



Previous year outbreak conditions and spring climate predict spruce budworm population changes in the following year

Mingke Li*, David A. MacLean, Chris R. Hennigar, Jae Ogilvie

University of New Brunswick, Faculty of Forestry and Environmental Management, P.O. Box 4400, Fredericton, NB E3B 5A3, Canada



ARTICLE INFO

Keywords:

Choristoneura fumiferana
Population prediction
Outbreak initiation
Simultaneous autoregression
Spatial autocorrelation

ABSTRACT

We determined effects of local spruce budworm (*Choristoneura fumiferana* Clem.; SBW) population level, proximity to sites with high SBW populations, insecticide spray, and environmental variables on SBW populations from 2014 to 2018, the outbreak initiation period in northern New Brunswick, Canada. SBW second instar larvae (L2) per branch data collected at 1100–2000 sample points per year were used to create annual interpolated population rasters. Fishnet sample points extracted from these rasters were overlaid with georeferenced layers of 46 possible predictor variables including forest composition, climate, topography, site quality, and insecticide treatment. Results showed that local SBW population in the previous year, proximity to sites with high SBW populations, and early spring climate were consistently the most important predictors over the 5 study years. Simultaneous autoregressive models were used to address spatial autocorrelation when forecasting the SBW L2 population, and a linear mixed effects model was fit to aggregate data for 2015–2018. The models reduced spatial dependence in the residuals, and explained 68–79% of variance in annual L2 levels and 53% of variance over the 4 years combined. Sensitivity analysis showed that locations with 5–10 more SBW L2 per branch than observed values, or 20–40 km closer to high population sites in the previous year could have up to 24 more L2 in the current year. Cumulative degree days in April helped to estimate the upper and lower bounds of the population. Expansion and retraction of SBW outbreak initiation were mathematically described. Understanding which variables influence SBW outbreak initiation and population level assists in design of small area target-specific insecticide spray applications and helps focus SBW L2 sampling on predicted outbreak hot spots.

1. Introduction

Synchrony in insect population fluctuations often occurs at the landscape scale, and such large-scale synchronous increases in insect population density during one year and persistence over subsequent years can lead to rising insect populations and outbreak initiation (Royama et al., 2005; Liebhold et al., 2012; Bouchard et al., 2018). Several underlying mechanisms of spatial synchrony of insect outbreaks and local population fluctuations have been proposed, including regional stochasticity, dispersal, and trophic interactions (Régnière and Lysyk, 1995; Myers, 1998; Williams and Liebhold, 2000; Liebhold et al., 2012). Research on various insects has shown how seasonal drought can accelerate outbreak initiation (e.g., spruce beetle (*Dendroctonus rufipennis*); Hart et al., 2017), stands with more abundant host species attract insect initial attack (e.g., mountain pine beetle (*Dendroctonus ponderosae* Hopkins); Klutsch et al., 2009), and how parasitism produces regional insect pest population patterns during the outbreak

onset phase (e.g., western tussock moth (*Orgyia vetusta*); Maron et al., 2001). From a forest management viewpoint, spread of an insect outbreak can be controlled or slowed by altering forest characteristics that have major effects on outbreak development (Robert et al., 2018). Alternative strategies will be required if influence of biotic factors is weak (e.g., Bouchard and Auger, 2014).

Major spruce budworm (*Choristoneura fumiferana* Clem.; SBW) outbreaks in balsam fir (*Abies balsamea* L. Mill.) and spruce (*Picea* spp. A. Dietr.) forests are the dominant natural disturbance in eastern North America. SBW outbreaks beginning in the 1910s, 1940s, and 1970s damaged 10, 25, and 55 million ha, respectively, across eastern Canada (Blais, 1983). Another SBW outbreak began in about 2006 in northern Québec, Canada, reaching 8.2 million ha of defoliation by 2018 (QMRNF, 2018). SBW populations spread southward and reached New Brunswick, Canada in 2014. In this paper, we analyze the spatial patterns of SBW population increases in New Brunswick from 2014 to 2018, and determine the influence of over 40 variables on SBW

* Corresponding author.

E-mail address: Mingke.Li@unb.ca (M. Li).

<https://doi.org/10.1016/j.foreco.2019.117737>

Received 10 September 2019; Received in revised form 29 October 2019; Accepted 30 October 2019

Available online 04 December 2019

0378-1127/ © 2019 Elsevier B.V. All rights reserved.

increases, assessed using overwintering second instar larvae (L2) samples on 4500–6000 tree branches per year.

Previous research has shown that forest composition, climate, topography, and site quality all can influence SBW outbreak dynamics, but with varying relative importance and consistency of relationships. Forest composition influences include high SBW population levels occurring when eggs were deposited in dense, mature softwood stands rather than young and open mixed stands (Greenbank, 1957). Also, higher abundance of balsam fir corresponded to more severe SBW defoliation (Bouchard and Auger, 2014), while higher abundance of hardwoods (broadleaved trees) resulted in less severe defoliation (Su et al., 1996; Zhang et al., 2018), reduced tree volume loss (Needham et al., 1999), and decreased radial growth reduction (Campbell et al., 2008). Climate influences include dry, warm summers increasing larval survival, potentially leading to higher larval populations during the flight season (Greenbank, 1957; Royama, 1984; Régnière and Nealis, 2007). High frequencies of defoliation were associated with dry June and cool spring periods (Candau and Fleming, 2005). Spring and summer degree days showed a strong influence on outbreak duration, severity, and spatial variability; locations with warmer spring temperatures and higher cumulative degree-days generally experienced shorter outbreak duration and lower severity (Gray, 2008, 2013). Proximity to previously defoliated areas positively influenced, while elevation negatively influenced SBW defoliation during the onset of the latest SBW outbreak in Québec (Bouchard and Auger, 2014). Stands on moist/rich sites had 19% higher defoliation than stands on wet/poor sites (MacKinnon and MacLean, 2003). Site quality has been shown to influence SBW-caused mortality, with higher mortality on xeric (85% mortality) and hydric (75%) than on mesic (45%) and sub-hydric (27%) fir stands (Dupont et al., 1991). Research on a SBW outbreak in northern British Columbia, Canada noted showed that increasing stand volume, higher current needle biomass, and proximity to the nearest river or nearest defoliation led to increased likelihood of onset of outbreaks (Magnussen et al., 2004).

An experiment to suppress rising SBW populations before major defoliation occurs, termed an ‘early intervention strategy’ (EIS), has been underway in New Brunswick, Canada from 2014 to 2019 (MacLean et al., 2019). This is the first attempt of area-wide management (all areas within the jurisdiction of the province of New Brunswick) of an endemic forest insect population. The EIS approach includes intensive monitoring of overwintering SBW to detect ‘hot spots’ of low but rising populations and targeted insecticide treatment to prevent spread. Following 5 years of over 420,000 ha of EIS treatments of low but increasing SBW populations, carried out in June of each year, SBW L2 levels across northern New Brunswick were considerably lower than populations in adjacent Québec (MacLean et al., 2019). SBW populations in blocks treated with *Bacillus thuringiensis* or tebufenozide insecticide were consistently reduced and generally did not require treatment in the subsequent year. SBW populations observed in a given year are the basis for insecticide treatment strategies in the following year (MacLean et al., 2019). Following 5 years of tests, EIS appears, thus far, to be effective in reducing the SBW outbreak (MacLean et al., 2019). Forecasting regions with high-level SBW populations in the following year is a critical basis for sampling and insecticide or other pest management tactics, and understanding the relationship of stand, site, and climate factors to population increases is an important component.

When performing statistical tests on ecological data, such as SBW L2 population data, the assumption of independence of residuals is often violated, increasing Type I errors and biasing the estimation of regression parameters (Dormann et al., 2007; Beale et al., 2010). Lack of independence in the residuals can arise because either response or predictor variables are spatially autocorrelated, i.e., objects that are closer to each other have a tendency to be more similar than those that are further apart (Sokal and Oden, 1978; Dale and Fortin, 2014). Spatially-structured residuals may be due to: 1) omission of important

predictors in the model, 2) inappropriate model specification, or 3) a mismatch between spatial patterns of the response and predictors (Dale and Fortin, 2014). Accordingly, solutions to avoid statistical issues resulting from spatial autocorrelation include: 1) incorporating all necessary ecological predictors to fit the model, 2) adopting generalized linear mixed models where a random effect is used to account for the effects of location, and 3) spatial regression that includes ‘space’ explicitly as an additional variable (Beale et al., 2010; Dale and Fortin, 2014). Spatial regression takes into account spatial dependence in the data by adding a lagged response variable (autoregression; Anselin, 1988) or lagged covariates (i.e. autocovariate regression; Augustin et al., 1996).

The objectives of this study were to: 1) determine variables that influenced SBW L2 populations from 2014 to 2018, the outbreak initiation phase in northern New Brunswick, and 2) test if the L2 population in the subsequent year could be predicted by L2 distribution in the preceding year and environmental variables. Based on previous research, we included four categories of environmental/site variables that might influence SBW outbreak initiation: 1) forest composition, 2) spring and summer climatic conditions, 3) topographic characteristics, and 4) site quality. We also included local SBW L2 population levels in the previous year, proximity to sites with high SBW populations in the previous year, and insecticide spray treatments in the previous two years as influencing variables.

2. Methods

2.1. Study area

The study area was a 3,730,000 ha area in northern New Brunswick, Canada, spanning 64° 30′ to 69° 0′ W and 46° 30′ to 48° 0′ N (Fig. 1a). Species composition of New Brunswick forests is approximately 68% softwood and 32% hardwood species (Erdle, 2008). SBW host species spruce and balsam fir together comprise more than half of the forest (55%), followed by red maple (*Acer rubrum* L.) and sugar maple (*Acer saccharum* Marsh.) at about 15%. Roughly 20% of the forest is younger than 20 years old, resulting from harvesting; and 45% is older than 60 years old (Erdle, 2008). In northern New Brunswick, balsam fir and spruce are the dominant species groups (Erdle, 2008).

2.2. Data collection and preparation

SBW L2 data in New Brunswick from 2013 to 2018 were provided by the New Brunswick Department of Energy and Resource Development (NBERD) and the Healthy Forest Partnership EIS Research group. In fall of each year, NBERD and forest industry staff from Forest Protection Limited, J.D. Irving, Limited, Acadian Timber Corporation, and Fornebu Lumber Company Inc. jointly collected L2 data from a large number of sample points, each of which included one mid-crown branch from each of three balsam fir or three spruce trees (2013–2015), or a mix of balsam fir and spruce representing the stand condition (2016–2018). From 2013 to 2018, 1136, 1503, 1561, 1649, 1964, and 1851 L2 points were sampled per year, respectively. NBERD processed the branch samples by washing them with a sodium hydroxide solution to destroy the hibernacula, and filtering and counting the number of SBW L2 (Miller et al., 1971; MacLean et al., 2019). The average number of L2 per sample point was classified into six classes: nil, trace, low, moderate, high, and extreme (0, 0.1–3.5, 3.6–6.5, 6.6–20.5, 20.6–40.5, and > 40.5 L2 per branch). Moderate or higher classes (> 6.5 L2 per branch) are particularly of concern, since this is the threshold used for insecticide intervention under the EIS approach (MacLean et al., 2019). Across northern New Brunswick, L2 populations increased consistently up to 2017, but declined dramatically in 2018 (Fig. 1b). SBW L2 data in the adjacent Gaspé-Bas St. Laurent region of Québec from 2013 to 2018, provided by Gouvernement du Québec, Ministère des Forêts, de la Faune et des Parcs and La Société de Protection des Forêts contre les

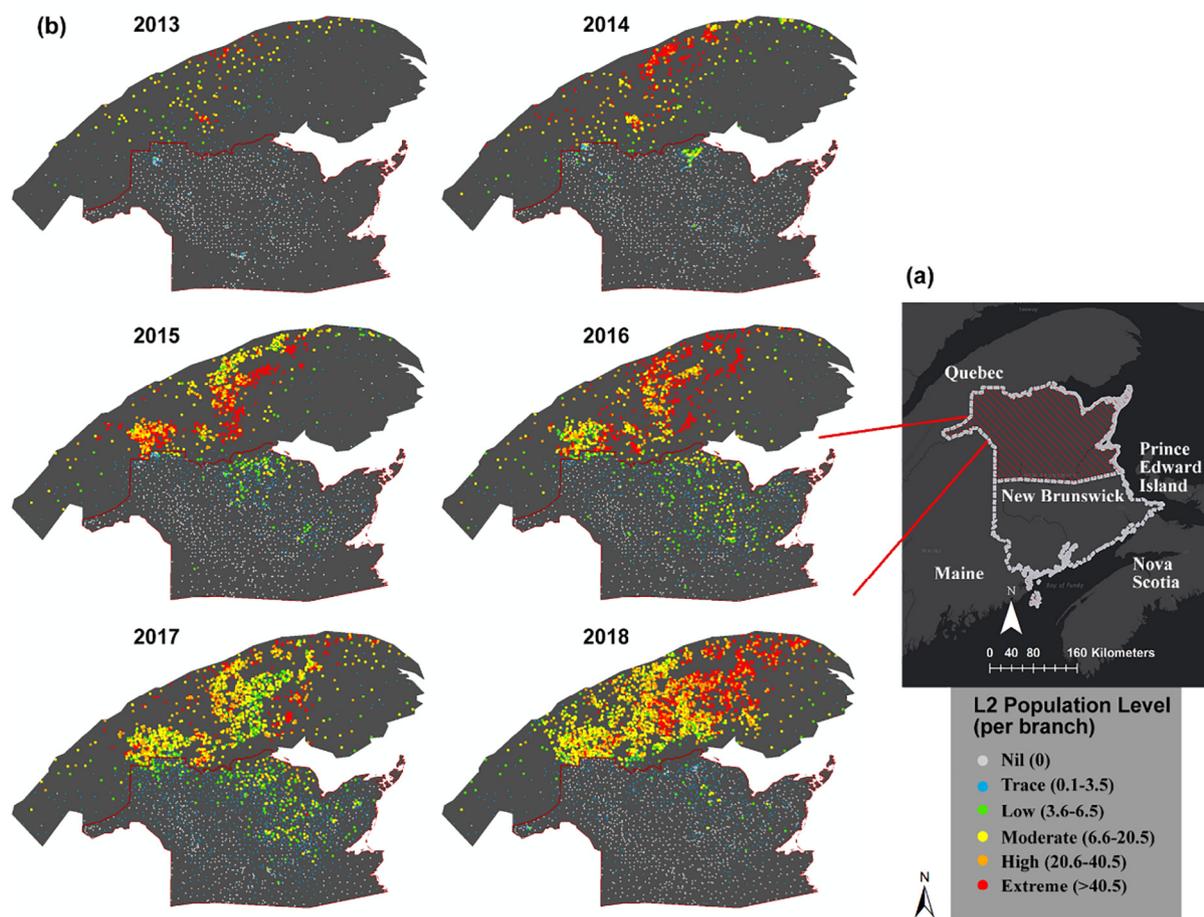


Fig. 1. Location of (a) the study area in northern New Brunswick and (b) distribution of annual SBW L2 populations in sample points within the study area and the adjacent Gaspé-Bas St. Laurent, Québec, from 2013 to 2018.

Insectes et Maladies, were also included in the analyses. L2 data in Québec were sampled on one < 75 cm mid-crown branch from each of three balsam fir or spruce trees. L2 populations in the Gaspé-Bas St. Laurent region continuously increased from 2013 to 2018 (Fig. 1b).

Point layers of the raw L2 sample data were interpolated into raster layers at 20 m resolution using Inverse Distance Weighted (IDW) methods (Watson and Philip, 1985). We tested interpolation methods using the nearest 6, 12, or 24 points in calculating the interpolated cell, and found that results were similar in Gradient Boosting Machine (GBM) tests and statistical models, with a tendency for models with fewer interpolated points to perform slightly poorer in terms of goodness-of-fit indicators. Therefore, we based interpolation on 12 points. Fishnet L2 points, which are points at the crosses of a systematic lattice, for each year from 2013 to 2018 were extracted from the interpolated raster layers at a 2 km interval. A total of 9183 fishnet points were extracted within the study area in each year. We tested both raw L2 point layers and fishnet L2 layer data in analyses.

Datasets for all the other influencing variables were available through various sources, and necessary preprocessing was done before analyses, as described in Table 1. These included forest composition, climate, proximity to high L2 population sites in the previous year, topography and site quality variables, the previous year SBW L2 population, and insecticide spray treatments in the previous two years (Table 1). All raster data used in the research were at 20 m resolution.

2.3. Spatial overlay analysis

Annual SBW L2 point layers from 2014 to 2018, including both the raw sample points and fishnet points, were overlaid with geo-

referenced layers of 46 possible influencing variables (Table 1). Variables included tree species proportions of balsam fir, spruce, and hardwood within 50 m circular buffer areas for each L2 point, using the dominant forest layer and an area-based weighted average of species proportions; 33 climate variables including monthly and periodic temperatures, cumulative degree days, and precipitation; previous local SBW population for Year N extracted from the interpolated raster in Year N-1; proximity to moderate or higher levels of SBW L2 populations in Year N-1 computed as distance to the nearest raw sample point with L2 > 6.5 per branch (ProxRaw), distance to the nearest fishnet point with L2 > 6.5 per branch (ProxGrid), and distance to the nearest cell with L2 > 6.5 per branch (ProxCell; Table 1). SBW L2 data from the Gaspé-Bas St. Laurent, Québec region were incorporated in the proximity calculation, although only those L2 points within the northern New Brunswick study area were used in the statistical analyses. Insecticide spray history was included as a dummy variable, with four numbers representing whether a sample point fell within the spray blocks in the previous two years, in one of the two preceding years, or none of two years. All spatial analyses were done using ArcMap 10.4 (ESRI, Redlands, CA, USA).

2.4. Variable importance analyses

With all 46 predictor variables included (Table 1), GBM analysis (Ridgeway, 2007) was used to determine the most influential predictors in forecasting the SBW L2 population, based on both raw sample points and fishnet points. As a machine learning technique, gradient boosting determines performance of decision trees using gradients in the loss function in a sequential fashion (Friedman, 2001). GBM analysis reports

Table 1

Predictor variables (abbreviations, units, and descriptions) included in Gradient Boosting Machine analysis to determine their relative importance in predicting L2 population in Year N.

Influencing factor	Variable abbreviation	Unit	Variable description
Forest composition ^a	BF%	%	Weighted average of balsam fir proportion within 50 m buffer zone
	SP%	%	Weighted average of spruce proportion within 50 m buffer zone
	HW%	%	Weighted average of hardwood proportion within 50 m buffer zone
Climate conditions ^b	Temp_04 to _09	°C	Monthly temperature (Apr. to Sept.)
	Temp_av45	°C	Average temperature in Apr. and May
	Temp_av67	°C	Average temperature in Jun. and Jul.
	Temp_av89	°C	Average temperature in Aug. and Sept.
	Temp_av456	°C	Average temperature in Apr. to Jun.
	Temp_av789	°C	Average temperature in Jul. to Sept.
	DD_04 to _09	°C·d	Cumulative degree days (> 5°C) (Apr. to Sept.)
	DD_45	°C·d	Cumulative degree days (> 5°C) in Apr. and May
	DD_67	°C·d	Cumulative degree days (> 5°C) in Jun. and Jul.
	DD_89	°C·d	Cumulative degree days (> 5°C) in Aug. and Sept.
	DD_456	°C·d	Cumulative degree days (> 5°C) in Apr. to Jun.
	DD_789	°C·d	Cumulative degree days (> 5°C) in Jul. to Sept.
	Prcp_04 to _09	mm	Monthly total precipitation (Apr. to Sept.)
	Prcp_45	mm	Total precipitation in Apr. and May
	Prcp_67	mm	Total precipitation in Jun. and Jul.
	Prcp_89	mm	Total precipitation in Aug. and Sept.
	Prcp_456	mm	Total precipitation in Apr. to Jun.
Prcp_789	mm	Total precipitation in Jul. to Sept.	
Proximity to high L2 population sites in the previous year	ProxRaw	km	Distance to the nearest raw sample point with L2 > 6.5 per branch in Year N-1
	ProxGrid	km	Distance to the nearest fishnet point with L2 > 6.5 per branch in Year N-1
	ProxCell	km	Distance to the nearest raster cell with L2 > 6.5 per branch in Year N-1
Topography ^c	DEM	m	Elevation
	Slope	°	Slope values
	Aspect	°	Aspect values
Site quality	DTW ^c	m	Depth to water
	BGI ^d	kg ha ⁻¹ yr ⁻¹	Biomass growth index
Previous local SBW population ^e	L2_PreYear	per branch	L2 population at certain sample sites in the preceding year
Spray treatment history ^f	TreatCode		If the point falls in the spray treatment blocks

0: not treated either in Year N-1 or Year N-2
1: treated in Year N-1, not treated in Year N-2
2: treated in Year N-2, not treated in Year N-1
3: treated both in Year N-1 and Year N-2

^a Forest data was obtained from New Brunswick Energy and Resource Development as a LandBase polygon layer. The shapefile was intersected with 50 m circular buffer zones around each L2 fishnet point. Species composition was calculated as the weighted average of species proportion based on the area of intercepted polygons.

^b Climate data was generated by BioSIM11 (Régnière et al., 2014) which used four climate databases: Canadian Climate Normals database; New Brunswick Fire Weather; New Brunswick Agriculture Weather; and Ministère du Développement Durable, de l'Environnement et de la Lutte contre les Changements Climatiques (MDDELCC). Processes under BioSIM11 included enquiring daily weather data for New Brunswick, transforming daily data into monthly data, and generating climate data for sample points by interpolation.

^c Elevation and depth to water data were provided by Forest Watershed Research Centre at the University of New Brunswick (Murphy et al., 2007; Furze et al., 2017), and were resampled into 20 m resolution for this study. Slope and aspect raster layers were generated from the elevation raster, and smoothed by focal statistics with 3 × 3 cells as neighborhood.

^d Biomass growth index was created by FORUS Research (Hennigar et al., 2017).

^e Previous local SBW population was extracted from the Year N-1 interpolated L2 raster.

^f Spray treatment shapefiles in the previous two years were provided as polygon layers by Early Intervention Strategy Research group. Whether a sample point was previously treated or not was identified by intersecting the point with the treatment blocks in each year.

the relative influence of analyzed variables according to frequencies that a variable was selected over all splits, weighted by the squared improvement to the model in each split (Elith et al., 2008). The “caret” package (Kuhn, 2008) within R version 3.4.3 (R Development Core Team, 2018) was used for the GBM analysis. Correlation tests were also done for all predictor variables, and highly correlated predictor variables ($r \geq 0.7$) were avoided in the models. The generally most important uncorrelated variables ($r < 0.7$) across all years selected from the GBM-resulted variable ranks and their interactions were used in fitting models.

2.5. Spatial regression models

Simultaneous autoregressive (SAR) models (Wall, 2004) were implemented to predict SBW L2 population levels each year from 2014 to 2018. SAR models assume that the response variable depends not only

on the predictor variables, but also the spatial neighborhood relationships among all samples in the study area, which were implemented as an $n \times n$ weights matrix in the models (Haining, 2003). The specification of the spatial weights usually starts from a binary neighbors list in which objects are either listed as neighbors or are absent, then is further weighted to give less-distant neighbors more weight (Hoef et al., 2018). The weights matrix consists of zeros on the diagonal and weights for the neighbors in the off-diagonal positions (Hoef et al., 2018). Based on results of Kissling and Carl (2008), who compared SAR model performance with 3240 different combinations of model settings, row-standardization (i.e., “W” coding style) was used to code the spatial weights matrix in this study.

We fitted SAR models using only the L2 fishnet points because, compared to the irregular distribution of L2 raw sample points, fishnet sample points have a regular lattice of point sites, which guarantees that all orders of neighborhood sizes remain the same for all samples.

This benefits the comparison between models across different years, since in this case, a common neighborhood size can be used in different-year SAR models. In addition, predicting regions with high L2 populations, instead of points with high L2 populations, has more guiding value in insect sampling and management. Given that our fishnet points were at 2 km intervals, we defined the upper neighbor boundary as 5 km to include eight adjacent neighbors, i.e., first order neighbors, for each sample point, as advised by Kissling and Carl (2008). SAR models take three forms: 1) lagged response model which assumes that the autoregressive processes merely occurs in the response variable, 2) lagged mixed model which assumes such processes occur in both response and predictor variables, and 3) spatial error model which assumes that the process only occurs in the error terms (Anselin, 1988). We applied the second form, SAR_{mix} model, because the response variable SBW L2 was likely to have both inherent spatial autocorrelation (i.e., derived from interactions within the observed variable itself) and induced spatial autocorrelation (i.e., the observed variable is functionally dependent on an underlying variable which is autocorrelated). SAR_{mix} models take the form (Anselin, 1988):

$$Y = \rho WY + WX\gamma + X\beta + \varepsilon$$

where $X\beta + \varepsilon$ are the standard terms in ordinary least squares (OLS) regression: β is a vector representing the slopes associated with the predictor-matrix X , and ε is a vector of the error terms. ρW is a term describing spatial autocorrelation in the response variable vector Y , where ρ is the autoregression parameter, and W is the spatial weights matrix. $WX\gamma$ describes the regression coefficients γ of the spatially lagged predictors WX . SAR models were implemented using the “spdep” package (Bivand, 2006), and the R script used to fit SAR models was derived from the appendix of Dormann et al. (2007). The top predictor variables determined by GBM analysis and their interactions were included in the full models, and then any non-significant variables or interactions, assessed by t tests ($\alpha = 0.05$), were dropped in the reduced models.

Spatial dependence of SAR model residuals was investigated using correlograms of Moran’s I and maps of residuals. Since parameters of SAR models are estimated by the maximum likelihood method (Wall, 2004), the likelihood-based measure of goodness-of-fit, log likelihood, was appropriate for SAR models (Lichstein et al., 2002; Tognelli and Kelt, 2004). For the same reason, we used the Nagelkerke pseudo R^2 (Nagelkerke, 1991), which is based on likelihood, instead of the traditional R^2 , to compare model performance. Nagelkerke pseudo R^2 was computed according to the formula (Nagelkerke, 1991):

$$R^2 = 1 - \exp[-2/n(l_A - l_0)]$$

where l_A is the log likelihood of the model to be tested, l_0 is the log likelihood of the null model that merely contains the intercept, and n is the sample size.

In each year from 2014 to 2018, OLS models were fitted with the same predictor variables and interactions to compare with the SAR_{mix} models in terms of spatial dependence as well as overall model performance. The Nagelkerke pseudo R^2 formula above yields the identical value as the traditional R^2 for OLS models (Lichstein et al., 2002). All correlograms in this study were created using functions in the “nfc” package (Bjørnstad et al., 1999) with 500 permutations for each test. Then the significance levels of the coefficient at each lag distance were adjusted by progressive Bonferroni correction, where the Bonferroni-corrected significance level was computed for each distance class separately (Legendre and Legendre, 2012):

$$\hat{a}(d) = a/d$$

where a is the commonly used significance level, i.e., $a = 0.05$, \hat{a} is the number of tests actually performed up to the specific distance class, \hat{a} is the adjusted significance level, and d is the distance class of interest.

Finally, we applied models forecasting the L2 population in Year N

to the datasets in the subsequent year using the “predict.sarlm” function (Goulard et al., 2017) in the package “spdep” (Bivand, 2006). We also carried out sensitivity analysis on each year’s model, by comparing the predicted L2 population under different scenarios systematically varying values of each of three predictor variables while controlling the other two predictors unchanged, to test the degree that each predictor variable influenced the response variable. The predicted fishnet point values were then interpolated as raster layers using the IDW method, and compared with the observed L2 population distributions.

2.6. Combined year model

To predict L2 population levels in the coming year using a generalized model formula, linear mixed effects (LME) models were fitted, combining all years’ data with year as a random effect and maximum likelihood as the parameter estimation using the “lme” function in the package “nlme” (Lindstrom and Bates, 1988). The effects of spatial autocorrelation were addressed using the “corr” argument in the model settings. The L2 population data were log-transformed to stabilize the variance and to improve normality of residuals. The full model was fitted using the same predictors and their interactions used in the SAR_{mix} models, with random intercepts and random slopes on all three predictors. Then a reduced model was generated by dropping non-significant variables and interactions, assessed by t tests ($\alpha = 0.05$), with fewer random components in slopes by comparing the full model and the reduced models using likelihood ratio tests. Assumptions of residual normality and homoscedasticity were evaluated using residual plots. To be comparable to each year’s SAR_{mix} model, goodness-of-fit of the combined-year (2015–2018) model was also evaluated by log likelihood and Nagelkerke pseudo R^2 . The reduced model was used to predict L2 populations from 2015 to 2018, with only fixed effects included, and compared with observed L2 population levels. L2 populations in 2019 were also predicted using the reduced combined-year model.

3. Results

3.1. Relative influence of predictor variables

GBM tests showed that local previous-year SBW L2 population levels, proximity to high population sites in the previous year, and climate conditions in spring (April or May) were the most important variables to predict SBW L2 population levels from 2015 to 2018 (Fig. 2). Specifically, distance to the nearest fishnet point with L2 > 6.5 per branch in the previous year was the most important predictor in 2015 (relative influence = 41%), 2017 (30%), and 2018 tests (16%; Fig. 2b, d and e). Distance to the nearest raw sample point with L2 > 6.5 per branch in the previous year, which was highly correlated with the distance to the nearest fishnet point ($r = 0.9$), was the most important predictor in 2016 (relative influence = 24%; Fig. 2c). Local previous-year L2 population level was the second most important variable in 2015 (relative influence = 8%) and 2016 (20%), and third in 2017 (14%) and 2018 (10%; Fig. 2c–e). Cumulative degree days in April ranked in the top five variables in all five years (Fig. 2). In 2014, results showed two other top predictor variables: monthly mean temperature in May (11%) and biomass growth index (10%; Fig. 2a). The L2 population was < 6.5 per branch for all sample points in New Brunswick in 2013, so proximity to high SBW populations in 2014 was entirely based on distances to populations in the Gaspé-Bas St. Laurent region. There were no insecticide spray treatments carried out in 2012 or 2013, so this variable was not included in the 2014 analysis. However, insecticide treatment was not among the top five influencing variables in any years.

To generalize the model over the years into a form that could be used for multi-year prediction, we used the variables local L2 population in the previous year, distance to the nearest fishnet point with

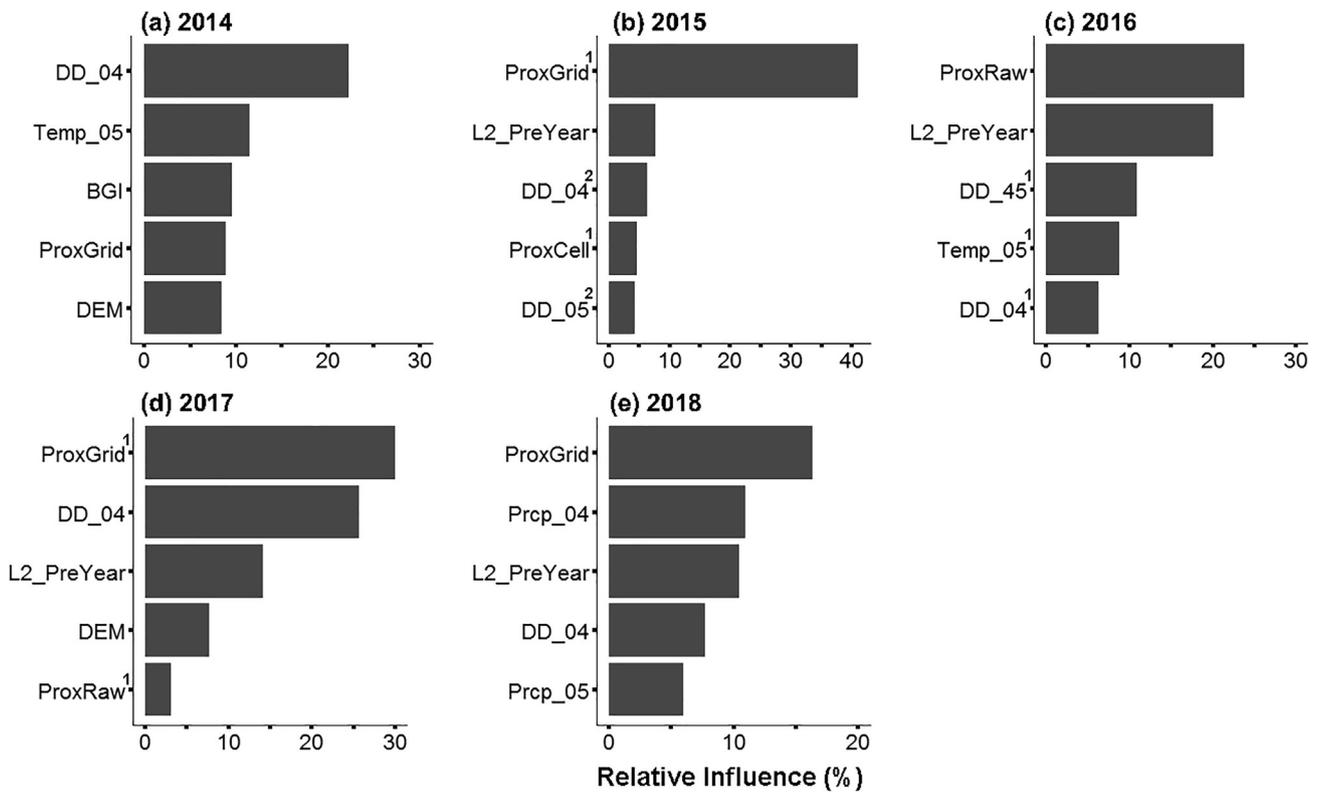


Fig. 2. Relative influence (%) of the five most important predictor variables based on Gradient Boosting Machine analysis to predict SBW L2 population each year from 2014 to 2018 (a–e). Predictor variable abbreviations are described in Table 1, and predictors with the same superscripts were highly and positively correlated with each other (correlation coefficient $r \geq 0.7$).

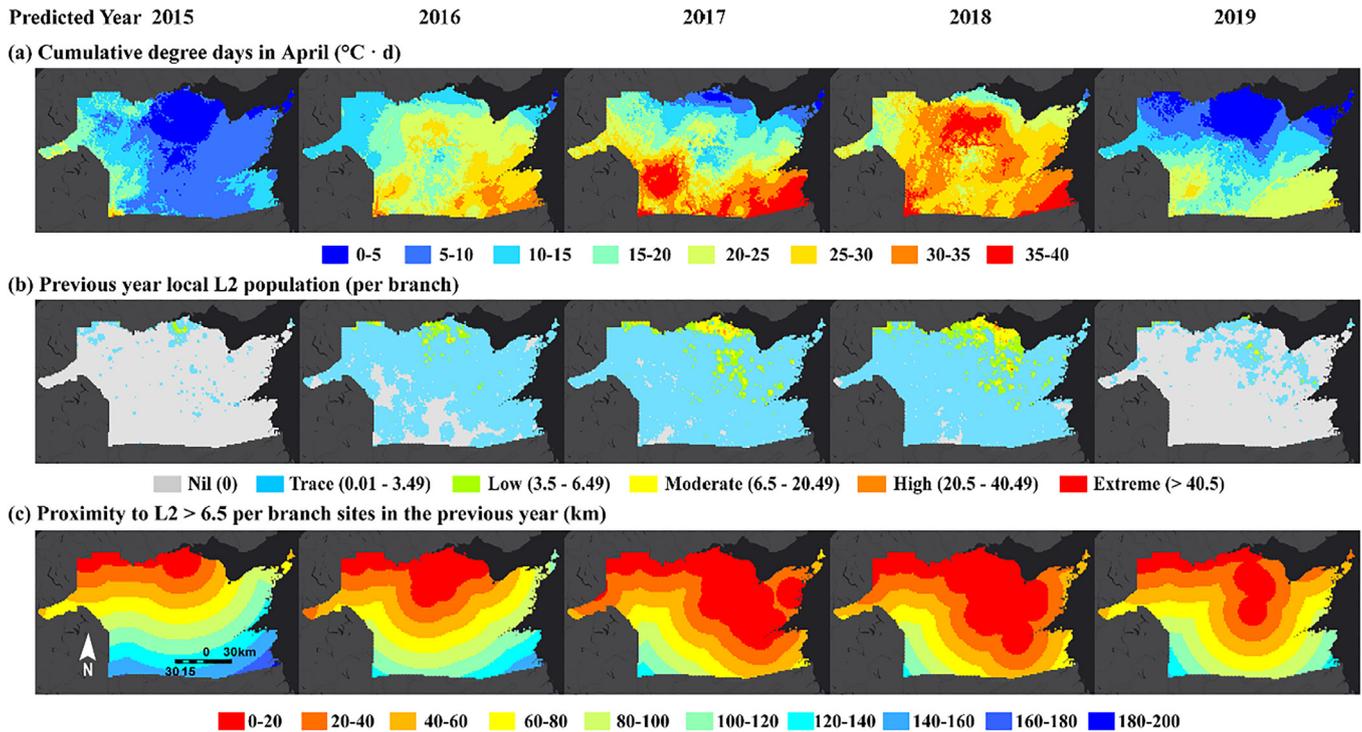


Fig. 3. Maps of (a) cumulative degree days in April ($^{\circ}\text{C} \cdot \text{d}$), (b) previous local SBW L2 population (per branch), and (c) proximity to L2 > 6.5 per branch sites in the previous year (km) from 2015 to 2019. These three variables were generally the most important predictors and were used in the modeling process.

L2 > 6.5 per branch in the previous year, and cumulative degree days in April and their interactions to fit the 2015–2018 models, predicting L2 population over the 4 years. Fig. 3 shows the trends of these three predictor variables from 2015 to 2019. Cumulative degree days in April

consistently increased from 2015 to 2018, but decreased in 2019 (Fig. 3a); 2015 and 2019 had noticeably cold springs. Previous-year L2 population levels (L2_PreYear) increased from 2015 to 2018 (Fig. 3b). With the L2 population expansion (Fig. 3b), there was increasing area

Table 2

Nagelkerke pseudo R^2 and log likelihood of L2 population predictions by ordinary least squares (OLS) and simultaneous autoregressive mixed (SARmix) models over the years from 2014 to 2018, and by the combined linear mixed effects (LME) model from 2015 to 2018. Predictor variable abbreviations are described in Table 1.

Year	Regression type	predictors ^a	Nagelkerke pseudo R^2 ^b	Log likelihood
2014	SAR _{mix}	DD_04 + Temp_05 + BGI	0.75	-1342
	OLS		0.06	-5676
2015	SAR _{mix}	L2_PreYear + ProxGrid + DD_04* L2_PreYear	0.72	-8648
	OLS		0.29	-12965
2016	SAR _{mix}	L2_PreYear + ProxGrid + DD_04* L2_PreYear + L2_PreYear* ProxGrid	0.79	-13528
	OLS		0.35	-18802
2017	SAR _{mix}	L2_PreYear + ProxGrid + DD_04* ProxGrid	0.77	-15367
	OLS		0.35	-20063
2018	SAR _{mix}	DD_04 + L2_PreYear + DD_04* L2_PreYear + DD_04* ProxGrid + L2_PreYear* ProxGrid	0.68	-3916
	OLS		0.14	-8437
2015–2018	LME	L2_PreYear + DD_04* L2_PreYear + L2_PreYear* ProxGrid	0.53	-13858
	OLS		0.22	-17833

^a Only significant variables and interactions were kept in the reduced models ($p < 0.05$).

^b For OLS models, Nagelkerke pseudo R^2 yields the identical value as the traditional R^2 , i.e., the proportion of the variance for a dependent variable explained by independent variables.

with low proximity values (0–20 km to the nearest fishnet point with $L2 > 6.5$ per branch) from 2015 to 2018 (Fig. 3c). For the 2014 model, we used cumulative degree days in April, monthly mean temperature in May, biomass growth index, and their interactions to predict L2 population levels.

3.2. Performance of spatial regression models and combined-year (2015–2018) LME models compared to OLS models

Over the 5 study years, SAR_{mix} models explained 68% to 79% of the variance in the SBW L2 populations, whereas OLS models explained only 6% to 35% of the variance (Table 2). SAR_{mix} models had higher log likelihood ratios than OLS models in each year, representing a better fit of the data by SAR_{mix} models (Table 2). The predictor cumulative degree days in April (DD_04) was non-significant in the 2015–2017 models ($\alpha = 0.05$), while the interaction term of DD_04 and L2_PreYear was significant in years 2015, 2016, and 2018 ($\alpha = 0.05$). Moreover, OLS residuals showed positive spatial autocorrelation up to a distance of 100 km, with the Moran's I ranging from 0 to about 0.6 over the 5 years (Fig. 4a), providing evidence that the assumption of independently distributed residuals was violated. OLS residuals were spatially clustered across all years, with clumps of positive residuals and negative residuals (Fig. 4b). L2 population level, the response variable, was also spatially autocorrelated up to a distance of 100 km across the years, suggesting that the spatially-correlated L2 population was one of the possible causes of the spatial-correlated residuals at roughly the same spatial scale (Fig. 4a). However, in contrast to OLS models, correlograms and maps of SAR_{mix} model residuals showed very low spatial autocorrelation: Moran's I was < 0.05 for all tested lag-distances in all study years (Fig. 4a), and much less spatial heterogeneity was exhibited (Fig. 4c). Being able to address the spatial autocorrelation in model residuals indicated the suitability of SAR_{mix} models for the ecological data analyzed in this study.

To aggregate multiple years' data and generate a generalized model form, the LME model was fit combining 2015–2018 data. Data in 2014 were excluded when fitting the 2015–2018 combined-year LME model because the top predictors for 2014 data differed from those for 2015–2018 data according to GBM tests (Fig. 2). The predictor variable L2 in the previous year, and interaction terms of it with cumulative degree days in April and distance to the nearest fishnet point with $L2 > 6.5$ per branch in the preceding year were kept in the reduced LME model, which explained 53% of the variance in SBW L2 population levels, versus only 22% of variance explained by the OLS model (Table 2). The log likelihood ratio of the LME model was also higher than the OLS model, suggesting a better goodness-of-fit (Table 2).

3.3. Relationships of SBW L2 population levels to top predictor variables

We used sensitivity analyses to systematically vary values of the three predictor variables and determine the influence on predicted SBW L2 populations. Higher L2 population forecasts were associated with lower degree days in 2016 and 2017 (Fig. 5a), during which years the observed values of DD_04 averaged 21.2 °C·d and 23.3 °C·d, respectively. In 2016, the observed DD_04 -20 °C·d and -10 °C·d scenarios averaged 5.2 and 2.7 L2 per branch increases, while $+10$ °C·d and $+20$ °C·d scenarios averaged 1.1–1.3 L2 per branch decreases (Fig. 5a). In 2017, scenarios of the observed DD_04 -20 °C·d and -10 °C·d, and $+10$ and $+20$ °C·d averaged 1.8 and 1.0 L2 per branch increases, and 0.6 and 1.1 L2 per branch decreases, respectively (Fig. 5a). In 2015, a year with a much cooler spring ($\bar{x} = 9.1$ °C·d), predicted L2 populations were more sensitive to increases in L2_PreYear than to DD_04 (Fig. 5a and b).

Sensitivity analyses of L2 in the previous year indicated that locations with higher L2 in Year N-1 had more L2 in Year N, although the effect was reduced in 2017 and even lower in 2018 (Fig. 5b). Because L2 in the previous year was set to zero in the sensitivity analyses when varying the observed L2 by -5 or -10 produced a negative value, many values were zeros in the sensitivity tests. Hence, sensitivity was mostly reflected in the original L2_PreYear $+5$ or $+10$ cases; systematically increased L2_PreYear of five larvae per branch averaged 9.8 L2 per branch increases in 2015 L2 forecasts, 12.1 in 2016, 2.1 in 2017, and 0.1 in 2018 (Fig. 5b). Systematically increased L2_PreYear of 10 larvae per branch averaged 19.5, 24.1, 4.1, and 0.2 per branch increases in L2 forecasts in 2015, 2016, 2017, and 2018, respectively (Fig. 5b).

In 2015, 2017, and 2018, locations closer to high L2 sites in Year N-1, i.e. with lower values of ProxGrid, were generally associated with higher predicted L2 population in Year N (Fig. 5c). Systematically increasing ProxGrid by 40 km resulted in decreases of 0.05, 1.0, and 0.2 L2 per branch in 2015, 2017, and 2018, respectively whereas in 2016, increasing ProxGrid by $+20$ and $+40$ km led to small increases, by 0.08–0.13 L2 per branch. A 40 km decrease in ProxGrid contributed to at most an average of 1.2 L2 per branch increase. Predicted L2 populations in 2018 showed less sensitivity to all three predictor variables, probably because of the relatively low and evenly-distributed L2 values observed ($\bar{x} = 0.3$, $\sigma = 0.7$; Fig. 1b).

3.4. Comparison of observed and predicted SBW L2 populations

Generally, the Year N SAR_{mix} models and combined-year LME model underestimated SBW L2 population levels in Year N, especially at high population levels, from 2015 to 2018. The Year N-1 SAR_{mix} models

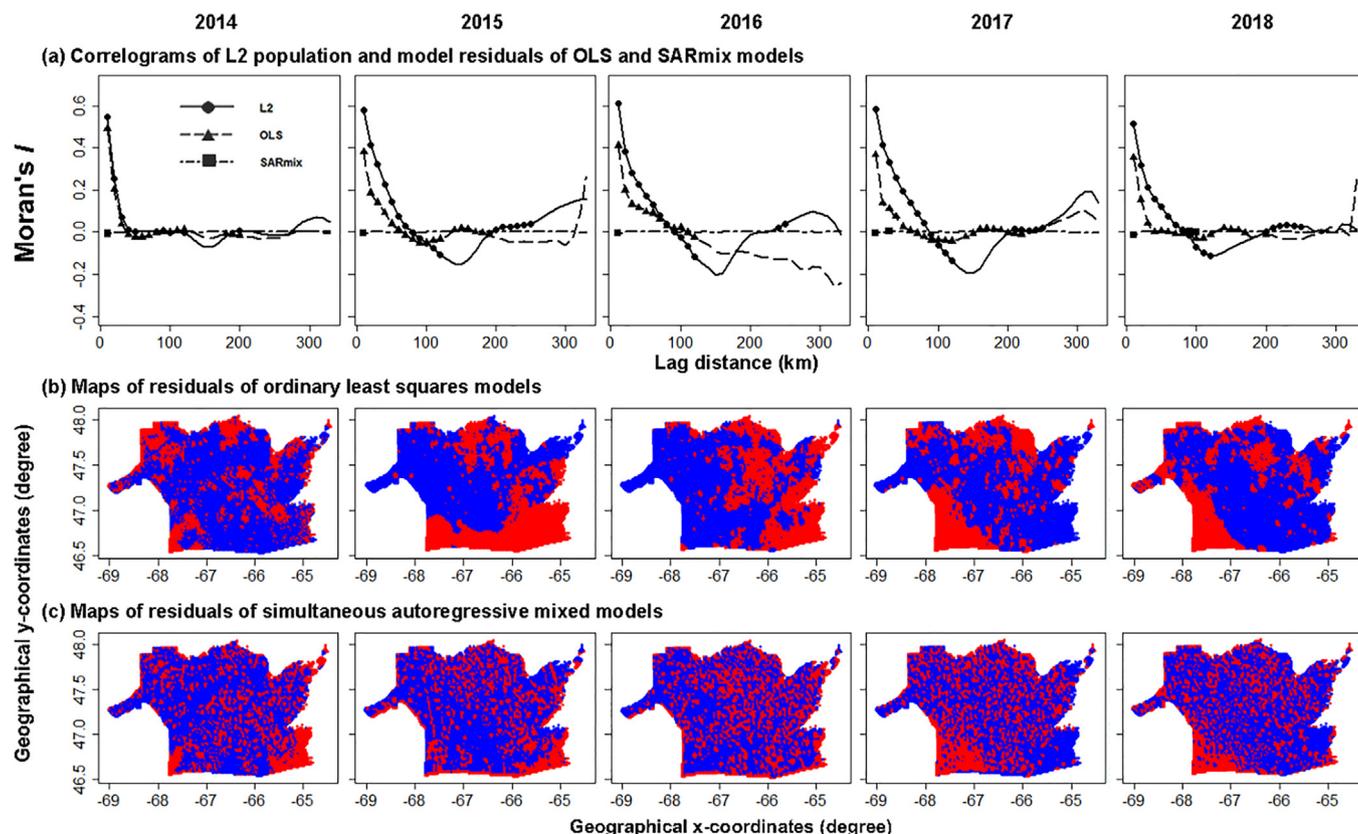


Fig. 4. Spatial autocorrelation of SBW L2 population and model residuals of ordinary least squares (OLS) models and simultaneous autoregressive mixed (SARmix) models over 5 years from 2014 to 2018, shown by (a) correlograms, where solid black symbols indicate significant coefficients after progressive Bonferroni corrections ($\alpha = 0.05$, 500 permutations), and spatial error maps of (b) OLS models and (c) SARmix models, where red area indicates positive residuals and blue area indicates negative residuals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

were more accurate in predicting Year N L2 populations than the combined-year (2015–2018) LME model, which had a greater bias in L2 population forecasts. In comparison with the observed SBW L2 population distribution (Fig. 6a), areas with moderate or higher levels of L2 population ($L2 > 6.5$ per branch) were underestimated by Year N SAR_{mix} models by 14,620 (72%), 8140 (7%), 15,470 (11%), and 2750 ha (100%) less than the observed area from 2015 to 2018, respectively (Fig. 6b). In 2015 and 2016, areas with moderate or higher levels of L2 population were underestimated by Year N-1 SAR_{mix} models by 18,856 ha in 2016, and no area with $L2 > 6.5$ per branch was predicted in 2015 (the SAR_{mix} model in 2014 used different predictor variables from the other three years; Table 2, Fig. 6a and c). However, area with moderate or higher levels of L2 populations were substantially overestimated, by 510,382 and 29,278 ha, by Year N-1 SAR_{mix} models in 2017 and 2018 (Fig. 6a and c).

The combined-year (2015–2018) LME model did a poor job of predicting moderate or higher L2 populations in any of the sample years (Fig. 6d). All areas (3,730,000 ha) were predicted to have SBW L2 populations < 6.5 per branch in 2015 and 2016 by the combined-year (2015–2018) LME model, whereas actual observed areas were 20,191 ha and 122,738 ha, respectively. The area with moderate or higher L2 population was predicted to be 138,560 ha less than observed in 2017 and 568 ha more than the observed area in 2018 (Fig. 6a and d), and in 2018 the predicted location differed from the actual. We used both the 2015–2018 combined-year LME model and the annual 2015–2018 SAR_{mix} models to predict moderate or higher SBW L2 populations in 2019 (Fig. 6d and e). Except for the 2016 SAR_{mix} model (which hugely overestimated area with $L2 > 6.5$ per branch at 450,000 ha), the area of moderate or higher L2 populations predicted for 2019 was low, 0–256 ha, because input values of the predictor variable L2_{PreYear}, i.e., L2 populations in 2018, were almost all at the

nil or trace level.

4. Discussion

4.1. Forecasts of SBW population levels in the coming year

For predicting SBW L2 population levels in the following year, local previous-year insect population was the most important variable. Locations with higher SBW L2 populations, located closer to high population sites, had higher predicted L2 levels in the subsequent year. Sensitivity analyses showed that increasing the L2 population by five to ten per branch at a site in a given year would result in the L2 population at that site in the subsequent year increasing by 12 to 24 per branch. Being 40 km closer to or further from locations with > 6.5 L2 per branch resulted in only having 1 more or 1 less L2 per branch in the next year. Rising populations due to large-scale pulses of resource or important variables in a given year would potentially remain at the same level over the following years, which can contribute to an outbreak initiation (Bouchard et al., 2018). SBW populations in our study area experienced over 90% reductions unexpectedly in 2018, the reason for the decline is currently unknown, although study is continuing, and it may well represent a temporary annual reduction in an increasing population trend (MacLean et al., 2019). The quantified relationships found, however, were not consistent across years. Thus, a model fit in Year N-2 could not be used to reliably predict L2 populations in Year N. Moreover, populations in Year N predicted by the Year N-1 SAR_{mix} models were generally more accurate than the population predicted by the combined-year (2015–2018) LME model, which consistently underestimated L2 populations. The effect of cumulative degree days in April in the predicted year models may help to characterize the upper and lower population bounds when forecasting the SBW population.

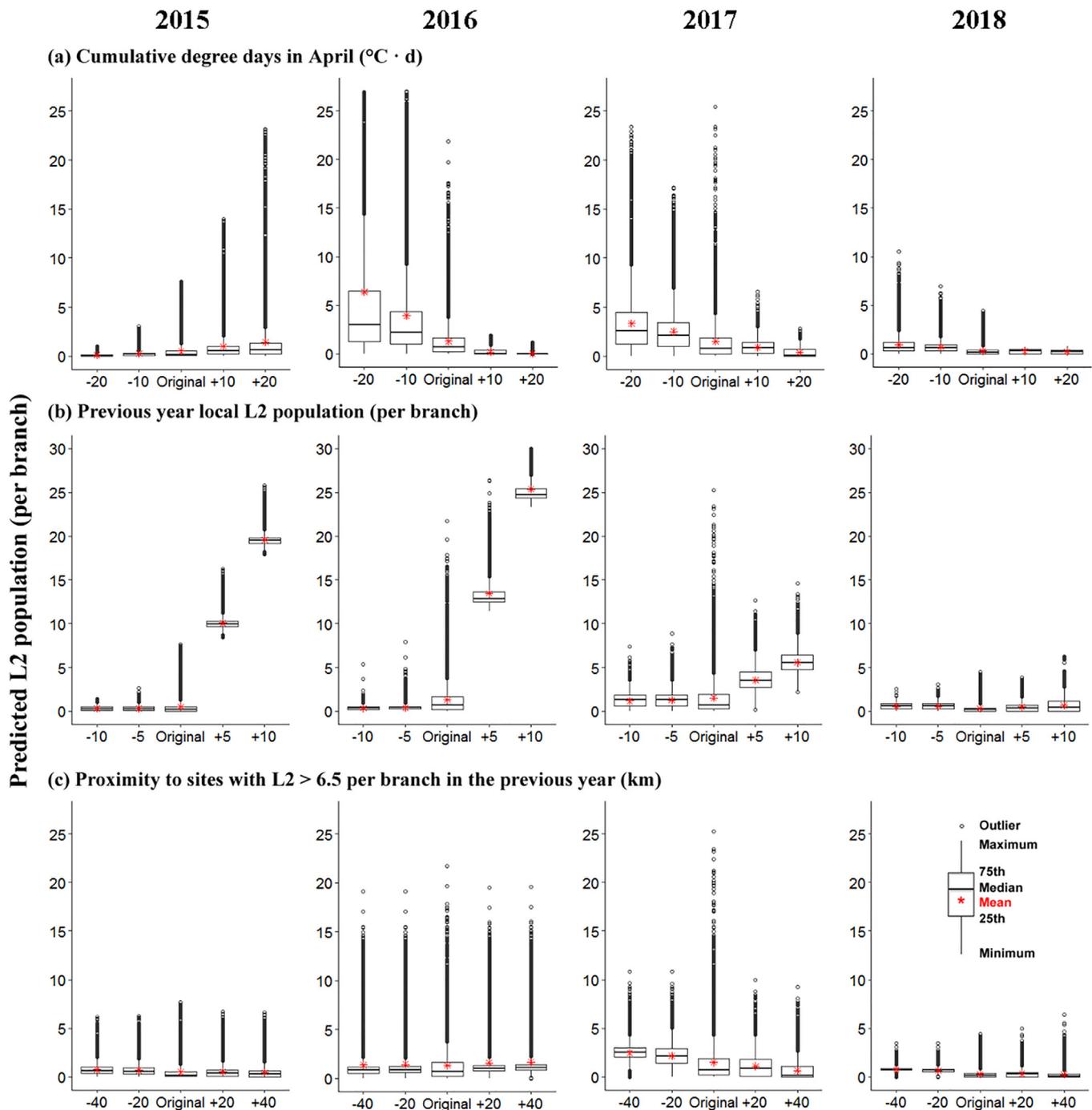


Fig. 5. Sensitivity analyses of SBW L2 population (per branch) predicted by simultaneous autoregressive mixed models from 2015 to 2018 under three scenarios: (a) original values of cumulative degree days in April compared to original values -20 , -10 , $+10$, and $+20$ °C·d, (b) original values of previous local L2 population compared to original values -10 , -5 , $+5$, and $+10$ per branch, and (c) original values of proximity to L2 > 6.5 per branch sites in the previous year compared to original values -40 , -20 , $+20$, and $+40$ km. Previous local L2 population and proximity with values < 0 in (b) and (c) were set to zero before prediction.

4.2. Effects of environmental variables on outbreak initiation

Early spring climate conditions were generally the most important influencing environmental variable analyzed. Overall, northern New Brunswick experienced increasingly warm springs from 2015 to 2018, the years when observed L2 populations increased. Spring temperatures (April degree days) in 2015 and 2019 were cool. Warmer climate conditions lead to a better phenological match between the larvae emerging from overwintering and the young developing foliage (Bouchard et al., 2018), and this results in lower larval starvation, less need to disperse, and lower larval mortality (Miller, 1958). A cooler

April, typically colder than about 15 °C, can potentially lead to late L2 emergence from hibernacula with longer duration remaining as L2 before reaching the third instar (Rose and Blais, 1954).

However, our sensitivity analyses showed that higher L2 populations were associated with lower early spring degree days in 2016 and 2017, which conflicted with the broad-scale climate conditions effect, i.e., that warm springs associated with SBW population increase at the broad scale. Nonetheless, our study suggested that effects of spring degree days on L2 population differed for different previous-year outbreak conditions, including local population and proximity to high population locations, as suggested by the statistically significant

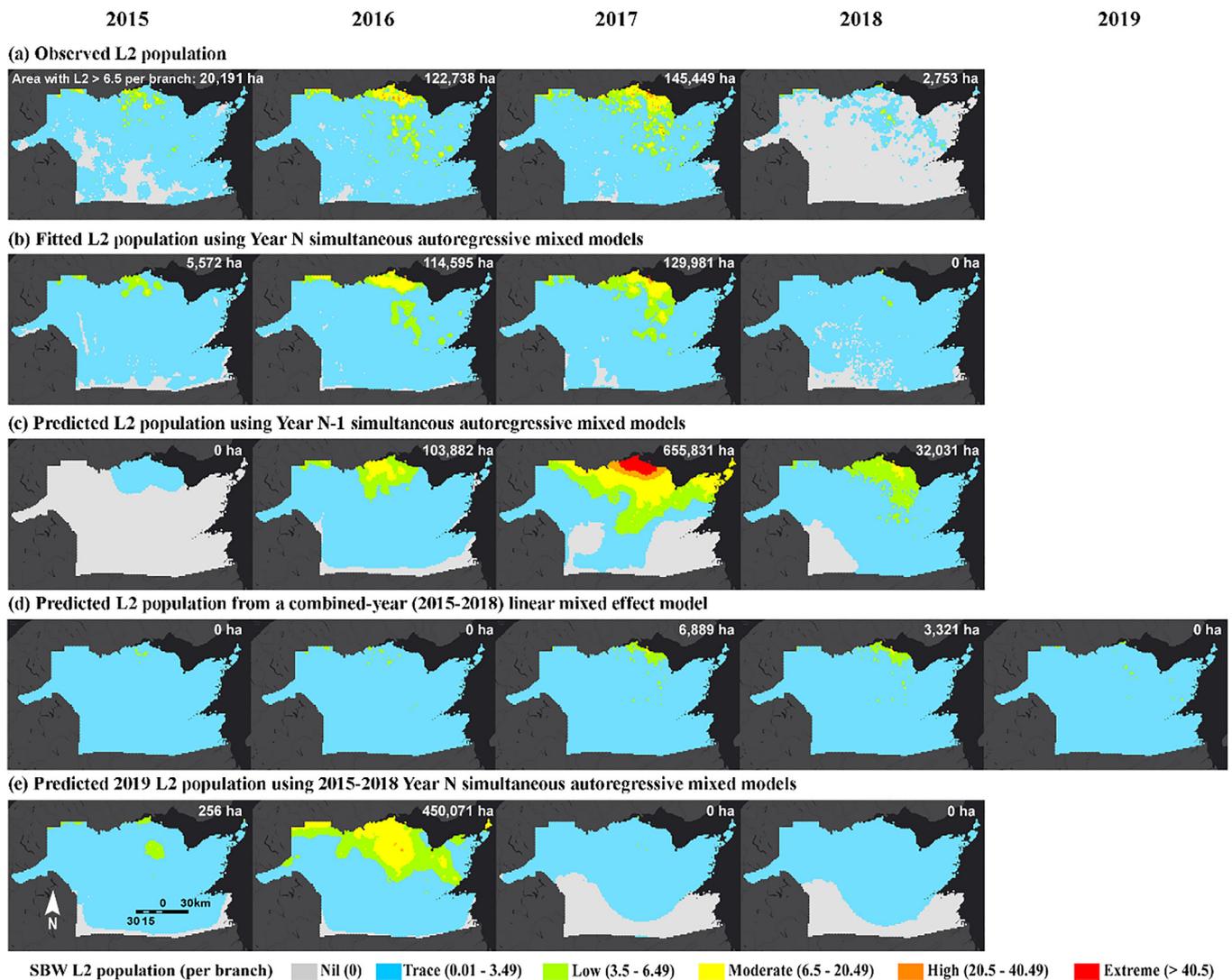


Fig. 6. Comparison of (a) observed L2 population spatial distribution, and the forecast spatial distribution, i.e., interpolated raster from L2 population point of (b) fitted values from Year N simultaneous autoregressive mixed models, (c) predicted values from Year N-1 simultaneous autoregressive mixed models, (d) predicted values from a combined-year (2015–2018) linear mixed effect model, and (e) 2019 L2 population predictions using annual Year N simultaneous autoregressive mixed models from 2015 to 2018. Areas (ha) with L2 > 6.5 per branch were estimated based on raster cells.

interaction terms. In other words, effects of spring degree days on L2 population depended on the degree to which previous-year outbreak conditions have influenced L2 population.

However, the findings about early spring climate in this study may not apply to other regions or different SBW outbreak phases. The current SBW outbreak onset in New Brunswick expanded from the north, where the climate is relatively colder. Previous SBW outbreaks in New Brunswick beginning in 1912 and 1949 tended to occur several years after outbreaks occurred in adjacent Québec, and the role of moth dispersal was believed to be important (Greenbank, 1957). In our case, because of the lack of moth dispersal data, the possible effect of moth dispersal from the current SBW epidemic in Québec was incorporated by including proximity to high L2 population sites in both Québec and the northern New Brunswick study area. If detailed spatial data on SBW moth dispersal were available, from radar monitoring (Boulanger et al., 2017) or modeling (Régnière et al., 2019), it might well supersede effects of spring temperature.

Some previous studies also found negative relationships between spring climate and SBW outbreak conditions: locations that experienced cooler springs had higher defoliation frequencies from 1967 to 1998 in Ontario, and longer outbreak duration and higher severity from 1961 to 1990 in Québec (Candau and Fleming, 2005; Gray, 2013). The lower

frequencies of defoliation with warm spring temperatures (Candau and Fleming, 2005) were thought to be caused by early emerged larvae suffering from heavy mortality later due to the exposure to late frosts and a frozen-ground-caused phenological mismatch.

Extreme cold winter temperature, typically below $-40\text{ }^{\circ}\text{C}$, can cause mortality of overwintering L2 (Blais, 1958). We evaluated this possible effect, but meteorological data showed that very few of the sample points in our study area experienced this low threshold temperature in winter during the study years.

Although forest composition has also been viewed as an important variable influencing SBW survival by affecting the diversity and populations of SBW parasitoids (Zhang et al., 2018), or the phenological matches between host trees and the insects (Nealis and Régnière, 2004), species composition did not appear as an important variable influencing SBW population in this study. This may have been because the L2 data used were collected from branches from each of three balsam fir or spruce trees, and the fishnet L2 data was interpolated, with the resulting values highly dependent on distances to surrounding raw L2 points. In addition, effects of forest composition might be detected at a different scale by using different neighborhood sizes.

Topography and site quality did not show any clear influences on SBW population during this analysis. Effects of insecticide spray

treatments also were not significant in our analysis. This was surprising because insecticide treatments have effectively reduced SBW within treated blocks in the study area (MacLean et al., 2019). The lack of stronger overall effects may have occurred because nearly all SBW populations > 6.5 L2 per branch in the study area were treated in insecticide spray blocks in the EIS project, so that there was effectively no control for comparison. When other factors were controlled, SBW populations within treated blocks were consistently reduced (MacLean et al., 2019). Because of this inherent confounding issue of ‘always treat high L2’, spray treatment may be more of a background effect (factor) that becomes a condition of the model (the model assumes protection will always be applied in a similar way each year). The study did not have an effective design to examine effects of treatment. Furthermore, this result may have been influenced by the relatively small area treated with insecticide spraying (which from 2015 to 2018 increased from 15,000 ha to 56,000 ha, 147,000 ha, to 199,000 ha per year), relative to the overall study area size of 3.7 million ha. We analyzed various spray treatment (or not) combinations in years N-1 and N-2, and also tried only year N-1 spray treatment effects, but neither had important effects on overall population trends.

4.3. Regression analysis of spatial population data

Better performance and less spatial autocorrelation in residuals demonstrated the suitability of SAR models compared to classic OLS models for analyzing these ecological data. These results suggest that modeling forest insect population dynamics should account for spatial structures in both target-insect distribution and underlying environmental factors. Given the spatial dynamics and periodicity of defoliator systems, it is essential to carry out outbreak phase-dependent analyses and to be careful about the study scale selection. There are many different statistical approaches to studies of spatial ecological data, depending on study objectives, data characteristics, sample sizes, model types, and sources of spatial autocorrelation (Dormann et al., 2007; Beale et al., 2010). Performance of different spatial models has been estimated and compared in several studies, which provide evidence of model application in different conditions (e.g., Dormann et al., 2007; Kissling and Carl, 2008; Bini et al., 2009; Beale et al., 2010).

4.4. Applications of the study results

The onset of the current SBW outbreak from 2015 (first forecast of defoliation by NBERD) to 2018 in New Brunswick gave us a good opportunity to explore possible factors influencing the initiation of SBW outbreaks, and to determine which variables contribute to population increases. Understanding the relationships among variables, and possible mechanisms, can help forest and pest managers to forecast future insect outbreak situations and make better management decisions. Identifying subsequent-year high SBW population areas can focus and reduce the sampling effort required to accurately estimate regional L2 density. In theory, rather than having a consistent sample of L2 points, it would be more efficient to stratify sampling intensity by likelihood of upcoming year L2 populations predicted to exceed the EIS treatment threshold of 6.5 L2 per branch. Additionally, our results suggested that previous-year outbreak conditions were most important in determining current-year SBW population levels. Therefore, applying insecticides on those outbreak “hot spots” in the preceding year to prevent further growth of local SBW populations is advisable, which corresponds to the current strategy used by the EIS project (MacLean et al., 2019). Attention to such temporal correlation of populations should be given when defining management strategies for other insect pests with similar population dynamics or mobility as SBW. However, factors influencing insect outbreaks are phase-dependent (e.g., Bouchard and Auger, 2014), and relationships between factors determined during outbreak initiation can differ during the development or collapse phases of an outbreak.

5. Conclusions

Using spatial regression models, we determined the variables that influenced increases of SBW populations in northern New Brunswick from 2014 to 2018. Variables important in determining upcoming year SBW L2 population were previous year local L2 population, proximity to high L2 population locations, and early spring degree days. However, the relationships quantified were inconsistent across years, so a model fit in a given year may not be applicable to predict L2 populations two or more years afterwards. Also, variables that are important during outbreak initiation may weaken or vary in spread and collapse phases of the outbreak. The combined-year (2015–2018) LME model performed the worst and underestimated L2 populations in forecasts. Other variables including forest species composition, topography, site quality, and insecticide spray history were not important. High SBW L2 populations tended to be at locations closer to or at the sites where the previous year local SBW population was high. The current SBW outbreak in New Brunswick expanded from the north, adjacent Gaspé-Bas St. Laurent region of Québec, and our results mathematically describe the expansion and retraction of the current SBW outbreak over the initial 5 years. Our results help determine which variables influenced SBW outbreak initiation and relationships between the SBW population and other influencing variables, and also demonstrate the suitability of spatial regression models for analysis of ecological data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by the Spruce Budworm Early Intervention Strategy project supported by the Healthy Forest Partnership, Atlantic Innovation Fund, Canada, and by Natural Resources Canada, with grant number 203544 to Dr. David A. MacLean. We gratefully acknowledge contributions in providing the research data from the New Brunswick Department of Energy and Resource Development, Spruce Budworm Early Intervention Strategy Research group, Gouvernement du Québec, Ministère des Forêts, de la Faune et des Parcs, La Société de Protection des Forêts contre les Insectes et Maladies, and the Forest Watershed Research Centre within the Faculty of Forestry and Environmental Management at the University of New Brunswick. We also thank Luke Amos-Binks and Jeremy Gullison of the New Brunswick Department of Energy and Resource Development who provided access to SBW L2 and insecticide spray treatment data.

References

- Anselin, L., 1988. *Spatial econometrics: methods and models*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Augustin, N.H., Mugglestone, M.A., Buckland, S.T., 1996. An autologistic model for the spatial distribution of wildlife. *J. Appl. Ecol.* 33, 339–347.
- Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J., Elston, D.A., 2010. Regression analysis of spatial data. *Ecol. Lett.* 13, 246–264.
- Bini, L.M., Diniz-filho, J.A.F., Rangel, T.F.L.V.B., Akre, S.B., Albaladejo, R.G., Albuquerque, F.S., Aparicio, A., Araujo, M.B., Baselga, A., Beck, J., Bellocq, M.I., Böhning-gaese, K., Borges, P.A.V., Castro-parga, I., Chey, V.K., Chown Jr, S.L., David, P.D.M., Ferrer-castán, S., Field, D., Filloy, R., Fleishman, J., Jose, E., Albaladejo, F., Albuquerque, R.G., Aparicio, F.S., Araujo, A., Baselga, M.B., Beck, A., Bellocq, J., Bohning-gaese, M.I., Borges, K., Castro-parga, P.A.V., Chey, I., Chown, V.K., Marco, S.L., Dobkin, P.D., Ferrer-castain, D.S., Field, D., Filloy, R., Fleishman, J., Gmez, E., Hortal, J.F., Iverson, J., Kerr, J.B., Kissling, J.T., Kitching, W.D., Len-cortes, J., Lobo, J.L., Montoya, J.M., Morales-castilla, D., Moreno, I., Oberdorff, J.C., Olalla-t, T., Pausas, M.A., Qian, J.G., Rahbek, H., Rodriguez, C., Rueda, M.A., Ruggiero, M., Sackmann, A., Sanders, P., Terribile, N.J., Vetaas, L.C., Hawkins, O.R., B.A., 2009. Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography* 32, 193–204.
- Bivand, R., 2006. Implementing spatial data analysis software tools in R. *Geogr. Anal.* 38,

- 23–40.
- Bjørnstad, O.N., Ims, R.A., Lambin, X., 1999. Spatial population dynamics: analyzing patterns and processes of population synchrony. *Trends Ecol. Evol.* 14, 427–432.
- Blais, J.R., 1958. Effects of 1956 spring and summer temperatures on spruce budworm populations (*Choristoneura fumiferana* Clem.) in the Gaspé peninsula. *Can. Entomol.* 90, 354–361.
- Blais, J.R., 1983. Trends in the frequency, extent, and severity of spruce budworm outbreaks in eastern Canada. *Can. J. For. Res.* 13, 539–547.
- Bouchard, M., Auger, I., 2014. Influence of environmental factors and spatio-temporal covariates during the initial development of a spruce budworm outbreak. *Landsc. Ecol.* 29, 111–126.
- Bouchard, M., Régnière, J., Therrien, P., 2018. Bottom-up factors contribute to large-scale synchrony in spruce budworm populations. *Can. J. For. Res.* 48, 277–284.
- Boulanger, Y., Fabry, F., Kilambi, A., Pureswaran, D.S., Sturtevant, B.R., Saint-Amant, R., 2017. The use of weather surveillance radar and high-resolution three dimensional weather data to monitor a spruce budworm mass exodus flight. *Agric. For. Meteorol.* 234, 127–135.
- Campbell, E.M., MacLean, D.A., Bergeron, Y., 2008. The severity of budworm-caused growth reductions in balsam fir/spruce stands varies with the hardwood content of surrounding forest landscapes. *For. Sci.* 54, 195–205.
- Candau, J.-N., Fleming, R.A., 2005. Landscape-scale spatial distribution of spruce budworm defoliation in relation to bioclimatic conditions. *Can. J. For. Res.* 35, 2218–2232.
- Dale, M.R., Fortin, M.J., 2014. *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press, Cambridge, UK.
- Dormann, C.F., McPherson, J.M., Arau, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hitzel, A., Jetz, W., Kissling, W.D., Ohlemu, R., Peres-neto, P.R., Schurr, F.M., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* 30, 609–628.
- Dupont, A., Bélanger, L., Bousquet, J., 1991. Relationship between balsam fir vulnerability to spruce budworm and ecological site conditions of fir stands in central Québec. *Can. J. For. Res.* 21, 1752–1759.
- Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. *J. Anim. Ecol.* 77, 802–813.
- Erdle, T.A., 2008. *Management Alternatives for New Brunswick's Public Forest: Report of the New Brunswick Task Force on Forest Diversity and Wood Supply*. New Brunswick Department of Natural Resources, Fredericton, NB, Canada.
- Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 29, 1189–1232.
- Furze, S., Ogilvie, J., Arp, P.A., 2017. Fusing digital elevation models to improve hydrological interpretations. *J. Geogr. Inf. Syst.* 9, 558–575.
- Goulard, M., Laurent, T., Thomas-Agnan, C., 2017. About predictions in spatial autoregressive models: optimal and almost optimal strategies. *Spatial Econ. Anal.* 12, 304–325.
- Gray, D.R., 2008. The relationship between climate and outbreak characteristics of the spruce budworm in eastern Canada. *Clim. Change* 87, 361–383.
- Gray, D.R., 2013. The influence of forest composition and climate on outbreak characteristics of the spruce budworm in eastern Canada. *Can. J. For. Res.* 43, 1181–1195.
- Greenbank, D.O., 1957. The role of climate and dispersal in the initiation of outbreaks of the spruce budworm in New Brunswick II. The role of dispersal. *Can. J. Zool.* 35, 385–403.
- Haining, R., 2003. *Spatial Data Analysis: Theory and Practice*. Cambridge University Press, Cambridge, UK.
- Hart, S.J., Veblen, T.T., Schneider, D., Molotch, N.P., 2017. Summer and winter drought drive the initiation and spread of spruce beetle outbreak. *Ecology* 98, 2698–2707.
- Hennigar, C., Weiskittel, A., Allen, H.L., MacLean, D.A., 2017. Development and evaluation of a biomass increment-based index for site productivity. *Can. J. For. Res.* 47, 400–410.
- Hoef, J.M.V., Hanks, E.M., Hooten, M.B., 2018. On the relationship between conditional (CAR) and simultaneous (SAR) autoregressive models. *Spat. Stat.* 25, 68–85.
- Kissling, W.D., Carl, G., 2008. Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecol. Biogeogr.* 17, 59–71.
- Klutsch, J.G., Negrón, J.F., Costello, S.L., Rhoades, C.C., West, D.R., Popp, J., Caissie, R., 2009. Stand characteristics and downed woody debris accumulations associated with a mountain pine beetle (*Dendroctonus ponderosae* Hopkins) outbreak in Colorado. *For. Ecol. Manage.* 258, 641–649.
- Kuhn, M., 2008. Building predictive models in R using the caret package. *J. Stat. Softw.* 28, 1–26.
- Legendre, P., Legendre, L., 2012. *Numerical Ecology*. Elsevier, Amsterdam, Netherlands.
- Lichstein, J.W., Simons, T.R., Shiner, S.A., Franzreb, K.E., 2002. Spatial autocorrelation and autoregressive models in ecology. *Ecol. Monogr.* 72, 445–463.
- Liebold, A.M., Haynes, K.J., Bjørnstad, O.N., 2012. Spatial synchrony of insect outbreaks. In: Barbosa, P., Letourneau, D.K., Agrawal, A.A. (Eds.), *Insect Outbreaks*. Revisited. Blackwell Publishing Ltd., Chichester, UK, pp. 113–125.
- Lindstrom, M.J., Bates, D.M., 1988. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *J. Am. Stat. Assoc.* 83, 1014–1022.
- MacKinnon, W.E., MacLean, D.A., 2003. The influence of forest and stand conditions on spruce budworm defoliation in New Brunswick. *Canada. For. Sci.* 49, 657–667.
- MacLean, D.A., Amirault, P., Amos-Binks, L., Carleton, D., Hennigar, C., Johns, R., Régnière, J., 2019. Positive results of an early intervention strategy to suppress a spruce budworm outbreak after five years of trials. *Forests* 10, 448.
- Magnussen, S., Boudewyn, P., Alfaro, R., 2004. Spatial prediction of the onset of spruce budworm defoliation. *For. Chron.* 80, 485–494.
- Maron, J.L., Harrison, S., Greaves, M., 2001. Origin of an insect outbreak: escape in space or time from natural enemies? *Oecologia* 126, 595–602.
- Miller, C.A., 1958. The measurement of spruce budworm populations and mortality during the first and second larval instars. *Can. J. Zool.* 36, 409–422.
- Miller, C.A., Kettela, E.G., McDougall, G.A., 1971. *A Sampling Technique for Overwintering Spruce Budworm and its Applicability to Population Surveys*. Canadian Forestry Service, Fredericton, NB, Canada. Inf. Rep. M-X-25.
- Murphy, P.N., Ogilvie, J., Connor, K., Arp, P.A., 2007. Mapping wetlands: a comparison of two different approaches for New Brunswick, Canada. *Wetlands* 27, 846–854.
- Myers, J.H., 1998. Synchrony in outbreaks of forest Lepidoptera: a possible example of the Moran effect. *Ecology* 79, 1111–1117.
- Nagelkerke, N.J.D., 1991. A note on a general definition of the coefficient of determination. *Biometrika* 78, 691–692.
- Nealis, V.G., Régnière, J., 2004. Insect–host relationships influencing disturbance by the spruce budworm in a boreal mixedwood forest. *Can. J. For. Res.* 34, 1870–1882.
- Needham, T., Kershaw, J.A., MacLean, D.A., Su, Q., 1999. Effects of mixed stand management to reduce impacts of spruce budworm defoliation on balsam fir stand-level growth and yield. *North. J. Appl. For.* 16, 19–24.
- QMRNF, 2018. Aires Infestées par la Tordeuse des Bourgeons de l'épinette au Québec en 2018. Québec Ministère des Ressources Naturelles et de la Faune. http://www.mffp.gouv.qc.ca/publications/forets/fimaq/insectes/tordeuse/TBE_2018_P.pdf (accessed on 2 January 2019).
- R Development Core Team, 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Régnière, J., Lysyk, T., 1995. Population dynamics of the spruce budworm, *Choristoneura fumiferana*. In: Armstrong, J.A., Ives, W.G.H. (Eds.), *Forest Insect Pests in Canada*. Canadian Forest Service, Science and Sustainable Development Directorate, Ottawa, ON, Canada, pp. 95–105.
- Régnière, J., Nealis, V.G., 2007. Ecological mechanisms of population change during outbreaks of the spruce budworm. *Ecol. Entomol.* 32, 461–477.
- Régnière, J., Saint-Amant, R., Béchard, A., 2014. *BioSIM 10 User's Manual*. Canadian Forest Service, Québec, QC, Canada Inf. Rep. LAU-X-137E.
- Régnière, J., Delisle, J., Sturtevant, B.R., García, M., Saint-Amant, R., 2019. Modeling migratory flight in the spruce budworm: temperature constraints. *Forests* 10, 802.
- Ridgeway, G., 2007. *Generalized boosted models: a guide to the GBM package*. R Package Vignette.
- Robert, L.-E., Sturtevant, B.R., Cooke, B.J., James, P.M.A., Fortin, M.-J., Townsend, P.A., Wolter, P.T., Kneeshaw, D., 2018. Landscape host abundance and configuration regulate periodic outbreak behavior in spruce budworm *Choristoneura fumiferana*. *Ecography* 41, 1556–1571.
- Rose, A.H., Blais, J.R., 1954. A relation between April and May temperatures and spruce budworm larval emergence. *Can. Entomol.* 86, 174–177.
- Royama, T., 1984. Population dynamics of the spruce budworm *Choristoneura fumiferana*. *Ecol. Monogr.* 54, 429–462.
- Royama, T., MacKinnon, W.E., Kettela, E.G., Carter, N.E., Hartling, L.K., 2005. Analysis of spruce budworm outbreak cycles in New Brunswick, Canada, since 1952. *Ecology* 86, 1212–1224.
- Sokal, R.R., Oden, N.L., 1978. Spatial autocorrelation in biology. I Methodology. *Biol. J. Linn. Soc.* 10, 199–228.
- Su, Q., MacLean, D.A., Needham, T.D., 1996. The influence of hardwood content on balsam fir defoliation by spruce budworm. *Can. J. For. Res.* 26, 1620–1628.
- Tognelli, M.F., Kelt, D.A., 2004. Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. *Ecography* 27, 427–436.
- Wall, M.M., 2004. A close look at the spatial structure implied by the CAR and SAR models. *J. Stat. Plan. Inference.* 121, 311–324.
- Watson, D.F., Philip, G.M., 1985. A refinement of inverse distance weighted interpolation. *Geoprocessing* 2, 315–327.
- Williams, D.W., Liebold, A.M., 2000. Spatial synchrony of spruce budworm outbreaks in eastern North America. *Ecology* 81, 2753–2766.
- Zhang, B., MacLean, D.A., Johns, R.C., Eveleigh, E.S., 2018. Effects of hardwood content on balsam fir defoliation during the building phase of a spruce budworm outbreak. *Forests* 9, 530.