (24% and 26%, respectively) (Table 3-1). RF predictions reflected the observational data trends, indicating generally lower DO with higher defoliation intensity (**Figure 3-7**). In all but the 1,000 - 10,000 km^2 group, predicted DO was lower with defoliation regardless of discharge. In 1,000 - 10,000 km^2 watersheds, the RF algorithm predicted higher DO with defoliation regardless of discharge, the opposite trend as observed in all other watershed groups. In the largest watersheds (>10,000 km^2), there were marginal differences between undefoliated and defoliated predictions (0.86% - 1.37%). The largest differences were predicted in 100 - 1,000 km^2 watersheds, with predicted temperature approximately 5% lower with 20% defoliation than 0%, and approximately 6% lower with 40% defoliation.

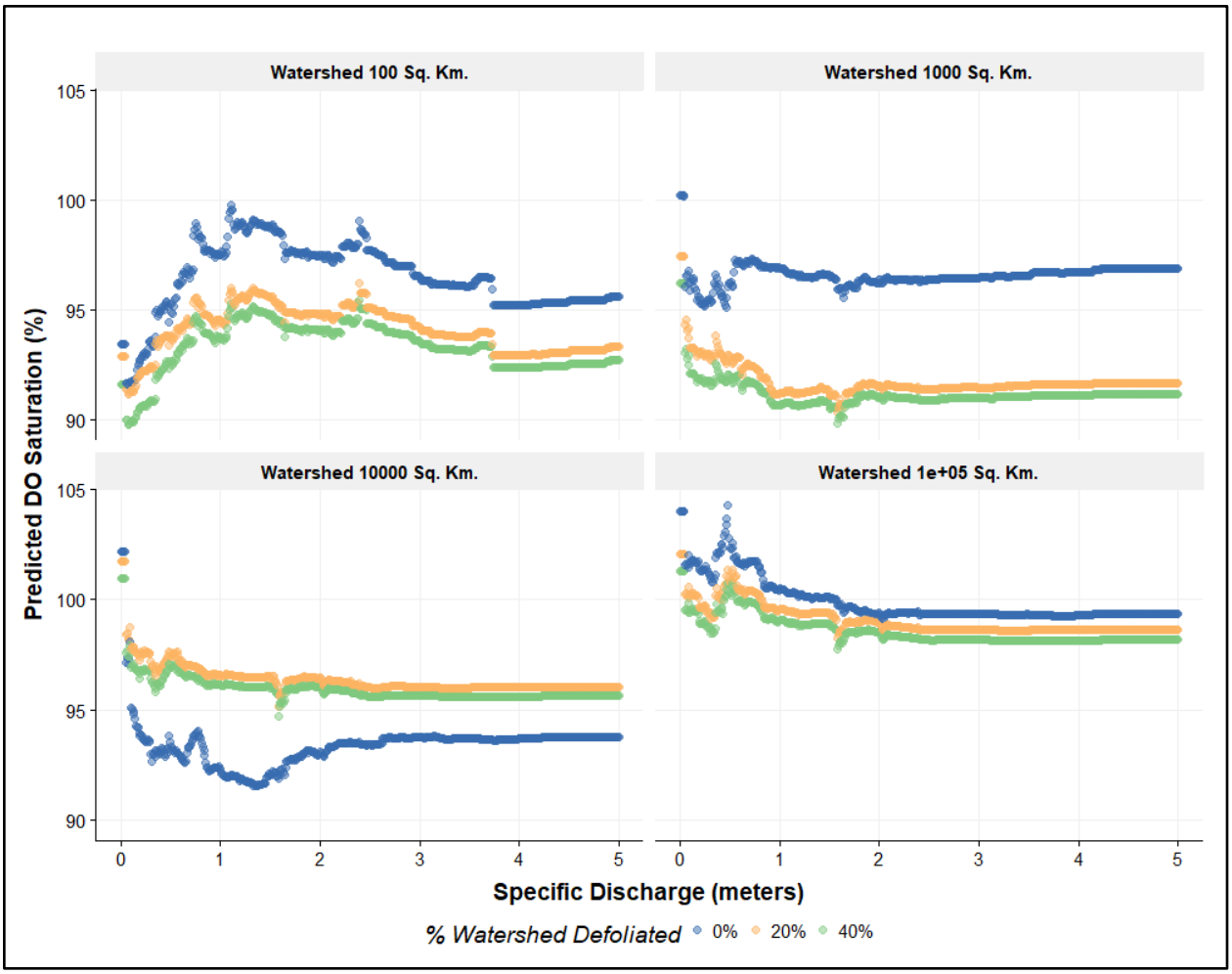

Figure 3-7. Dissolved oxygen predictions. Dissolved oxygen saturation (%) predicted from the random forest algorithm, using a set range of precipitation, watershed area, temperature, and gypsy moth defoliation values**.**

3.4 Discussion

We observed a negative relationship between defoliation intensity and discharge, the opposite of the expected result based on our hypotheses. Both observational data and RF predictions showed lower discharge with higher defoliation. Similarly, there was a negative trend between defoliation intensity and DO in both observational data and RF predictions, again opposite of our expected result. We saw small increases in temperature with higher defoliation intensity, consistent with our hypothesis. While our results largely did not support our hypotheses, we propose additional considerations below.

3.4.1 Discharge

We observed a negative relationship between discharge and defoliation intensity across all watershed groupings, months, and in the most intensely defoliated sites. Tukey tests of observational discharge data also indicated discharge was lower at sites with defoliation than those without defoliation in several watershed groupings. In addition, the RF predictions indicated the same relationship regardless of watershed size while accounting for variation in relevant hydrologic conditions and watershed characteristics. The results of our study indicate across much of the US gypsy moth range, there is generally a decrease in discharge with defoliation. We estimate the decrease as about 0.03 runoff ratio units in watersheds from 100 - 10,000 km^2, and there may be as much as 25% decrease in discharge in 40% defoliated watersheds (**Table 3-2**). This result is contrary to our original hypothesis on the effects of defoliation on discharge.

However, several studies have shown similar or no patterns between defoliation and discharge. Lewis and Liken (2007) found that following defoliation by elm spanworm (*Ennomos subsignarius* Hübner), two watersheds experienced drier conditions. In addition, Addy et al. (2018) were unable to identify trends between discharge and gypsy moth defoliation during an extreme defoliation event (35% watershed defoliation) and a major drought. Therefore, the effects of defoliation on water quantity likely vary by many factors, including region, forest composition, soil and geology type, intensity and duration of defoliation, and compounded stress from additional disturbances (such as drought, high wind, etc.).

In studies of the effect of forest clearing on discharge, several authors concluded that in general, changes in discharge were observable only when at least 20% of a catchment was disturbed (Bosch and Hewlett 1982, Stednick 1996). Clearcutting is a more intense disturbance than gypsy moth defoliation, given that entire trees are removed from the system. In addition, many trees grow new leaves after defoliation in the same growing season. Unless trees are repeatedly defoliated, the impact of lost foliage on ET can be minimal due to regrowth of foliage and epicormic branching compensating for lost ET (Schafer et al. 2014). We can assume a larger percentage of a watershed must be defoliated before seeing increases in discharge with an outbreak event similar to that observed following clearcutting.

However, even in heavily defoliated watersheds (>60% of the watershed defoliated), we did not observe a positive effect on discharge, and instead saw slight negative relationships. One potential explanation for this observed negative relationship, even in heavily defoliated sites, is

that drier soils absorb water more readily. Our hypothesis was that more water would enter the system as trees transpire less during and after defoliation. However, if more solar radiation reaches the ground via gaps in the canopy due to defoliation, temperature at the forest floor will be higher (Lance et al. 1987), and evaporation and thus water lost from the system can also increase, counteracting the effect of lost transpiration and also drying soils further.

3.4.2 Temperature

Though we did not find statistically significant differences in temperature among defoliation intensities, and the RF predicted temperature values similarly showed small differences in temperature between defoliation intensities, we saw a slight positive trend between defoliation and temperature in all but the largest watersheds. The impact of defoliation on temperature may be smaller than can be detected statistically, while still having ecologically relevant impacts. Even very small increases in water temperature can negatively impact aquatic life, including fish (Tarzwell 1970) and aquatic insects (Li et al. 2011). Studies of the impact of hemlock woolly adelgid on water quality have also found increased stream temperature resulting from loss of canopy cover, though as we demonstrated here, the changes are often small (Roberts et al. 2009, Webster et al. 2012).

There are similar interactions between drought, gypsy moth outbreaks, and temperature as with discharge. High air temperature and dry conditions (such as droughts) are directly linked (Trenberth and Guillemot 1996). Higher air temperature in conjunction with increased solar radiation during drought-induced defoliation events can lead to higher water temperatures than during dry conditions or defoliation events alone. Climate events, such as drought, in conjunction with forest disturbances are shown to have additive effects on water quality (Kaushal et al. 2018). For example, in a study of the effect of gypsy moth defoliation on water quality and metabolism, Addy et al. (2018) demonstrated the additive effect of drought and defoliation. They found increased stream temperature during a gypsy moth outbreak in which canopy cover in the riparian area of the watershed was reduced by 51% and by 35% in the entire watershed. This study also demonstrates that a high level of defoliation may be necessary to observe impacts to stream temperature, especially in riparian areas, and even in small watersheds. The watershed investigated by Addy et al. (2018) was 4.4 km^2, within the range of watershed area the RF in our results predicted to have higher temperature with defoliation (**Figure S7**). However, we suspect that due

to low sample size in small watersheds, predictions of temperature below about 100 km^2 are less reliable than the watershed sizes with more samples in the observational data. Based on the results of the RF presented here and existing literature, we conclude that defoliation could increase water temperature minimally, but with biologically meaningful implications.

The positive trend we observed between defoliation intensity and water temperature in the observational data was not verified by the RF predictions. This difference may be due to unexplored relationships between discharge and temperature, where at lower discharge, especially in smaller watersheds, the impact of defoliation on water temperature may be exaggerated.

3.4.3 Dissolved Oxygen

While the results of Tukey tests with observational data showed no significant relationship between defoliation and DO across groupings by month, watershed size, and major water basin, we observed a slight negative trend overall. Both results of linear mixed models and RF predictions indicate that defoliation intensity has a negative relationship with DO. One notable exception to the negative trend we observed was in RF predictions for watersheds 1,000 - 10,000 km^2 in size. In these watersheds, the RF predictions for DO were 2-3% higher with defoliation than without, potentially indicating a relationship between stream size (represented by watershed size) and the impacts of defoliation on DO.

During gypsy moth defoliation events, higher levels of turbidity and lower light penetration in the water can result in decreased autotrophic activity, which results in lower DO (Kortmann and Cummins 2018). These effects are stream-specific, depending on the amount of detritus that enters the stream, and are more likely in headwater streams where generally higher amounts of canopy cover and smaller stream sizes contribute to increased turbidity. In contrast, increased solar radiation from loss of canopy cover via defoliation can stimulate autotrophic activity in streams, thus increasing DO (Mulholland et al. 2001, Bernot et al. 2010). There is also evidence that gypsy moth defoliation events increase nutrient loading to streams (Webb et al. 1995, Kortmann and Cummins 2018), which can also stimulate autotroph activity, again increasing DO. This stimulation of stream metabolism is common in larger streams, where low turbidity and long water residence times, particularly during droughts, enhances photosynthesis (Hosen et al. 2019). The results of our study may show the effect of this drought-metabolism relationship, in which larger streams experience increases in metabolism during droughts while small streams do not. Our

results support this relationship, given that we observed lower DO in smaller watersheds during defoliation (which contributes to turbidity and often co-occurs with drought), but saw increased DO in larger downstream waters. Therefore, if defoliation by gypsy moth is stimulating photosynthesis by increasing nutrient input (e.g., from release of nitrogen in frass) and reducing canopy cover, the effects may be masked until waters travel downstream where turbidity lessens, more solar radiation reaches surface waters, and autotroph activity is stimulated.

Additionally, low flows, which are common during drought conditions, are related to both low DO and higher water temperature (Woltemade 2017, Danladi Bello et al. 2017), and warmer waters can hold less oxygen. We observed both decreased discharge and saw slightly increased predicted temperature at sites with defoliation, which could in turn decrease the DO levels in those sites despite increases in either autotroph activity or nutrient loadings. Furthermore, there may be additional variables influencing DO flux in surface water, as indicated by the relatively low amount of variability explained by the RF algorithm predicting DO. We anticipate additional hydrological and metabolism modeling is required to determine differences in DO with varying defoliation intensity.

3.5 Conclusions

In summary, we did not find support for our hypothesis that defoliated watersheds experience increased discharge and in fact found support for the opposite trend. In our analysis over a large geographic area, we found that the overall relationship between defoliation and discharge was negative, regardless of several watershed characteristics and even in highly defoliated watersheds. We suggest several potential causes for this negative relationship, and there may be other factors contributing to reduced discharge beyond the scope of our modeling efforts. In addition, we did not find evidence to support the hypothesis that DO is higher in watersheds with defoliation and found that watershed area may play a significant role in identifying defoliation impacts. We did, however, find evidence via trends extracted from mixed effects models that defoliation increases water temperature marginally. The results of our study indicate that broad-scale relationships between gypsy moth defoliation and water quantity and quality are largely different from trends observed on a single- or several-watershed scale.

It is also worth noting that gypsy moth defoliation reduces defenses of trees and increases tree stress, thus making them more susceptible to other secondary pests and more likely to be killed

by additional disturbances (Twery 1991). Trees impacted by gypsy moth defoliation may die a year or several years after the gypsy moth outbreak via secondary mechanisms, though this occurs in a relatively small fraction of trees. This delayed mortality could impact water resources long after the initial gypsy moth defoliation occurred in a watershed. Several studies suggest changes in water chemistry following defoliation are strongest a year after the defoliation event, and can last for several years, depending on the duration of defoliation (Webb et al. 1995, Eshlemann et al. 1998, Lewis and Liken 2007). These delayed impacts may be a result of secondary mechanisms resulting in increased tree mortality due to the weakening or stressing of trees by gypsy moth defoliation, and could potentially be captured with additional hydrological time-series modeling.

Though the effects of gypsy moth defoliation on ecosystems have been widely studied, there are few syntheses of gypsy moth defoliation impacts to water quantity on a broad scale. Here, we identified trends and quantified impacts between defoliation and several measures of water quality and quantity across a broad spatial scale. This research highlights the importance of quantifying relationships at the landscape scale when adequate multi-scale data are available.

3.6 Supplementary Materials

Table S3-1. USGS site table. USGS NWIS sites used in analyses, included in a separate file in appendix.

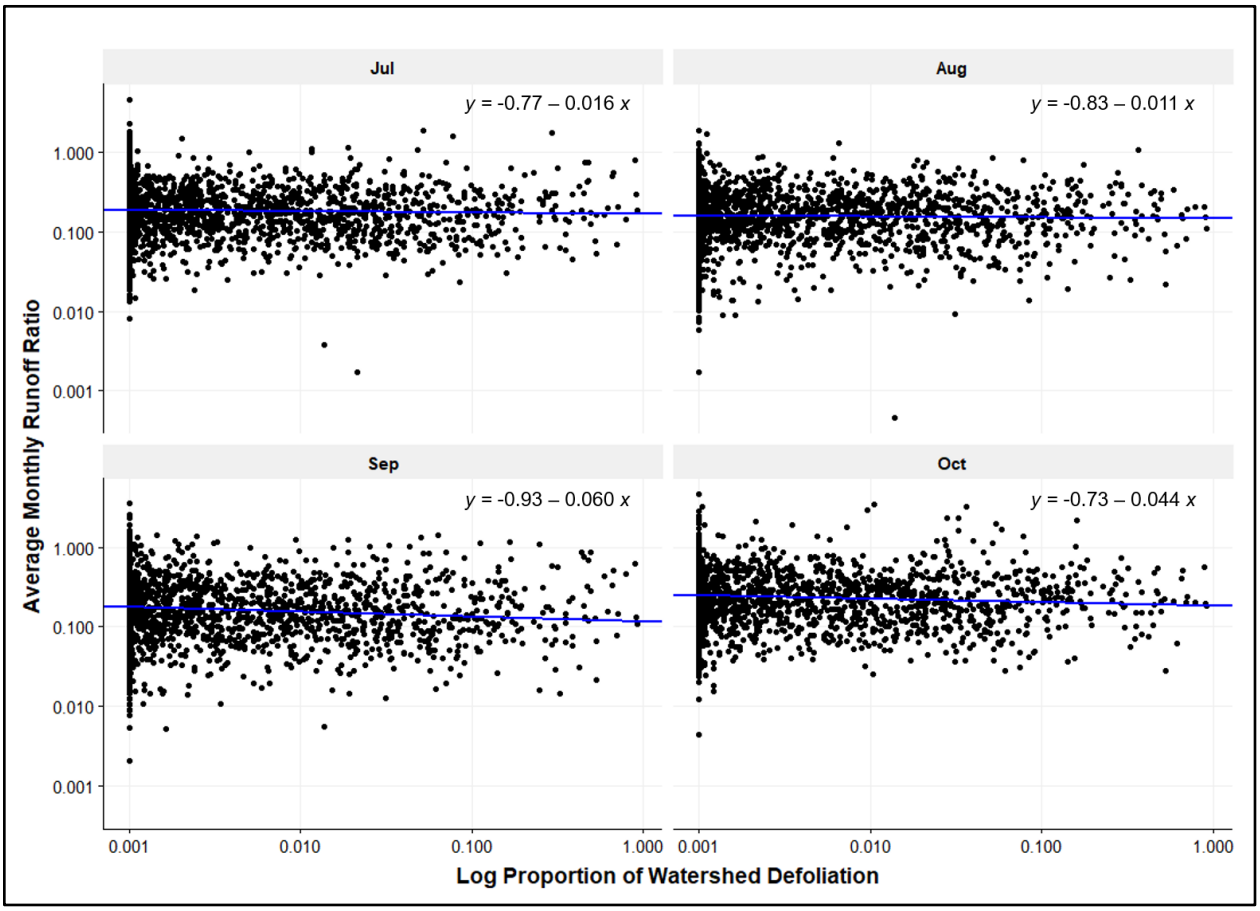

Figure S3-1. Runoff ratio by month. Average monthly runoff ratio by proportion of watershed with defoliation by month. Reported equation extracted from mixed effect model.

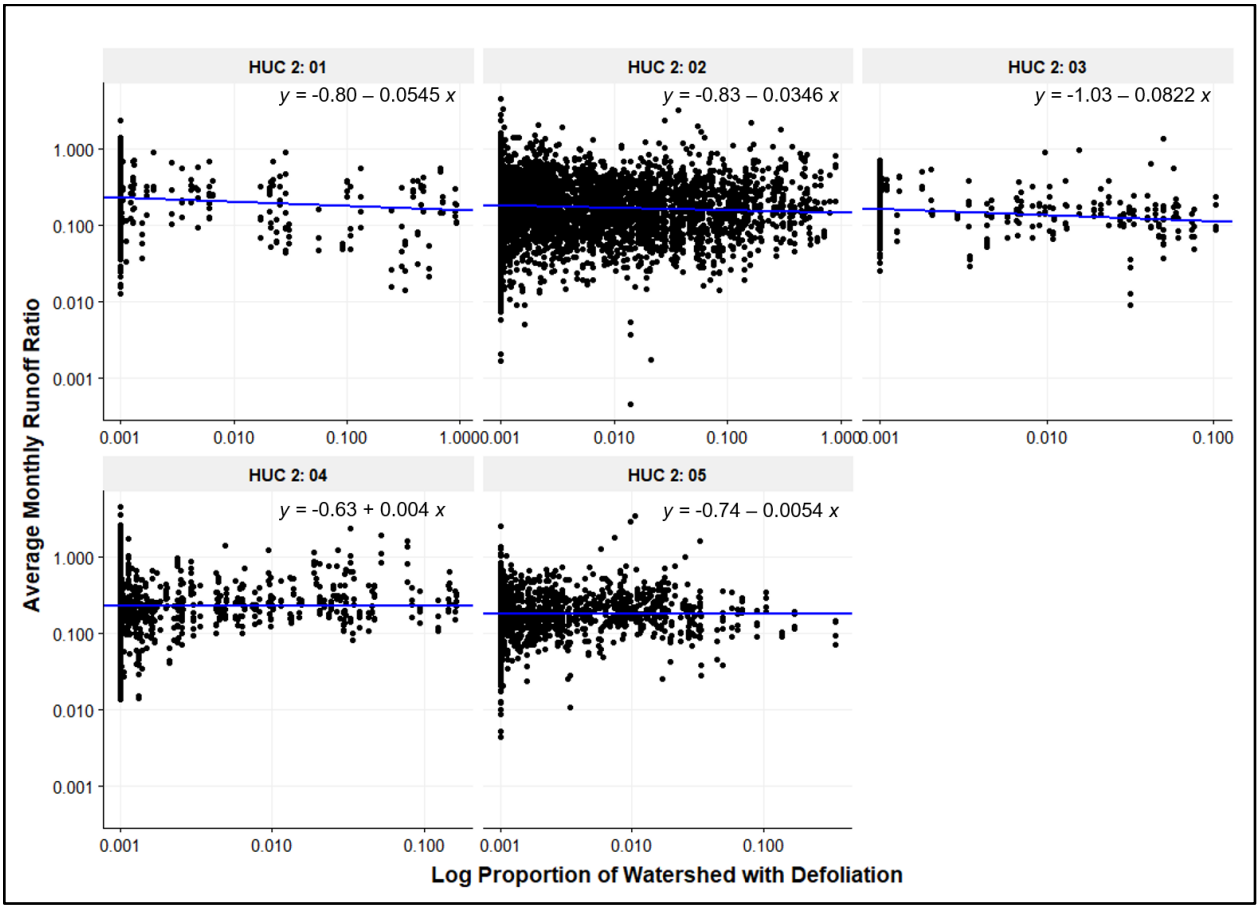

Figure S3-2. Runoff ratio by major water basin. Average monthly runoff ratio by proportion of watershed with defoliation by major water basin (2-digit HUC). Reported equation extracted from mixed effect model.

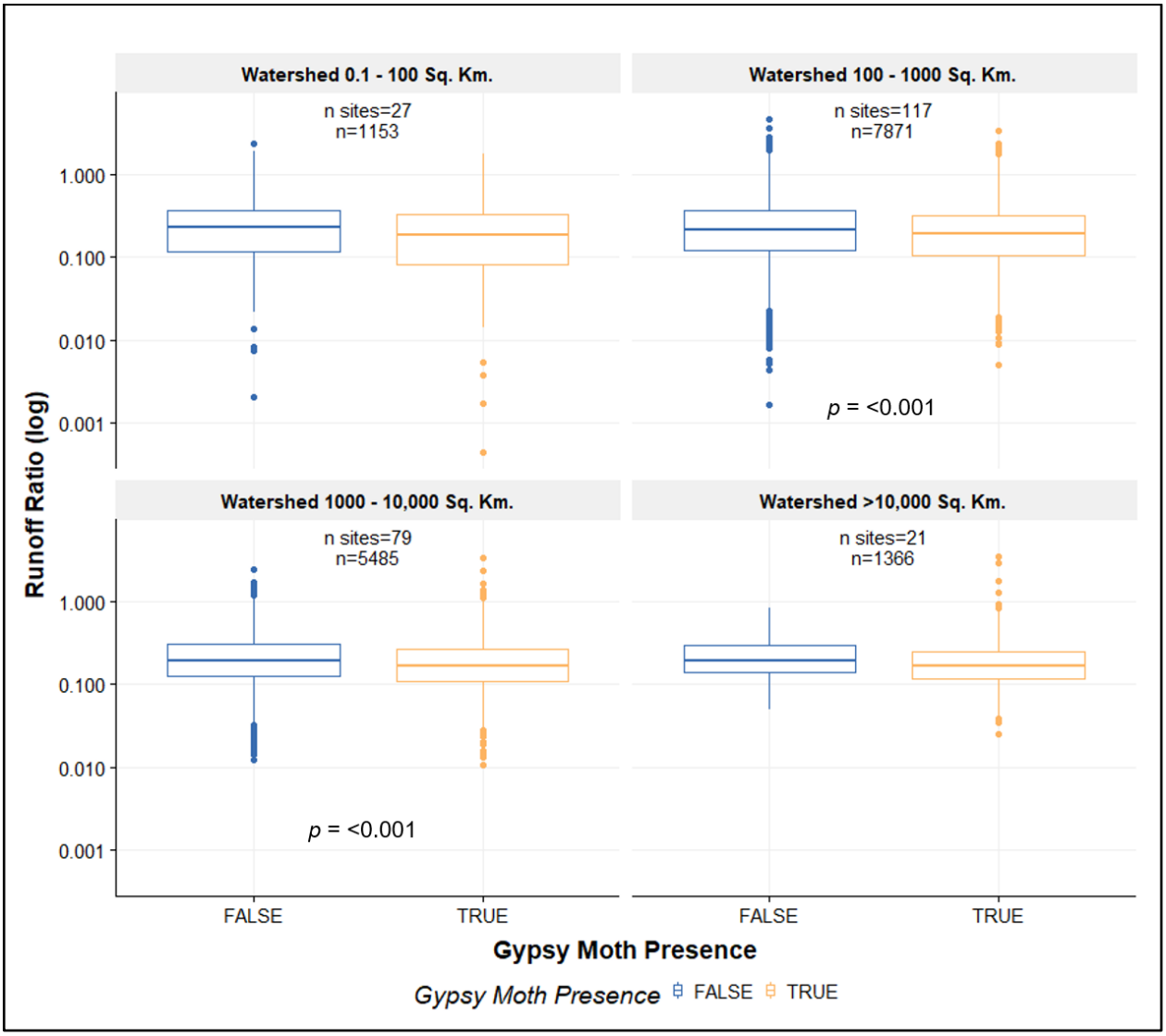

Figure S3-3. Runoff ratio by gypsy moth presence and watershed size. Runoff ratio by gypsy moth presence, split by watershed area. Associated *p*-values for significant Tukey tests included.

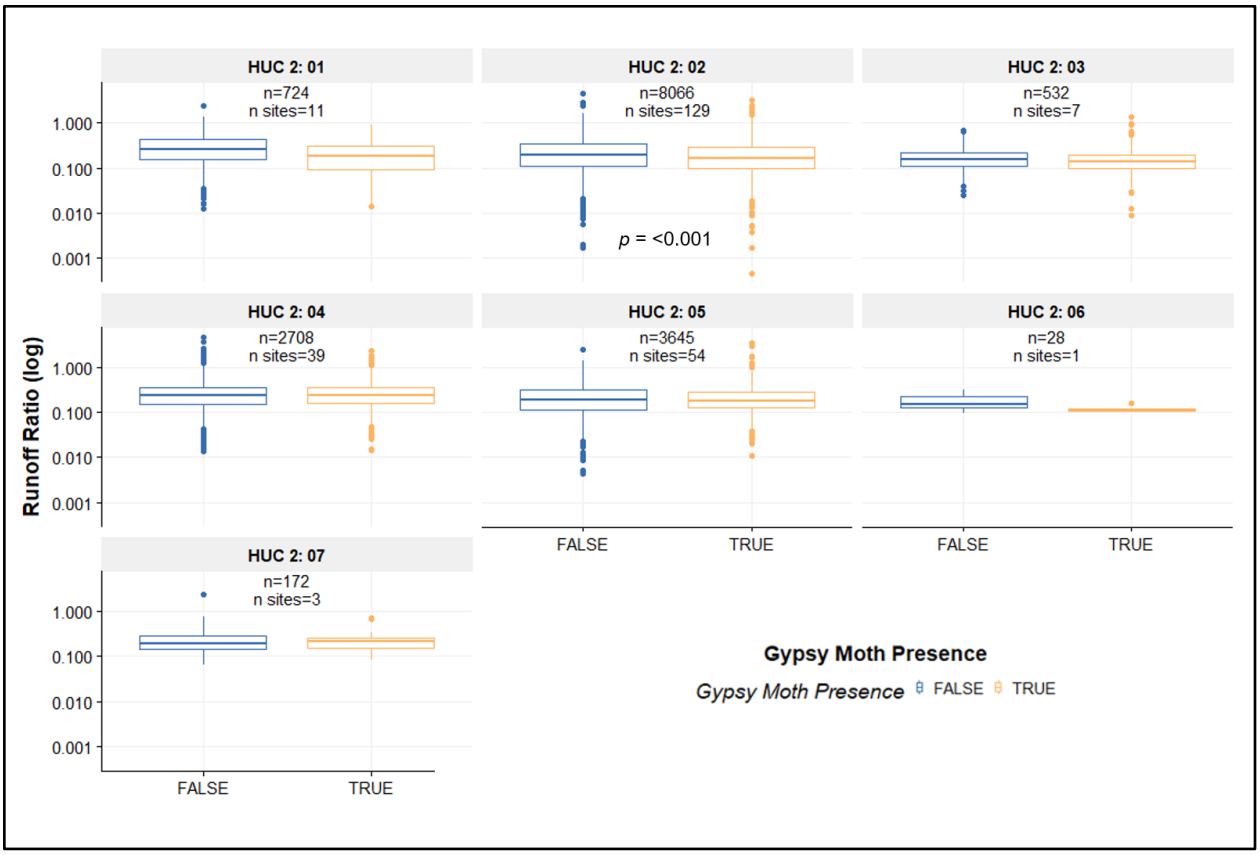

Figure S3-4. Runoff ratio by gypsy moth presence and major water basin. Runoff ratio by gypsy moth presence, split by major water basin (HUC 2). Associated *p*-values for significant Tukey tests included.

Table S3-2. Tukey test results. Results of Satterthwaite-approximated Tukey tests for each variable, split by month, major water basin (HUC), and watershed area. Dashes (-) indicate there were not enough records to conduct the Tukey test.

	Contrast	Estimate	SE	df	*t*-ratio	*p*-value
Discharge						
Month						
July	Gypsy Moth False - Gypsy Moth True	0.0234	0.0065	15758.11	3.62	0.007
August	Gypsy Moth False - Gypsy Moth True	0.0135	0.0065	15760.06	2.09	0.423
September	Gypsy Moth False - Gypsy Moth True	0.0304	0.0065	15759.23	4.69	0.0001
October	Gypsy Moth False - Gypsy Moth True	0.0348	0.0065	15758.93	5.32	0.000
2-Digit HUC						
01	Gypsy Moth False - Gypsy Moth True	0.0409	0.0184	15817.40	2.22	0.613
02	Gypsy Moth False - Gypsy Moth True	0.0309	0.0050	15826.66	6.22	0.000
03	Gypsy Moth False - Gypsy Moth True	0.0184	0.0181	15695.16	1.02	0.999
04	Gypsy Moth False - Gypsy Moth True	0.0092	0.0102	15842.18	0.90	1.000
05	Gypsy Moth False - Gypsy Moth True	0.0201	0.0083	15779.67	2.41	0.469
06	Gypsy Moth False - Gypsy Moth True	0.0586	0.1039	15597.21	0.56	1.000
07	Gypsy Moth False - Gypsy Moth True	0.0101	0.0399	15828.88	0.25	1.000
Watershed Size						
<100	Gypsy Moth False - Gypsy Moth True	-0.0071	0.0143	15319.71	-0.50	1.000
100-1000	Gypsy Moth False - Gypsy Moth True	0.0284	0.0055	15834.82	5.19	0.000

Table S3-2 continued

1000-10,000	Gypsy Moth False - Gypsy Moth True	0.0293	0.0061	15842.43	4.84	0.000
>10,000	Gypsy Moth False - Gypsy Moth True	0.0201	0.0133	15767.56	1.51	0.803
Temperature						
Month						
July	Gypsy Moth False - Gypsy Moth True	-0.04	0.11	4869.84	-0.40	1.000
August	Gypsy Moth False - Gypsy Moth True	0.02	0.11	4868.91	0.17	1.000
September	Gypsy Moth False - Gypsy Moth True	0.02	0.11	4869.15	0.15	1.000
October	Gypsy Moth False - Gypsy Moth True	0.13	0.11	4869.19	1.20	0.933
2-Digit HUC						
01	Gypsy Moth False - Gypsy Moth True	0.04	0.83	3850.67	0.05	1.000
02	Gypsy Moth False - Gypsy Moth True	-0.30	0.22	4393.54	-1.38	0.984
03	Gypsy Moth False - Gypsy Moth True	0.83	2.56	225.42	0.32	1.000
04	Gypsy Moth False - Gypsy Moth True	0.49	0.33	4862.95	1.49	0.970
05	Gypsy Moth False - Gypsy Moth True	-0.26	0.37	4778.86	-0.70	1.000
06	Gypsy Moth False - Gypsy Moth True	-1.44	2.13	4687.74	-0.68	1.000
07	Gypsy Moth False - Gypsy Moth True	-	-	-	-	-

Table S3-2 continued

Watershed Size						
<100	Gypsy Moth False - Gypsy Moth True	0.12	0.49	3611.45	0.24	1.000
100-1000	Gypsy Moth False - Gypsy Moth True	-0.11	0.25	4235.78	-0.44	1.000
1000-10,000	Gypsy Moth False - Gypsy Moth True	-0.22	0.23	4742.66	-0.94	0.982
>10,000	Gypsy Moth False - Gypsy Moth True	0.03	0.66	4985.32	0.05	1.000
DO						
Month						
July	Gypsy Moth False - Gypsy Moth True	-0.84	0.60	2318.99	-1.41	0.854
August	Gypsy Moth False - Gypsy Moth True	0.67	0.60	2318.95	1.13	0.951
September	Gypsy Moth False - Gypsy Moth True	0.39	0.59	2313.27	0.65	0.998
October	Gypsy Moth False - Gypsy Moth True	1.00	0.60	2311.18	1.66	0.710
2-Digit HUC						

Table S3-2 continued

01	Gypsy Moth False - Gypsy Moth True	-	-	-	-	-
02	Gypsy Moth False - Gypsy Moth True	0.71	0.66	2346.97	1.07	0.987
03	Gypsy Moth False - Gypsy Moth True	-	-	-	-	-
04	Gypsy Moth False - Gypsy Moth True	0.17	0.53	2314.42	0.31	1.000
05	Gypsy Moth False - Gypsy Moth True	-0.01	0.73	2331.67	-0.01	1.000
06	Gypsy Moth False - Gypsy Moth True	-	-	-	-	-
07	Gypsy Moth False - Gypsy Moth True	-	-	-	-	-
Watershed Size						
<100	Gypsy Moth False - Gypsy Moth True	1.29	1.73	2304.70	0.74	0.996
100-1000	Gypsy Moth False - Gypsy Moth True	1.10	0.64	2342.79	1.73	0.666
1000-10,000	Gypsy Moth False - Gypsy Moth True	-0.18	0.50	2335.01	-0.36	1.000
>10,000	Gypsy Moth False - Gypsy Moth True	0.20	1.01	2310.57	0.20	1.000

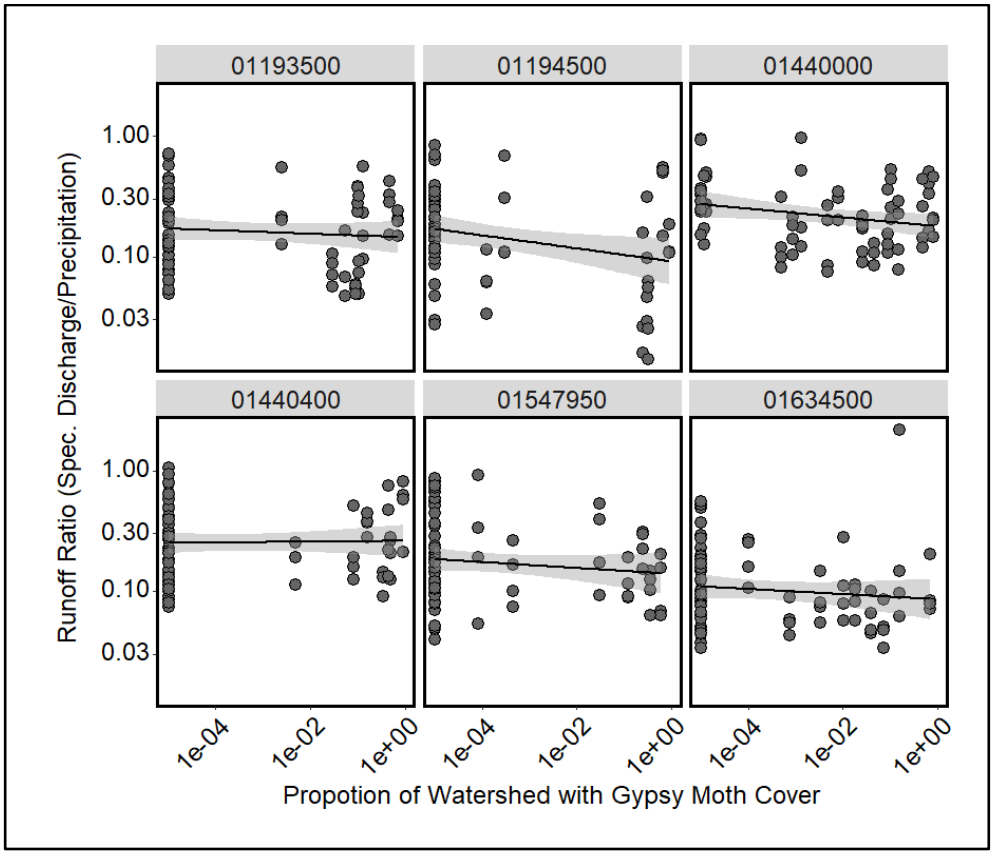

Figure S3-5. Runoff ratio by gypsy moth cover in the most defoliated sites. Runoff ratio versus proportion of watershed with gypsy moth defoliation at sites with at least 60% of watershed defoliated.

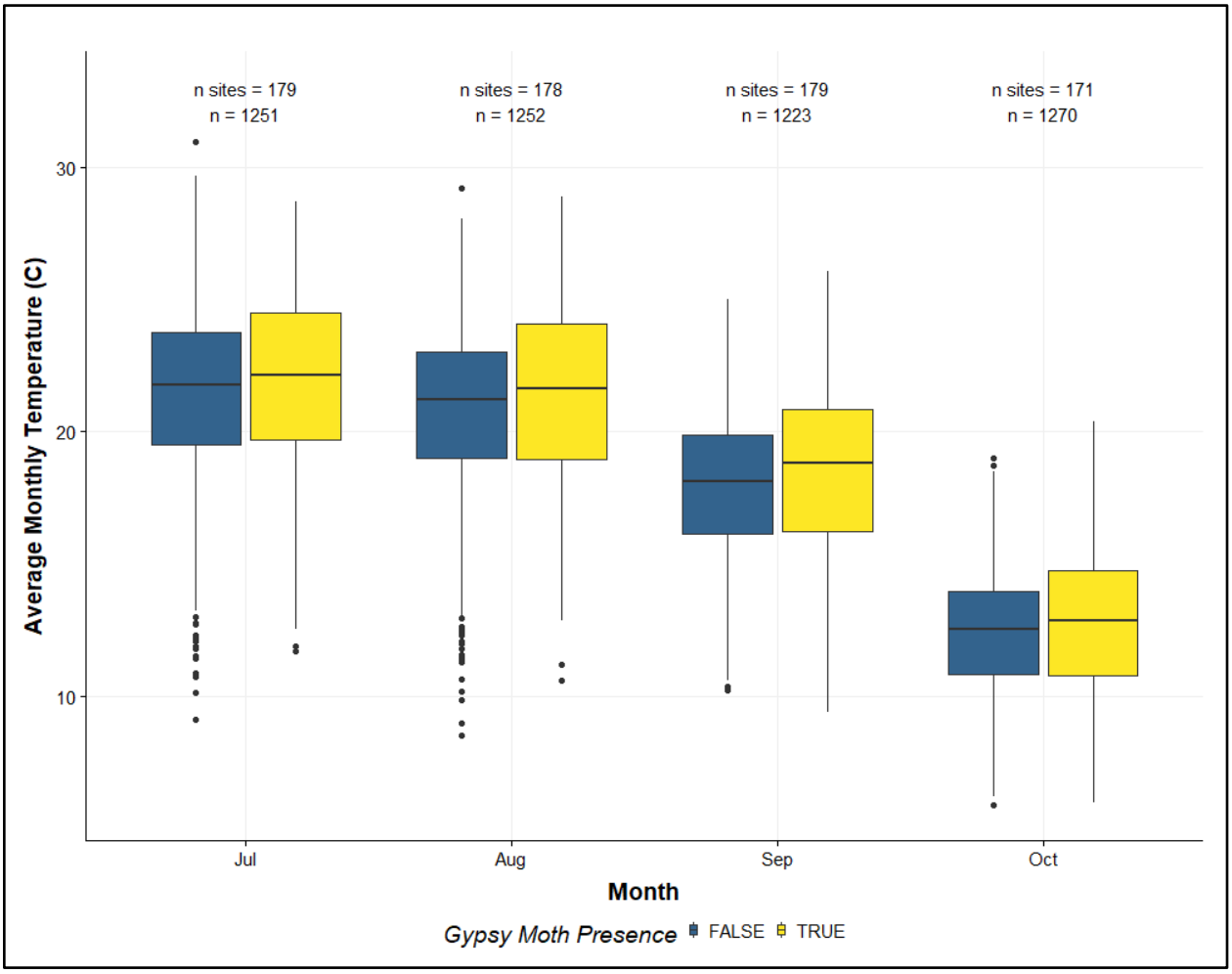

Figure S3-6. Temperature by month. Difference in observed temperature in watersheds with and without defoliation by month.

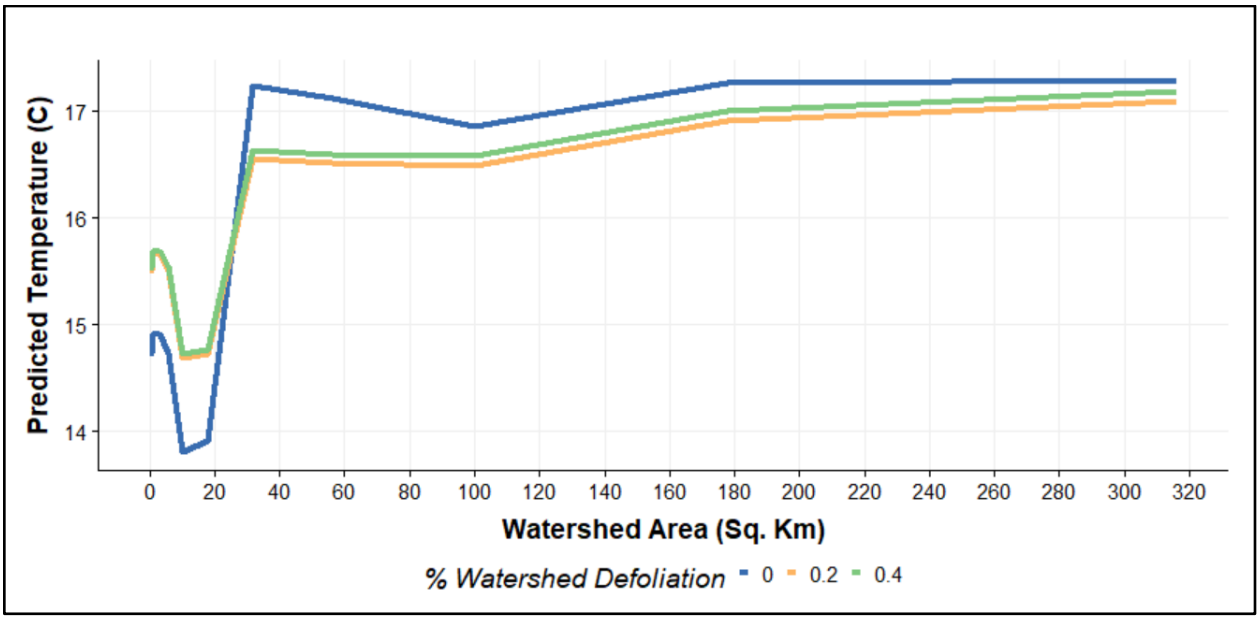

Figure S3-7. Expanded temperature predictions by watershed area. Random forest predicted temperature values at a range of smaller watershed sizes with three levels of defoliation intensity: 0%, 20%, and 40%.

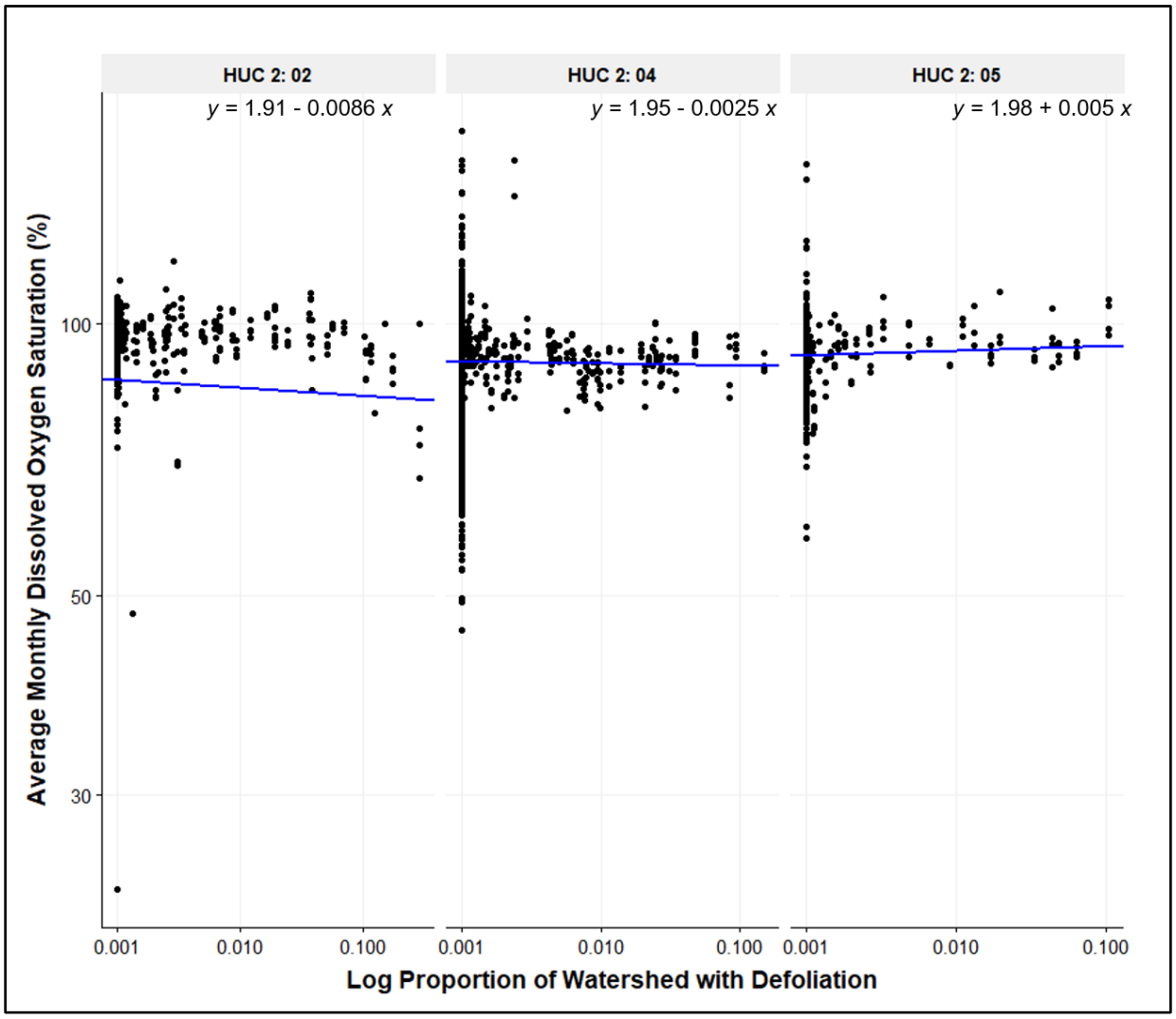

Figure S3-8. Dissolved oxygen by major water basin. Average monthly dissolved oxygen saturation (%) by defoliation intensity in major water basins with adequate dissolved oxygen samples to run mixed effects models. Reported equation extracted from mixed effect model

CHAPTER 4. CONCLUSIONS

The purpose of my research was to investigate two major components of invasion ecology - spread dynamics and ecosystem impacts**.** By utilizing large data sets, I was able to investigate both how a new non-native pest spreads and how defoliation from a well-known damaging forest pest affects the nation's water resources. I concluded from this research that human populations are an important driver of the spread of a new pest, even more so than the availability of host plants. This conclusion has important implications about the relationship between humans and invasive pests. We know that more non-native pests are being introduced due to travel and trade, and the knowledge that humans are also driving the spread of new pests post-establishment provides another link between human dynamics and spread dynamics. These pests then have complicated impacts to ecosystems that provide services humans rely heavily on. Thus, there is an inextricable link between human-mediated introduction and spread, and our own ability to utilize ecosystem services, such as availability of clean water.

These studies also highlight the importance of utilizing broad-scale data to investigate multi-regional trends and patterns when the data are available. In a world of increasing data collection and storage capacity, there are millions of records of ecological data on invasive pests, factors potentially impacting spread, and ecosystem resources. In Chapter 2, we identified trends not captured by small-scale studies, showing that regional and ecosystem-level patterns may differ. Thus, it is important to investigate multi-scale relationships and discover what we might miss at smaller scales. The topics I studied have rich available datasets for further investigation, and I hope to see future research on both the spread and invasion risk of spotted lanternfly, and broad-scale impacts of gypsy moth on water resources.

REFERENCES

Abdi H, Williams LJ (2010) Newman-Keuls test and Tukey test. Encyclopedia of Research Design, Thousand Oaks: 1-11.

Addy K, Gold AJ, Loffredo JA, Schroth AW, Inamdar SP, Bowden WB, Kellogg DQ, Birgand F (2018) Stream response to an extreme drought-induced defoliation event. Biogeochemistry 140: 199-215.

Arnell NW (1999) Climate change and global water resources. Global Environmental Change 9: S31-S49. https://doi.org/10.1016/S0959-3780(99)00017-5

Aukema JE, Leung B, Kovacs K, Chivers C, Britton KO, Englin J, Frankel SJ, Haight RG, Holmes TP, Liebhold AM, McCullough DG, von Holle B (2011) Economic impacts of non-native forest insects in the continental United States. PLoS ONE 6(9): e24587. doi:10.1371/journal.pone.0024587.

Aukema, J. E., McCullough, D. G., Von Holle, B., Liebhold, A. M., Britton, K., & Frankel, S. J. (2010). Historical accumulation of nonindige- nous forest pests in the continental United States. BioScience, 60, 886–897.

Barringer LE, Ciafre CM (2020) Worldwide feeding host plants of spotted lanternfly, with significant additions from North America. Environmental Entomology 49(5): 999-1011. https://doi.org/10.1093/ee/nvaa093

Barringer LE, Donovall LR, Spichiger SE, Lynch D, Henry D (2015) The first new world record of Lycorma delicatula (Insecta: Hemiptera: Fulgoridae). Entomological News 125: 20-23. DOI: 10.3157/021.125.0105.

Bearup LA, Maxwell RM, Clow DW, McCray JE (2014) Hydrological effects of forest transpiration loss in bark beetle-impacted watersheds. Nature Climate Change 4: 481-486. https://doi.org/10.1038/nclimate2198

Bechtold WA, Patterson PL (2005) The enhanced forest inventory and analysis program – national sampling design and estimation procedures. General Technical Report SRS-80. U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville. DOI: https://doi.org/10.2737/SRS-GTR-80.

Bernot MJ, Sobota DJ, Hall RO, Mulholland PJ, Dodds WK, Webster JR, Tank JL, Ashkenas LR, Cooper LW, Dahm CN, Gregory SV, Grimm NB, Hamilton SK, Johnson SL, McDowell WH, Meyer JL, Peterson B, Poole GC, Valett HM, Arango C, Beaulieu JJ, Burgin AJ, Crenshaw C, Helton AM, Johnson L, Merriam J, Niederlehner BR, O'Brien JM, Potter JD, Sheibley RW, Thomas SM, Wilson K (2010) Inter-regional comparison of land-use effects on stream metabolism. Freshwater Biology 55:1874–1890. https://doi.org/10.1111/j.1365-2427.2010.02422.x

Blackburn TM, Pysek P, Bacher S, Carlton JT, Duncan RP, Jarosik V, Wilson JRU, Richardson DM (2011) A proposed unified framework for biological invasions. Trends in Ecology & Evolution 26(7): 333-339. DOI: https://doi.org/10.1016/j.tree.2011.03.023.

Blodgett D (2019) nhdplusTools: Tools for Accessing and Working with the NHDPlus. https://code.usgs.gov/water/nhdplusTools

Bosch JM, Hewlett JD (1982) A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration. Journal of Hydrology 55: 3-23. https://doi.org/10.1016/0022-1694(82)90117-2

Breiman L (2001) Random forests. Machine Learning 45: 5-32. https://doi.org/10.1023/A:1010933404324

Campbell RW, Sloan RJ (1977) Forest stand responses to defoliation by the gypsy moth. Forest Science 23(S2) :a0001 -z0001. https://doi.org/10.1093/forestscience/23.s2.a0001

Ciesla WM (2000) Remote sensing in forest health protection. US Department of Agriculture, Forest Service, Forest Health Technology Enterprise Team. FHTET Report 00-03.

Collins S (1961) Benefits to the understory from canopy defoliation by gypsy moth larvae. Ecology 42: 836-838. https://doi.org/10.2307/1933521

Corbett ES, Lynch JA (1987) The gypsy moth - does it affect soil and water resources? Proceedings of Coping with the Gypsy Moth in the New Frontier, August 4-6.

Cornell (2021) New York State Integrated Pest Management: Spotted Lanternfly. https://nysipm.cornell.edu/environment/invasive-species-exotic-pests/spotted-lanternfly/

Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. International Journal of Climatology 28: 2031-2064. https://doi.org/10.1002/joc.1688

Danladi Bello AA, Hashim NB, Mohd Haniffah MR (2017) Predicting impact of climate change on water temperature and dissolved oxygen in tropical rivers. Climate 5: 58. https://doi.org/10.3390/cli5030058

Dara SK, Barringer L, Arthurs SP (2015) Lycorma delicatula (Hemiptera: Fulgoridae): A new invasive pest in the United States. Journal of Integrated Pest Management 6(1): 1-6. DOI: https://doi.org/10.1093/jipm/pmv021.

De Cicco LA, Hirsch RM, Lorenz D, Watkins WD (2018) dataRetrieval: R packages for discovering and retrieving water data available from Federal hydrologic web services. doi:10.5066/P9X4L3GE

Ding J, Wu Y, Zheng H, Fu W, Reardon R, Liu M (2006) Assessing potential biological control of the invasive plant, tree-of-heaven, Ailanthus altissima. Biocontrol and Technology 16: 547-566. DOI: https://doi.org/10.1080/09583150500531909.

Dingman LS (2015) Physical Hydrology: Third Edition. Waveland Press.

EDDMapS (2021) Early Detection & Distribution Mapping System. The University of Georgia Center for Invasive Species and Ecosystem Health. https://www.eddmaps.org/

Eddy D (2018) Spotted Lanternfly Found in Virginia Grapes. Growing Produce. https://www.growingproduce.com/fruits/grapes/spotted-lanternfly-found-virginia-grapes/#:~:text=For%20the%20first%20time%2C%20spotted,in%20Frederick%20County%20in%20January

Eshleman KN, Morgan RP, Webb JR, Deviney FA, Galloway JN (1998) Temporal patterns of nitrogen leakage from mid-Appalachian forested watersheds: Role of insect defoliation. Water Resources Research 34(8): 2005-2016. https://doi.org/10.1029/98WR01198

Fahrner S, Aukema BH (2018) Correlates of spread rates for introduced insects. Global Ecology and Biogeography 27(6): 734-743. https://doi.org/10.1111/geb.12737

Fei S, Morin RS, Oswalt CM, Liebhold AM (2019) Biomass losses resulting from insect and disease invasions in US forests. Proceedings of the National Academy of Sciences 116(35): 17371-17376. DOI: https://doi.org/10.1073/pnas.1820601116.

Feret PP (1985) Ailanthus: variation, cultivation, and frustration. Journal of Arboriculture 11(12): 361-368.

Francese JA, Cooperband MF, Murman KM, Cannon SL, Booth EG, Devine SM, Wallace MS (2020) Developing traps for the spotted lanternfly, Lycorma delicatula (Hemiptera: Fulgoridae). Environmental Entomology 49(2): 269-276. DOI: https://doi.org/10.1093/ee/nvz166.

Gilbert M, Gregoire JC, Freise JF, Heitland W (2004) Long-distance dispersal and human population density allow the prediction of invasive patterns in the horse chestnut leafminer Cameraria ohridella. Journal of Animal Ecology 73(3): 459-68. DOI: https://doi.org/10.1111/j.0021-8790.2004.00820.x.

Gilbert M, Liebhold, A (2010) Comparing methods for measuring the rate of spread of invading populations. Ecography 33(5): 809-817. DOI: https://doi.org/10.1111/j.1600-0587.2009.06018.x.

Hart EM, Bell K (2015) prism: Download data from the Oregon prism project. R package version 0.0.6. http://github.com/ropensci/prism. DOI: 10.5281/zenodo.33663

Hastings A, Cuddington K, Davies KF, Dugaw CJ, Elmendorf S, Freestone A, Harrison S, Holland M, Lambrinos J, Malvadkar U, Melbourne BA, Moore K, Taylor C, Thomson D (2005) The spatial spread of invasions: new developments in theory and evidence. Ecology Letters 8(1): 91-101. DOI: https://doi.org/10.1111/j.1461-0248.2004.00687.x.

Hellmann JJ, Byers JE, Bierwagen BG, Dukes JS (2008) Five potential consequences of climate change for invasive species. Conservation Biology 22(3): 534-543. https://doi.org/10.1111/j.1523-1739.2008.00951.x

Hosen JD, Aho KS, Appling AP, Creech EC, Fair JH, Hall Jr. RO, Kyzivat ED, Lowenthal RS, Matt S, Morrison J, Saiers JE, Shanley JB, Weber LC, Yoon B, Raymond PA (2019) Enhancement of primary production during drought in a temperate watershed is greater in larger rivers than headwater streams. Limnology and Oceanography 64: 1458-1472. https://doi.org/10.1002/lno.11127

Huebner CD (2003) Vulnerability of oak-dominated forests in West Virginia to invasive exotic plants: temporal and spatial patterns of nine exotic species using herbarium records and land classification data. Castanea 68(1): 1-14.

Hulme PE (2009) Trade, transport and trouble: Managing invasive species pathways in an era of globalization. Journal of Applied Ecology 46: 10-18.

Johnson EW, Wittwer D (2006) Aerial detection surveys in the United States. Proceedings of Monitoring Science and Technology Symposium: Unifying Knowledge for Sustainability in the Western Hemisphere. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. RMRS-P-42CD: 809-811.

Jules ES, Kauffman MJ, Ritts WD, Carroll AL (2002) Spread of an invasive pathogen over a variable landscape: A nonnative root rot on Port Orford cedar. Ecology 83(11): 3167-3181. https://doi.org/10.1890/0012-9658(2002)083[3167:SOAIPO]2.0.CO;2

Kaushal SS, Gold AJ, Bernal S, Tank JL (2018) Diverse water quality responses to extreme climate events: an introduction. Biogeochemistry 141: 273-279. https://doi.org/10.1007/s10533-018-0527-x

Kortmann RW, Cummins E (2018) Climate change in the Northeast: What might it mean to water quality management? The Journal of New England Water Works Association: 236-237.

Kovacs KF, Haight RG, McCullough DG, Mercader RJ, Siegert NW, Liebhold AM (2010) Cost of potential emerald ash borer damage in U.S. communities, 2009-2019. Ecological Economics 69(3): 569-578. DOI: 10.1016/j.ecolecon.2009.09.004.

Kreuzwieser J, Gessler A (2010) Global climate change and tree nutrition: influence of water availability. Tree Physiology 30(9): 1221-1234. https://doi.org/10.1093/treephys/tpq055

Lance DR, Elkinton JS, Schwalbe CP (1987) Microhabitat and temperature effects explain accelerated development during outbreaks of the gypsy moth (Lepidoptera: Lymantriidae). Environmental entomology 16(1): 202-205.

Lenth RV (2021) emmeans: Estimated marginal means, aka least-squares means. R package version 1.5.4. https://CRAN.R-project.org/package=emmeans

Lewis GP, Likens GE (2007) Changes in stream chemistry associated with insect defoliation in a Pennsylvania hemlock-hardwoods forest. Forest Ecology and Management 238: 199-211. https://doi.org/10.1016/j.foreco.2006.10.013

Li JL, Johnson SL, Sobota JB (2011) Three responses to small changes in stream temperature by autumn-emerging aquatic insects. Journal of the North American Benthological Society 30: 474- 484.

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2(3): 18-22.

Liebhold AM (2012) Forest pest management in a changing world. International Journal of Pest Management 58(3): 289-295. DOI: https://doi.org/10.1080/09670874.2012.678405.

Liebhold AM, Bascompte J (2003) The Allee effect, stochastic dynamics and the eradication of alien species. Ecology Letters 6: 133-140. DOI: 10.1046/j.1461-0248.2003.00405.x.

Liebhold AM, Kean JM (2019) Eradication and containment of non-native forest insects: successes and failures. Journal of Pest Science 92(1): 83-91. DOI: https://doi.org/10.1007/s10340-018-1056-z.

Liebhold AM, MacDonald WL, Bergdahl D, Mastro VC (1995) Invasion by exotic forest pests: A threat to forest ecosystems. Forest Science 41: 1-58.

Liebhold AM, Mastro V, Schaefer PW (1989) Learning from the legacy of Leopold Trouvelot. Bulletin of the Entomological Society America 35: 20-21.

Liebhold AM, McCullough D, Blackburn L, Frankel S, Von Holle B, Aukema J (2013) A highly aggregated geographical distribution of forest pest invasions in the USA. Diversity and Distributions 19(9): 1208-1216. DOI: https://doi.org/10.1111/ddi.12112.

Liebhold AM, Tobin PC (2008) Population ecology of insect invasions and their management. Annual Review of Entomology 53: 387-408. DOI: https://doi.org/10.1146/annurev.ento.52.110405.091401.

Lovett GM, Canham CD, Arthur MA, Weathers KC, Fitzhugh RD (2006) Forest ecosystem responses to exotic pests and pathogens in eastern North America. BioScience 56(5): 395-405. DOI: https://doi.org/10.1641/0006-3568(2006)056[0395:FERTEP]2.0.CO;2.

Moulton TL (2018) rMR: Importing data from Loligo Systems Software, calculating metabolic rates and critical tensions. R package version 1.1.0. https://CRAN.R-project.org/package=rMR

Mulholland PJ, Fellows CS, Tank JL, Grimm NB, Webster JR, Hamilton SK, Martí E, Ashkenas L, Bowden WB, Dodds WK, McDowell WH, Paul MJ, Peterson BJ (2001). Inter-biome comparison of factors controlling stream metabolism. Freshwater Biology 46. https://doi.org/10.1046/j.1365-2427.2001.00773.x

Orlova-Bienkowskaja MJ, Bienkowski AO (2018) Modeling long-distance dispersal of emerald ash borer in European Russia and prognosis of spread of this pest to neighboring countries within next 5 years. Ecology and Evolution 8(18): 9295-9304. https://doi.org/10.1002/ece3.4437

Parra G, Moylett H, Bulluck R (2017) Technical working group summary report: spotted lanternfly, Lycorma delicatula (White, 1845). USDA APHIS Plant Protection and Quarantine Center for Plant Health Science and Technology. https://www.agriculture.pa.gov/Plants_Land_Water/PlantIndustry/Entomology/spotted_lanternfly/research/Documents/SLF%20TWG%20Report%20020718%20final.pdf

PRISM Climate Group (2004) Oregon State University. http://prism.oregonstate.edu.

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

Roberts SW, Tankersley Jr. R, Orvis KH (2009) Assessing the potential impacts to riparian ecosystems resulting from hemlock mortality in Great Smoky Mountains National Park. Environmental Management 44(2): 335-345. 10.1007/s00267-009-9317-5

Rüegg J, Conn CC, Anderson EP, Battin TJ, Bernhardt ES, Boix Canadell M, Bonjour SM, Hosen JD, Marzolf NS, Yackulic CB (2020) Thinking like a consumer: Linking aquatic basal metabolism and consumer dynamics. Limnology and Oceanography Letters 6(1): 1-17. https://doi.org/10.1002/lol2.10172

Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG (2001) The population biology of invasive species. Annual Review of Ecology and Systematics 32: 305-332. DOI: https://doi.org/10.1146/annurev.ecolsys.32.081501.114037.

Schäfer KVR, Renninger HJ, Clark KL, Medvigy D (2014) Hydrological responses to defoliation and drought of an upland oak/pine forest. Hydrological Processes 28: 6113-6123. https://doi.org/10.1002/hyp.10104

Scheid L (2020) Spotted lanternfly egg masses found in Maine were traced to Pennsylvania. Reading Eagle, [online] https://www.readingeagle.com/news/spotted-lanternfly-egg-masses-found-in-maine-were-traced-to-pennsylvania/article_85071664-12e0-11eb-837e-a7e0cd5a2f28.html

Schwarze R, Beudert B (2009) Analyses of flood generation and water budget in a forest catchment impacted by a bark-beetle outbreak. Hydrologie und Wasserbewirtschaftung 53: 236-249.

Scoles S, Anderson S, Turton D, Miller E (1996) Forestry and water quality: a review of watershed research in the Ouachita Mountains. Oklahoma State University, Oklahoma Cooperative Extension Service, Division of Agricultural Sciences and Natural Resources: 1-29.

Senay GB, Bohms S, Singh RK, Gowda PH, Velpuri NM, Alemu H, Verdin, JP (2013) Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the SSEB approach. JAWRA Journal of the American Water Resources Association 49(3): 577-591. https://doi.org/10.1111/jawr.12057

Sharov AA, Liebhold AM (1998). Bioeconomics of managing the spread of exotic species with barrier zones. Ecological applications, 8: 833-845. DOI: 10.1111/j.0272-4332.2004.00486.x.

Shigesada N, Kawasaki K, Takeda Y (1995) Modeling stratified diffusion in biological invasions. The American Naturalist 146: 229-251. DOI: https://doi.org/10.1086/285796.

Simberloff D, Von Holle B (1999) Positive interactions of nonindigenous species: Invasional meltdown? Biological Invasions 1: 21-32. DOI: https://doi.org/10.1023/A:1010086329619.

Skellam JG (1951) Random dispersal in theoretical populations. Biometrika 38: 196-218.

Smith-Tripp S, Griffith A, Pasquarella V, Matthes JH (in review) Impacts of a regional multi-year insect defoliation event on seasonal runoff ratios and instantaneous streamflow characteristics. https://doi.org/10.31223/X5901K

Stednick JD (1996) Monitoring the effects of timber harvest on annual water yield. Journal of Hydrology 176: 79-95. https://doi.org/10.1016/0022-1694(95)02780-7

Tarzwell CM (1970) Thermal requirements to protect aquatic life. Journal of the Water Pollution Control Federation 42 (5): 824-828.

Thomas L, Reyes EM (2014) Tutorial: Survival estimation for Cox regression models with time-varying coefficients using SAS and R. Journal of Statistical Software 61(CS1): 1-23.

Trenberth KE, Guillemot CJ (1996) Physical processes involved in the 1988 drought and 1993 floods in North America. Journal of Climate 9: 1288-98. https://doi.org/10.1175/1520-0442(1996)009<1288:PPIITD>2.0.CO;2

Twery MJ (1991) Effects of defoliation by gypsy moth. Proceedings of USDA Interagency Gypsy Moth Research Review 1990. USDA Forest

Urban JM (2019) Perspective: Shedding light on spotted lanternfly impacts in the USA. Pest Management Science 76(1): 10-17. DOI: https://doi.org/10.1002/ps.5619.

U.S. Congress, Office of Technology Assessment (1993) Harmful Non-Indigenous Species in the United States. OTA-F-565. Washington, DC: U.S. Government Printing Office.

US Geological Survey (2019) Simplified Surface Energy Balance Actual Evapotranspiration. https://earlywarning.usgs.gov/ssebop/modis

USDA Animal and Plant Health Inspection Services (USDA APHIS) (2018) Spotted lanternfly eradication program in select counties of Pennsylvania, Supplemental Environmental Assessment. U. S. Department of Agriculture, Animal and Plant Health Inspection Service, Plant Protection and Quarantine, Riverdale, MD. https://www.aphis.usda.gov/plant_health/ea/downloads/2018/slf-pa-supplemental-ea.pdf

USDA Forest Service (2016) USFS Percent Tree Canopy (Cartographic Version). Edition 1.0. Salt Lake City, UT. https://data.fs.usda.gov/geodata/rastergateway/treecanopycover/

USDA Forest Service (2019) Forest Inventory and Analysis National Core Field Guide. Volume I: Field Data Collection Procedures for Phase 2 Plots, Version 9.0. https://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/2019/core_ver9-0_10_2019_final_rev_2_10_2020.pdf

Uyi O, Keller JA, Johnson A, Long D, Walsh B, Hoover K (2020) Spotted lanternfly (Hemiptera: Fulgoridae) can complete development and reproduce without access to the preferred host, Ailanthus altissima. Environmental Entomology 49(5): 1185-1190. DOI: https://doi.org/10.1093/ee/nvaa083.

Wakie TT, Neven LG, Yee WL, Lu Z (2020) The establishment risk of Lycorma delicatula (Hemiptera: Fulgoridae) in the United States and globally. Journal of Economic Entomology 113: 306-314. DOI: https://doi.org/10.1093/jee/toz259.

Ward SF, Fei S, Liebhold AM (2020) Temporal dynamics and drivers of landscape-level spread by emerald ash borer. Journal of Applied Ecology 57: 1020-1030. DOI: https://doi.org/10.1111/1365-2664.13613.

Webb JR, Cosby BJ, Deviney FA, Eschleman Jr. KN, Galloway JN (1995) Change in the acid-base status of an Appalachian mountain catchment following forest defoliation by the gypsy moth. Water, Air, and Soil Pollution 85: 535-540.

Webster JR, Morkeski K, Wojculeswki CA, Niederlehner BR, Benfield EF, Elliott KJ (2012) Effects of hemlock mortality on streams in the southern Appalachian mountains. The American Midland Naturalist 168: 112-131.

Woltemade CJ (2017) Stream temperature spatial variability reflects geomorphology, hydrology, and microclimate: Navarro River watershed, California. The Professional Geographer 69: 177-190. https://doi.org/10.1080/00330124.2016.1193032

Yemshanov D, Koch FH, Ducey M, Koehler K (2012) Trade-associated pathways of alien forest insect entries in Canada. Biological Invasions 14: 797812. DOI: https://doi.org/10.1007/s10530-011-0117-5.