

Quantifying the sources of epistemic uncertainty in model predictions of insect disturbances in an uncertain climate

David R. Gray¹

Received: 4 November 2016 / Accepted: 16 May 2017 / Published online: 21 June 2017
© INRA and Springer-Verlag France 2017

Abstract

• **Key message** Natural disturbance can disrupt the anticipated delivery of forest-related ecosystem goods and services. Model predictions of natural disturbances have substantial uncertainties arising from the choices of input data and spatial scale used in the model building process, and the uncertainty of future climate conditions which are a major driver of disturbances. Quantifying the multiple contributions to uncertainty will aid decision making and guide future research needs.

• **Context** Forest management planning has been able, in the past, to rely on substantial empirical evidence regarding tree growth, succession, frequency and impacts of natural disturbances to estimate the future delivery of goods and services. Uncertainty has not been thought large enough to warrant consideration. Our rapidly changing climate is casting that empirical knowledge in doubt.

• **Aims** This paper describes how models of future spruce budworm outbreaks are plagued by uncertainty contributed by (among others): selection of data used in the model building process; model error; and uncertainty of the future climate and forest that will drive the future insect outbreak. The contribution of each to the total uncertainty will be quantified.

• **Methods** Outbreak models are built by the multivariate technique of reduced rank regression using different datasets. Each

model and an estimate of its error are then used to predict future outbreaks under different future conditions of climate and forest composition. Variation in predictions is calculated, and the variance is apportioned among the model components that contributed to the epistemic uncertainty in predictions.

• **Results** Projections of future outbreaks are highly uncertain under the range of input data and future conditions examined. Uncertainty is not uniformly distributed spatially; the average 75% confidence interval for outbreak duration is 10 years. Estimates of forest inventory for model building and choice of climate scenario for projections of future climate had the greatest contributions to predictions of outbreak duration and severity.

• **Conclusion** Predictions of future spruce budworm outbreaks are highly uncertain. More precise outbreak data with which to build a new outbreak model will have the biggest impact on reducing uncertainty. However, an uncertain future climate will continue to produce uncertainty in outbreak projections. Forest management strategies must, therefore, include alternatives that present a reasonable likelihood of achieving acceptable outcomes over a wide range of future conditions.

Keywords Decision making · Ecosystem good and services · Pest management · Forest planning · Spruce budworm · Insect outbreak projections

Handling Editor: Aurélien Salle

Contribution of the co-authors: no co-authors

✉ David R. Gray
david.gray@canada.ca

¹ Natural Resources Canada, Canadian Forest Service - Atlantic Forestry Centre, P.O. Box 4000, Fredericton, NB E3B 5P7, Canada

1 Introduction

Boreal forests provide valued ecosystem goods and services (G&S) at multiple scales, from the community to the global level. These ecosystem G&S include fibre production (for lumber and paper), biodiversity, clean air and water, hunting and fishing, recreation and others (Gauthier et al. 2015). Boreal forests can also act as important carbon sinks

(Dymond et al. 2010). Ecosystem G&S are delivered by the heterogeneity of the forest landscape—the matrix of forest “types”—which is a result of climate, physical environment, natural disturbances and human disturbances (Grondin et al. 2014). Forest management planning (and policy) for the delivery of the desired ecosystem G&S involves a prediction of the future matrix of forest types (species mixtures and ages), given the current matrix, a set of management practices, and the assumed species responses to the management practices and to the future abiotic conditions such as soil, temperature and precipitation. These predictions are uncertain, and the delivery of desired ecosystem G&S may be jeopardized. Minimizing the uncertainty will aid decision making in policy and forest management planning. Understanding and quantifying the sources of the uncertainty will guide future research to reduce the uncertainty.

Natural disturbances are important drivers of the boreal forest. The spruce budworm (*Choristoneura fumiferana* Clemens) (SBW) is arguably the second most impactful natural disturbance agent (after wildfire) in Canada’s boreal forest. Spruce budworm is a native defoliator in North America that feeds principally on *Abies balsamea* (L.) (Mill.), *Picea glauca* (Moench) Voss, *P. rubens* Sarg. and *P. mariana* (Mill.) Brittoni, Sterns & Poggenb. Populations undergo more or less regular 30–40 cycles of abundance. During periods of high population levels (outbreaks), a few hundred larvae may be found on a single branch of a host tree. Between the outbreak periods, populations may be so low as to make it difficult to find a single larva among several hundred branches (Royama 1992). Outbreaks occur somewhat synchronously over extensive areas (Royama 1984; Candau et al. 1998; Gray et al. 2000), but outbreak duration varies regionally from as few as one to as many as 20 years. Mortality and growth loss are typically very high over the course of an outbreak. Approximately 45% of the host trees in eastern Canada were killed during an outbreak of the 1910s and 1920s (Swaine et al. 1924). Between 1977 and 1987, the annual timber loss to SBW was $40\text{--}100 \times 10^6 \text{ m}^3$, greater than the $25 \times 10^6 \text{ m}^3$ annual loss to fire during the same period (Natural Resources Canada 2014).

Given the very large impact of a future SBW outbreak on the forest matrix, there is considerable interest in predicting future outbreaks and their impacts. It is generally accepted that climate and forest composition are key factors in outbreak dynamics (see Candau and Fleming (2005) and Gray (2013) for a brief review), but the precise manner in which they exert an influence on the initiation, severity (annual defoliation level) and duration of an outbreak is not well understood. More than 43 primary parasitoids and 21 entomopathogens are associated with SBW populations on *A. balsamea* alone (Eveleigh et al. 2007). Temperature and/or precipitation affect many life history traits (e.g. aggregation, developmental rates, phenology, fecundity and survival and dispersal) of the SBW, its natural enemies and its hosts. Forest composition

(including host abundance) is affected by climate and has a demonstrated effect on SBW and its natural enemies (Fig. 1).

The complexity of the SBW outbreak system has caused authors to resort to correlation-based (as opposed to process-based) models to model one or more SBW outbreak characteristics. All of the techniques rely on a similar combination of input data for the modelling exercise: spatially referenced records of annual defoliation for an extended time period (an outbreak cycle) from which the outbreak characteristic/characteristics is/are derived (the response variable(s)) and estimates of forest composition and climatic characteristics thought to be relevant to the SBW life cycle (the explanatory variables) at the same spatial scale. Throughout this paper, “spatial scale” denotes the size of the quadrat over which the input data are aggregated and in which the projected outcomes are estimated (also known as “spatial resolution”). Candau and Fleming (2005, 2011) used a regression tree and the Random Forests classification technique to model the outbreak duration (the frequency of recorded annual defoliation in an outbreak cycle (1967–1998)) in Ontario at a 1-km (Candau and Fleming 2005) or a 10-km (Candau and Fleming 2011) scale. Boulanger et al. (2015) also modelled the outbreak duration. They added a spruce budworm population “growth potential” variable derived from a process-based model (Régnière et al. 2012) to the climate and forest composition matrix, and then built a consensus model from a mix of correlative techniques at a 10-km scale for eastern Canada (east of the Manitoba-Ontario border). Gray (2008, 2013) used constrained ordination (reduced rank regression) because of its ability to simultaneously model the two response variables—outbreak duration and outbreak severity (defoliation level during the outbreak)—that he argues is necessary to estimate impacts of the outbreak. He used a 30,000-ha (approx. 17.3 km) spatial scale for eastern Canada.

Despite the importance and implications of the predictions of future outbreaks on the choice of a management plan that is expected to deliver the desired ecosystem G&S, the uncertainty (that is, any departure from the unachievable ideal of complete determinism (Walker et al. 2003)) of the predictions has never been adequately examined. Candau and Fleming (2011) reported the mean square error of their outbreak model, and they predicted future outbreak durations by applying their model to the predicted future climates of three climate models (CGCM2 (Environment Canada), HadCM3 (Met Centre) and CSIRO Mk2) and two climate scenarios (SRES A2 and B2) of the International Panel on Climate Change (IPCC). But they did not quantify the prediction variability (the uncertainty) in outbreak duration that would be generated by the combination of their model error and the choices of climate model and climate scenario. Gray (2013) reported the R^2 of his outbreak model (duration and defoliation level), but applied his model to the predicted future climates of only one climate model (CGCM3 (Environment Canada)) and only one climate

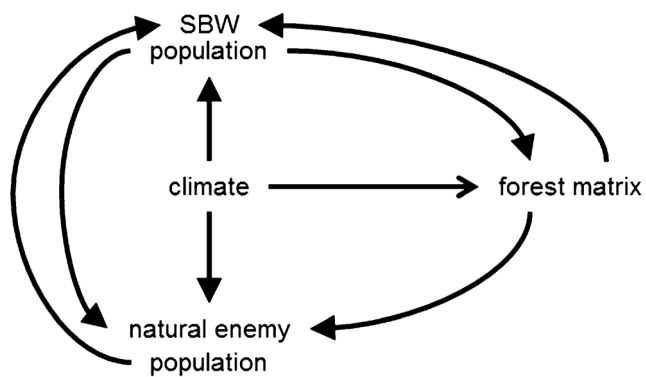


Fig. 1 The interrelationships among populations of the spruce budworm and its natural enemies, climate and matrix of forest compositions

scenario (A2). He did not report the prediction variability (uncertainty) in outbreak duration or severity that would be generated from his model uncertainty and did not consider the contributions to uncertainty of alternative climate models and climate scenarios. Boulanger et al. (2015) used three climate scenarios (RCPs 2.6, 4.6 and 8.5) of just one climate model (CanESM2 (Environment Canada)). They analysed the uncertainty in outbreak prediction only and contributed by the choice of data for building their outbreak model, the correlative technique used to build their model, and climate scenario. They did not include alternative climate models. In summary, one or more key contributors to uncertainty have been missing in previous investigations.

There are numerous typologies of model uncertainty (Walker et al. 2003; Refsgaard et al. 2007). Model uncertainty may be stochastic, arising from the inherent variability of the system and the impossibility of completely capturing the variability, or epistemic, arising from an incomplete knowledge of the system or an inability to adequately model the system. Stochastic uncertainty is nonreducible; epistemic uncertainty can be reduced by better knowledge and/or better data. There are numerous ways to further subdivide prediction uncertainty into its many epistemic sources. In this paper, I will refer to prediction uncertainty from (Fig. 2):

1. Model data sources. These include the following:
 - a. historical SBW defoliation (the response variable), which may be limited, inaccurate or imprecise;
 - b. forest composition and climate (the explanatory variables), which may be limited, inaccurate or imprecise; and
 - c. the spatial scale at which the data were collected/aggregated.
2. Model structure. This includes the following:
 - a. model type, which is the choice from many alternative algorithmic strategies applicable in the context of a

correlation question (e.g. robust linear model, general additive model, multivariate adaptive regression splines, random forests and redundancy analysis); and

- b. parameter estimates.

3. Future conditions that drive the future SBW outbreaks. These include the following:
 - a. future climate scenario (Representative Concentration Pathway (RPC) of the IPCC) run by the climate model;
 - b. climate model running the climate scenario;
 - c. forest inventory; and
 - d. consistency in spatial scale between the model data spatial sources and the future conditions (forest composition and climate).

Residual error, which is the discrepancy between predicted and observed outcomes under the individual combinations of model data sources and parameter values (1 and 2, earlier), will be considered as stochastic uncertainty inherent in the natural system.

In this paper, I describe a Monte Carlo analysis that generates multiple versions of predicted future SBW outbreak characteristics (duration and defoliation) at each location by varying the model data sources (forest composition), spatial scale, parameter estimates and future conditions (climate and forest inventory). A variety of model types is not included because reduced rank regression (see below) is the only model type available that will simultaneously model multiple response variables. The contribution of each of the other sources of epistemic uncertainty is quantified and discussed.

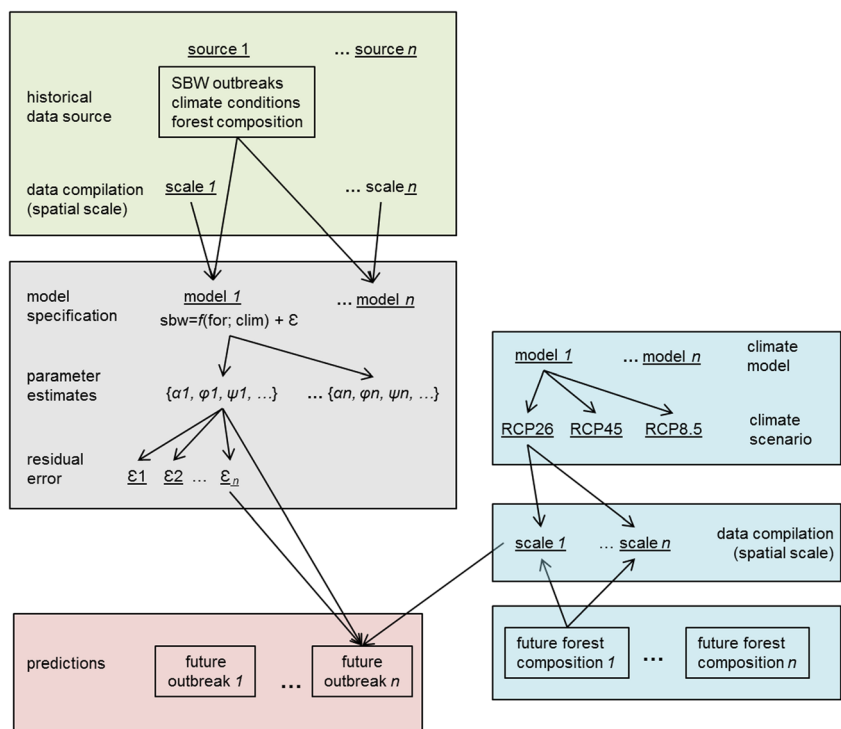
2 Materials and methods

The methodology follows these steps: compile datasets of outbreak characteristics, forest composition and climate at common spatial scales; build multiple models of outbreak characteristics as functions of forest composition and climate at each spatial scale by repeatedly sampling the data sources; generate future climates for the 2031–2060 time period at each spatial scale using a variety of climate models and climate scenarios; and simulate (predict) future outbreaks for the 2031–2060 time period based on each combination of model and future condition at each spatial scale. This will produce multiple, equally likely predictions at each location of a future outbreak that together constitute the uncertainty at the location.

2.1 Model building data sources (historic conditions)

Spatial scale The building of a correlation-based model needs response (SBW outbreak characteristics) and explanatory

Fig. 2 A division of prediction uncertainty into three broad categories: the input data used in model building and calibration (*green box*); model structure (*grey box*) and the future conditions that will drive the model (*blue box*). Final prediction uncertainty (*red box*) is an accumulation of the uncertainties from the three categories. For simplicity, not all combinations are shown with arrows



(forest composition and climate) data compiled at a common spatial scale. Data were compiled, and models were built at the two spatial scales, 10×10 km and 15×15 km, that roughly match the two scales used in previous examinations of outbreak characteristics in this area (Gray 2008; Gray 2013; Boulanger et al. 2015). A uniform grid pattern of cells at each scale was placed over eastern Canada (east of the Manitoba-Ontario border). Spruce budworm outbreak characteristics, forest composition and climatic conditions were summarized within each cell of each spatial scale. A total of 11,825 (15×15 km) and 25,701 (10×10 km) grid cells occurred on land within the eastern Canada study area.

Spruce budworm outbreak characteristics The National Forest Information System (Natural Resources Canada 2014) compiled the results of the extensive annual surveys of SBW defoliation done by individual provinces. Surveys were (and are still) done according to provincial protocols, and those protocols vary slightly between years and among jurisdictions. The variations were reduced to the common codes of nil (0), light-medium (1) and severe (2) defoliation. The codes were transformed into a ratio scale to reflect the difference between the midpoints of the range of defoliation within each code: 0 (nil), 1 (light-medium (20%)) and 3.25 (severe (65%)). The defoliation polygons from each year (1939–2008) were intersected with the 10×10 km and the 15×15 km grids using ArcGIS (ESRI 2006). An area-weighted average defoliation code was calculated for each year in each grid cell at each scale from the portions of the defoliation polygons in the cell. Outbreaks begin in a staggered fashion across

the landscape. Therefore, the most recent and complete outbreak event was extracted from the 1931–2008 series in each grid cell. Outbreak duration (the length of the extracted data string) and defoliation code (average of the ratio scale defoliation codes in the data string) describe the outbreak in each grid cell.

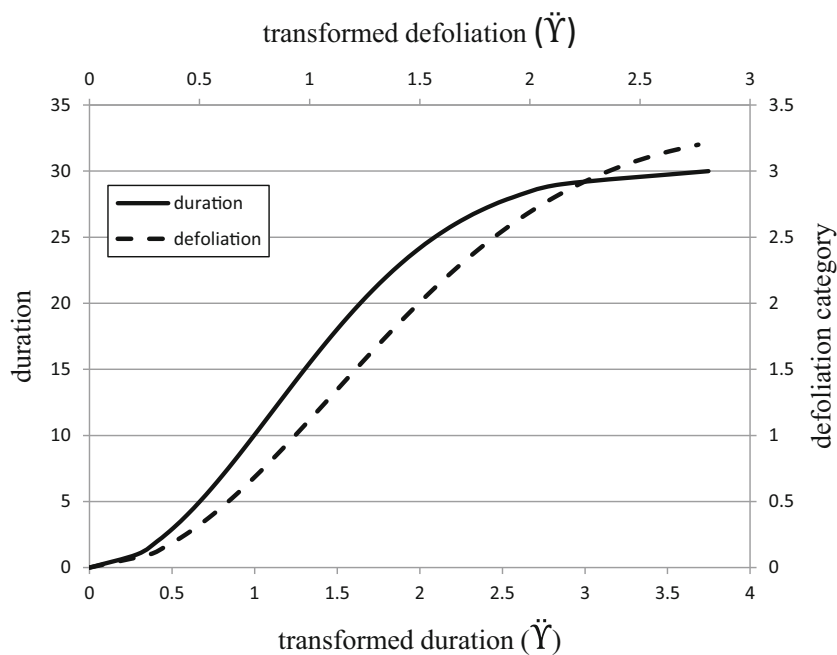
Assuming that future SBW populations will still exhibit their cyclic behaviour (outbreak periods alternating with periods of endemic populations), there is an upper limit to outbreak duration. Similarly, the defoliation code cannot exceed 3.25. Therefore, duration and defoliation codes were transformed to variables whose linear responses (as per the linear reduced rank regression) to climate and forest composition would produce a mildly sigmoidal increase (Fig. 3) in the back-transformed duration and defoliation code:

$$\tilde{Y} = \sqrt{17 \times \frac{\ln\left(\frac{-1}{\frac{y}{\Phi} - 1}\right)}{3}} \tag{1}$$

where y is the original duration or defoliation code, Φ is a constant (30.0 and 3.25 for the transformation of duration and defoliation code, respectively), and \tilde{Y} is the transformed duration or defoliation code that is modelled together by the reduced rank regression.

Forest composition The Canadian Forest Inventory (CanFI) is a compilation of 48 individual inventories from independent

Fig. 3 The transformation of an outbreak characteristic (duration or defoliation (y)). The transformed variable (\check{Y}) responds linearly to increases in climate and forest composition (as per the reduced rank regression method) and the back-transformed duration and defoliation variables exhibit mildly sigmoidal increases toward an asymptote



jurisdictions (Gray and Power 1997), 17 of which are in the study area, that contain detailed descriptions including area (ha) and species volumes ($\text{m}^3 \text{ha}^{-1}$) of individual forest stands. Multiple (often many hundreds) individual records, each constituting a single forest stand, are georeferenced to a single cell (hereafter CanFI mapsheet); CanFI mapsheet size varies within and among jurisdictions. A composite forest type was estimated for each CanFI mapsheet by aggregating species into four species types (Table 1) and calculating the average volume (m^3) per hectare of each species type. The CanFI mapsheets were intersected with 10×10 km and the 15×15 km grids (above) using ArcGIS, and an area-weighted average $\text{m}^3 \text{ha}^{-1}$ of each species type was calculated for each cell of each grid scale.

An alternative version of forest composition in each 10×10 km and 15×15 km grid cells was derived from the National Forest Inventory (NFI) (Canadian Council of Forest Ministers 2010). The collaborative, multijurisdictional NFI project produced a set of 2×2 km photo plots on a 20×20 km spacing from which Beaudoin et al. (2014) estimated forest cover at a 250-m resolution by k -nearest neighbours interpolation. The same four species types as used for the CanFI database (earlier) were used to calculate average volumes ($\text{m}^3 \text{ha}^{-1}$) per species type in the NFI database. The 250-m resolution estimates of species type volumes were spatially intersected with the 10×10 km and the 15×15 km grid cells of this study using ArcGIS, and an average $\text{m}^3 \text{ha}^{-1}$ of each species type (Table 1) was calculated for each cell of each grid scale. The two versions of forest composition were compared by canonical correlation (vegan package (Oksanen et al. 2015) of R (R Development Core Team 2008)).

There is no forest inventory that can be a perfect temporal match for the spruce budworm outbreaks across the large study area of eastern Canada. The location-specific beginning of the

most recent outbreak varied by 38 years (1960–1998). The dates of the 17 inventories in the study area vary between 1971 and 1994. The photo plots from which the NFI inventory is derived were created between 2000 and 2006. However, summation of forest composition at the 10×10 km and 15×15 km scales used here will result in estimates of forest cover that are close to temporally constant within the cells because forest harvesting in Canada must be done on small blocks (usually <5 ha) which makes a substantial change in cell forest composition unlikely (Boulanger et al. 2015).

Climatic conditions The historic climate was derived from the 1961–1990 Canadian Climate Normals database (available from the Meteorological Service of Canada, Environment Canada). The daily minimum (T_n) and maximum (T_x) temperatures ($^{\circ}\text{C}$) were simulated from the normals for 20 individual years in each cell of each spatial grid using the stochastic weather generator of BioSIM (Régnière and Bolstad 1994; Régnière and St-Amant 2004). BioSIM simulates the T_n , and the T_x at each cell centroid by matching georeferenced sources of weather data (the climate normals) to the cell centroids, adjusting the weather data for differences in latitude, longitude and elevation between the source of weather data and the centroid. BioSIM restores stochastic daily variation to the normals while maintaining the autocorrelation between daily variables that occurs naturally (Régnière and Bolstad 1994).

From the 20 time series of daily weather generated in each cell, the following five climate variables were chosen a priori to reflect conditions associated with SBW life cycle events that are thought important to the progression of an outbreak of a poikilothermic organism (Gray 2013):

Table 1 Four forest species types compiled from the Canadian Forest Inventory (CanFI) and the Canadian National Forest Inventory (NFI) databases

Species type	Species	Spruce budworm host (Y/N)
Balsam fir + red spruce + white spruce	<i>Abies balsamea</i> (L.) Mill.	Y
	<i>P. rubens</i> Sarg. &	
	<i>P. glauca</i> (Moench) Voss	
Black spruce	<i>Picea mariana</i> (Mill.) B. S. P.	Y
Other softwood		N
Hardwood		N

$$\text{October to March "winter days"} = \sum_{m=\text{October}}^{\text{March}} \left[\frac{\sum_{y=1}^{20} wd_{ym}}{20} \right] \quad (2)$$

where wd_{ym} is the number of degree days below $-5\text{ }^{\circ}\text{C}$ in the month;

spring extreme daily maximum temperature

$$= \sum_{m=\text{April}}^{\text{May}} \left[\frac{\sum_{y=1}^{20} \max(Tx_{ym})}{20} \right] \quad (3)$$

$$\text{spring degree days} = \sum_{m=\text{April}}^{\text{May}} \left[\frac{\sum_{y=1}^{20} dd_{ym}}{20} \right] \quad (4)$$

where dd_{ym} is the number of degree days above $5\text{ }^{\circ}\text{C}$ in the month;

summer extreme minimum temperature

$$= \sum_{m=\text{June}}^{\text{August}} \left[\frac{\sum_{y=1}^{20} \min(Tn_{ym})}{20} \right] \quad (5)$$

and

summer extreme maximum temperature

$$= \sum_{m=\text{June}}^{\text{August}} \left[\frac{\sum_{y=1}^{20} \max(Tx_{ym})}{20} \right] \quad (6)$$

2.2 Model structure

Model type Many correlative model types have been examined for their utility in predicting *one* characteristic of future

outbreaks (Boulanger et al. 2015) (e.g. generalized additive model, robust linear model and regression tree (including Random Forests)). But following the logic that two characteristics provide a fuller description of an outbreak than does one, I use redundancy analysis as the model type. Redundancy analysis is a form of constrained ordination in which the principal axes are constrained to be linear combinations of the explanatory variables (Lepš and Šmilauer 2014). Thus, multiple response variables are modelled together in a reduced rank regression model (ter Braak 1994), $Y = MX + E$, where Y is constructed from the matrix of response variables (transformed outbreak duration and defoliation code, in the specific case here), X is the matrix of explanatory variables (five climate variables (Eqs. (2)–(6))) and forest composition ($\text{m}^3 \text{ha}^{-1}$ of the four species types of Table 1), M is the matrix of regression coefficients and E is the error matrix. Model fitting was done with the vegan package (Oksanen et al. 2015) of R (R Development Core Team 2008).

Regression estimates Ten independent samples of 25% of the $10 \times 10 \text{ km}$ and the $15 \times 15 \text{ km}$ grid cells were randomly selected without replacement. Regression coefficients of the $Y = MX + E$ model were estimated for each independent sample of each grid scale.

Residual errors A residual error ($\check{Y}_{\text{observed}} - \check{Y}_{\text{predicted}}$) was calculated for each sampled cell used in each of the 10 iterations of parameter estimation of each spatial scale. A visual examination of residuals did not detect a noticeable spatial trend within an iteration. Therefore, all residuals from an iteration of parameter estimation were grouped, and an empirical frequency distribution was constructed for later sampling during the prediction process for future outbreaks (below).

2.3 Future conditions

Climate models and future (projected) climate scenario

The climate projections of three general circulation models (Table 2) under three future climate scenarios of the IPCC (RCP2.6, RCP4.5 and RCP8.5) were used as the source for the estimates of the five climate variables used in this study (Eqs. (2)–(6)). The “peak and decline” RCP2.6 scenario

represents a future trajectory where radiative forcing peaks at $\sim 3 \text{ W m}^{-2}$ before 2100 and declines to 2.6 W m^{-2} by 2100. The “stabilization without overshoot” RCP4.5 scenario represents a future trajectory where radiative forcing stabilizes at 4.5 W m^{-2} after 2100 without overshooting the 6 W m^{-2} targeted limit. In the “high concentration” RCP8.5 scenario, radiative forcing reaches 8.5 W m^{-2} by 2100. McKenney et al. (2013) summarized these outputs into climate normals at a 10-km resolution. The climate normals formed the input for the simulation of daily weather from which the five climate variables were estimated (see Eqs. (2)–(6)).

Climate variables The daily minimum (T_n) and maximum (T_x) temperatures ($^{\circ}\text{C}$) were simulated for 20 individual years in each cell of each spatial grid ($10 \times 10 \text{ km}$ and $15 \times 15 \text{ km}$) with the climate normals from each combination of climate model and RCP scenario using the stochastic weather generator of BioSIM (see above). The five climate variables (Eqs. (2)–(6)) were calculated from the 20-year BioSIM output.

Forest inventory Previous work (Candau and Fleming 2011; Boulanger et al. 2013; Gray 2013) has assumed that at the temporal and spatial scales used here, there will be no significant changes in forest composition from the historical conditions because forest harvesting in Canada must be done on small blocks ($< 5 \text{ ha}$) which makes a substantial change in the cell forest composition unlikely (Boulanger et al. 2015). There does not exist a forest succession model that is sensitive to climate and can project future forest compositions for the large Canadian study area used here. Therefore, the sensitivity of projected spruce budworm projections to the assumption of an unchanged forest composition was tested by using the same forest composition for historic and future conditions, and by using different forest compositions in the historic and future conditions (i.e. CanFI changing to NFI; or NFI changing to CanFI).

2.4 Simulation procedure

The different combinations of a forest composition (2), scale (2) and independent parameter estimates (10) produced 40 independent models, each with 10 estimates of error distribution. The parameter estimates of each model were used with a combination of future conditions (climate and forest) to predict future SBW transformed outbreak characteristics in a sample of cells. One thousand six hundred cells were randomly selected from the $15 \times 15 \text{ km}$ scale grid, and the cell centroids were cross-referenced with the centroids of the cells of the $10 \times 10 \text{ km}$ scale grid in order to select 1600 locations from the $10 \times 10 \text{ km}$ scale grid that were as close as possible. An error estimate was added to each prediction by randomly selecting from the error distribution of the model. Ten independent selections from the error distribution were made, producing 10 independent estimates (predicted + error) of the two

transformed outbreak characteristic in each of the 1600 cells. Thus, there was a total of 14,400 estimates of transformed outbreak characteristics in each of 1600 cells:

- 40 independent models (from the combinations of a forest inventory (CanFI or NFI) and a spatial scale ($10 \times 10 \text{ km}$ or $15 \times 15 \text{ km}$) and 10 parameter estimations);
- $\times 10$ error distributions for each model
- $\times 36$ future conditions (3 RCP scenarios \times 3 climate models \times 2 forest inventories \times 2 spatial scales).

The projected outbreak characteristics were back-transformed before uncertainty estimation.

2.5 Uncertainty estimation

Type III sums of squares (SS) and mean squares (MS) of the epistemic and stochastic sources of variability in projected outcomes (back-transformed duration and defoliation) were calculated by SAS v9.2 (SAS Institute n.d.) from general linear models (Table 3). All sources were treated as random effects. Each of the 10 iterations of model parameters was specific to a combination of inventory and scale used in the model, and was therefore treated as a nested effect. The model error variance is best estimated by the MS_{error} reported by SAS; similarly, the expected mean square (EMS) of effect i is best estimated by the MS_i reported by SAS (Hicks 1982, pp. 55–57). After formulating the equations for EMS by the usual rules for random and nested effects (Steel and Torrie 1980, pp. 357–358), the estimated variance of each effect (s_i^2) is obtained by substitution and simple re-arrangement. The percent contribution of source i to the total epistemic uncertainty is $s_i^2 / \sum_{\text{source}} s_{\text{source}}^2 \times 100$. Uncertainty in the location-specific predictions of outbreak duration and defoliation level is expressed in (1) absolute terms, as the distance between the 0.125th–0.875th quantiles of the duration and defoliation distributions of each cell (i.e. 75% of the outcomes of each cell); and (2) relative terms, as the ratio of absolute uncertainty divided by median prediction.

3 Results

Median-projected outbreak duration (back-transformed) was 4.8 years. Median-projected outbreak defoliation code (back-transformed) was 0.5. Projected outbreaks are relatively short (4.8 years) and of low defoliation principally because the study area extends to northern areas where future climates and forest composition are still not predicted to be highly supportive of spruce budworm outbreaks. Both projections were highly variable across the landscape as is naturally expected. Projected outbreaks were longest, and defoliation was

Table 2 Three global circulation models used to produce climate projections under the three RCP scenarios

Acronym	Model name	Source	URL
CanESM2	Second Generation Earth System Model	Canadian Centre for Climate Modelling and Analysis	http://www.cccma.ec.gc.ca/data/data.shtml
HadGEM2-ES	Hadley Global Environment Model 2	Met Office, Hadley Centre	https://verc.enes.org/models/earthsystemmodels/metoffice-hadley-centre/hadgem2es
MIROC_ESM	Model for Interdisciplinary Research on Climate, Earth System Model	JAMSECTEC/UT/NIES ^a	http://www.geosci-modeldev.net/4/845/2011/gmd-4-845-2011.html

^a Japan Agency for Marine-Earth Science and Technology/Atmosphere and Ocean Research Institute (The University of Tokyo)/National Institute for Environmental Studies

the most severe in an east-west band through a southern portion of the study area. To the north and to the south of this band, the projected outbreak became shorter (Fig. 4) and defoliation became less severe (Fig. 5).

Projected outbreak characteristics were very uncertain under the range of model building inputs and projection conditions used here. An average of 9.8 years is required to encompass 75% of the projected outcomes of outbreak duration (absolute uncertainty). An average of 1.1 on the defoliation scale is required to encompass 75% of the projected outcomes of outbreak defoliation (absolute uncertainty). The spatial pattern of absolute uncertainty followed roughly that of the projected median of the corresponding outbreak characteristic: absolute uncertainty

was higher where the projected median was higher. The spatial pattern of relative uncertainty was roughly the inverse of the projected median of the corresponding outbreak characteristic: relative uncertainty was higher where the projected median was lower (Figs. 4 and 5). The area of lower relative uncertainty (duration and defoliation) corresponds roughly to the current eastern outbreak range of the spruce budworm.

Projections of future outbreak durations were most uncertain due to the choice of the forest inventory database (69%) for historical conditions (i.e. for model building). The choice of spatial resolution for model building was the second largest contributor (11%) to the uncertainty in future outbreak durations. The choice of climate model,

Table 3 Sources of uncertainty in projections of future SBW outbreak characteristics (transformed duration and defoliation) and the estimated mean squares (EMS) formula, the mean square (MS), estimated variance (s^2) and variance component (VC) of each epistemic source for each outbreak characteristic

Source	n	Designation	EMS	Duration			Defoliation		
				MS _i	s _i ²	VC (%)	MS	s _i ²	VC (%)
Inventory ^a	2	A	$s_{\text{error}}^2 + 57600 \times s_{C(A \times B)}^2 + 1.152E7 \times s_A^2$	18,661,891	1.62	69.25	140,438	0.0122	72.56
Model scale	2	B	$s_{\text{error}}^2 + 57600 \times s_{C(A \times B)}^2 + 1.152E7 \times s_B^2$	2,986,714	0.26	11.02	7952	0.0007	4.05
Model parameters ^b	10	C	$s_{\text{error}}^2 + 57600 \times s_{C(A \times B)}^2$	18,403	0.03	1.36	119	0.0002	1.22
Inventory change ^c	2	D	$s_{\text{error}}^2 + 1.152E7 \times s_D^2$	1,640,066	0.14	6.09	16,363	0.0014	8.46
Scale change ^c	2	E	$s_{\text{error}}^2 + 1.152E7 \times s_E^2$	600	0.00	0.00	6	0.0000	0.00
Climate scenario	3	F	$s_{\text{error}}^2 + 7.68E6 \times s_F^2$	1,014,672	0.13	5.65	17,423	0.0023	13.51
Climate model	3	G	$s_{\text{error}}^2 + 7.68E6 \times s_G^2$	1,188,156	0.15	6.62	244	0.0000	0.19
Predictions error	10								
Cells	1600								
Model error				29.61			0.39		

^a Forest inventory used in model building data

^b Model parameter effect was nested within inventory and model scale

^c Forest composition was changed (CanFI↔NFI), or scale was changed (10 × 10 km↔15 × 15 km) between model building and projected outbreaks

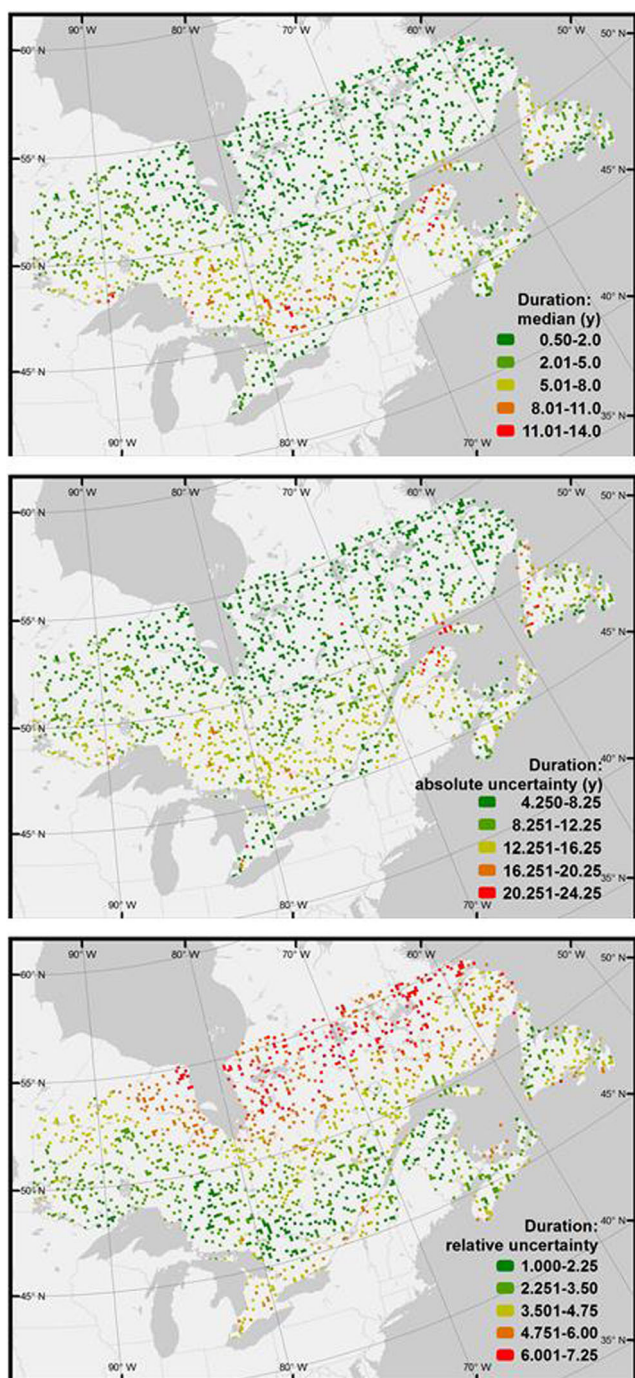


Fig. 4 Median-projected outbreak duration (*top*); absolute uncertainty of projected duration (*middle*) and relative uncertainty of projected duration (*bottom*). Absolute uncertainty is the span between the 0.125 and 0.875 quantiles of the distribution (i.e. 75% of the predicted outcomes). Relative uncertainty is the absolute uncertainty divided by the projected median

RCP scenario and future forest inventory database contributed roughly equally to uncertainty in future outbreak durations (6.6, 5.6 and 6.1%, respectively) (Table 3).

As for future outbreak durations, projections of future outbreak defoliation levels were most uncertain due to the choice of forest inventory for model building (73%).

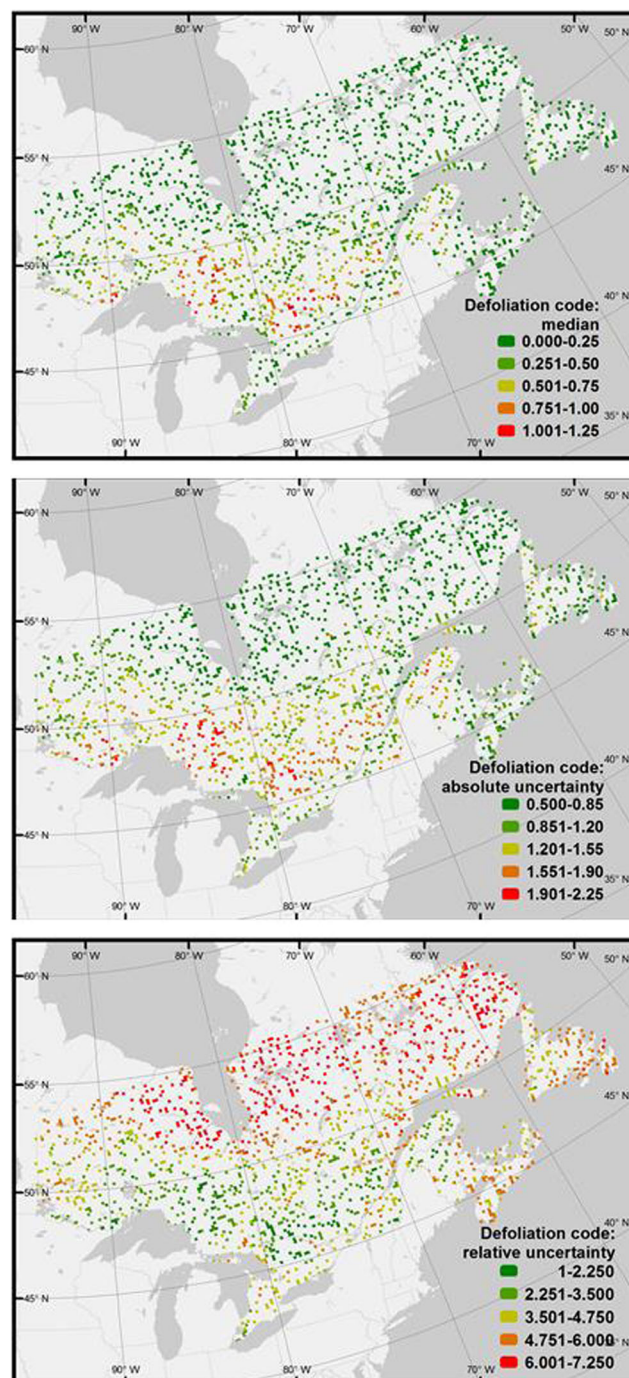


Fig. 5 Median-projected outbreak defoliation level (*top*); absolute uncertainty of projected defoliation level (*middle*) and relative uncertainty of projected defoliation level (*bottom*). Absolute uncertainty is the span between the 0.125 and 0.875 quantiles of the distribution (i.e. 75% of the predicted outcomes). Relative uncertainty is the absolute uncertainty divided by the projected median

The choice of RCP scenario (13.5%) and future forest inventory database (8.5%) were the second and third largest contributors to uncertainty of future outbreak defoliation levels (Table 3).

4 Discussion

Forests have, for millennia, been able to adapt to the slowly changing climatic conditions (Davis et al. 2005). But the unprecedented rapidity of the current changes in climate conditions is challenging, or perhaps overwhelming, the adaptive capacity of forests. Natural disturbances are important determinants of forest composition; climate change is altering the frequency and severity of fires (Flannigan et al. 2009; Boulanger et al. 2013) and the occurrence and severity of insect outbreaks (Dukes et al. 2009). As natural disturbance regimes and postdisturbance successional pathways change under the future climate, the anticipated ecosystem G&S may not materialize. Ochuodho et al. (2012) estimate the economic impact of climate change on Canada's forests at \$2–17 bil. year⁻¹. One suggested strategy of dealing with the effects of climate change is termed “adaptation”, which refers to the “adjustment in natural or human systems in response to actual or expected stimuli or their effects, which moderates harm or exploits beneficial opportunities” (IPCC 2001). These might include modifications to harvesting regimes and species selection at the time of regeneration.

The Canadian government (among others) has recognized the importance of an adaptive management strategy for its forests (CCFM 2008). But part of the adaptive strategy requires predicting the effects of future climate change, and an essential component of the prediction is an assessment of its uncertainty. Adaptation in anticipation of a future effect that does not materialize and failure to adapt to the future effect that *does* materialize—both because of a poor prediction—can have negative effects on their own. The immense range of the predicted annual impact of climate change on Canada's forests (Ochuodho et al. 2012) reflects the immense uncertainty of how a complex natural system will respond to a changed climate (where the specifics of the climate changes are themselves uncertain). Eddy et al. (2014) illustrate (conceptually) an adaptive management strategy wherein data analysis and scenario modelling produce a prescriptive decision, which in turn leads to an iterative process of actions → evaluations of the actions → adjustments to the actions in light of the evaluation and new information. But this may not give sufficient consideration to the very lengthy cycle of a forest stand. Decisions we make today cannot easily be modified until the next rotation (often 60+ years).

Candau and Fleming (2011), Gray (2013) and Boulanger et al. (2015) have produced predictions of one or two characteristics of future SBW outbreaks under one or more future climates. Candau and Fleming (2011) used the Random Forest (RF) R package (Liaw and Wiener 2002) to predict outbreak duration in Ontario under two climate scenarios (IPCC “story lines”) and three climate models. They qualitatively described differences in outcome predictions but did not determine the source (climate model or climate scenario) of the differences

in outcomes. They did not generate the additional projections that would arise from the uncertainty in the model data sources or model parameter estimates that would produce additional uncertainty in projected outcomes.

Boulanger et al. (2015) built a consensus model from two model data sources (explanatory dataset) and six model types (correlative techniques, including RF) and projected future outbreak duration in three time periods under three climate scenarios (RCPs) with one climate model. Their study area was roughly equivalent to the one used here. The RF model type was the biggest contributor to the consensus model; the two model data sources (explanatory datasets) were approximately equal contributors. They compared only three sources of uncertainty (model dataset, climate scenario and model type). Model data source (of the explanatory variables) was a greater contributor to projected uncertainty than the model type and the climate scenario in all time periods. Model type was responsible for more uncertainty than climate scenario in all time periods. They did not include the uncertainty from parameter estimation nor from model error in their summary; and the map of projected uncertainties includes only the uncertainty from the climate scenario.

Gray (2013) argued in favour of the reduced rank regression model type because of its ability to model multiple outbreak characteristics simultaneously: logically, an object is better described by two characteristics than by one. He criticized the RF technique because of its willingness to produce a discontinuous relationship (Gray 2013, Fig. 8)—obviously good for classifications, but at odds with biological systems whose behaviours are ultimately driven by mechanisms that may be wildly nonlinear but rarely discontinuous. His study area was roughly equivalent to the one described here. But there was no exploration of projection uncertainty, and he projected future outbreak characteristics (duration and defoliation) using only one IPCC story line and only one climate model. Candau and Fleming (2011) justifiably criticised reduced rank regression for its linear restriction. The transformation of duration and defoliation code (Fig. 3) in the work described here addresses that restriction. Still, there remains considerable debate regarding the application of highly flexible or less flexible model types to alternative (e.g. future) conditions (Elith et al. 2002; Thuiller 2003; Araújo et al. 2005; Randin et al. 2006).

This work has compared more sources of uncertainty than were considered in previous work and it has quantified the separate contributions of each source to the total uncertainty. Similar to the study of Boulanger et al. (2015), climate scenario was not a leading contributor to uncertainty of projected outbreak duration (5.6%; Table 3). However, climate scenario was the second largest contributor (13.5%) to uncertainty of projected defoliation code. Climate models did not contribute substantially to uncertainty 6.6 and <1% to duration and defoliation code, respectively. Model data source (forest

composition used as explanatory dataset) was by far the greatest contributor to uncertainty to projected duration and defoliation code 69 and 73%, respectively.

Boulanger et al. (2015) found model data source (of the explanatory variables) to be the biggest contributor to uncertainty in their comparison of three contributors. They assembled two alternative explanatory datasets for model building: (1) host biomass (from the NFI dataset) and climate variables and (2) host biomass (from NFI) multiplied by a overwintering survival factor from a process-based phenology model (Régnière et al. 2012). The choice here was between two alternative forest inventories. The CanFI and NFI estimates of forest composition were derived by very different methodologies, and they result in different “pictures” of essentially the same forest: the first canonical correlation between the CanFI and NFI databases was 0.8 (both spatial resolutions) but decreased to 0.6 and 0.4 by the third and fourth canonical correlations (both spatial resolutions). It is assumed that despite the difference in when the databases were compiled, the true forest compositions will not have changed significantly at the very large spatial scale used here, and differences between the databases are predominantly due to methodologies. Both studies illustrate the importance that estimates of initial conditions (i.e. the status of the explanatory data during model building) have on projections of future outbreaks. Changing inventory datasets between model building and outbreak projections also serves as a sensitivity test of the assumption (Candau and Fleming 2011; Gray 2013; Boulanger et al. 2015) that projections of future outbreaks do not need to be done under the conditions of a climate-changed forest composition. We see here that changing the estimated forest composition between model building and outbreak projection contributed 6 and 8% to uncertainty of outbreak duration and defoliation, respectively (Table 3). Unfortunately, at this time, there does not exist a quantitative, climate-sensitive model of forest succession that can project future forest compositions at the required national and subnational levels for this Canadian study area.

Estimates of initial conditions (specifically, the datasets chosen to represent forest composition) made the largest contribution to predictions of future outbreaks in this study (Table 3). Gray (2008, 2013) used the CanFI database because it was the only database available at the time that covered the study area. Candau and Fleming (2005, 2011) used a more precise provincial database because they restricted their study area to Ontario. Boulanger et al. (2015) used the NFI database because its format lends itself to more accurate compilation at one’s chosen spatial resolution. Similarly, Candau and Fleming’s (2005, 2011) estimates of defoliation were restricted to Ontario, where defoliation categories were somewhat consistent over time. But Gray (2008, 2013, and here) and Boulanger et al. (2015) created broader (and, therefore, fewer and less precise) defoliation categories in order to incorporate the differences among the multiple provincial categories of their study area. It seems highly likely that projections of

future outbreaks could be improved by building a new outbreak model with more precise outbreak data. The new and ongoing outbreak of spruce budworm in eastern Canada (Natural Resources Canada 2014) presents an opportunity to collect defoliation observations using uniform and temporally consistent categories that can be used with the regularly updated NFI data to build a new outbreak model with less uncertainty.

There seems less opportunity to reduce the uncertainty in outbreak projections that are attributable to our uncertainty in the future climate (climate model plus climate scenario). If this uncertainty cannot be reduced (and it may not be possible to substantially reduce it), alternative decision strategies may be required. McInerney et al. (2012) define strategy robustness as “trading a small decrease in a strategy’s expected performance... for a significant increase in a strategy’s performance in the worst case”. McDaniels et al. (2012) employ a similar definition. McInerney et al. (2012) compare three decision strategies under conditions of deep uncertainty.

Policy decisions invariably rely on projections of future outcomes that have some amount of uncertainty. As projections become more uncertain, they become less useful in decision making (Thuiller et al. 2004). Uncertainty may be “a fact of life” (Walker et al. 2003), but a better understanding of the types of uncertainty, their sources, their sizes and their implications for policy choices will ultimately lead to better policy (Walker et al. 2003). In addition to better policy decisions, assessing the magnitude of the various uncertainty sources suggests an efficient allocation of limited resources to reduce uncertainty (e.g. better parameter estimates, better model construct and better input data).

Acknowledgements The author is extremely grateful to two anonymous reviewers, whose diligence greatly improved the manuscript.

The author is especially grateful to Drs. G. Heuvelink and K. Keller who provided valuable comments to the presentation that formed the basis of this work, and to Mr. I. DeMerchant for his assistance with data preparation and ArcGIS procedures.

Compliance with ethical standards

Funding All funding was from the Natural Resources Canada (government of Canada).

References

- Araújo MB, Whittaker RJ, Ladle RJ, Erhard M (2005) Reducing uncertainty in projections of extinction risk from climate change. *Glob Ecol Biogeogr* 14:529–538
- Beaudoin A, Bernier PY, Guindon L, Villemaire P, Guo XJ, Stinson G, Bergeron T, Magnussen S, Hall RJ (2014) Mapping attributes of Canada’s forests at moderate resolution through kNN and MODIS imagery. *Can J For Res* 44:521–532. doi:10.1139/cjfr-2013-0401

- Boulanger Y, Gauthier S, Gray DR, Le Goff H, Lefort P, Morissette J (2013) Fire regime zonation under current and future climate over eastern Canada. *Ecol Appl* 23:904–923. doi:10.1890/12-0698.1
- Boulanger Y, Gray DR, Cooke BJ, De Grandpré L (2015) Model specification uncertainty in current and future spruce budworm (*Choristoneura fumiferana* [Clem.]) outbreak duration. *Glob Chang Biol* 22:1595–1607. doi:10.1111/gcb.13142
- Canadian Council of Forest Ministers (2010) National Forest Inventory <https://nfi.nfis.org>
- Candau J-N, Fleming RA (2005) Landscape-scale spatial distribution of spruce budworm defoliation in relation to bioclimatic conditions. *Can J For Res* 35:2218–2232
- Candau J-N, Fleming RA (2011) Forecasting the response of spruce budworm defoliation to climate change in Ontario. *Can J For Res* 41:1948–1960
- Candau J-N, Fleming RA, Hopkin A (1998) Spatiotemporal patterns of large-scale defoliation caused by the spruce budworm in Ontario since 1941. *Can J For Res* 28:1733–1741
- Davis MB, Shaw RG, Etterson JR (2005) Evolutionary responses to changing climate. *Ecology* 86:1704–1714
- Dukes JS, Pontius J, Orwig D, Garnas JR, Rodgers VL, Brazee N, Cooke BJ, Theoharides KA, Stange EE, Harrington R, Ehrenfeld J, Gurevitch J, Lerdau M, Stinson K, Wick R, Ayres MP (2009) Responses of insect pest, pathogens, and invasive plant species to climate change in the forests of northeastern North America: what can we predict? *Can J For Res* 39:231–248
- Dymond CC, Neilson ET, Stinson G, Porter K, MacLean DA, Gray DR, Campagna M, Kurz WA (2010) Future spruce budworm outbreak may create a carbon source in eastern Canadian forests. *Ecosystems* 13:917–931. doi:10.1007/s10021-010-9364-z
- Eddy BG, Hearn B, Luther JE, van Zyll de Jong M, Bowers W, Parsons R, Piercey D, Strickland G, Wheeler B (2014) An information ecology approach to science–policy integration in adaptive management of social-ecological systems. *Ecol Soc* 19
- Elith J, Burgman MA, Regan HM (2002) Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecol Model* 157:313–329. doi:10.1016/S0304-3800(02)00202-8
- ESRI (2006) ARC-Info, Redlands, CA. 9.2
- Eveleigh ES, McCann KS, McCarthy PC, Pollock SJ, Lucarotti CJ, Morin B, McDougall GA, Strongman DB, Huber JT, Umbanhowar J, Faria LDB (2007) Fluctuations in density of an outbreak species drive diversity cascades in food webs. *Proc Natl Acad Sci U S A* 104:16976–16981
- Flannigan MD, Krawchuk MA, De Groot WJ, Wotton BM, Gowman LM (2009) Implications of changing climate for global wildland fire. *Int J Wildland Fire* 18:483–507. doi:10.1071/WF08187
- Gauthier S, Bernier P, Kuuluvainen T, Shvidenko AZ, Schepaschenko DG (2015) Boreal forest health and global change. *Science* 349:819–822. doi:10.1126/science.aaa9092
- Gray DR (2008) The relationship between climate and outbreak characteristics of the spruce budworm in eastern Canada. *Clim Chang* 87:361–383. doi:10.1007/s10584-007-9317-5
- Gray DR (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. doi:10.1139/cjfr-2013-0240
- Gray SL, Power K (1997) Canada's forest inventory 1991: the 1994 version—technical supplement. Inf rep. Natural Resources Canada. Canadian Forest Service - Pacific Forestry Centre, Victoria, p 159
- Gray DR, Régnière J, Boulet B (2000) Analysis and use of historical patterns of spruce budworm defoliation to forecast outbreak patterns in Quebec. *For Ecol Manag* 127:217–231
- Grondin P, Gauthier S, Borcard D, Bergeron Y, Tardif P, Hotte D (2014) Drivers of contemporary landscape vegetation heterogeneity in the Canadian boreal forest: integrating disturbances (natural and human) with climate and physical environment. *Ecoscience* 21:340–373. doi:10.2980/21-(3-4)-3696
- Hicks CR (1982) Fundamental concepts in design of experiments. Holt, Reinhart and Winston, New York
- IPCC (2001) Climate change 2001. IPCC third assessment report. Intergovernmental Panel on Climate Change. http://www.grida.no/publications/other/ipcc_tar/. Accessed October 2016
- Lepš J, Šmilauer P (2014) Multivariate analysis of ecological data using CANOCO 5. Cambridge University Press, Cambridge
- Liaw A, Wiener M (2002) Classification and regression by randomForest. *R News* 2:18–22
- McDaniels T, Mills T, Gregory R, Ohlson D (2012) Using expert judgments to explore robust alternatives for forest management under climate change. *Risk Anal* 32:2098–2112. doi:10.1111/j.1539-6924.2012.01822.x
- McInerney D, Lempert R, Keller K (2012) What are robust strategies in the face of uncertain climate threshold responses?: robust climate strategies. *Clim Chang* 112:547–568. doi:10.1007/s10584-011-0377-1
- McKenney D, Pedlar J, Hutchinson M, Papadopol P, Lawrence K, Campbell K, Milewska E, Hopkinson RF, Price D (2013) Spatial climate models for Canada's forestry community. *For Chron* 89:659–663. doi:10.5558/frc2013-118
- Natural Resources Canada (2014) National Forest Pest Strategy Information System. Natural Resources Canada, Canadian Forest Service, Atlantic Forestry Centre, Fredericton, NB. <https://afc-fs.cfsnet.nfis.org/NFPS-SNLRF/>
- Ochudho TO, Lantz VA, Lloyd-Smith P, Benitez P (2012) Regional economic impacts of climate change and adaptation in Canadian forests: a CGE modeling analysis. *For Policy Econ* 25:100–112
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2015) Vegan: community ecology package. R package version 2.2–1
- R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Randin CF, Dimböck T, Dullinger S, Zimmermann NE, Zappa M, Guisan A (2006) Are niche-based species distribution models transferable in space? *J Biogeogr* 33:1689–1703. doi:10.1111/j.1365-2699.2006.01466.x
- Refsgaard JC, van der Sluijs JP, Højberg AL, Vanrolleghem PA (2007) Uncertainty in the environmental modelling process—a framework and guidance. *Environ Model Softw* 22:1543–1556. doi:10.1016/j.envsoft.2007.02.004
- Régnière J, Bolstad P (1994) Statistical simulation of daily air temperature patterns in eastern North America to forecast seasonal events in insect pest management. *Environ Entomol* 23:1368–1380
- Régnière J, St-Amant R (2004) BioSIM 8.0 User's manual. Canadian Forest Service
- Régnière J, St-Amant R, Duval P (2012) Predicting insect distributions under climate change from physiological responses: spruce budworm as an example. *Biol Invasions* 14:1571–1586. doi:10.1007/s10530-010-9918-1
- Royama T (1984) Population dynamics of the spruce budworm. *Ecol Monogr* 54:429–462
- Royama T (1992) Analytical population ecology. Chapman and Hall, London
- SAS Institute (n.d.) SAS v9.2. SAS Institute, Cary, NC, U.S.A.
- Steel RG, Torrie JH (1980) Principles and procedures of statistics. A biometrical approach. McGraw-Hill, Inc., New York
- Swaine JM, Craighead FC, Bailey IW (1924) Studies on the spruce budworm (*Cacoecia fumiferana* Clem.). Part I. A general account of the outbreaks, injury and associated insects. *Tech Bull.*

- Natural Resources Canada. Canadian Forest Service - HQ, Ottawa, p 27
- ter Braak CJF (1994) Canonical community ordination. Part I: Basic theory and linear methods *Ecoscience* 1:127–140
- Thuiller W (2003) BIOMOD—optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob Chang Biol* 9:1353–1362
- Thuiller W, Araújo MB, Pearson RG, Whittaker RJ, Brotons L, Lavorel S (2004) Uncertainty in predictions of extinction risk. *Nature* 430. doi: [10.1038/nature02716](https://doi.org/10.1038/nature02716)
- Walker WE, Harremoës P, Rotmans J, van der Sluijs JP, van Asselt MBA, Janssen P, Kreyer von Kraus MP (2003) Defining uncertainty a conceptual basis for uncertainty management in model-based decision support. *Integr Assess* 4:5–17