Biol Invasions (2021) 23:297–349
https://doi.org/10.1007/s10530-020-02372-9

ORIGINAL PAPER

Oh the places they'll go: improving species distribution modelling for invasive forest pests in an uncertain world

Vivek Srivastava · Amanda D. Roe · Melody A. Keena · Richard C. Hamelin · Verena C. Griess

Received: 14 February 2020 / Accepted: 20 September 2020 / Published online: 2 November 2020

Abstract Species distribution modelling (SDM) is a valuable tool for predicting the potential distribution of invasive species across space and time. Maximum entropy modelling (MaxEnt) is a popular choice for SDM, but questions have been raised about how these models are developed. Without biologically informed baseline assumptions, complex default SDM models could be selected, even though alternative settings may be more appropriate. Here we explored the effects of various SDM design strategies on distribution mapping of four forest invasive species (FIS) in Canada. We found that if we ignored the underlying FIS biology such as use of biologically relevant predictors, appropriate feature selection and inclusion of dispersal and biotic interactions when we developed our SDMs, we obtained complex SDMs (default) that provided an incomplete picture of the potential FIS invasion. We recommend simplifying SDM complexity and including biologically informed assumptions to achieve more accurate dispersal predictions, particularly when projecting FIS spread across time. We strongly encourage SDM users to perform species-specific tuning when modeling FIS distributions with MaxEnt to determine the best SDM design.

Keywords Forest invasive species · Species distribution models · Niche modelling · Asian gyspy moth · Asian longhorned beetle · Dutch elm disease · Sudden oak death

V. Srivastava (✉) · V. C. Griess
Department of Forest Resources Management, University of British Columbia, Vancouver, BC, Canada
e-mail: vivek.srivastava@ubc.ca

V. C. Griess
e-mail: verena.griess@ubc.ca

A. D. Roe
Great Lakes Forestry Center, Natural Resources Canada, Sault Ste. Marie, ON, Canada
e-mail: amanda.roe@canada.ca

M. A. Keena
Northern Research Station, USDA Forest Service, Hamden, CT 06514, USA
e-mail: melody.keena@usda.gov

R. C. Hamelin
Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, Canada
e-mail: richard.hamelin@ubc.ca

Introduction

Invasive species pose significant threats to the economic and ecological stability of our forests. The United Nations identifies invasive species as one of the greatest threats to biodiversity (Groombridge 1992) and cause billions of dollars of damage every year (Pejchar and Mooney 2009; Vilà and Hulme 2017).

Given the potential impact of invasive species, tools are needed that will help prevent invasions, or enable effective early responses through robust interception and surveillance frameworks (Barbosa et al. 2012; Jiménez-Valverde et al. 2011; Lafond et al. 2019). Unfortunately, these programs are costly and require regional prioritization. Species distribution models (SDMs) are one approach used to identify areas at risk of invasion. SDMs are a combination of tools that translate environmental conditions from a species' known distribution to predict its potential distribution in a new habitat. These can be combined with climatic models that forecast future climate scenarios, providing further information on the future potential for invasion and spread. The information provided by SDMs is critical for conservation and management planning and for understanding invasive species ecology and behavior under changing climatic conditions (Padalia et al. 2014).

SDMs can be broadly classified into two groups: correlative models and process-based/mechanistic models (Peterson et al. 2015). Correlative SDMs are trained with species occurrence data and associated environmental layers from a known distribution, which is then used to identify suitable habitats in a new range for a given invasive species (Elith et al. 2010). The process is complex and must manage uncertainties within the modelling process (Gould et al. 2014). Recent advances in iterative model development, model fitting, evaluation and improvement have led to an increase in accuracy; however, questions remain around the practice of model building. Without biologically informed baseline data, complex default SDM models could be selected, even though alternative settings may be more appropriate (Merow et al. 2013). Baseline data should take the dispersal capability, effects of climate on growth and reproduction, and interactions with hosts and symbiotes into account when selecting appropriate variables and settings for models. Additional scrutiny has come to other aspects of SDM modelling, such as estimating FIS distributions accurately within a new geography (transferability), the choice of bioclimatic variables, and the effect of varying model-fitting parameters on the resulting distribution predictions (Jiménez-Valverde et al. 2008; Srivastava et al. 2019). Given the prevalence of SDMs within the invasion literature, addressing these criticisms is critical.

One of the critical assumptions in these models is that the species be in equilibrium with its environment in the area that is used for training the model. In the native habitat the invasive species should be in equilibrium where it has a long history of dispersal and colonization, but even in native regions if large areas that are poorly sampled are used for training the model it will perform more poorly than if a small native range that is well sampled is used (Zhu et al. 2016). Invasive non-native species are typically in disequilibrium with their environment (not fully occupying all the suitable environmental niche and temporarily occupying unsuitable habitat), especially during the early stages of invasion and establishment (Elith et al. 2010; Uden et al. 2015). This results from a colonization lag time while they must buildup and adapt locally before they begin to disperse, other limitations to dispersal, chance movement into areas where they can't establish and insufficient time to occupy all the potential suitable habitat. Data for invasive species are also prone to two forms of sampling bias, which can produce poorly predictive models. First, invasive species are often unevenly surveyed (intense survey in new areas and poorly surveyed in native or already naturalized regions) which can affect the model unless the spatial data is filtered (Bean et al. 2012). When invasive species are first discovered little may be known about their biology and they may not be a past in the native range so little to no information on their distribution is available. Second, short sampling time frames in invaded areas reduce the length of time the model is predictive and require the models to be adaptive and updated with new data from novel environments often to ensure their continuing usefulness (Dimson et al. 2019). Another potential issue is the possibility that the invasive species has expanded its realized niche in the invaded region. If this were the case, then the predicted range would be inaccurate and no improvement would be realized by improving the data from the native range that the model is based on.

In the past, studies have shown that model complexity also plays a major role in transferability of SDMs in novel environments (Moreno-Amat et al. 2015; Warren and Seifert 2011); when an SDM is overfit it underestimates the species potential habitat whereas when it is under fitted SDM it tends to overestimate it. Studies recommend to optimally balance the model complexity and accuracy by fine

tuning model parameters (Moreno-Amat et al. 2015; Warren et al. 2014). The practice of fine tuning SDM parameters includes calibrating several initial models with a wide array of model parameters, selecting the optimal set of parameters that results in the "best model", and then further calibrating the model with the chosen parameters (Cobos et al. 2019; Warren and Seifert 2011). MaxEnt, one of the most popular correlative SDMs (Morales et al. 2017; Potts and Elith 2006) enables users to map potential distributions while making a number of modelling assumptions and choosing a number of model settings (Barry and Elith 2006). This includes choice of background samples or pseudo absences (PAs), selection of appropriate features and regularization (β) multiplier (Elith et al. 2011). The choice of background impacts the transferability of SDMs, thus it becomes important to modify the background sample so that there is a clear ecological justification for the selection (Chapman et al. 2019; Liang et al. 2018). It is recommended to constrain the PA locations to the same geographic range as presences, while also considering dispersal constraints for accurate predictions (Liang et al. 2018). MaxEnt is a powerful SDM capable of incorporating complex and highly non-linear response curves using various feature classes and it is also equally vital to select appropriate feature shape prior to model development along with optimal regularization value to reduce over fitting (Anderson and Gonzalez 2011; Merow et al. 2013). Regularization penalizes the model in proportion to the magnitude of the coefficients and consequently shrinks many coefficients toward zero while setting others to zero, thereby putting off many features from the model (Merow et al. 2013; Tibshirani 1996).

The use of MaxEnt has grown regularly every year since 2008, in part due to increasing focus on invasive species. Accessibility to software platforms that implement MaxEnt, as well as forest invasive species (FIS) distribution data have further accelerated its use within the literature. With this accessibility there is increased need for methodological studies that ideally analyze effects of various SDM design strategies implementing MaxEnt. Evaluations of SDMs are not rare (Guisan et al. 2017; Liu et al. 2011; Potts and Elith 2006; Senay and Worner 2019), but studies evaluating the effects of SDM design strategies on model performance with FIS are limited. Furthermore, a recent review (Srivastava et al. 2019), suggested that SDM outputs should address prediction uncertainty, biotic interactions, and link species dispersal traits with projections of species distributions, details which are often missing in many SDM studies (Araújo and Guisan 2006; Engler et al. 2012). Also, SDM studies should account for the effects of sampling bias in the occurrence data, critical yet rarely reported details for models based on presence-only datasets (Phillips et al. 2009). Failing to correct for sampling bias may lead to distribution projects reflecting the sampling bias rather than the true potential distribution of a species (Støa et al. 2018; Syfert et al. 2013). These different aspects of SDMs need to be addressed in an accurate modelling framework to ensure that the FIS distribution predictions match future invasion scenarios, especially under the inherently unpredictable changing global climate (Venette et al. 2010).

Here we focused on four FIS to evaluate the effects of various SDM design strategies on FIS distribution predictions in Canada, as well as their overall global distributions. We chose four FIS to serve as case studies i: two insects (Asian longhorned beetle [ALB], *Anoplophora glabripennis* (Motschulsky); Asian gypsy moth [AGM], *Lymantria dispar asiatica* Vnukovskij *and L. d. japonica* Motschulsky) and two pathogens (sudden oak death [SOD], *Phytophthora ramorum* Werres; Dutch elm disease [DED], *Ophiostoma ulmi* (Buisman) Melin & Nannf. and *O. novo-ulmi* Brasier) (Fig. 1). All four FIS pose significant threats to Canadian forests (Hamelin and Roe 2019). The two insects are considered high risk invasives and subject to regulatory action (e.g., removal of tree hosts for ALB eradication, denying vessel entry if AGM is detected) by the Canadian Food Inspection Agency, the national regulatory body responsible for protecting Canada's plant resources from invasive species. SOD has been found associated with nursery plants in the southern coastal area of BC, but not yet in urban or forest environments in Canada. The CFIA conducts annual surveys for SOD in nurseries, and if detected, the nursery is placed under quarantine and infected plant material is destroyed resulting in economic losses to the owner. DED has devastated elms across most of Canada since its arrival in the 1940's but does not yet occur in Alberta or British Columbia (BC). There are various provincial and municipal groups across the country supporting programs to protect remaining elms from the deadly disease. Despite these efforts from various plant

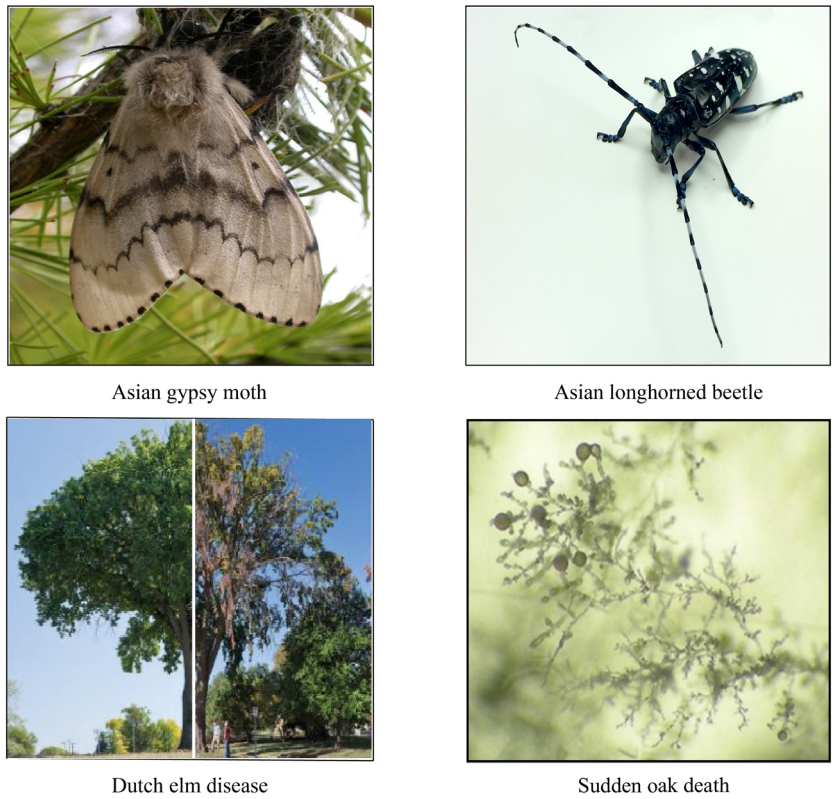


Fig. 1 Selected FIS to serve as case studies

protection groups, these species continue to pose a risk to Canada. For example, a new ALB population was detected in Toronto in 2013 after the first population, detected in 2003, was successfully eliminated (Turgeon et al. 2015). Similarly, the CFIA continues to detect AGM on vessels coming from Asian ports (recent find in 2019) after being eradicated around the Vancouver port in 1992 (Nealis 2009). In addition to focusing on these four species, we also chose to explore the sensitivity of SDM modelling in two major ports in Canada: Vancouver, British Columbia and Toronto, Ontario. The ports of Vancouver and Toronto are two of the major ports in Canada wherein cargo volume reached a record high of 147 million tons and 2.2 million metric tons in 2018 respectively. In addition to high trade volumes, CFIA continue to detect FIS around these two ports. Continued detections suggest that these ports are high risk entry points (Paini et al. 2018). Thus, we hypothesize that the ports of Vancouver and Toronto are likely to serve as points of entry for the FIS, so we produced dispersal restricted projections of individual FIS distributions in various climate change scenarios that also accounted for anthropogenic factors governing the species spread.

By using a case study approach, we show how distribution predictions for each FIS can vary quite dramatically and we highlight the need for biologically informed modelling.

Materials and methods

In general, we took the following steps below to develop SDMs for our target species and evaluate the effects of model design on distribution predictions. We summarize the technical workflow in Fig. 2 and break down each step-in further detail below:

1. Obtain occurrence records for each target species (see "Occurrence data" section).

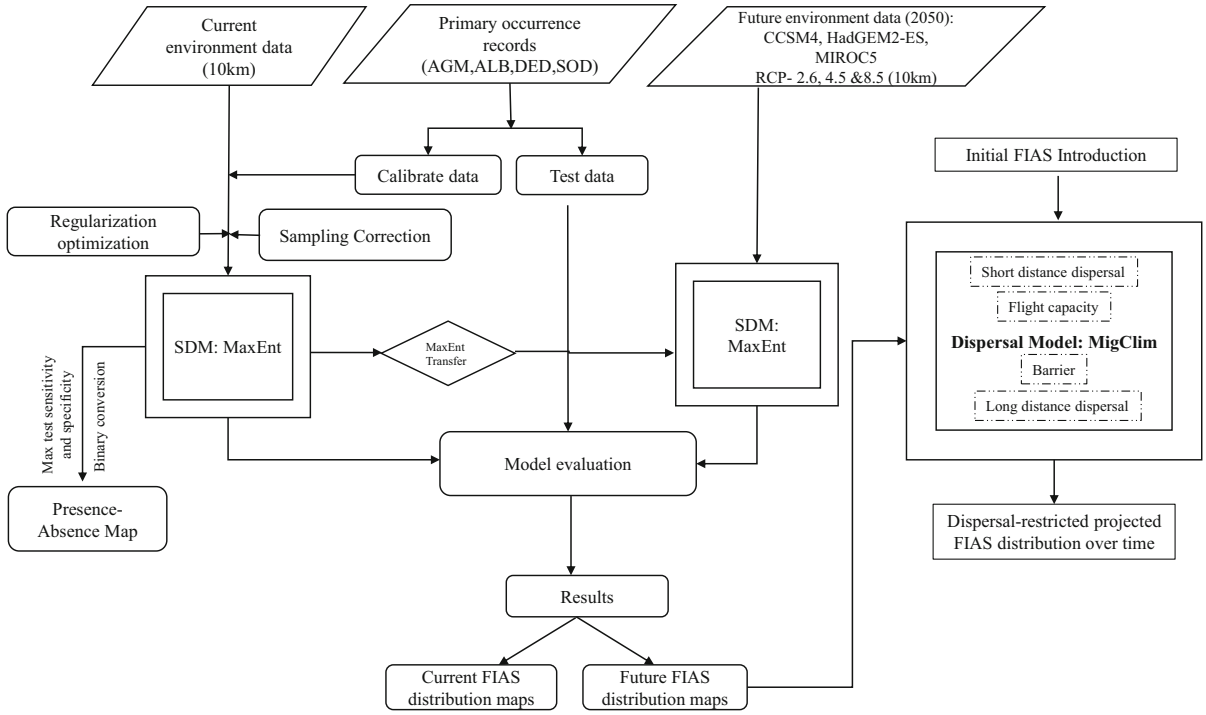


Fig. 2 Flowchart representing the modelling flow used to model FIS distribution, dispersal and uncertainty in this study

2. Develop spatial datasets representing current environmental conditions and future climate scenarios; alongside spatial information on the human footprint in the area of interest (see "Environmental variables" section).
3. Build MaxEnt models to predict the potential distribution of selected FIS in Canada. (See "Species distribution model" section)
4. Create dispersal-limited projections of future FIS distributions under selected climate change scenarios using MigClim. (See "Dispersal mapping" section).

Occurrence data

We collected presence records of AGM, ALB, DED and SOD from various sources to map the known distributions of our selected FIS Commonwealth Agricultural Bureaux International (CABI). The sources included (1) Records provided by the Canadian Food Inspection Agency (CFIA); (2) Global Biodiversity Information Facility database (2017a, b, 2018, 2019), an online database for species occurrences; (3) CABI invasive species compendium (Commonwealth Agricultural Bureaux International 2019a, b, c, d) and (4) Scientific articles and maps. We screened erroneous occurrences and deleted duplicate records such that each observation falls inside a separate 10 km grid cell. Spatial filtering was carried out to mitigate sampling bias in the data (Dimson et al. 2019). A total of 186, 198, 193 and 95 distinct occurrence records were finalized for AGM, ALB, DED and SOD respectively (Fig. 3).

The occurrence data was partitioned randomly into training and evaluation sets (30% for the AGM and DED models and 20% for ALB and SOD). 10,000 background locations were generated within an area defined by a buffer of 100 km around FIS occurrences using SDM toolbox (Brown et al. 2017). The choice of buffer radius was made with respect to the dispersal abilities of the focal FIS (See Appendix 7). To account for sampling bias in the FIS occurrence data we gathered occurrence records for the target group of FIS and generated a bias grid that up-weights occurrence data points with fewer neighbors in the geographic landscape using the Gaussian kernel density of sampling localities tool in SDMToolbox (Brown, 2014) as recommended by Phillips et al. (2009). We

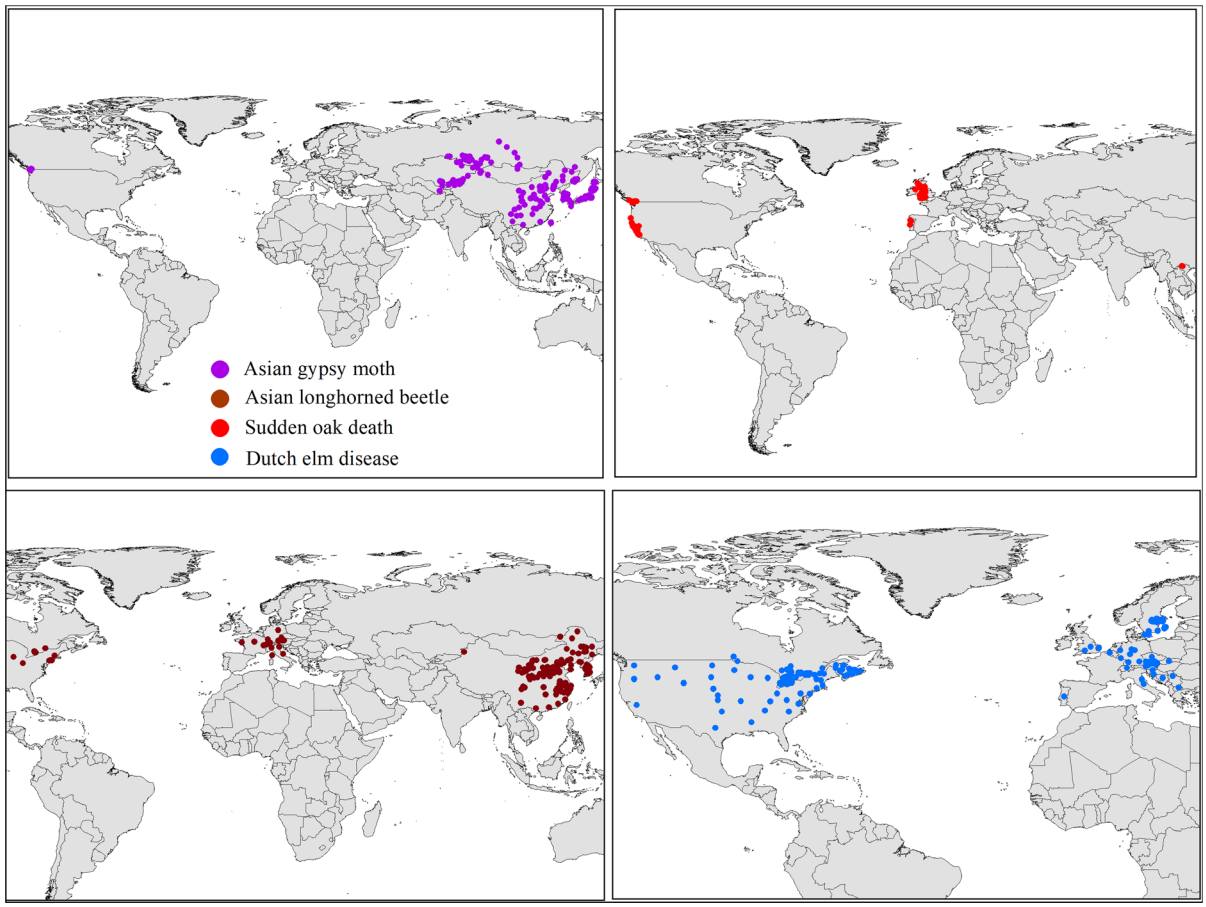


Fig. 3 Selected FIS occurrences in their respective native and introduced ranges

selected families of selected FIS i.e. Erebidae, Cerambycidae, Ophiostomataceae and Peronosporaceae for AGM, ALB, DED and SOD respectively as the target group. The motive behind generating bias grid with species target group approach was to select background points with the same bias as occurrence data.

Environmental variables

We downloaded 19 bioclimatic variables from the WorldClim database version 1.4 (http://www.worldclim.org/) (Hijmans et al. 2005), averaged for the 1950–2000 period, at a spatial resolution of five arc minutes (approximately 9 km resolution at the equator). For future 2050 climate projections, we used three representative concentration pathways (RCPs) of the IPCC—RCP 2.6 (greenhouse gas emissions peak in 2010–2020 and declining after), 4.5 (emissions peak around 2040 and then decline), and 8.5. (rise temperature throughout the twenty first century). We chose three general circulation models (GCMs) of physical climate processes for which the predicted values of each of the bioclimatic variables were available: Community Climate System Model (CCSM4), Hadley Global Environment Model 2-Earth System (HadGEM2-ES) and Model for Interdisciplinary Research On Climate (MIROC5). The selected GCMs are being used in Climate BC/WNA/NA model for producing future grid projections of climate and species range in Canada and North America (Wang et al. 2012). In order to incorporate responses of dispersal and human footprint on FIS distributions we incorporated additional data on human footprint "Human Influence Index-HII" at 1 km from SEDAC and resampled to match the native climate grid resolution. The Human Influence Index (HII) is a measure of direct human influence on terrestrial ecosystems, derived from nine

global variables including land cover, population density, built-up areas, roads, navigable rivers and nighttime lights ("Socioeconomic Data and Applications Center|SEDAC"). For inclusion in the model, these predictors were subjected to multicollinearilty test using Pearson's correlation coefficient (r). The Pearson's correlation coefficient threshold was set to 0.8 and only one variable from each set of highly correlated variables depending on its biological importance to the species and its relative contribution to the overall model was selected. Regularization multipliers ranging from one to four were tested for each FIS and the modelled distribution and AUC (area under curve, or area under the receiver operating characteristic (ROC) curve) was analyzed, as over fitted models are poorly transferable in novel environments (Elith et al. 2010).

Once the predictor set of variables and a regularization parameter value was chosen for all selected FIS, the model was evaluated using the training data set. The training data was portioned into ten random subsets using k-fold cross validation function in MaxEnt. This was done to evaluate the average behavior of the model. In order to produce simple models with smooth fitted functions we used only hinge features for the tuned models (Elith et al. 2010). Jackknife resampling was used to identify those variables that contributed most to the model. The method provides systematic resampling and leads to improved estimates of the sample parameter and a lower sampling bias (Tukey 1958).

For AGM we selected four variables (bio1-Annual Mean Temperature, bio 17-Precipitation of Driest Quarter, bio 19-Precipitation of Coldest Quarter and HII-Human influence index) and for ALB, we included five variables (bio10-Mean Temperature of Warmest Quarter, bio13-Precipitation of Wettest Month, bio 07-Temperature Annual Range (BIO5-BIO6), bio09-Mean Temperature of Driest Quarter and HII). A temperature variable chosen for AGM is not surprising since temperature has a huge impact on both the egg diapause and larval development (Limbu et al. 2017; Gray et al. 2001). Precipitation is also important for AGM because of the effects on the host and because egg masses under snow can survive colder conditions than they otherwise could (Andresen et al. 2001). Temperature is the most important factor in ALB development for all stages (Trotter and Keena 2016) and impacts dispersal (Keena 2018) so multiple variables involving temperature is to be expected. A precipitation variable for ALB was also chosen likely because of host requirements and because drier conditions in the wood can accelerate pupation (MK unpublished data). For DED we had six variables (bio 11-Mean Temperature of Coldest Quarter, bio15-Precipitation Seasonality, bio 18-Precipitation of Warmest Quarter, bio 2-Mean Diurnal Range, bio 7-Temperature Annual Range and HII) and for SOD seven variables (bio1-Annual Mean Temperature, bio4-Temperature Seasonality, bio14-Precipitation of Driest Month, bio15-Precipitation Seasonality, bio18-Precipitation of Warmest Quarter, bio19-Precipitation of Coldest Quarter and HII). The variables chosen for SOD are tied to conditions required for the fungus to develop and spread; temperature and moisture conditions must be right for fungal development (Tooley et al. 2009) and rain aids in spore dispersal (Kuske 1983). The variables chosen for DED are tied both to conditions that promote fungal and beetle vector development or spread (Jones et al. 2019). The selected predictor set for each FIS is also shown in Table 1.

Species distribution model

MaxEnt

MaxEnt version 3.3.3k (Phillips et al. 2006) was used to map the potential distribution of the selected FIS due to unavailability of FIS absence data. MaxEnt being presence—background model has been successfully used in mapping the potential distribution of FIS in past (Kumar et al. 2016; Lira-Noriega et al. 2018). MaxEnt is a machine learning algorithm used for describing probability distributions following the principle of maximum entropy, subject to restraints imposed by the presence of species and their surrounding environment (Phillips and Dudík 2008). MaxEnt model for each FIS was built separately using available training data from native as well as introduced ranges and was later projected onto Canada to map potential suitable areas for FIS establishment. Additionally, the fade by clamping function was used to limit extrapolations beyond the environmental range of the training data. Multivariate environmental similarity surface (MESS) analysis was performed to identify regions of extrapolation for both current and future FIS distribution maps.

Table 1 Summary of individually designed MaxEnt models along with their predictors

Model	Number of predictors AGM ALB DED SOD				Model details
Type i	20 (BIO1–BIO19 + HII)	20 (BIO1–BIO19 + HII)	20 (BIO1–BIO19 + HII)	20 (BIO1–BIO19 + HII)	MaxEnt default + all environmental variables (climatic predictors + HII)
Type ii	19 (BIO1–BIO19)	19 (BIO1–BIO19)	19 (BIO1–BIO19)	19 (BIO1–BIO19)	MaxEnt default + all climatic variables (here only climatic predictors)
Type iii	4 (BIO1,17,19,HII)	6 (BIO7,9,10,13,HII)	8 (BIO2,7,11,15,18,HII)	5 (BIO1,4,14,15,18,19,HII)	MaxEnt default + selected variables (ref. above section on variable selection, only selected environmental variables were kept)
Type iv	4 (BIO1,17,19)	3 (BIO7,9,10,13)	4 (BIO2,7,11,15,18)	3 (BIO1,4,14,15,18,19)	MaxEnt default + selected climatic variables (ref. above section on variable selection, here only climatic predictors were used)
Type v	4 (BIO1,17,19,HII)	6 (BIO7,9,10,13,HII)	8 (BIO2,7,11,15,18,HII)	5 (BIO1,4,14,15,18,19,HII)	MaxEnt tuned (selected environmental variables along with tuned regularization value)
Type vi	4 (BIO1,17,19,HII)	6 (BIO7,9,10,13,HII)	8 (BIO2,7,11,15,18,HII)	5 (BIO1,4,14,15,18,19,HII)	MaxEnt tuned—sampling correction (selected environmental variables along with tuned regularization value plus restricted background. Here there was no sampling correction implemented)
Type vii	4 (BIO1,17,19,HII)	6 (BIO7,9,10,13,HII)	8 (BIO2,7,11,15,18,HII)	5 (BIO1,4,14,15,18,19,HII)	MaxEnt tuned + sampling correction (selected environmental variables along with tuned regularization value plus restricted background with sampling bias grid)

Model comparisons with different SDM approaches

In order to evaluate the effects of various SDM design strategies and to find the best SDM design strategy for our case-based FIS we individually designed and evaluated seven different MaxEnt models (Table 1). The seven MaxEnt models were (1) MaxEnt model with default parameters and all environmental variables (climatic predictors + HII) (2) MaxEnt model with default parameters and all climatic variables (here only climatic predictors) (3) MaxEnt model with default parameters and selected variables (ref. above section on variable selection, only selected environmental variables were kept) (4) MaxEnt model with default parameters and selected climatic variables (ref. above section on variable selection, here only climatic predictors were used) (5) MaxEnt model with tuned parameters (selected environmental variables along with tuned regularization value) (6) MaxEnt model with no sampling correction (selected environmental

variables along with tuned regularization value plus restricted background. Here there was no sampling correction implemented) (7) MaxEnt model with sampling correction (selected environmental variables along with tuned regularization value plus restricted background with sampling bias grid).

Model evaluation

Model evaluation was performed using the withheld presence data for the selected FIS. True skill statistic (TSS) (difference between the rate of successes and errors), sensitivity (fraction of correctly predicted presences), correct classification rate (fraction of correctly predicted points of presence and pseudo absence) and omission error (presence is outside the area predicted, also called False Negative Rates) scores at maximizing test sensitivity and specificity threshold were used to evaluate the models. We extracted the same number of pseudo absences (PAs) as testing presences in order to calculate the evaluation scores. PAs were extracted in the same spatial range as the presences. The same set of pseudo absences and testing data was used for all compared models and was excluded from all model fits. TSS was used here due to being independent of prevalence. However, metrics such as TSS and correct classification rate make use of pseudo-absences rather than true absences for presence only modelling approaches. Hence, PAs should be selected cautiously for reliable estimates on false positives and true negatives (Leroy et al. 2018). The TSS ranges from -1 to $+1$, where values of 0 or less indicate a model performance no better than random, and a value of $+1$ indicates perfect performance (Allouche et al. 2006). The evaluation scores were calculated using NicheToolBox (Osorio-Olvera et al. 2020).

Dispersal mapping

In order to include selected FIS specific dispersal constraints into projections of their potential distributions under climate change, we used MigClim (Engler et al. 2012). MigClim is a function library built in R software that allows implementation of species dispersal limitations in SDM predictions under climate change conditions. MigClim is a cellular automaton model so cells are the measured units and here each cell corresponds to 10 km pixel. Here a target cell becomes colonized with the combined probability *P col*:

$$Pcol = \left(1 - \prod_{i=1}^{n}(1 - PDispi \times PPropi\right)) \times Pinv$$

Here *P Disp i* is a probability function of the distance between the target cell and the source cell *i*. *P prop i* is a probability that is function of time since the source cell *i* became occupied and represents the propagule production potential of the source cell *i* over the time. *P prop is* specified via two parameters: initial maturity age [iniMatAge] and a vector indicating the probability of propagule production for each age between initial and full maturity [propaguleProd]. *P inv* denotes the habitat invisibility of the target cell. Functions of *P Disp* are provided in Appendix 8.

A dispersal kernel, which is the dispersal probability as a function of *P Disp and P prop,* was created based on a negative exponential, with values ranging from 1 for one cell distance, to 0.06 for a distance of 4 cells for AGM. AGM females can fly from less than 1 km up to 20–40 km (Keena et al. 2001; Srivastava et al. 2020). For the other FIS a dispersal kernel was set to 1 with maximum probability since the short distance dispersal was found to be limited (< 10 km) (Dunn 2012; Grünwald et al. 2012, 2019; Smith et al. 2001). We fixed *P prop* as 1 assuming the probability of the source cell to produce propagules to be 100% as our focal FIS are socially active. Additional, random long-distance dispersal events were generated at a frequency of 0.1 at min–max distance range of 100 (10 cells) and 200 km (20 cells) since the selected FIS are capable of dispersal through various means of transport, such as human-assisted transportation (Koch et al. 2013). Since, DED dispersal is limited to the presence of Elm trees, a strong barrier was implemented in the model to simulate dispersal events only in the pixels having Elm presence. Additional information on considered FIS biology and ecology is provided in Appendix 7. A reclassification threshold was selected based upon maximum test sensitivity and specificity for each FIS along with their respective dispersal kernel.

Since MigClim does not generate habitat suitability maps itself, we used MaxEnt to generate the required inputs. Future FIS distribution maps for the year 2050 were produced for climate change scenarios using MaxEnt for three RCPs (2.6, 4.5, and 8.5) and three

GCMs (CCSM4, HadGEM2-ES and MIROC5). These maps were used as an input along with an initial distribution map of the FIS. We assumed two initial infestation points i.e. Vancouver port and Toronto port, since the chosen FIS have been intercepted at these introduction points in the past (Hamelin and Roe 2019; Nealis 2009). We had one environment change step (2050) where in total 40 dispersal steps were simulated [envChgSteps] × [dispSteps], here 40, which corresponds to 40 years from 2010 to 2050. The simulations were repeated for 10 times producing dispersal limited future distribution maps of FIS from 2010 to 2050 under selected climate change scenarios.

Results

Effects of SDM design on predictive performance

The performance of the seven SDM designs for individual FIS varied. When the models were evaluated using the withheld presence data, we found the best SDM design strategy to be model type vii with "sampling correction" for each FIS. The performance of models created with default parameters was found to improve by tuning of the model parameters and correcting the sampling bias and further improved when both were implemented together. Upon comparing the model type ii (classical default SDM) with model type vii (tuned model with sampling correction), we found that the TSS score increased from 0.482 to 0.528 for the AGM model. Likewise, for ALB, DED and SOD it increased from 0.450 to 0.555, 0.672 to 0.734 and 0.684 to 0.842 respectively. The detailed model evaluations for comparison between different SDM designs for DED is shown in Fig. 4 and for other selected FIS in Appendix 1.

Potential distribution of FIS

Modelled predictions selected from best SDM design practice for each FIS matched closely with the observed FIS observations in their respective environment (Appendix 3). Predictions obtained from analyzing these FIS models to a nonnative range (Canada) highlighted areas at risk (available for their potential establishment). Similar provinces were predicted to be suitable for AGM, ALB and DED potential establishment; however, different suitability scores were recorded in these provinces for these three FIS. The identified suitable areas were in the provinces of British Columbia, Alberta, Saskatchewan, Manitoba, Ontario, Quebec, New Brunswick, Nova Scotia and Newfoundland. Whereas, highly suitable areas for SOD were predicted only for the province of British Columbia, near the western coast (Appendix 2). With climate change, FIS distributions in Canada expanded to the north and west (Fig. 5 and Appendix 2).

Upon comparing the current Canadian potential distributional ranges of FIS with their future suitability ranges in various climate change scenarios, we found that AGM, ALB and DED expanded their ranges in all considered GCMs and RCPs. The greatest range expansion was recorded in RCP 8.5 and the least in RCP 2.6 for all selected GCMs. AGM, ALB and DED expanded their ranges by 89,259, 119,024,104,293 km^2 respectively in CCSM4. Whereas, in GCM-HadGEMES, the range was expanded by 102,456, 109,822 and 199,175 km^2. Lastly, in MIROC5, 93,822, 11,190 and 130,025 km^2 expansion was recorded for AGM, ALB and DED respectively.

Interestingly, SOD potential distributional range was observed to contract with warming climate change conditions. Range contraction for SOD was highest in RCP 8.5 and lowest in RCP 2.6. Current SOD range got shrunk by 1306.50, 5850.87 and 7725.43 km^2 in RCP 2.6, 4.5 and 8.5 respectively for CCSM4 GCM. Similarly, for other two GCMs i.e. HADGEM2_ES and MIROC5 it shrunk by 10,111.22, 11,644.95, 5 and 5282.831, 3578.69 and 6986.97 respectively in RCP 2.6, 4.5 and 8.5 climate change scenarios.

Environmental responses and variable contribution

Suitable conditions (probability of presence > 0.45, based on maximum test sensitivity and specificity threshold) for AGM were related to areas with annual temperatures between 5 and 27 °C which agrees with the temperature responses of AGM and that its populations may struggle in areas that experience longer periods of temperatures $\geq$ 30 °C (Limbu et al. 2017). AGM distributions were found to expand into areas where temperatures warmed to acceptable levels and to decline in areas where temperatures began to exceed the 30 °C level for longer periods. The occurrence data and modelled distributions also

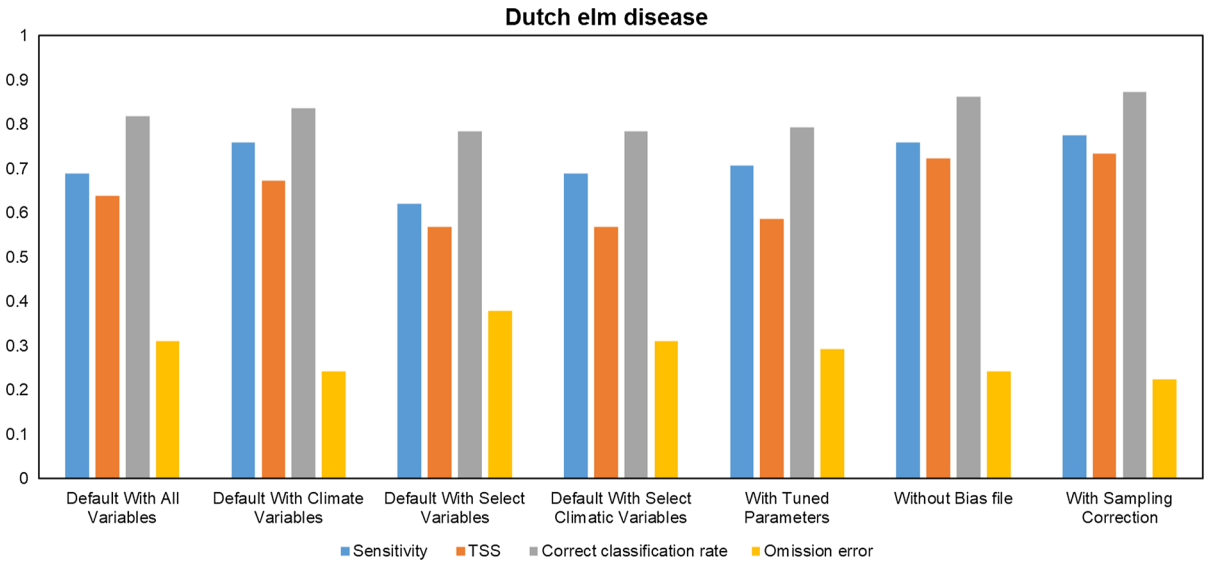


Fig. 4 Evaluation summary of DED models using TSS, correct classification rates, omission error and sensitivity metrics. Summary of other FIS models are shown in Appendix 1

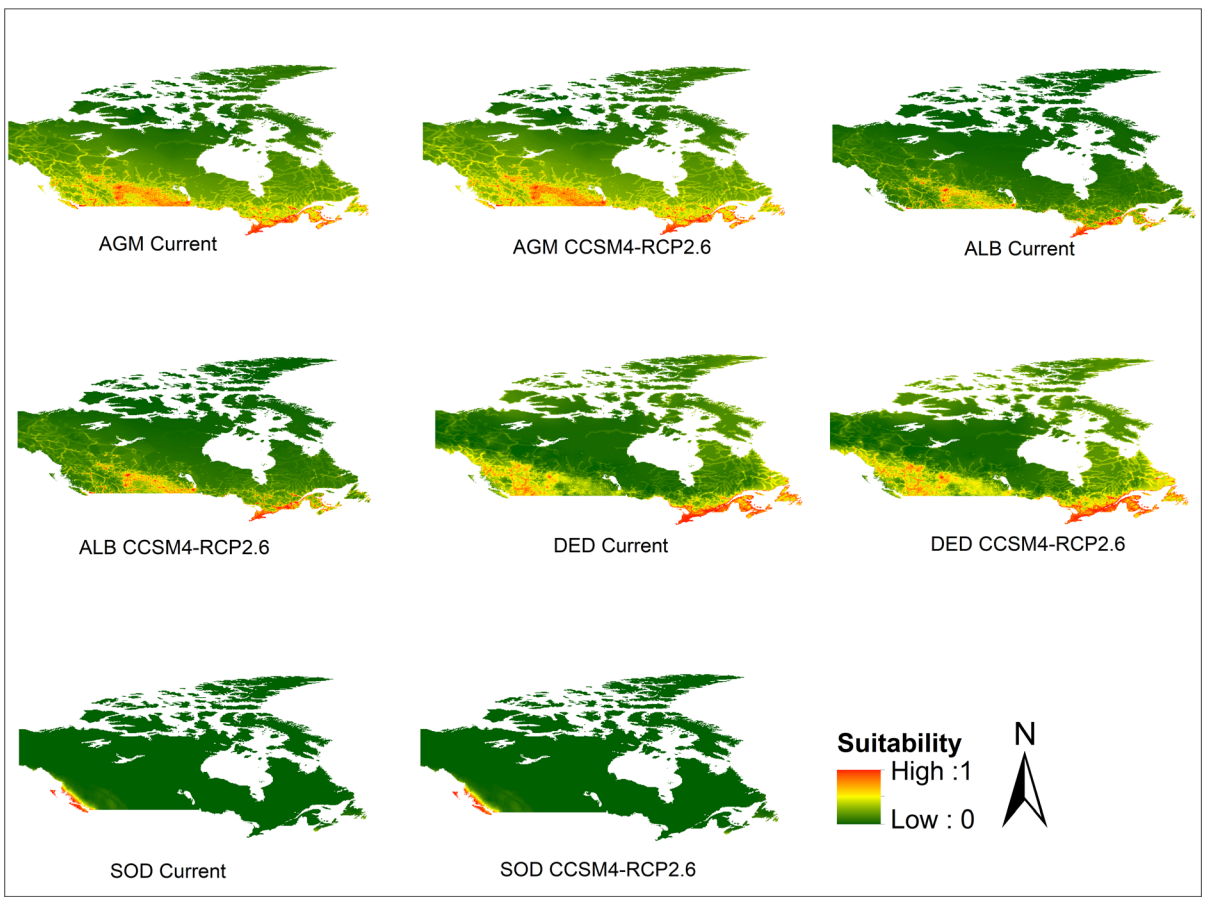


Fig. 5 Potential distribution of selected FIS in current and future climate change scenario

indicate that areas receiving precipitation in the coldest quarter between 20 and 870 mm are suitable for AGM establishment. ALB had the highest suitability (probability of presence > 0.49, based on maximum test sensitivity and specificity threshold) in areas receiving annual temperature range of 16–44 °C and mean temperature of warmest quarter between 20 and 35 °C with 50–500 mm precipitation in the wettest month. This agrees with ALB's ability to develop at temperatures between 10 and 35 °C and the fact that larvae are freeze tolerant (Keena and Moore 2010, Torson et al. in prep, Roe. unpublished data). Suitable areas for DED (probability of presence > 0.521) were found in areas having mean temperature of coldest quarter and annual temperature range between − 8 and − 1 °C and temperature annual range between 20 and 44 °C. This supports the findings of Brasier and Mehrotra (1995) where DED was found to be adapted to a subtropical environment due to its high optimum growth temperature. Sporulation of DED fungus are inhibited by prolonged exposure to high summer temperatures and low moisture content (Webber 1990). SOD potential distribution was associated with moderate temperature variability and high variability in precipitation. We showed that the SOD populations can be strongly influenced by variability in precipitation which is more confined to regions typical to coastal areas and islands. Most of the SOD infestation sites in California and Oregon are located within 30 km of the Pacific coastline or San Francisco bay (Rizzo and Garbelotto 2003). Also, moisture is critical for the germination of spores and fungal growth (see Appendices 4 and 7).

The jackknife test identified Human Influence Index (HII) and Precipitation of Driest Quarter (Bio 17) as the most important predictors of AGM distribution, while identifying HII and Mean Temperature of Warmest Quarter (Bio 10) for ALB distribution. Jackknife test found Human Influence Index (HII), Mean Diurnal Range (Bio 2) and Mean Temperature of Coldest Quarter (Bio 11) for DED and Temperature Seasonality (Bio 4) and Precipitation of Coldest Quarter (Bio 19) for SOD. HII made the largest contribution to the MaxEnt model of AGM and ALB distribution when used in isolation and reduced the model's predictive ability the most when omitted. Similarly, Mean Temperature of Coldest Quarter (Bio 11) was identified for DED and Temperature Seasonality Bio 4 for SOD.

FIS dispersal

Inclusion of FIS specific dispersal limitations into projections of FIS distributions under climate change conditions limited the distribution range when compared with scenarios of unlimited dispersal for all selected FIS (Appendix 5). Figure 6 shows the total number of cells found to be colonized at the end of a simulation for each FIS under multiple climate change conditions starting from two infestation points. The increase in the number of cells colonized by AGM and ALB was found to be highest in RCP 8.5 and when the infestation started from port of Toronto. No cells were colonized for DED when infestation started from port of Vancouver and similarly, for SOD when infestation started from port of Toronto. The number of cells colonized by DED was higher in RCP 8.5 than RCP 2.6 and 4.5 but for SOD the colonization was lowest in RCP 8.5 and highest in RCP 2.6. Detailed dispersal limited output maps for each FIS under each considered scenario are shown in Appendix 6 and an example output is shown in Fig. 7.

Discussion and conclusions

The results we generated are in agreement with the recent findings that have shown the effects of model complexity and varying parameters on SDM performance (Halvorsen et al. 2016; Morales et al. 2017; Rodda et al. 2011; Stolar and Nielsen 2015; Syfert et al. 2013). Our analysis clearly found that the use of default settings for the distribution modelling of FIS were not adequate in each of the considered cases, instead high accuracy was achieved when model parameters were finely tuned and model complexity was optimally balanced. The method proposed by Brown et al. (2017) to correct sample selection bias greatly improved the predictive performance of SDMs when the collected data resulted from an unclear survey design or was derived from online databases. The importance of the background data selection strategy in SDMs have been highlighted for a long time (Barbet-Massin et al. 2012; Syfert et al. 2013) and here we validated its importance in working with FIS by improving the accuracy of our SDMs by following an accessible background while accommodating the dispersal constraints. We achieved both significant improvements in the predictive performances of

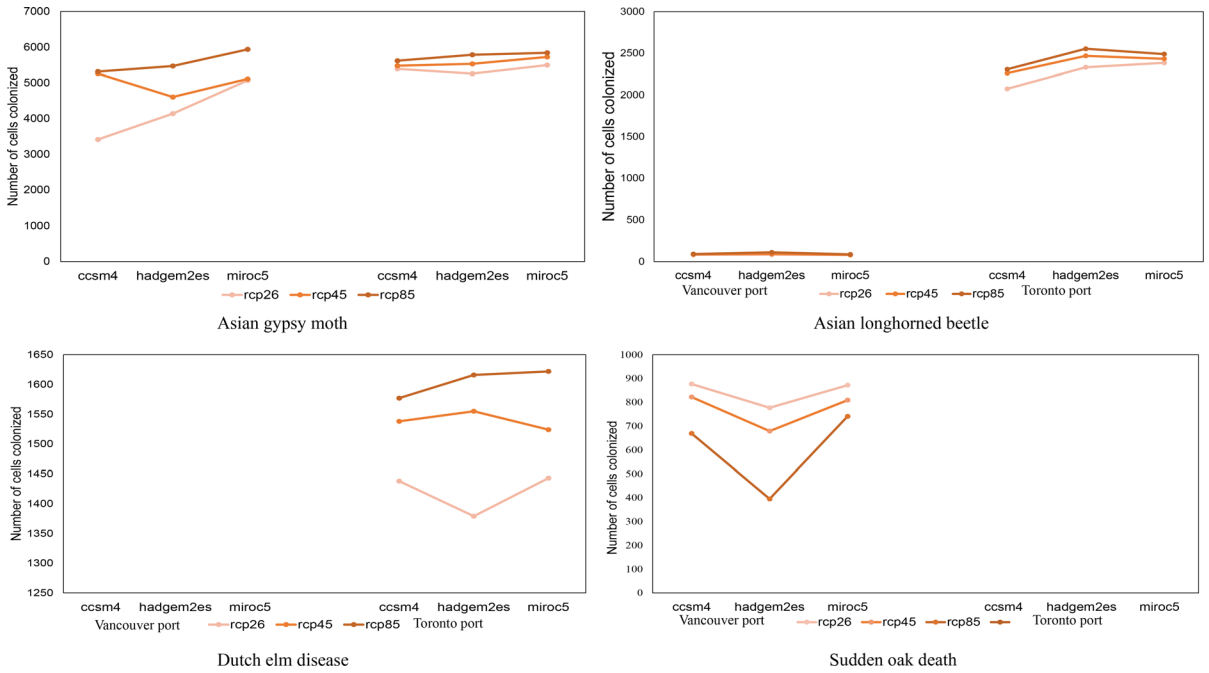


Fig. 6 Colonization in dispersal restricted future distribution of FIS under different climate change scenarios. We assumed Vancouver port and Toronto port as two initial infestation points

SDMs with less complex models and found them to be more accurate in providing predictions upon transferring to Canadian (i.e. novel) environments.

HII is a measure of direct human influence on terrestrial ecosystems, which includes access routes, navigable rivers, nighttime lights along with other important variables; it effectively contributed in correctly identifying the suitable areas for each FIS. Inclusion of human influence index "HII", in addition to climatic predictors to account for FIS dispersal and human footprint, increased the overall accuracy of the FIS model. HII also significantly contributed to each FIS model. The effectiveness of HII in the models can be directly linked to the biology of each FIS. For example, AGM can hitchhike on man-made objects and disperse along the transportation corridors, particularly as egg masses or pupae. Dispersal of adult moths along transportation corridors is further promoted by their attraction to light sources. The SOD pathogen shows similar association with human influence. Spores can move with infected plants, which helps explain why more infected trees were detected on public lands open to general recreation than on adjacent lands lacking public access (Cushman et al. 2008). Furthermore, the chances of an SOD infection increased when sites were within 50 km of human habitation (Cushman et al. 2008). DED and ALB, our remaining focal taxa, also showed higher dispersal in areas with an increased human footprint. In fact, most of the ALB invasions are located in or near urban areas (Appendix 5).

Climate has been considered a critical barrier for the establishment and spread of invasive species into temperate regions. However, climatic models predict that eastern Canada's average temperature will increase by 3–5 °C by 2100 (Dukes et al. 2009). This increase in temperature, particularly winter temperatures, could eventually lead to much higher probabilities of successful FIS establishments (Dukes et al. 2009; Huang et al. 2011). We chose to examine how our FIS distributions changed as we accounted for temperature increases associated with climate change. Three of the four species (AGM, ALB, and DED) showed greater distributions under future climate projections. FIS range is expected to expand with highest range expansion in rcp 8.5, thus leading to much higher probabilities of FIS establishments and spread. SOD range was observed to shrink possibly due to its specific moisture requirements. Given the combined threat of invasion and climate change, it is

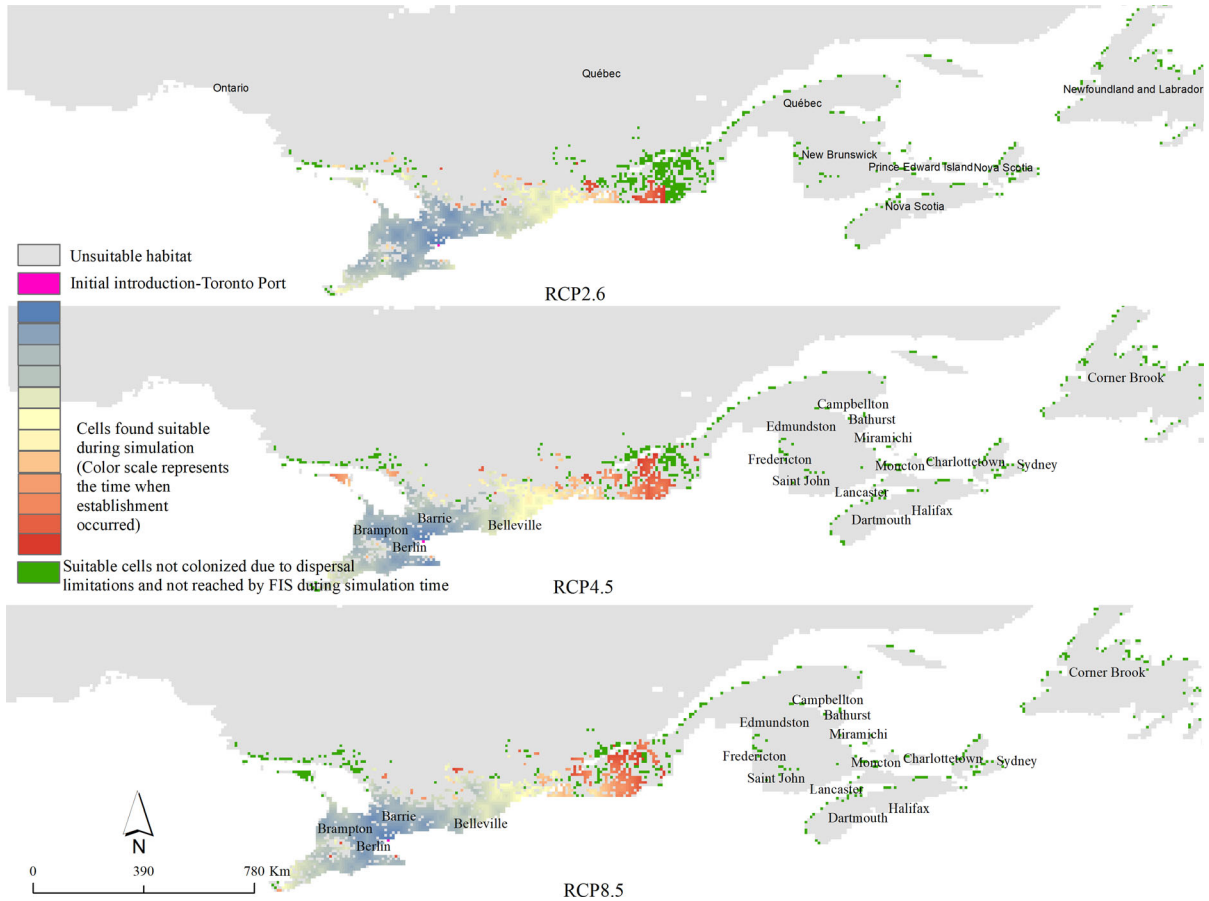


Fig. 7 Dispersal restricted future distribution of DED under GCM-CCM4 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation

critical to model future distributions to ensure resources are adequately allocated to at risk areas based on current or future climate regimes rather than historical estimations (Huang et al. 2011).

Correlative SDMs (MaxEnt) are focused primarily on the realized niche (actual distribution) so may underestimate the fundamental niche (potential distribution). This is because the model is representing only a portion of niche that is represented by the observed records and it is highly unlikely that a new FIS is at equilibrium with its current environmental conditions (Václavík and Meentemeyer 2012). This might provide an inaccurate assessment of overall species niche. Additionally, the habitat that is colonizable may differ from the potential habitat so including dispersal into habitat projections can greatly improve projections. This was shown for SOD when several types of models were evaluated (Václavík and Meentemeyer 2009). Uncertainty when deciding on resource allocation for FIS control measures can lead to poor targeting and unnecessary economic expenditures. Our approach of addressing dispersal limitations using MigClim integrates species-specific genetic traits (flight capacity, long dispersal distance, etc.) and allows for better simulating FIS spread under potential future climate conditions. We have found suitable areas where FIS likely to spread if it gets introduced and establishes in Vancouver and Toronto. Such information can be used by managers to more finely focus eradication efforts. However, it should be noted that

most of the FIS would continue to spread after year 2050 as our simulations did not reach stability and colonization was found to be increasing at the end of simulation. Moreover, suitable areas might be reached through human transport specifically in the case of ALB. Our findings suggest that the ALB control efforts in Vancouver must be focused on the accidental transport of the beetle.

The invasion history, ecology, and biology of each FIS must be considered when building the model. About 50 years ago, ALB native habitat in China was limited to the eastern provinces and the insect wasn't causing serious harm to the native maples it infested. Then, nonnative trees were widely planted and this insect became invasive, now being found throughout most of China (Hu et al. 2009). It was only after it became invasive in China that it spread to other countries. So, this insect may not be in complete equilibrium even in what is considered its native habitat which can adversely affect model predictions. The native range for SOD was only recently discovered, so few data points are available for it and its true native range is still poorly known. Factors that affect DED niche must include more than just the fungus, since a beetle is required to move it from host to host. Climatic factors that influence the fungus may be different from those affecting the beetles that vector it so more realistic predictions may be reached by including predictors specific to the vectors present in the region of interest once data on the effects of climate on the vectors is available. Also, the dispersal capability of the beetle is important to understanding how far and fast this fungus can spread. Dispersal capability has been shown to be an important factor in constructing ecological niche models (de Andrade et al. 2019). Finally, female flight capability in gypsy moths varies across its native range (Eurasia) with most AGM females having full flight capability, while female flight capability in EGM is variable and declines as you move both west and south (Keena et al. 2008). So maybe models should include different dispersal components in MigClim to account for the difference in female flight capability in different parts of the range or in invaded areas depending on weather the invasive gypsy moth strain has females capable of flight or not. More data and analysis of other model parameters should still be considered in future studies of these FIS to further refine the models.

We find that ignoring underlying FIS ecology and biology in SDMs and using complex (i.e. default) SDMs provide an incomplete picture of FIS invasion both in space and time. In these focal cases we recommend simplifying model complexity and including dispersal and biotic factors to achieve more accurate outputs for each species when projecting models across time. We strongly encourage SDM users to perform species-specific tuning when modeling FIS distributions with MaxEnt to determine the best SDM design, as suggested by other authors (Halvorsen et al. 2015; Moreno-Amat et al. 2015; Muscarella et al. 2014; Shcheglovitova and Anderson 2013). However, often the biological and ecological knowledge of new incoming FIS is unavailable. In such cases, climate suitability seems to be the most widely-accepted approach to delineate the probable target regions for the FIS (Srivastava et al. 2019). Yet, climate suitability alone cannot explain the niche requirements of the species, though occasionally it is the most important factor (Stohlgren and Schnase 2006). In such cases we suggest performing species specific model parameterization using recently developed tools like kuenm (Cobos et al. 2019) that offers more rigorous processes of model evaluation and selection and further linking the SDM with simplified dispersal models like KISSMig (Nobis and Normand 2014) which seems to be a sound alternative since it does not require information on species demography and dispersal processes.

Preventing FIS introductions completely is by far the best method to protect the forest resources of a country (Myers et al. 2000) and a key component of the strategy involves detection of infestation areas in the early stages of invasion by means of surveys and constant monitoring. Maps produced from this study provide information about the potential suitable distribution ranges of focal FIS. This type of information is useful in designing early pest surveys and setting of domestic quarantines. Additionally, these maps can be effectively used in making scientifically informed management choices and help to further inform related conservation priorities and trade decisions. However, the maps produced should be interpreted with caution as there is no best transferable SDM for all species and predictions differ with varying modelling assumptions (Qiao et al. 2015, 2019). Also, FIS infested material could arrive to any other vulnerable port or transportation destination, thus additional simulations for

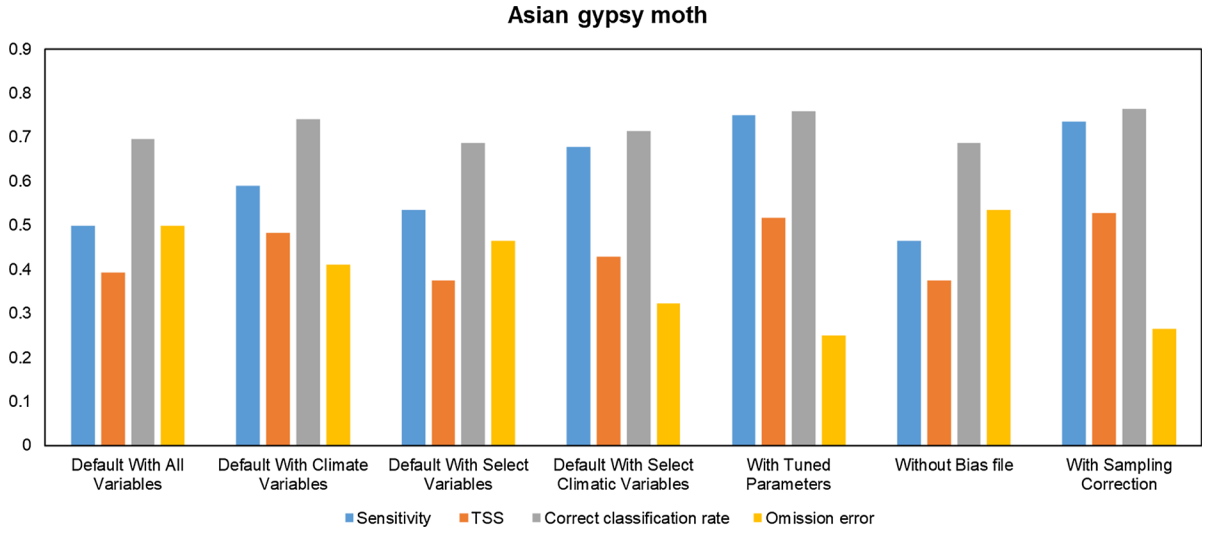
specific scenarios based on actual points of entry are still needed. The outputs will benefit Canada's forest resources ecologically and economically as the mid-range projected annual loss to industry for individual FIS are: US$16 M for ALB, US$625 M for DED, US$121M for the gypsy moth (Colautti et al. 2006) and US$25 M for SOD (Nelson et al. 2009).

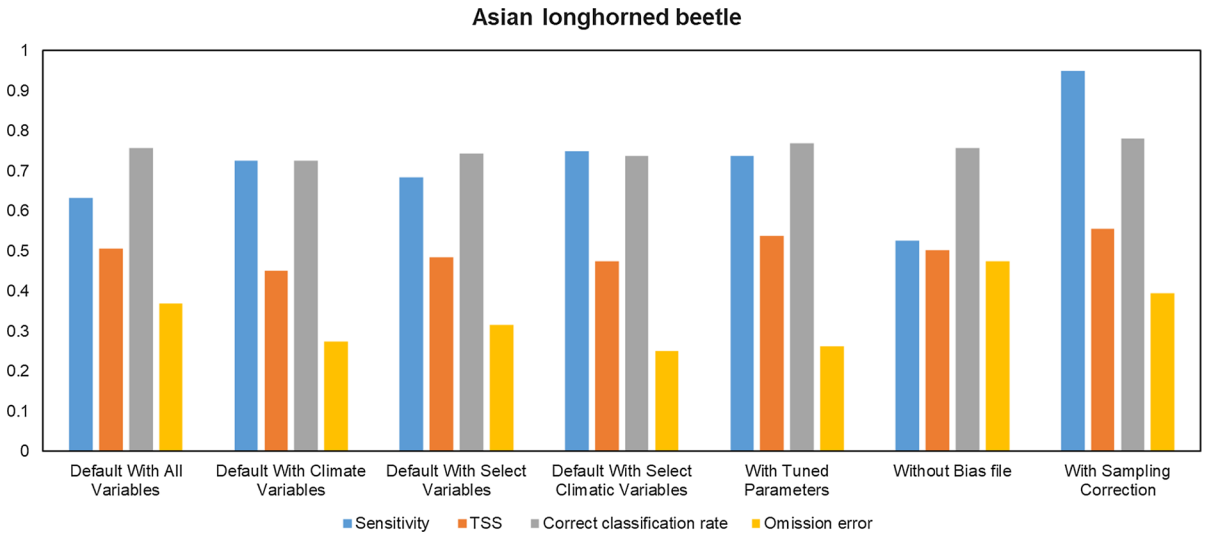
Acknowledgements This work was generously funded by Genome Canada, Genome British Columbia and Genome Quebec within the framework of project bioSAFE (Biosurveillance of Alien Forest Enemies, Project Number #10106) as part of a Large-Scale Applied Research Project in Natural Resources and the Environment. We are thankful to Dr A. Townsend Peterson and colleagues at Canadian Food Inspection Agency for reviewing and helping us to improve the manuscript with their insights.

Appendix 1

Evaluation summary of FIS models using TSS, correct classification rates, omission error and sensitivity metrics.

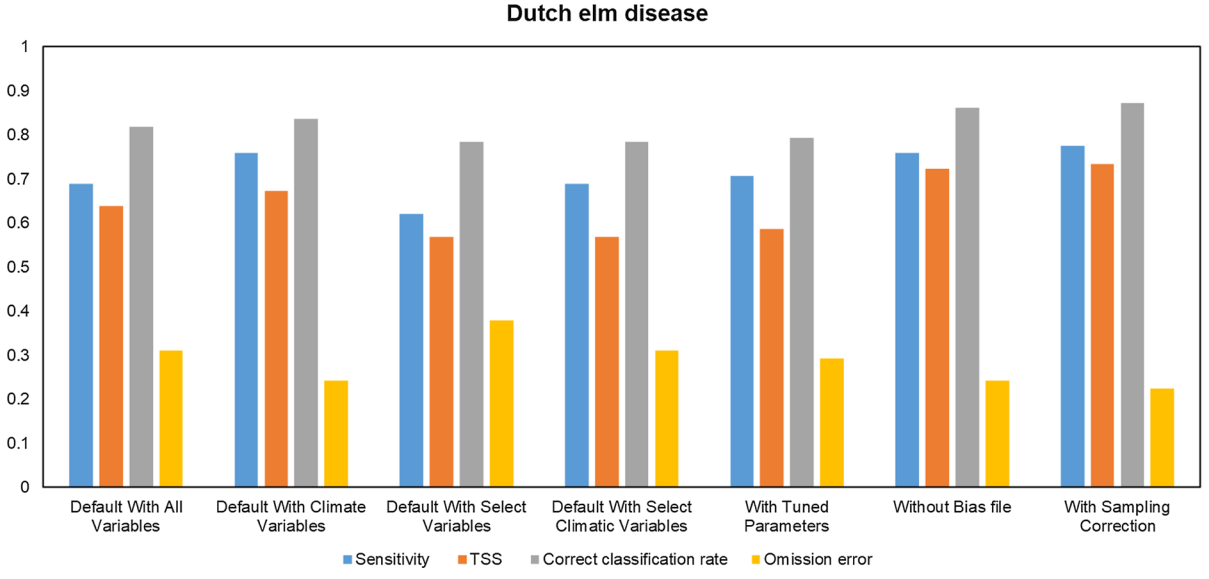
AGM				
Evaluation metric	Sensitivity	TSS	CCR	OR
Default with all variables	0.500	0.393	0.696	0.500
Default with climate variables	0.589	0.482	0.741	0.411
Default with select variables	0.536	0.375	0.688	0.464
Default with select climatic variables	0.679	0.429	0.714	0.321
With tuned parameters	0.750	0.518	0.759	0.250
Without bias file	0.464	0.375	0.688	0.536
With sampling correction	0.736	0.528	0.764	0.264



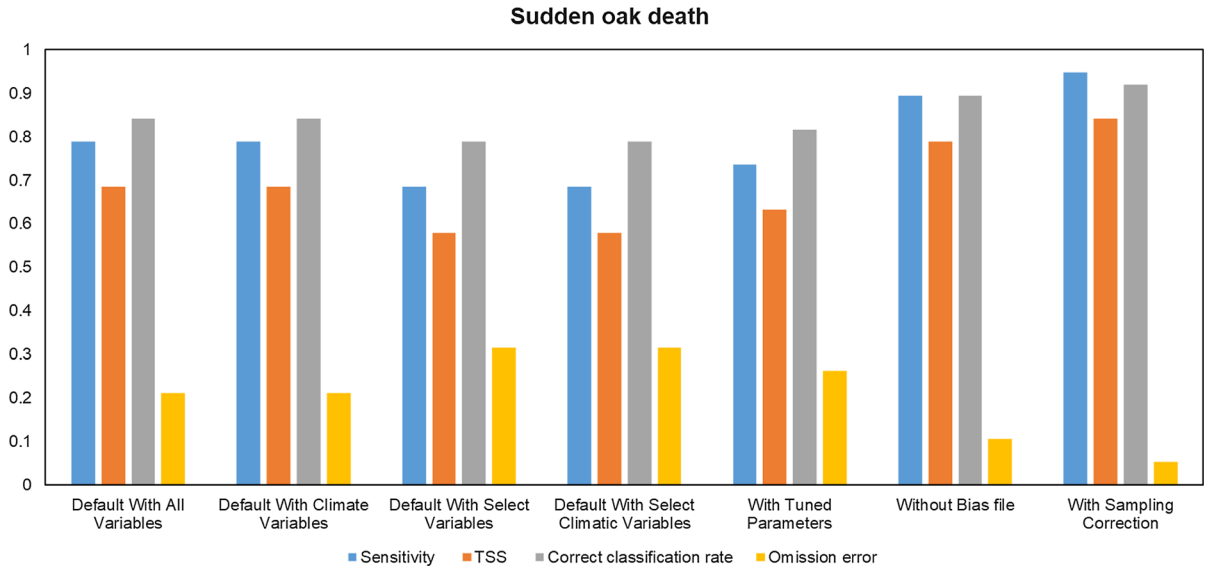


ALB

Evaluation metric	Sensitivity	TSS	CCR	OR
Default with all variables	0.632	0.507	0.756	0.368
Default with climate variables	0.725	0.450	0.725	0.275
Default with select variables	0.684	0.484	0.744	0.316
Default with select climatic variables	0.750	0.475	0.738	0.250
With tuned parameters	0.737	0.537	0.769	0.263
Without bias file	0.526	0.501	0.756	0.474
With sampling correction	0.950	0.555	0.782	0.395



DED				
Evaluation metric	Sensitivity	TSS	CCR	OR
Default with all variables	0.690	0.638	0.819	0.310
Default with climate variables	0.759	0.672	0.836	0.241
Default with select variables	0.621	0.569	0.784	0.379
Default with select climatic variables	0.690	0.569	0.784	0.310
With tuned parameters	0.707	0.586	0.793	0.293
Without bias file	0.759	0.724	0.862	0.241
With sampling correction	0.776	0.734	0.872	0.224

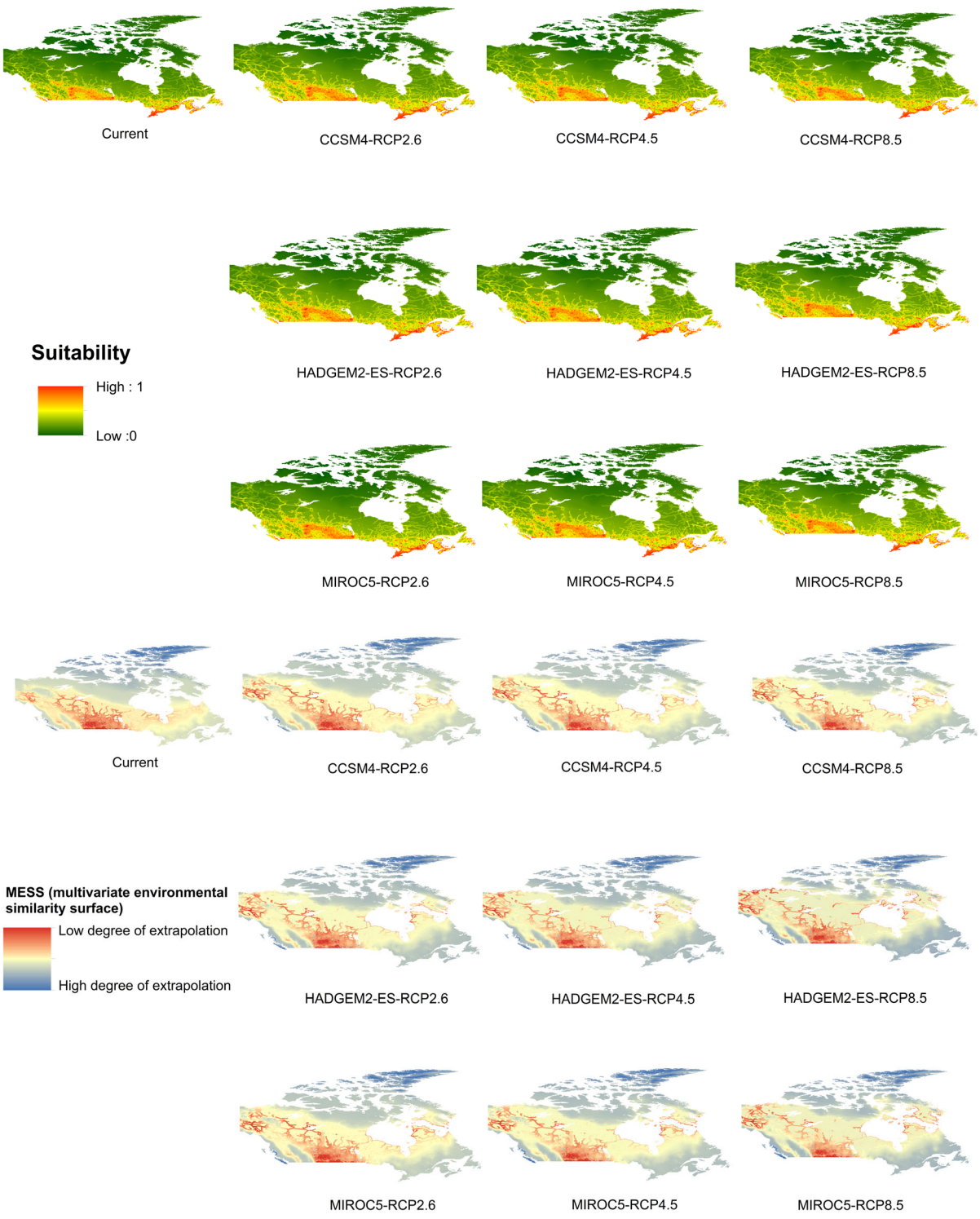


SOD				
Evaluation metric	Sensitivity	TSS	CCR	OR
Default with all variables	0.789	0.684	0.842	0.211
Default with climate variables	0.789	0.684	0.842	0.211
Default with select variables	0.684	0.579	0.789	0.316
Default with select climatic variables	0.684	0.579	0.789	0.316
With tuned parameters	0.737	0.632	0.816	0.263
Without bias file	0.895	0.789	0.895	0.105
With sampling correction	0.947	0.842	0.921	0.053

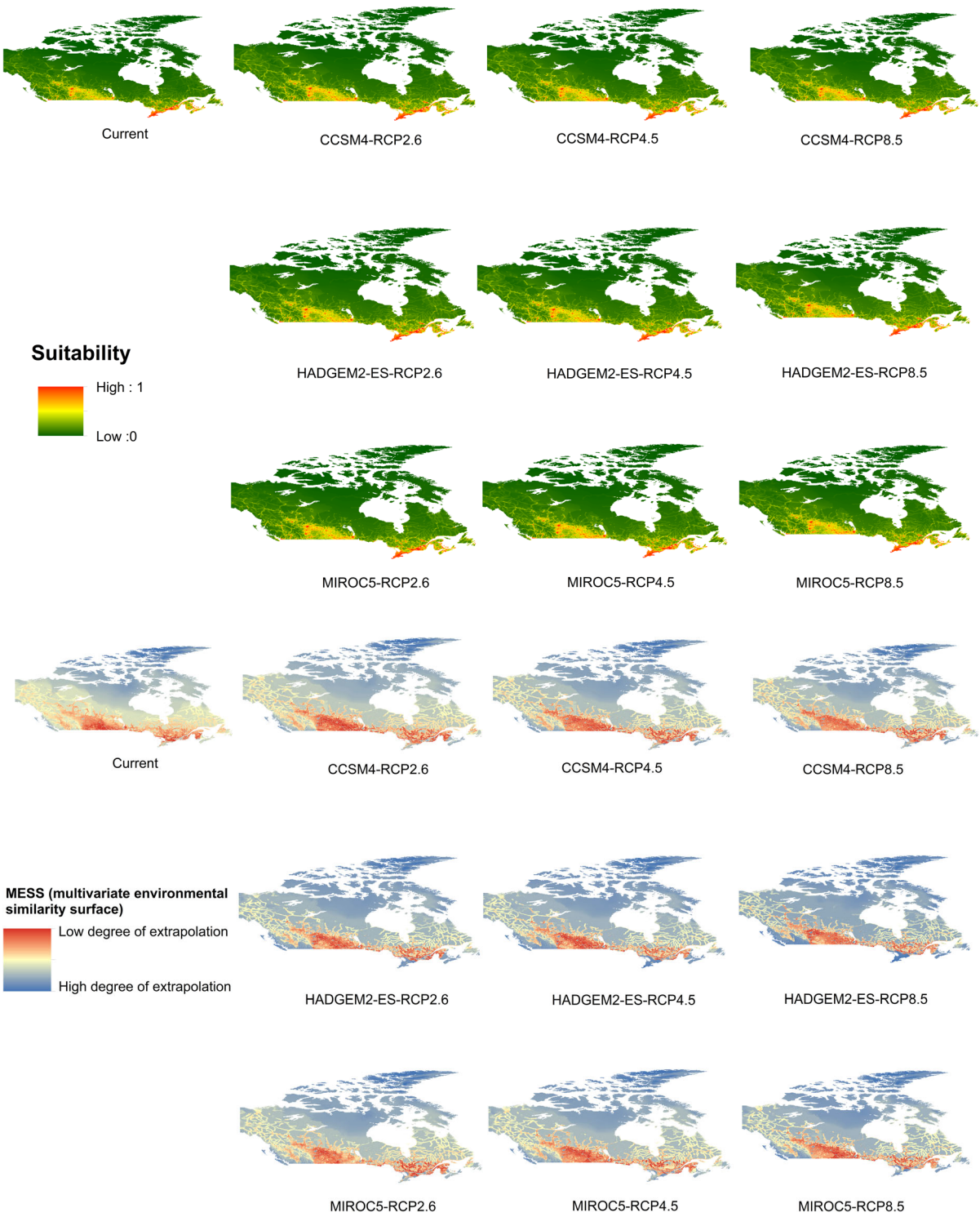
Appendix 2

Potential distribution of selected FIS in current and future climate change scenarios along with MESS (multivariate environmental similarity surface) maps. For suitability maps, higher probability (red colors) represent areas suitable for FIS. Zero probability or lower probability (dark green) indicates areas less suitable. For MESS maps, increase in blue tone denotes increasing degree of extrapolation on at least one variable. Suitability predictions in those areas (blue) should be treated with high caution.

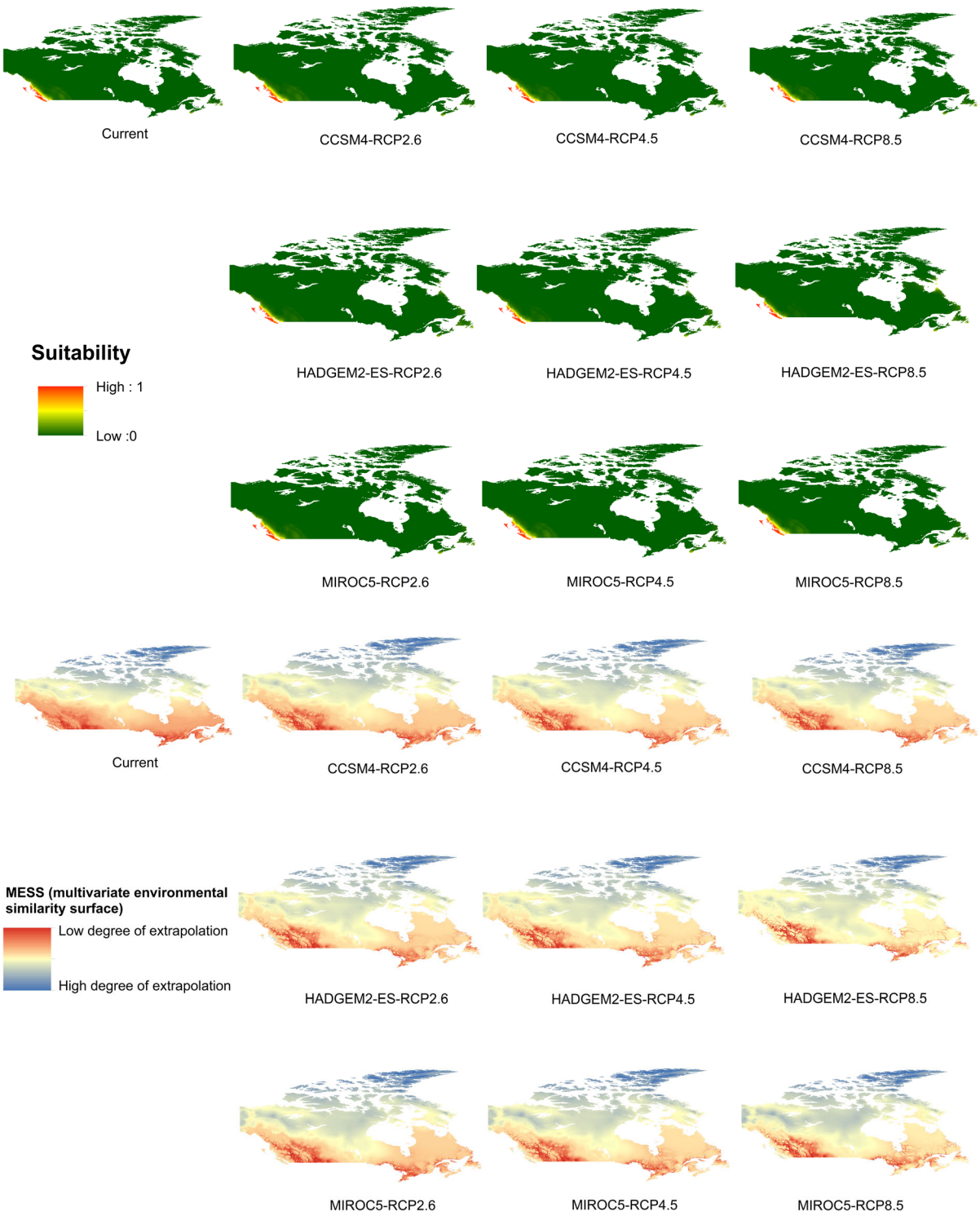
(a) Asian gyspy moth



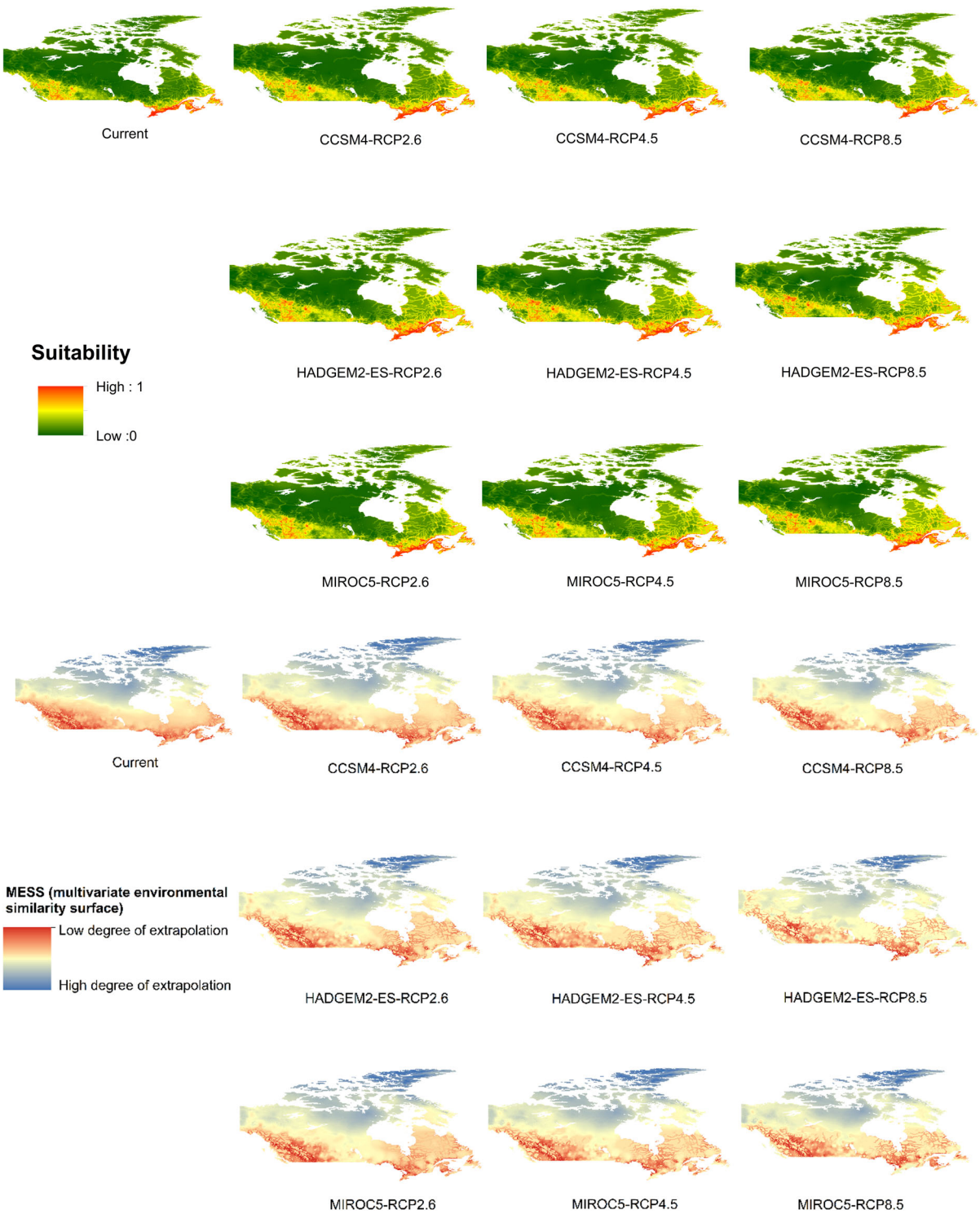
(b) Asian longhoned beetle



(c) Sudden oak death



(d) Dutch elm disease



Appendix 3

See Fig. 8.

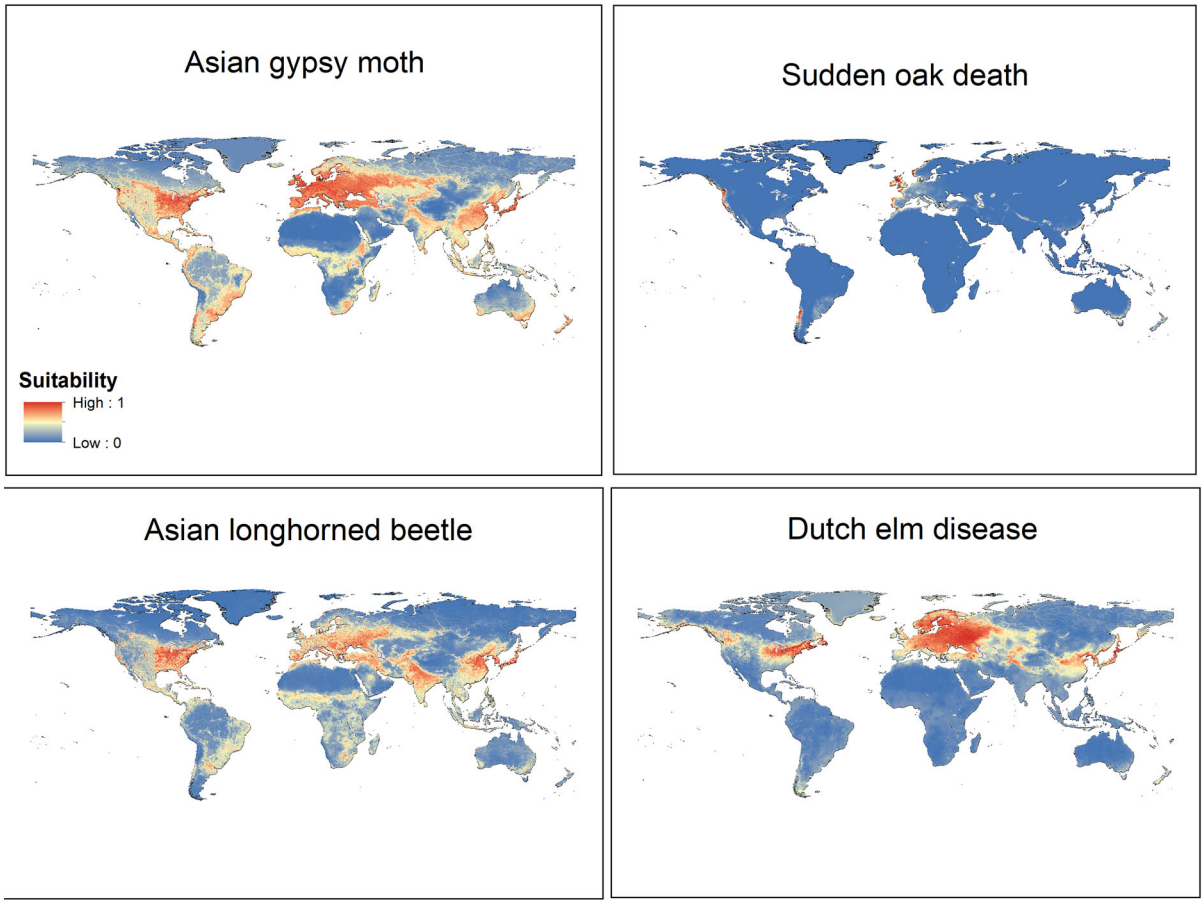


Fig. 8 Predicted potential distribution of selected FIS on a global scale. Higher probability (red colors) represent areas suitable for FIS. Zero probability or lower probability (dark blue) indicates areas less suitable

Appendix 4: environmental response curves

See Fig. 9.

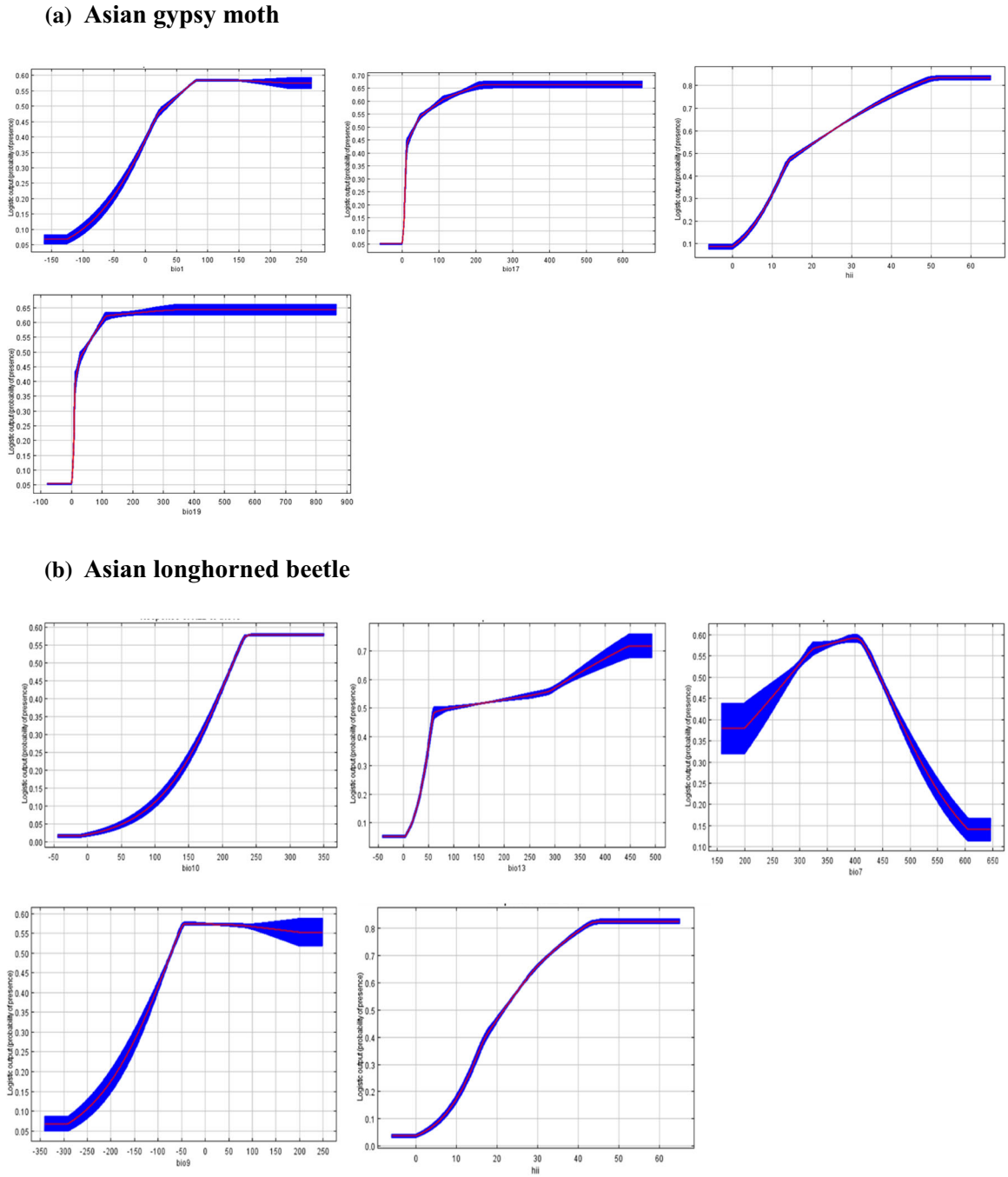
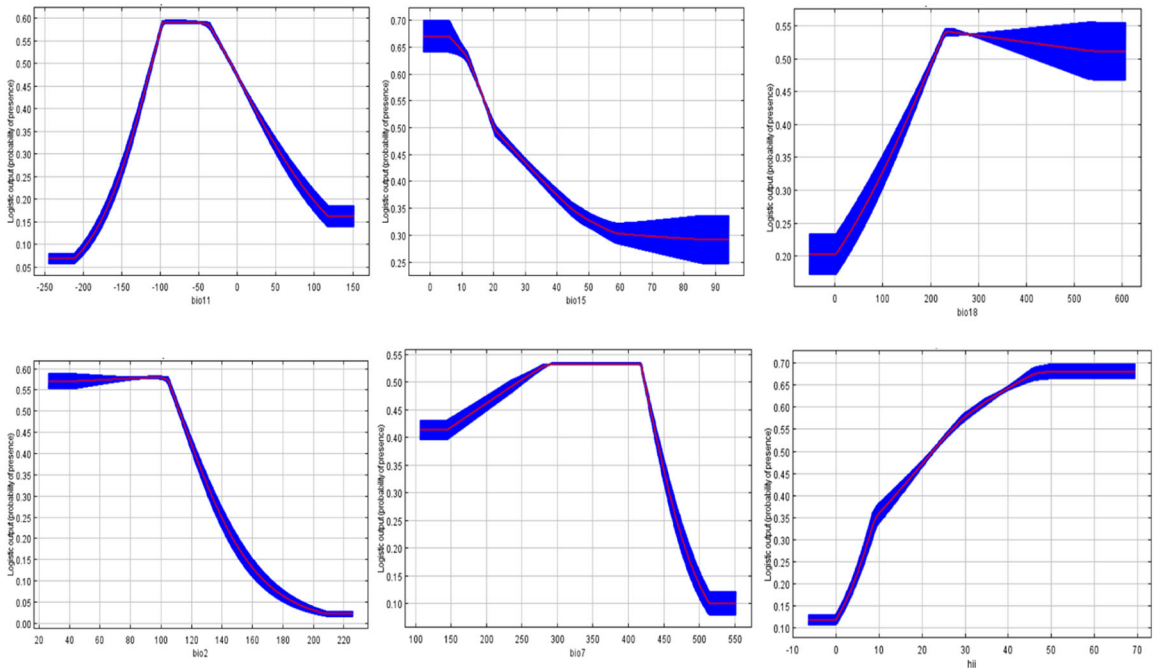


Fig. 9 Relationships between environmental predictors and the probability of the presence of FIS: red curves show the mean response and blue margins are ± 1 SD calculated over 10 replicates

(c) Dutch elm disease



(d) Sudden oak death

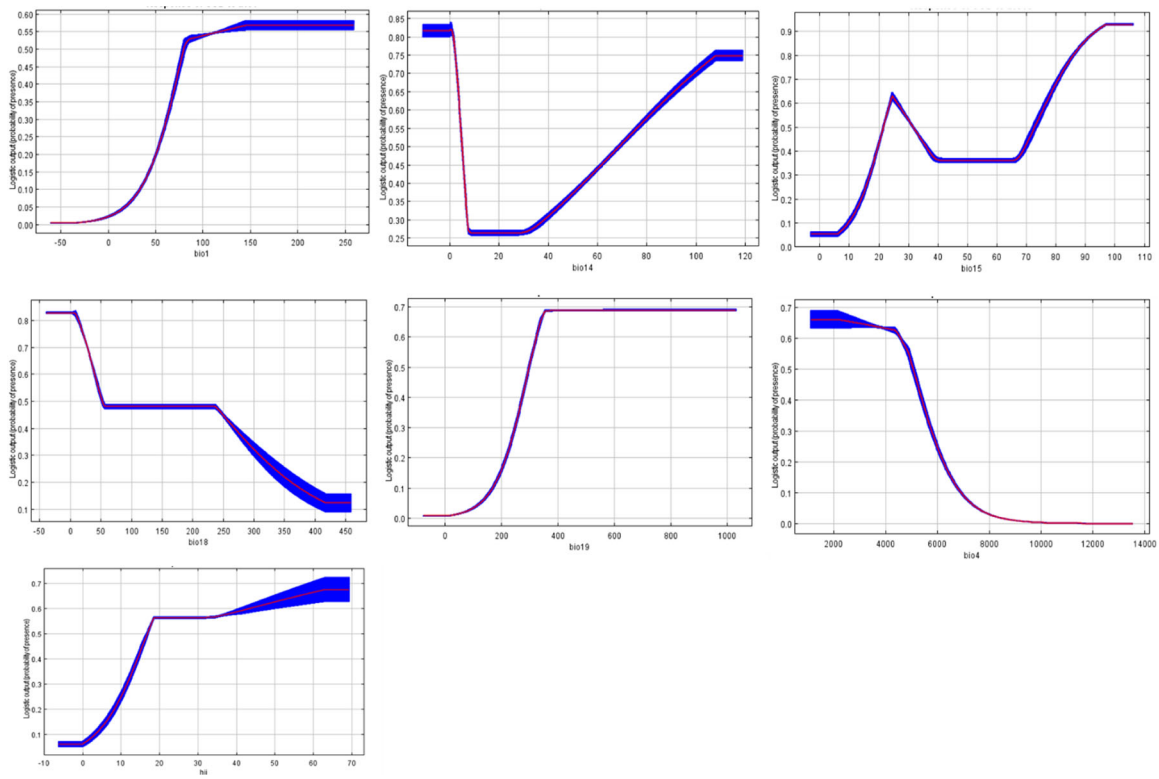


Fig. 9 continued

Appendix 5

Comparing dispersal limited to unlimited FIS dispersal projections under climate change conditions. Here numbers represents total number of cells colonized under each scenario.

(a) Asian gypsy moth

GCM/scenario	Infestation point-Vancouver port			Infestation point-Toronto port			Unlimited dispersal		
	ccsm4	hadgem2es	miroc5	ccsm4	hadgem2es	miroc5	ccsm4	hadgem2es	miroc5
rcp26	3416	4143	5074	5393	5260	5499	19,142	19,509	19,265
rcp45	5262	4608	5115	5479	5532	5726	19,383	19,592	19,445
rcp85	5318	5473	5939	5619	5788	5843	19,548	19,770	19,599

(b) Asian longhorned beetle

GCM/scenario	Infestation point-Vancouver port			Infestation point-Toronto port			Unlimited dispersal		
	ccsm4	hadgem2es	miroc5	ccsm4	hadgem2es	miroc5	ccsm4	hadgem2es	miroc5
rcp26	71	76	68	2482	2497	2446	7701	7366	7647
rcp45	75	82	72	2528	2539	2613	8117	7904	8116
rcp85	76	91	79	2567	2587	2663	8716	8510	8334

(c) Dutch elm disease

GCM/scenario	Infestation point-Toronto port			Unlimited dispersal		
	ccsm4	hadgem2es	miroc5	ccsm4	hadgem2es	miroc5
rcp26	1438	1379	1443	2051	2045	1992
rcp45	1538	1555	1524	2026	2061	2013
rcp85	1577	1616	1622	2039	2059	2040

(d) Sudden oak death

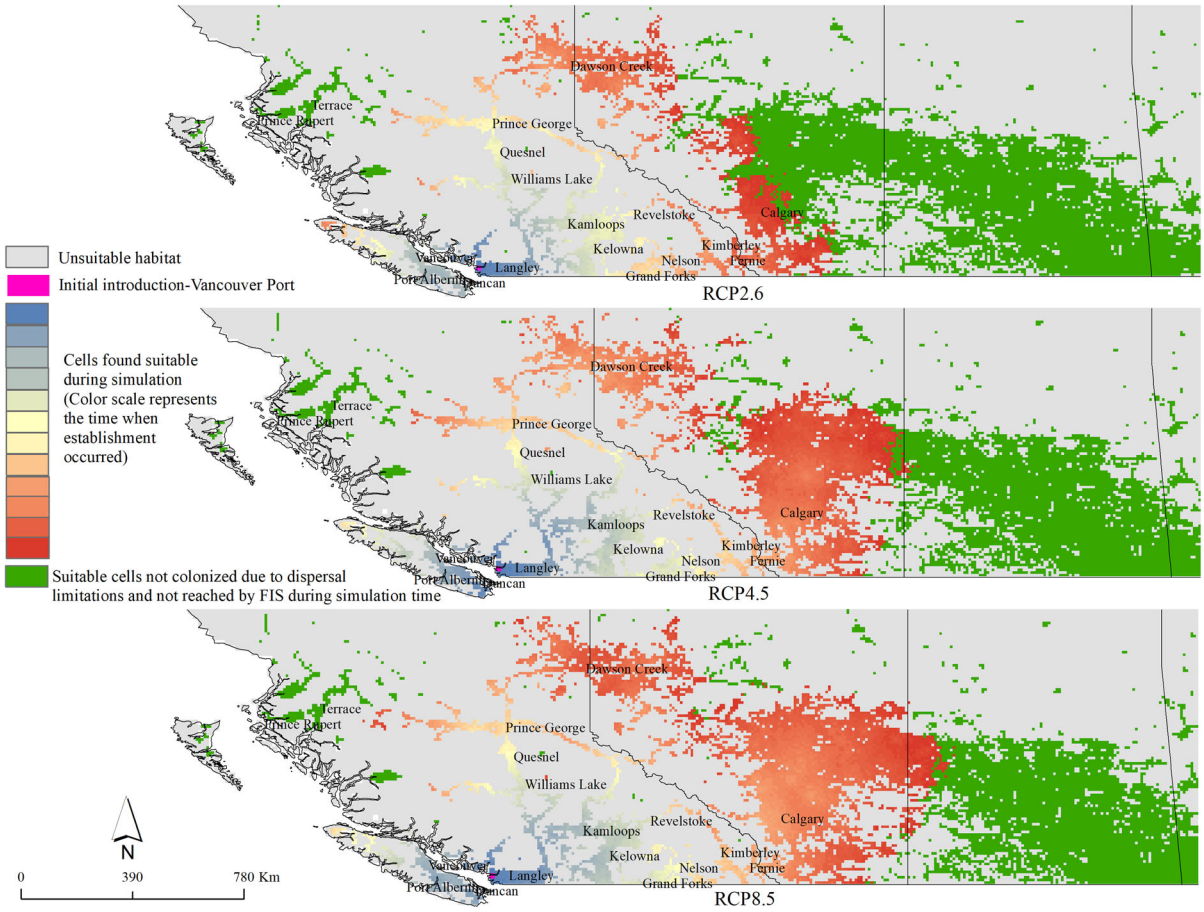
GCM/scenario	Infestation point-Vancouver port			Unlimited dispersal		
	ccsm4	hadgem2es	miroc5	ccsm4	hadgem2es	miroc5
rcp26	879	625	893	1166	977	1061
rcp45	606	650	848	1063	947	1097
rcp85	433	500	759	1033	911	1036

Appendix 6

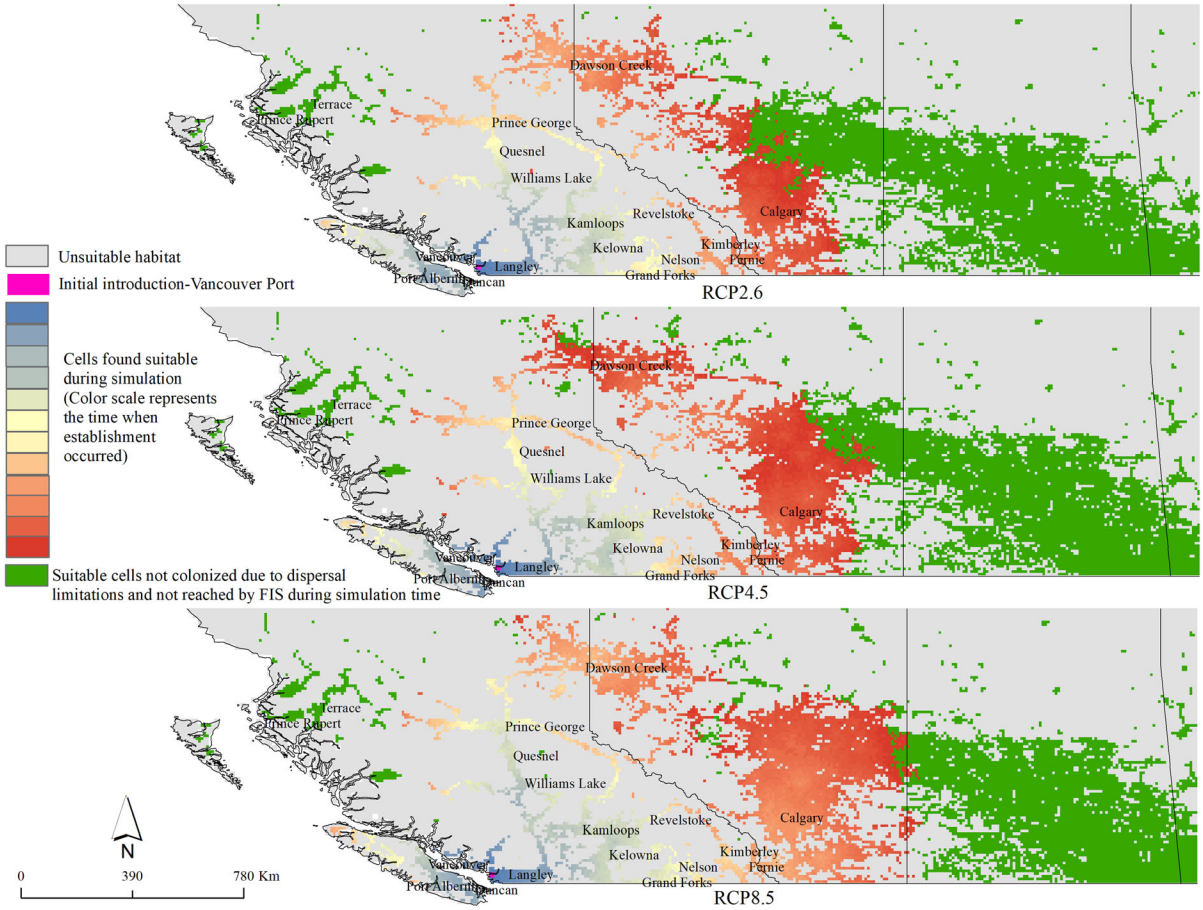
FIS dispersal limited distributions under different climate change scenarios and two hypothesized infestation points-

AGM (Infestation point-Vancouver port)

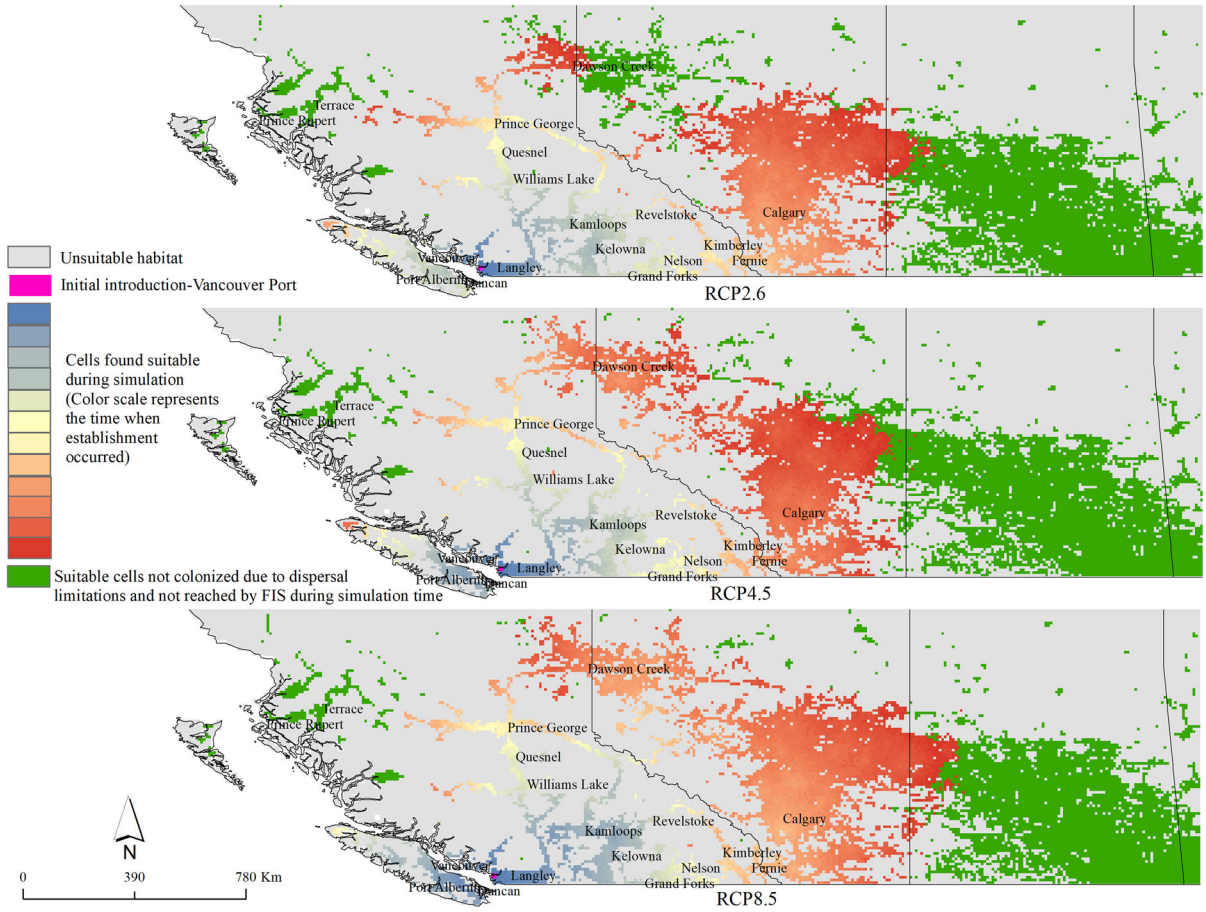
Dispersal restricted future distribution of AGM under GCM-CCM4 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next



10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

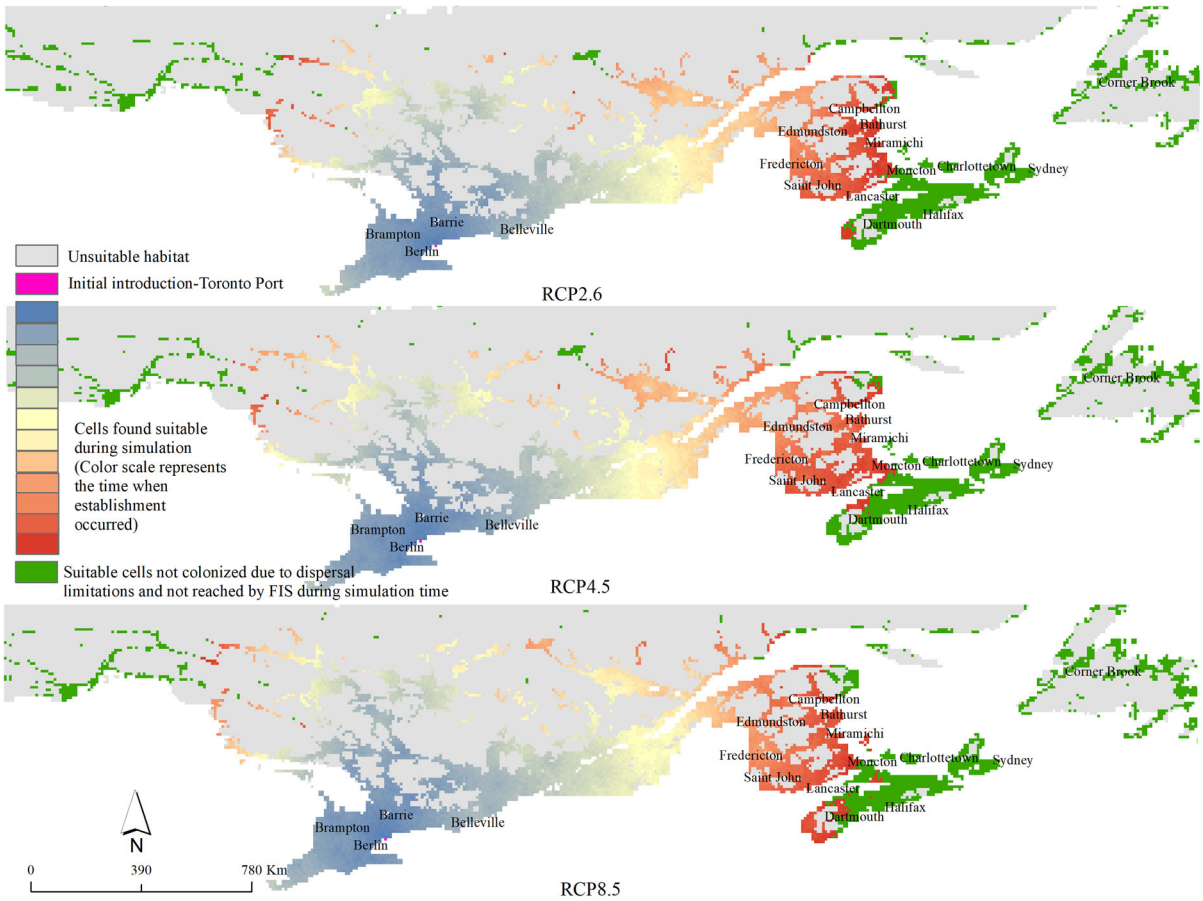


Dispersal restricted future distribution of AGM under GCM-HADGEM2ES and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

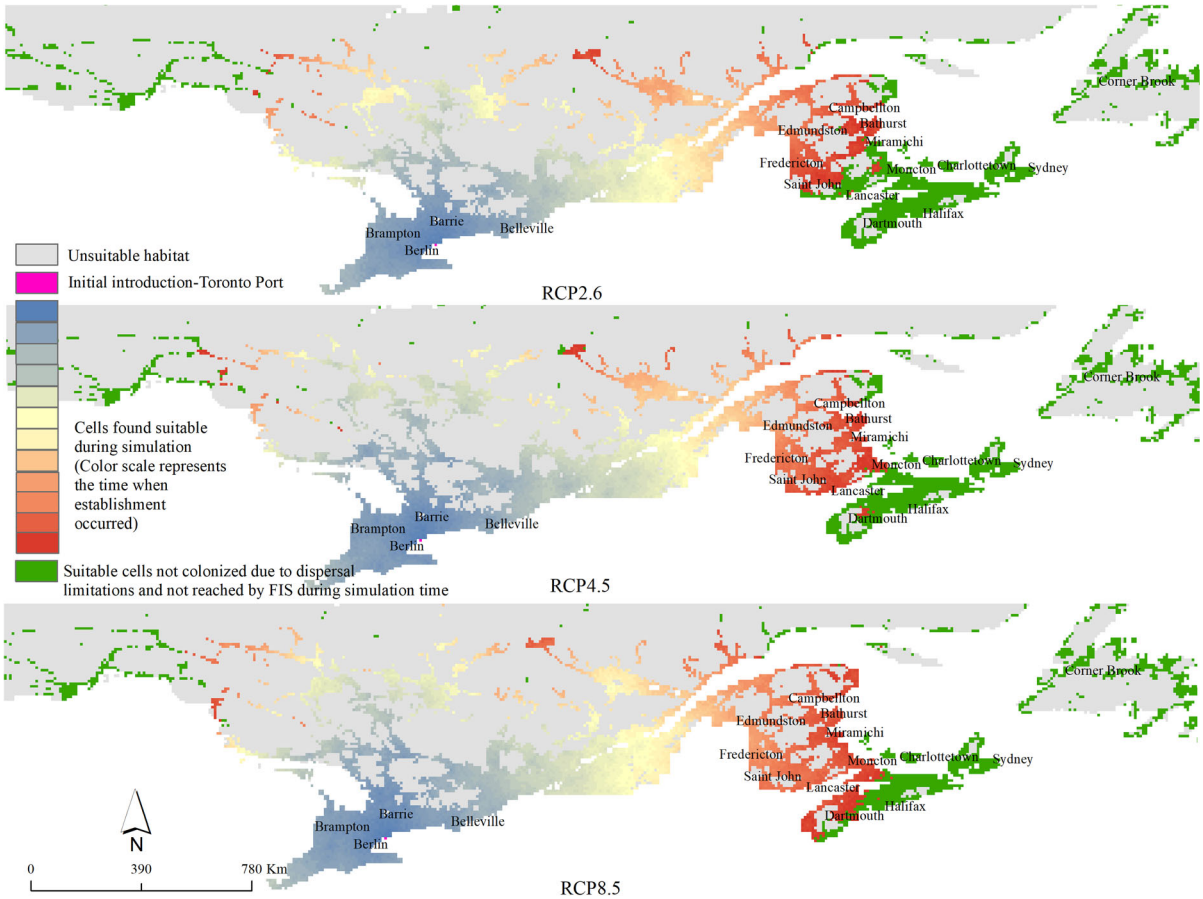


Dispersal restricted future distribution of AGM under GCM-MIROC5 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

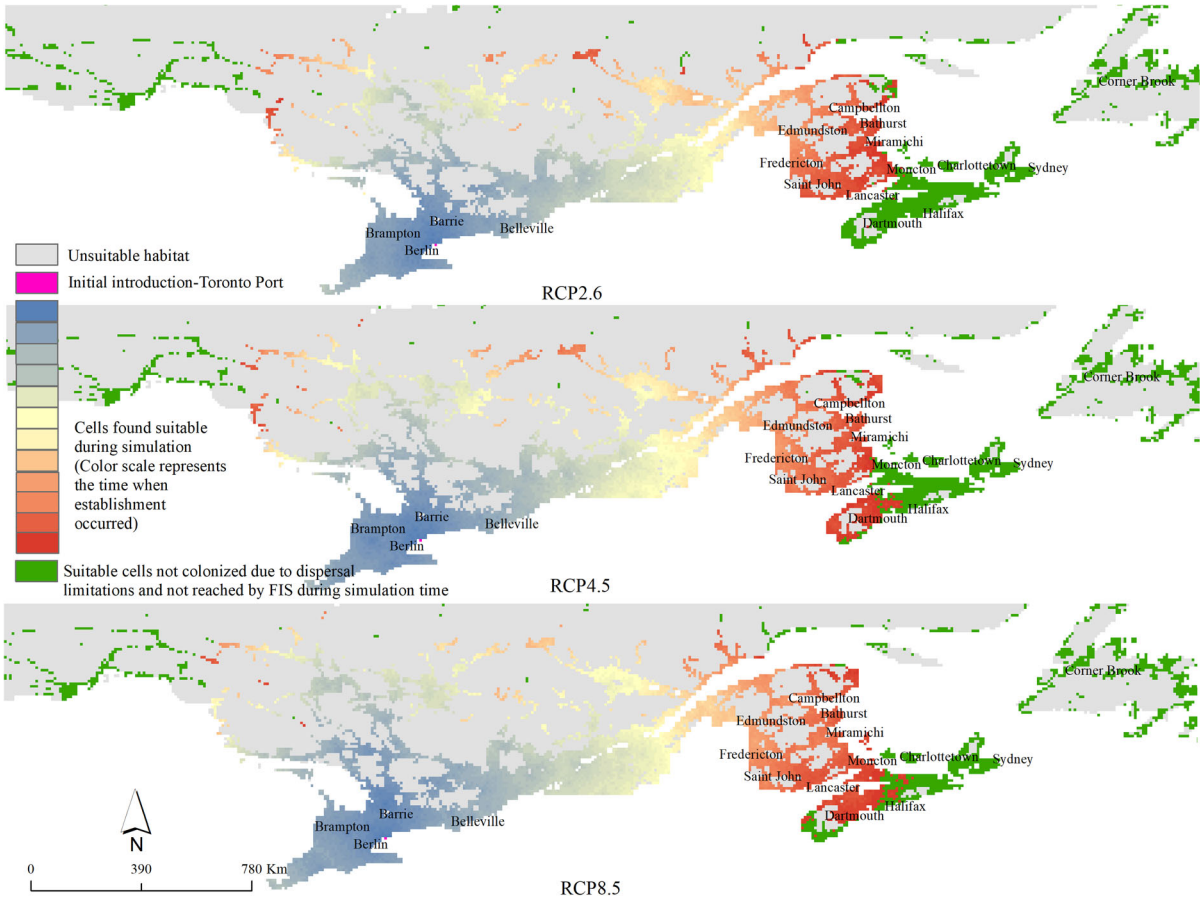
AGM (Introduction point-Toronto port)



Dispersal restricted future distribution of AGM under GCM-CCSM4 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

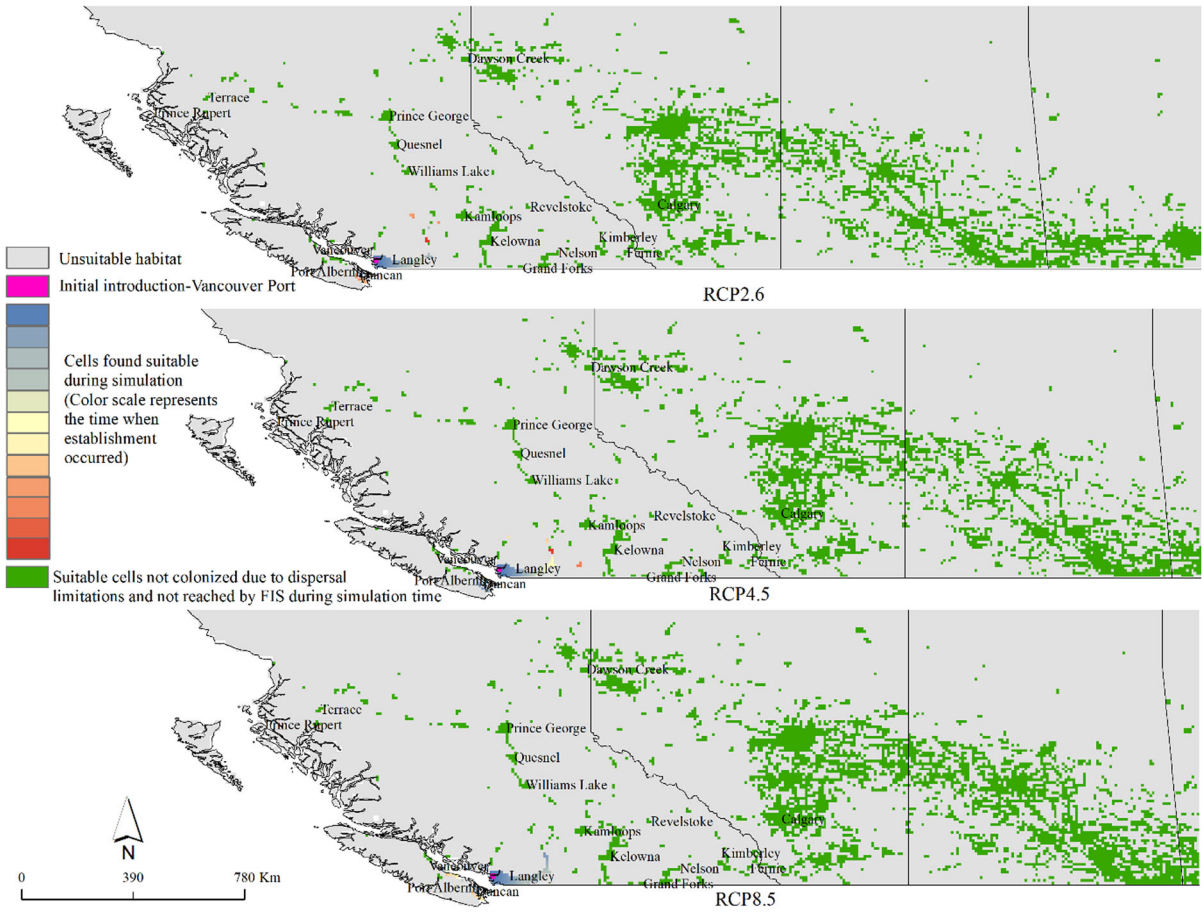


Dispersal restricted future distribution of AGM under GCM-HADGEM2ES and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

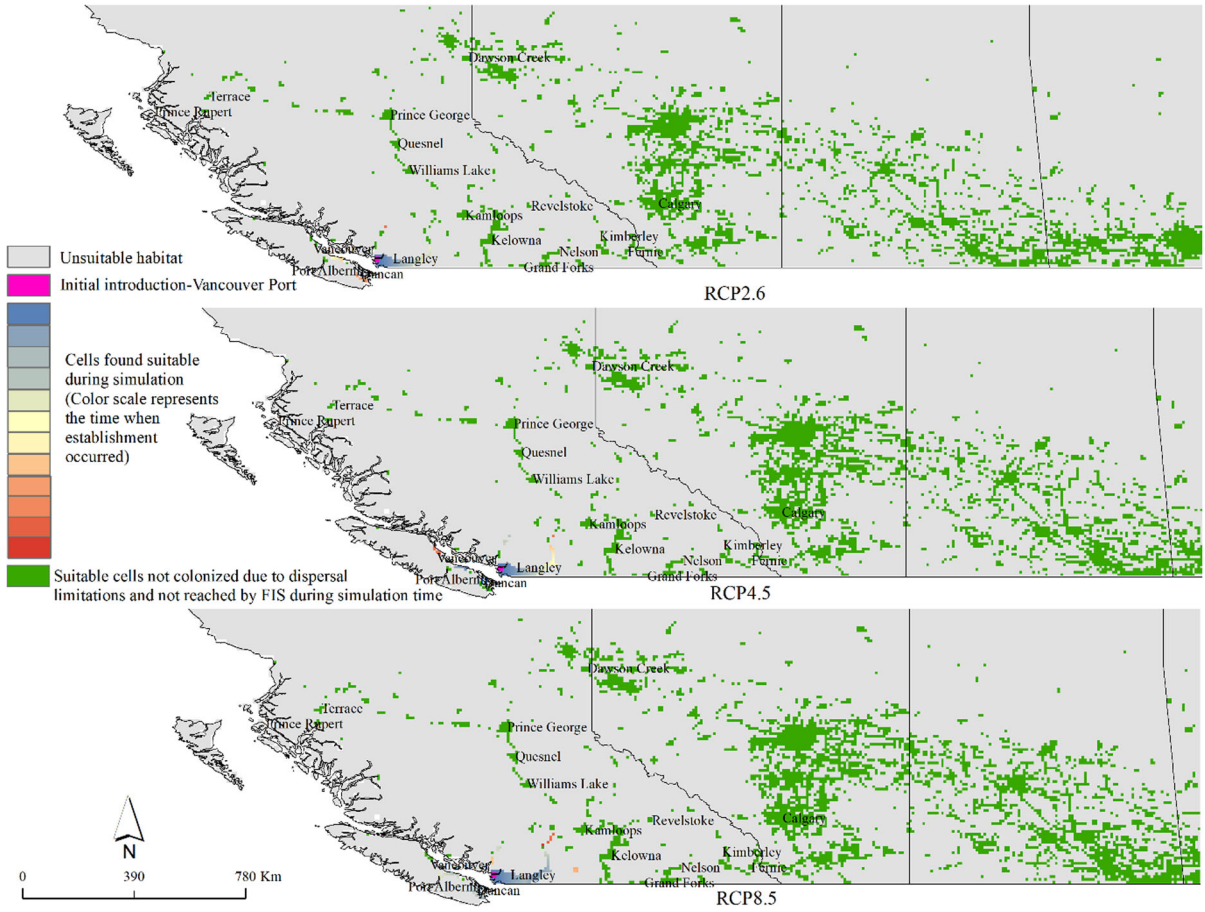


Dispersal restricted future distribution of AGM under GCM-MIROC5 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

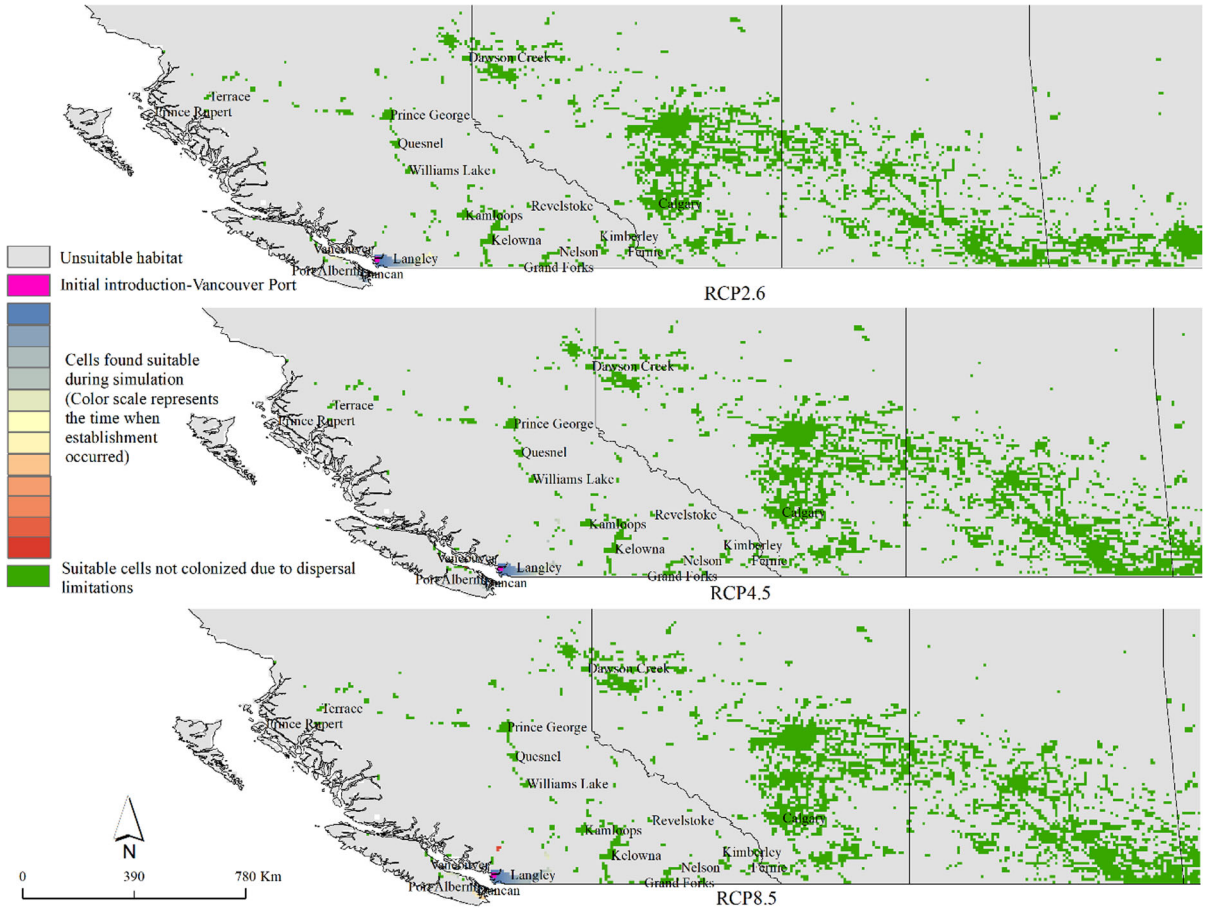
ALB (Introduction point-Vancouver port)



Dispersal restricted future distribution of ALB under GCM-CCSM4 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

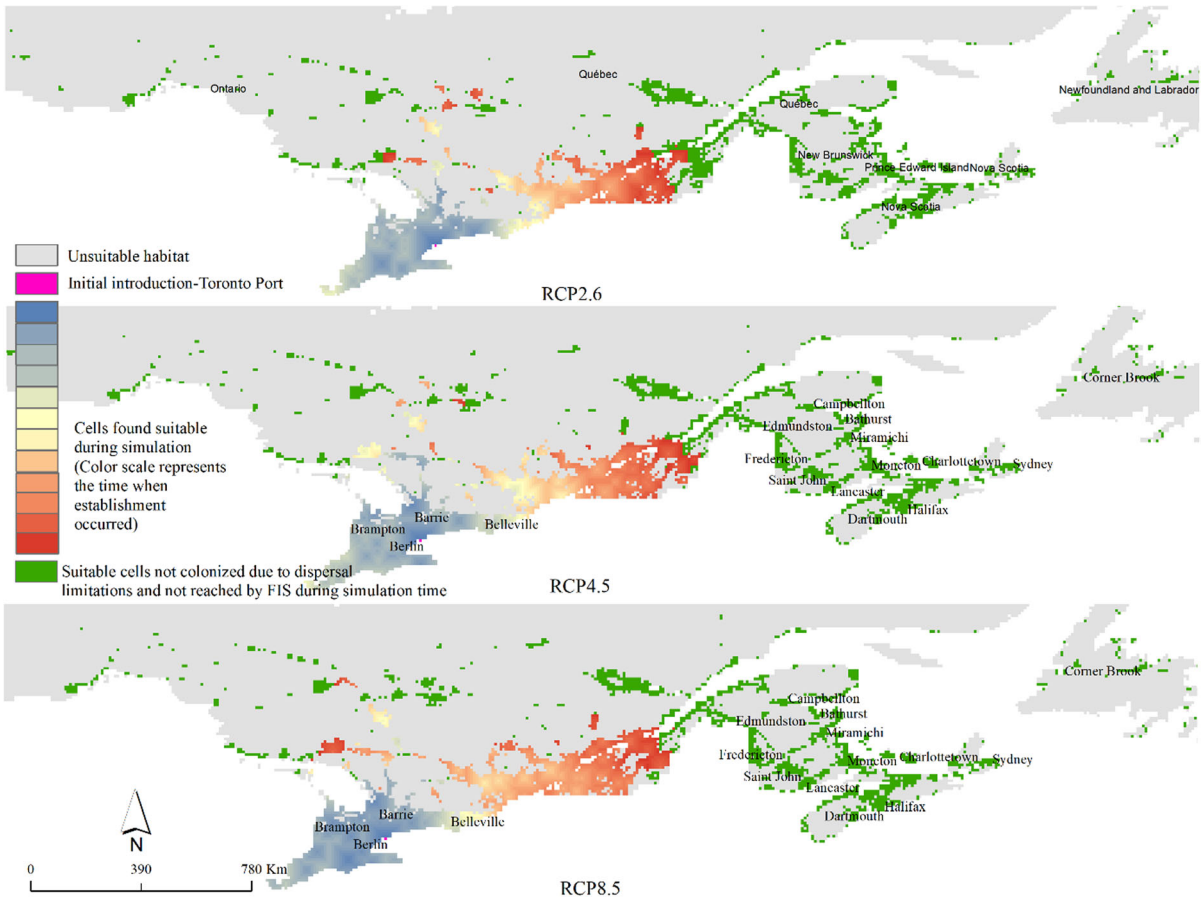


Dispersal restricted future distribution of ALB under GCM-HADGEM2ES and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

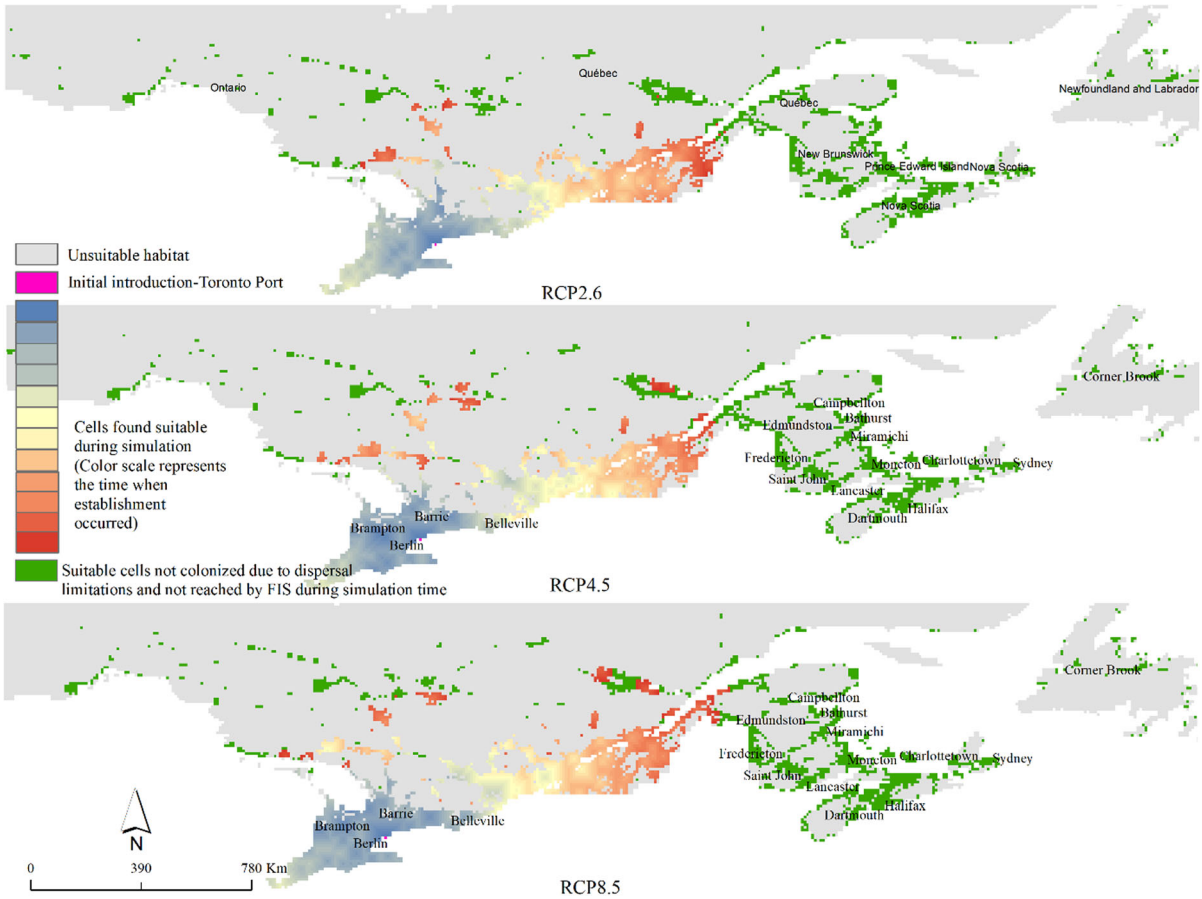


Dispersal restricted future distribution of ALB under GCM-MIROC5 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

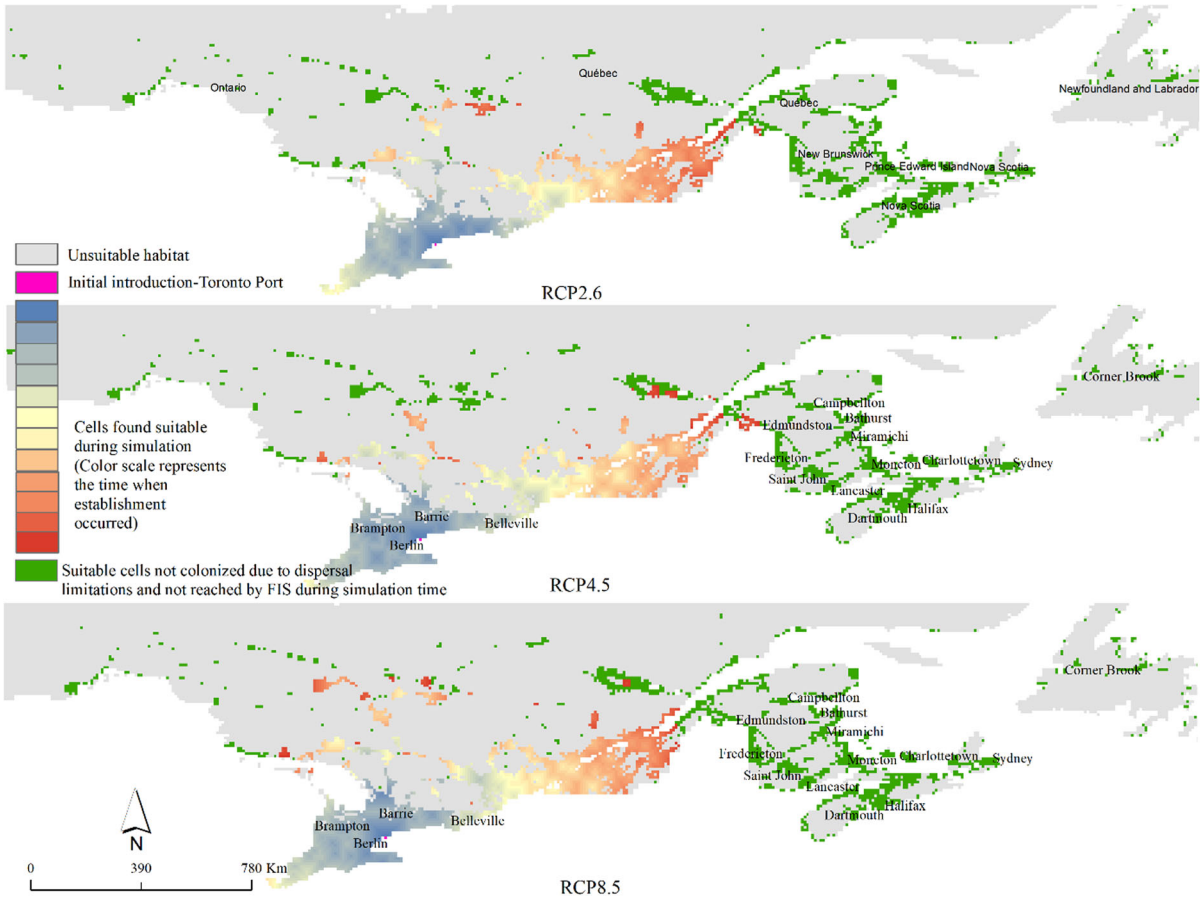
ALB (Introduction point-Toronto port)



Dispersal restricted future distribution of ALB under GCM-CCSM4 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

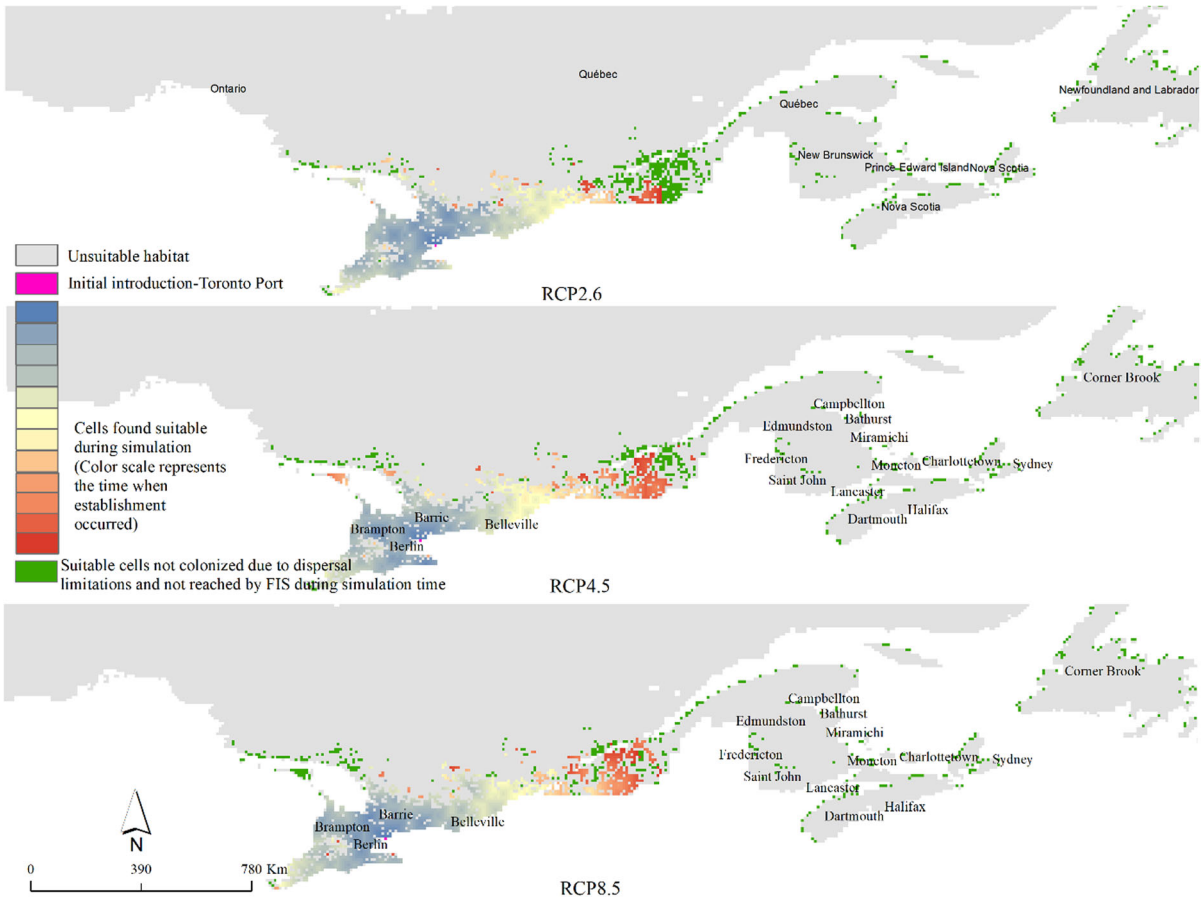


Dispersal restricted future distribution of ALB under GCM-HADGEM2ES and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

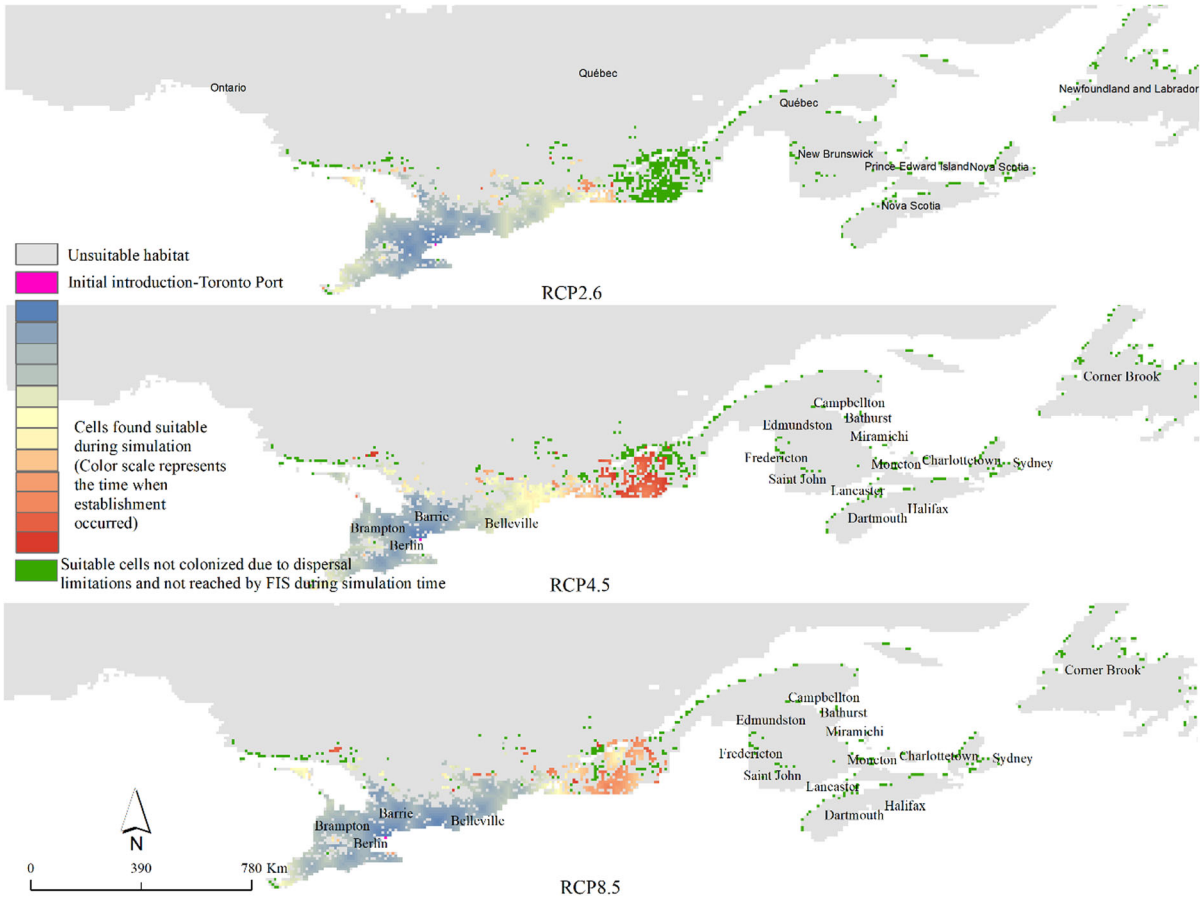


Dispersal restricted future distribution of ALB under GCM-MIROC5 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

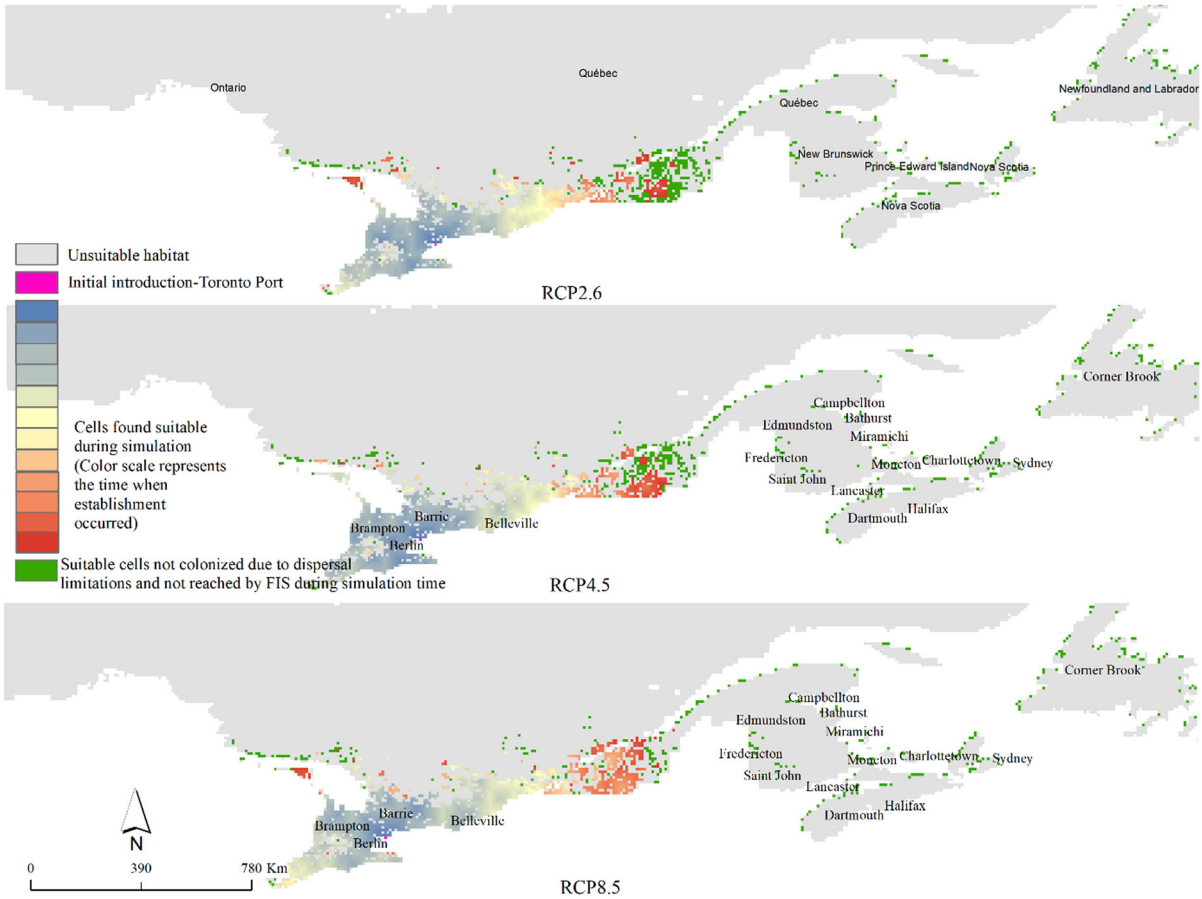
DED (Introduction point-Toronto port)



Dispersal restricted future distribution of DED under GCM-CCM4 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

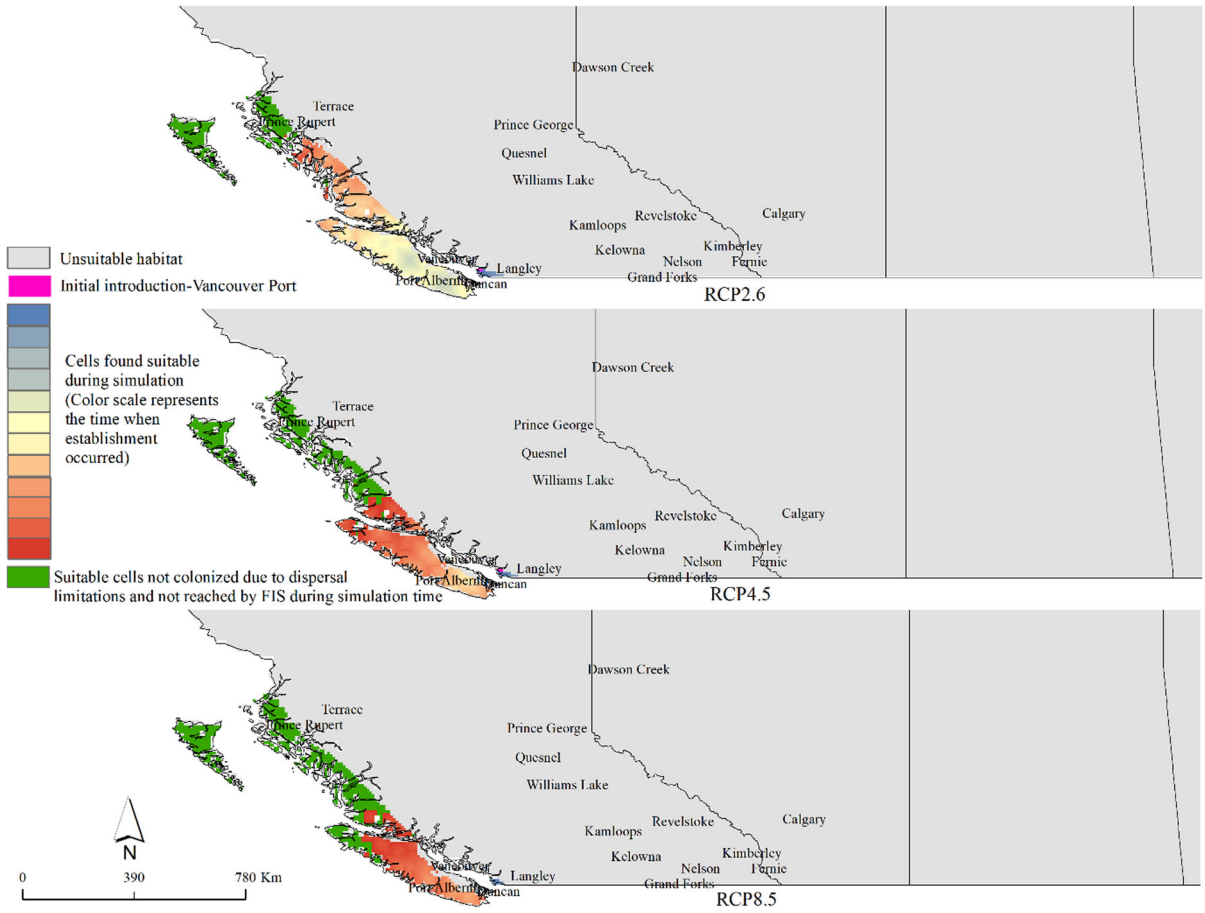


Dispersal restricted future distribution of DED under GCM-HADGEM2ES and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

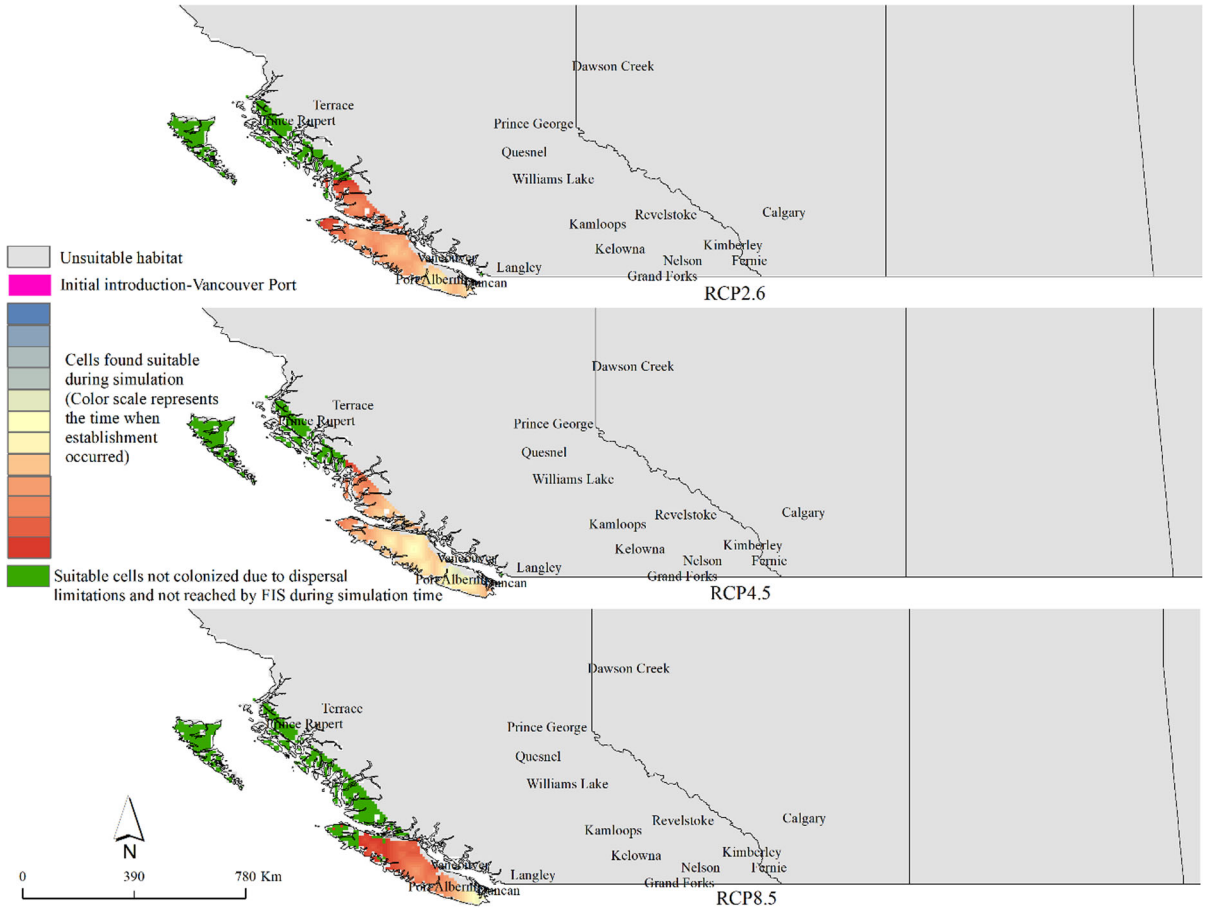


Dispersal restricted future distribution of DED under GCM-MIROC5 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

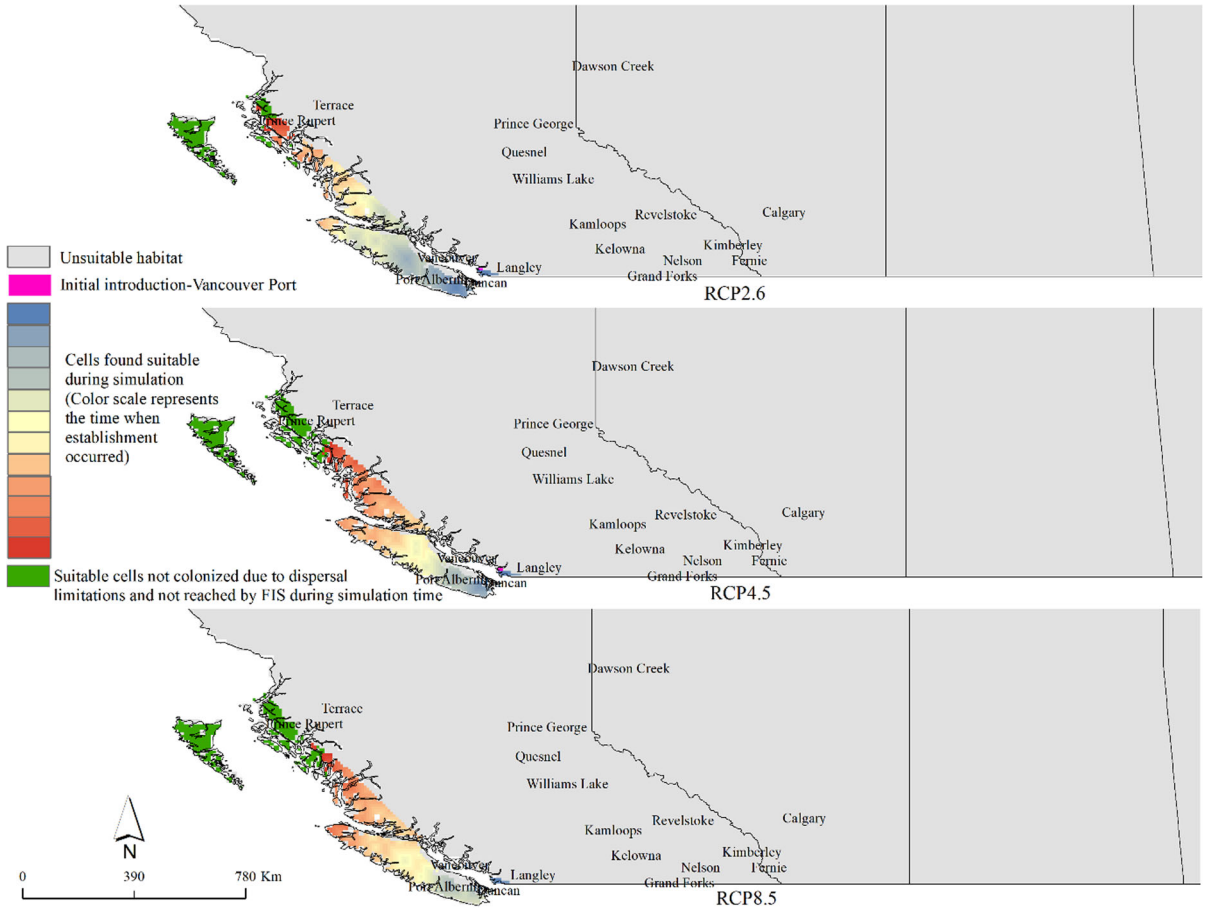
SOD (Introduction point-Vancouver port)



Dispersal restricted future distribution of SOD under GCM-CCM4 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.



Dispersal restricted future distribution of SOD under GCM-HADGEM2ES and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.



Dispersal restricted future distribution of SOD under GCM-MIROC5 and RCP 2.6, 4.5 and 8.5 climate change scenarios. Color gradient from blue to grey represents the first 10 years of the simulation time frame when colonization first occurred, the light grey to light yellow color gradient represents the next 10 years followed by orange and rose color gradients (years 2031–2050). Pink colored pixel indicates the hypothesized point of DED introduction (port of Toronto) while the green pixels represent suitable areas that were not colonized due to dispersal limitations and not reached by FIS during the time of simulation.

Appendix 7

AGM life history parameters and associated references

AGM is a potent invader with more than 600 known hosts. AGM females are capable of flight and can lay eggs on human-made objects.

- Generations per year
  - Univoltine—one generation per year (Elkinton and Liebhold 1990)

- Dispersal
  - Adult females disperse and spread their population naturally by sustained flight and windborne dispersal of first instars (Keena et al. 2008).
  - Attracted to lights at night (Montgomery and Wallner 1988; Schaefer and Strothkamp 2014)
- Dispersal distance
  - Frequent long distance dispersal flights (average less than 1 km to max range of 20–40 km) (Iwaizumi et al. 2010; Keena et al. 2008)
  - Russian females may fly distances up to 100 km and eastern Siberian females seen crossing mountain ranges in large groups during outbreaks (Rozhkov and Vasilyeva 1982)
  - Egg masses in Japanese cities found within 1 km of forests(Liebhold et al. 2008).
  - Average flight distance of 1 day old Chinese females in 8 h on flight mills was 5.65 km and maximum was 10.67 km (Yang et al. 2017)
- Reproductive capacity
  - Producing an average of 600–1000 eggs per egg mass (USDA)
- Distribution
  - Found throughout temperate Asia. Usually east of the Ural Mountains into Far East Russia, through most of Japan, China and Korea. It is not found east of the Himalayan range in India (USDA)
- Critical temp.
  - AGM populations may struggle in regions experiencing longer periods of temperatures $\geq$ 30 °C and survival rate is highest between 15and 25 °C (Limbu et al. 2017).

ALB life history parameters and associated references

- Sex ratio
  - 1-!:14 male–female (Bancroft and Smith 2005)
  - 1:1 male–female (Trotter et al. 2019)
  - Other papers use only the females to model spread since she drives the establishment of new infestations
- Generations per year
  - Temperature dependent
    - (Haack et al. 2010)
    - (Keena and Moore 2010)
    - (Faccoli and Gatto 2016)
    - (Favaro et al. 2015)
  - Not strictly univoltine (1 year); may take multiple years to develop
    - (Keena and Moore 2010; Trotter and Keena 2016) (in Finland may take 10+ years)
    - (Straw et al. 2015)
      - 3 years for Paddock Woods
    - (Kappel et al. 2017)
      - Do not use the Newtonian Cooling model to estimate within tree temps—may not accurately reflect temps within tree
      - But estimated that in northern states will take minimum 2–3 years to complete development, some areas up to 5–6 years
- Dispersal distance
  - Frequent short distance dispersal flights (< 1.5 km)
    - (Javal et al. 2018)
  - Tendency to remain on and reinfest natal tree
    - (Haack et al. 2010)
  - Dispersal occurs when tree host quality deteriorates
    - (Sawyer 2007)
  - Rare long distance flights (< 1.5 km)
    - Human mediated transport likely more significant at farther distances
      - (Fournier and Turgeon 2017)
  - $\sim$ 10 km (modeled and based on graph)
    - (Trotter et al. 2019)

- Longest single sustained flight on flight mill = 4006 m; median = 247.6 m
  - (Javal et al. 2018)
- Lifetime dispersal for a female = 14,060 m; median = 3964 m
  - (Javal et al. 2018)

- Spread rates
  - in England in one stand, mean rate of population spread 29.3 m/year
    - (Straw et al. 2016)
  - Jersey City spread 50 m/year
    - (Sawyer et al. 2011)
  - New Jersey spread 2.4–3.2 km in 5–6years
    - (Sawyer et al. 2011)
  - Italy spread 2 × 2 km in 5 years
    - (Faccoli et al. 2015)
- Probability to disperse
  - 55% of tethered test flights = no flight
    - Javal et al. 2018
  - < 50% took flight in a number of laboratory experiments, esp females
    - (Keena 2018)
- Critical temps
  - 10.2C-egg hatch
  - Temperature developmental model-(Trotter and Keena 2016)
  - Adult emergence in spring after 400-degree-days (10C threshold)
    - (Smith et al. 2004)
  - Dispersal ceases below 15 °C
    - (Keena 2018)
- Habitat preferences
  - Edge preference
    - (Williams et al. 2004)
    - (Shatz et al. 2013)

DED biology and vector life history

DED is vectored by several species of bark beetle: Hylurgopinus rufipes (native), Scolytus multistriatus (introduced-Europe), and Scolytpus schevyrewi (introduced-Asia).

- Surprising lack of dispersal information for above three species
  - (Harwood et al. 2011)
- Dispersal kernel
  - (Harwood et al., 2011)
    - DED vectors
      - Negative exponential kernel of 20 km (15–40 km)
      - Experts estimate max dispersal = 12.88 km
      - Most dispersal within 500 m of host
      - Median dispersal distance of 150 m for a negative square power law function for incorporating radial dispersal
      - Probability of 0.002 for dispersal > 12.88 km
      - Combined beetle and firewood kernel of 3:1 beetle:firewood movement gives a reasonable pattern of spread in early stages of epidemic
- History of DED in UK
  - (Tomlinson and Potter 2010)
- Review of factors influencing flight in bark beetles (Jones et al. 2019)
- Bark beetles (= Scolytinae) contain vectors of DED
  - Flight capacity versus dispersal—distinct
    - Capacity = physiological ability to fly
    - Dispersal = capacity + imapct of external factors (e.g. environment)
    - Long distance dispercal characterized by above canopy flight carried by wind (e.g. Mountain pine beetle dispersal over Rocky Mountains; 30–100 km/day via wind)
- Dispersal distance

- mean for beetles ranges from 500 m to 6 km; max distances can be > 25 km, but this is a long thin tail, bulk of the pop is short distance

- Fat-tailed dispersal kernel needed to capture potential for bark beetles to disperse long distances
- Min temp for flight initiation in bark beetles range from 10.6C–21C; mean = 15.6C
- Dispersal distance
  - Mark recapture-38% pop close to release site, 52% within 400–600 m from release site (Strobel and Lanier 1981)
  - 5–6 km dispersal (Wolfenbarger and Jones 1943)
  - Mark recapture-1 km (Pines and Westwood 2008)
  - 400–600 m dispersal (Wollerman 1979)

SOD biology and vector life history parameters and associated references

There are no known vectors of SOD other than humans but any organism that can move soil is potentially a vector of SOD.(Grünwald et al. 2012, 2019; Kliejunas 2010; Rizzo et al. 2005)

- Dispersal
  - Long range spread of disease through sporangia and chlamydospores, chlamydospores can survive for a week at a constant temperature of 55 °C.
  - Natural dispersal of SOD is by movement of plant material, waterborne and soilborne chlamydospores, and by waterborne, soilborne and wind-blown rain containing sporangia (Rizzo and Garbelotto 2003; Rizzo et al. 2005; Grünwald et al. 2019)
- Dispersal distance
  - Splash dispersal-propagules can travel up to 60 cm above infested surfaces (Kuske 1983).
  - Local spread < 1 km (ecological (Condeso et al. 2007; Ellis et al. 2010) and genetic (Mascheretti et al. 2008, 2009))
  - Most inoculum remains within 10 m of the host (Davidson et al. 2005)
  - Maximum dispersal distance < 8 km during rare storm events (apsnet.org).
  - Number of trees infected was higher on public lands that were open to recreation than on adjacent lands lacking public access and higher human population densities within 50 km increased chances of fungal infection (Cushman et al. 2008).
- Effects of temperature and moisture on growth and sporulation
  - Fungal growth occurs 10–31 °C (Tooley et al. 2009)
  - Exposure to temperatures over 30 °C decreases survival and a few minutes at 40 °C kills the fungus (Browning et al. 2008)
  - Sporangia production occurs over the temperature range of 16–22 °C (Englander et al. 2006)
  - A dew period of as little as 1 h was enough for fungal development but moisture for 24–48 h is required for maximal disease development in the laboratory (Tooley et al. 2009)
  - Most clonal hyphal colonies can survive 24 h exposure to − 5 C and some can withstand − 25 C for 24 h. (Browning et al. 2008).
- Distribution
  - SOD is distributed only in Europe and parts of North America, with three identified clonal lineages (EU1, NA1 and NA2), named for the continent where they were first found, followed by a number indicating the order of discovery (Grünwald et al. 2009)
- Habitat
  - Coastal forest types (Rizzo and Garbelotto 2003; Rizzo et al. 2002), moist and moderate climates (Rizzo et al. 2005).

Appendix 8 (MigClim: Calibration of P_{Disp})-based on source Engler and Guisan (2009)

P_{Disp}, is the colonization probability of a pixel given its distance from a source pixel. To calibrate P_{Disp} we defined a dispersal kernel used to model regular seed dispersal. The kernel was based on the following negative exponential seed dispersal probability distribution function (Eq. 1).

$$P_{seed}(x) = e^{(x-\text{pixelsize})\left(\frac{ln(1-k)}{DispDist}\right)x.\left(\frac{ln(1-k)}{DispDist}\right)} \quad (1)$$

Further simplified in more conventional simple negative exponential form (Eq. 2)

$$P_{seed}(x) = \left((1-k)^{-\frac{pixelsize}{DispDist}} - 1\right).e^{x.\left(\frac{ln(1-k)}{DispDist}\right)} \quad (2)$$

where P*seed* is the probability of a seed reaching distance x ≥ pixelsize, pixelsize is the one-dimensional size of a pixel, DispDist is the dispersal distance reached by the proportion k of the seeds.

Since the surface composed of pixels located at distance j from a source cell increases with distance from that source cell, the probability of a pixel to receive a seed is computed as (Eq. 3)

$$P_{seed}\left(Pixel_j\right) = {}^{P_{seed(x)}}/_{Surface_j} \quad (3)$$

where Surfacej is the number of pixels covered by all pixels belonging to a same distance class. Assuming that the distribution of successful seeds (i.e. seeds leading a pixel to become colonized) is proportional to the overall distribution of seeds (P_{seed}), P_{Disp} is computed as (Eq. 4):

$$P_{Disp}\left(Pixel_j\right) = 1 - \left(1 - P_{seed}\left(Pixel_j\right)\right)^{successfulseeds} \quad (4)$$

where P_{Disp} is the probability of colonisation for a target pixel with distance j from a source pixel and Successful Seeds the number of successful seeds produced by a fully mature source pixel.

References

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J App Ecol 43:1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Anderson RP, Gonzalez I (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: an implementation with Maxent. Ecol Model 222:2796–2811. https://doi.org/10.1016/j.ecolmodel.2011.04.011

Andresen JA, McCullough DG, Potter BE, Koller CN, Bauer LS, Lusch DP, Ramm CW (2001) Effects of winter temperatures on gypsy moth egg masses in the Great Lakes region of the United States. Agric For Meteorol 110(2):85–100. https://doi.org/10.1016/s0168-1923(01)00282-9

Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. J Biogeogr 33:1677–1688. https://doi.org/10.1111/j.1365-2699.2006.01584.x

Bancroft JS, Smith MT (2005) Dispersal and influences on movement for *Anoplophora glabripennis* calculated from individual mark-recapture. Entomol Exp Appl 116:83–92. https://doi.org/10.1111/j.1570-7458.2005.00320.x

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol Evol 3:327–338. https://doi.org/10.1111/j.2041-210x.2011.00172.x

Barbosa FG, Schneck F, Melo AS (2012) Use of ecological niche models to predict the distribution of invasive species: a scientometric analysis. Braz J Biol 72:821–829. https://doi.org/10.1590/s1519-69842012000500007

Barry S, Elith J (2006) Error and uncertainty in habitat models. J Appl Ecol 43:413–423. https://doi.org/10.1111/j.1365-2664.2006.01136.x

Bean WT, Stafford R, Brashares JS (2012) The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. Ecography 35:250–258

Brasier CM, Mehrotra MD (1995) *Ophiostoma himal-ulmi* sp. nov., a new species of Dutch elm disease fungus endemic to the Himalayas. Mycol Res 99:205–215. https://doi.org/10.1016/s0953-7562(09)80887-3

Brown JL (2014) SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods Ecol Evol 5:694–700. https://doi.org/10.1111/2041-210x.12200

Brown JL, Bennett JR, French CM (2017) SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ 5:e4095

Browning M, Englander L, Tooley PW, Berner D (2008) Survival of *Phytophthora ramorum* hyphae after exposure to temperature extremes and various humidities. Mycologia 100(2):236–245

Chapman D, Pescott OL, Roy HE, Tanner R (2019) Improving species distribution models for invasive non-native species with biologically informed pseudo-absence selection. J Biogeogr 46:1029–1040. https://doi.org/10.1111/jbi.13555

Cobos ME, Peterson TA, Barve N, Osorio-Olvera L (2019) kuenm: an R package for detailed development of ecological niche models using Maxent. PeerJ 7:e6281. https://doi.org/10.7717/peerj.6281

Colautti RI, Bailey SA, van Overdijk CDA, Amundsen K, MacIsaac HJ (2006) Characterized and projected costs of nonindigenous species in Canada. Biol Invasions 8:45–59. https://doi.org/10.1007/s10530-005-0236-y

Commonwealth Agricultural Bureaux International (CABI) (2019a) *Anoplophora glabripennis* (Asian longhorned beetle). https://www.cabi.org/isc/datasheet/5557#71A9CFAD-8527-4217-9678-41E0B01EE83A Accessed 26 June 2020

Commonwealth Agricultural Bureaux International (CABI) (2019b) *Lymantria dispar* (gypsy moth). https://www.cabi.org/isc/datasheet/31807#toDistributionMaps Accessed 26 June 2020

Commonwealth Agricultural Bureaux International (CABI) (2019c) *Ophiostoma ulmi* (Dutch elm disease). https://www.cabi.org/isc/datasheet/12165#D451998E−7E25-40C0-8BCB-7F9C88580E02 Accessed 26 June 2020

Commonwealth Agricultural Bureaux International (CABI) (2019d) *Phytophthora ramorum* (sudden oak death (SOD)). https://www.cabi.org/isc/datasheet/40991 Accessed 26 June 2020

Condeso TE, Emiko Condeso T, Meentemeyer RK (2007) Effects of landscape heterogeneity on the emerging forest disease sudden oak death. J Ecol 95:364–375. https://doi.org/10.1111/j.1365-2745.2006.01206.x

Cushman JH, Cushman J, Meentemeyer RK (2008) Multi-scale patterns of human activity and the incidence of an exotic forest pathogen. J Ecol 96:766–776. https://doi.org/10.1111/j.1365-2745.2008.01376.x

Davidson JM, Wickland AC, Patterson HA, Falk KR, Rizzo DM (2005) Transmission of *Phytophthora ramorumin* mixed-evergreen forest in California. Phytopathology 95:587–596. https://doi.org/10.1094/phyto-95-0587

de Andrade AFA, Velazco SJE, De Marco P (2019) Niche mismatches can impair our ability to predict potential invasions. Biol Invasions 21(10):3135–3150

Dimson M, Lynch SC, Gillespie TW (2019) Using biased sampling data to model the distribution of invasive shothole borers in California. Biol Invasions 21(8):2693–2712

Dukes JS, Pontius J, Orwig D, Garnas JR, Rodgers VL, Brazee N, Ayres M et al (2009) Responses of insect pests, pathogens, and invasive plant species to climate change in the forests of northeastern North America: what can we predict? This article is one of a selection of papers from NE Forests 2100: a synthesis of climate change impacts on forests of the Northeastern US and Eastern Canada. Can J For Res 39:231–248. https://doi.org/10.1139/x08-171

Dunn CP (2012) The elms: breeding, conservation, and disease management. Springer, Berlin

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Methods Ecol Evol 1:330–342. https://doi.org/10.1111/j.2041-210x.2010.00036.x

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. Divers Distrib 17:43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Elkinton JS, Liebhold AM (1990) Population dynamics of gypsy moth in North America. Ann Rev Entomol 35:571–596. https://doi.org/10.1146/annurev.en.35.010190.003035

Ellis AM, Václavík T, Meentemeyer RK (2010) When is connectivity important? A case study of the spatial pattern of sudden oak death. Oikos 119:485–493. https://doi.org/10.1111/j.1600-0706.2009.17918.x

Englander L, Browning M, Tooley PW (2006) Growth and sporulation of *Phytophthora ramorum* in vitro in response to temperature and light. Mycologia 98(3):365–373

Engler R, Guisan A (2009) MigClim: predicting plant distribution and dispersal in a changing climate. Divers Distrib 15(4):590–601

Engler R, Hordijk W, Guisan A (2012) The MIGCLIM R package—seamless integration of dispersal constraints into projections of species distribution models. Ecography 35:872–878. https://doi.org/10.1111/j.1600-0587.2012.07608.x

Faccoli M, Gatto P (2016) Analysis of costs and benefits of Asian longhorned beetle eradication in Italy. Forestry 89:301–309. https://doi.org/10.1093/forestry/cpv041

Faccoli M, Favaro R, Smith MT, Wu J (2015) Life history of the Asian longhorn beetle *Anoplophora glabripennis* (Coleoptera Cerambycidae) in southern Europe. Agric For Entomol 17(2):188–196

Favaro R, Wichmann L, Ravn HP, Faccoli M (2015) Spatial spread and infestation risk assessment in the Asian longhorned beetle, *Anoplophora glabripennis*. Entomol Exp Appl 155:95–101. https://doi.org/10.1111/eea.12292

Fournier RE, Turgeon JJ (2017) Surveillance during monitoring phase of an eradication programme against *Anoplophora glabripennis* (Motschulsky) guided by a spatial decision support system. Biol Invasions 19:3013–3035. https://doi.org/10.1007/s10530-017-1505-2

GBIF.org (2017a) GBIF Occurrence Download https://doi.org/10.15468/dl.kzwpyw

GBIF.org (2017b) GBIF Occurrence Download https://doi.org/10.15468/dl.nk4sfs

GBIF.org (2018) GBIF Occurrence Download https://doi.org/10.15468/dl.fnyusm

GBIF.org (2019) GBIF Occurrence Download https://doi.org/10.15468/dl.oq5kxy

Gould SF, Beeton NJ, Harris RMB, Hutchinson MF, Lechner AM, Porfirio LL, Mackey BG (2014) A tool for simulating and communicating uncertainty when modelling species distributions under future climates. Ecol Evol 4(24):4798–4811

Gray DR, William Ravlin F, Braine JA (2001) Diapause in the gypsy moth: a model of inhibition and development. J Insect Physiol 47(2):173–184. https://doi.org/10.1016/s0022-1910(00)00103-7

Groombridge B (1992) The convention on biological diversity. Glob Biodivers. https://doi.org/10.1007/978-94-011-2282-5_35

Grünwald NJ, Goss EM, Ivors K, Garbelotto M, Martin FN, Prospero S, Widmer TL et al (2009) Standardizing the nomenclature for clonal lineages of the sudden oak death pathogen, *Phytophthora ramorum*. Phytopathology 99(7):792–795

Grünwald NJ, Garbelotto M, Goss EM, Heungens K, Prospero S (2012) Emergence of the sudden oak death pathogen *Phytophthora ramorum*. Trends Microbiol 20:131–138. https://doi.org/10.1016/j.tim.2011.12.006

Grünwald NJ, LeBoldus JM, Hamelin RC (2019) Ecology and evolution of the sudden oak death pathogen *Phytophthora ramorum*. Annu Rev Phytopathol 57:301–321. https://doi.org/10.1146/annurev-phyto-082718-100117

Guisan A, Thuiller W, Zimmermann NE (2017) Habitat suitability and distribution models: with applications in R. Cambridge University Press, Cambridge

Haack RA, Hérard F, Sun J, Turgeon JJ (2010) Managing invasive populations of Asian longhorned beetle and citrus longhorned beetle: a worldwide perspective. Annu Rev Entomol 55:521–546

Halvorsen R, Mazzoni S, Bryn A, Bakkestuen V (2015) Opportunities for improved distribution modelling practice via a strict maximum likelihood interpretation of MaxEnt. Ecography 38:172–183. https://doi.org/10.1111/ecog.00565

Halvorsen R, Mazzoni S, Dirksen JW, Næsset E, Gobakken T, Ohlson M (2016) How important are choice of model selection method and spatial autocorrelation of presence data for distribution modelling by MaxEnt? Ecol Model 328:108–118. https://doi.org/10.1016/j.ecolmodel.2016.02.021

Hamelin RC, Roe AD (2019) Genomic biosurveillance of forest invasive alien enemies: a story written in code. Evol Appl. https://doi.org/10.1111/eva.12853

Harwood TD, Tomlinson I, Potter CA, Knight JD (2011) Dutch elm disease revisited: past, present and future management in Great Britain. Plant Pathol 60:545–555. https://doi.org/10.1111/j.1365-3059.2010.02391.x

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25:1965–1978. https://doi.org/10.1002/joc.1276

Hu JF, Angeli S, Schuetz S, Luo YQ, Hajek AE (2009) Ecology and management of exotic and endemic Asian longhorned beetle *Anoplophora glabripennis*. Agric For Entomol 11(4):359–375

Huang D, Haack RA, Zhang R (2011) Does global warming increase establishment rates of invasive alien species? A centurial time series analysis. PLoS ONE 6(9):e24733

Iwaizumi R, Arakawa K, Koshio C (2010) Nocturnal flight activities of the female Asian gypsy moth, *Lymantria dispar* (Linnaeus) (Lepidoptera: Lymantriidae). Appl Entomol Zool 45:121–128. https://doi.org/10.1303/aez.2010.121

Javal M, Roux G, Roques A, Sauvard D (2018) Asian Longhorned Beetle dispersal potential estimated in computer-linked flight mills. J Appl Entomol 142:282–286

Jiménez-Valverde A, Lobo JM, Hortal J (2008) Not as good as they seem: the importance of concepts in species distribution modelling. Divers Distrib 14(6):885–890

Jiménez-Valverde A, Peterson AT, Soberón J, Overton JM, Aragón P, Lobo JM (2011) Use of niche models in invasive species risk assessments. Biol Invasions 13:2785–2797. https://doi.org/10.1007/s10530-011-9963-4

Jones KL, Shegelski VA, Marculis NG, Wijerathna AN, Evenden ML (2019) Factors influencing dispersal by flight in bark beetles (Coleoptera: Curculionidae: Scolytinae): from genes to landscapes. Can J For Res 49(9):1024–1041

Kappel AP, Trotter RT, Keena MA, Rogan J, Williams CA (2017) Mapping of the Asian longhorned beetle's time to maturity and risk to invasion at contiguous United States extent. Biol Invasions 19:1999–2013. https://doi.org/10.1007/s10530-017-1398-0

Keena MA (2018) Factors that influence flight propensity in *Anoplophora glabripennis* (Coleoptera: Cerambycidae). Environ Entomol 47(5):1233–1241

Keena MA, Moore PM (2010) Effects of temperature on *Anoplophora glabripennis* (Coleoptera: Cerambycidae) Larvae and Pupae. Environ Entomol 39:1323–1335. https://doi.org/10.1603/en09369

Keena MA, Wallner WE, Grinberg PS, Cardé RT (2001) Female flight propensity and capability in *Lymantria dispar* (Lepidoptera: Lymantriidae) from Russia, North America, and their reciprocal F1 hybrids. Environ Entomol 30:380–387. https://doi.org/10.1603/0046-225x-30.2.380

Keena MA, Côté M-J, Grinberg PS, Wallner WE (2008) World distribution of female flight and genetic variation in Lymantria dispar (Lepidoptera: Lymantridae). Environ Entomol 37(3):636–649

Kliejunas JT (2010) Sudden oak death and *Phytophthora ramorum*: a summary of the literature. US Department of Agriculture, Albany, CA

Koch F, Yemshanov D, Haack R (2013) Representing uncertainty in a spatial invasion model that incorporates human-mediated dispersal. NeoBiota 18:173–191. https://doi.org/10.3897/neobiota.18.4016

Kumar S, Yee WL, Neven LG (2016) Mapping global potential risk of establishment of *Rhagoletis pomonella* (Diptera: Tephritidae) using MaxEnt and CLIMEX niche models. J Econ Entomol 109(5):2043–2053

Kuske CR (1983) Survival and splash dispersal of *Phytophthora parasitica*, causing dieback of *Rhododendron*. Phytopathology 73:1188. https://doi.org/10.1094/phyto-73-1188

Lafond V, Lingua F, Lumnitz S, Paradis G, Srivastava V, Griess VC (2019) Challenges and opportunities in developing decision support systems for risk assessment and management of forest invasive alien species. Environ Rev (ja) 28:218–245

Leroy B, Delsol R, Hugueny B, Meynard CN, Barhoumi C, Barbet-Massin M, Bellard C (2018) Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. J Biogeogr 45(9):1994–2002

Liang W, Papeş M, Tran L, Grant J, Washington-Allen R, Stewart S, Wiggins G (2018) The effect of pseudo-absence selection method on transferability of species distribution models in the context of non-adaptive niche shift. Ecol Model 388:1–9. https://doi.org/10.1016/j.ecolmodel.2018.09.018

Liebhold AM, Turcáni M, Kamata N (2008) Inference of adult female dispersal from the distribution of gypsy moth egg masses in a Japanese city. Agric For Entomol 10:69–73. https://doi.org/10.1111/j.1461-9563.2007.00359.x

Limbu S, Keena M, Chen F, Cook G, Nadel H, Hoover K (2017) Effects of temperature on development of *Lymantria dispar asiatica* and *Lymantria dispar japonica* (Lepidoptera: Erebidae). Environ Entomol 46(4):1012–1023

Lira-Noriega A, Soberón J, Equihua J (2018) Potential invasion of exotic ambrosia beetles *Xyleborus glabratus* and *Euwallacea* sp. in Mexico: a major threat for native and cultivated forest ecosystems. Sci Rep 8(1):10179

Liu C, White M, Newell G (2011) Measuring and comparing the accuracy of species distribution models with presence–absence data. Ecography 34:232–243. https://doi.org/10.1111/j.1600-0587.2010.06354.x

Mascheretti S, Croucher PJP, Vettraino A, Prospero S, Garbelotto M (2008) Reconstruction of the sudden oak death epidemic in California through microsatellite analysis of the pathogen *Phytophthora ramorum*. Mol Ecol 17(11):2755–2768

Mascheretti S, Croucher PJP, Kozanitas M, Baker L, Garbelotto M (2009) Genetic epidemiology of the sudden oak death pathogen *Phytophthora ramorumin* California. Mol Ecol 18:4577–4590. https://doi.org/10.1111/j.1365-294x.2009.04379.x

Merow C, Smith MJ, Silander JA (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36:1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Montgomery ME, Wallner WE (1988) The gypsy moth. In: Berryman AA (ed) Dynamics of forest insect populations. Springer, Boston, pp 353–375. https://doi.org/10.1007/978-1-4899-0789-9_18

Morales NS, Fernández IC, Baca-González V (2017) MaxEnt's parameter configuration and small samples: are we paying attention to recommendations? A systematic review. PeerJ 5:e3093

Moreno-Amat E, Mateo RG, Nieto-Lugilde D, Morueta-Holme N, Svenning J-C, García-Amorena I (2015) Impact of model complexity on cross-temporal transferability in Maxent species distribution models: an assessment using paleobotanical data. Ecol Model 312:308–317. https://doi.org/10.1016/j.ecolmodel.2015.05.035

Muscarella R, Galante PJ, Soley-Guardia M, Boria RA, Kass JM, Uriarte M, Anderson RP (2014) ENMeval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods Ecol Evol 5:1198–1205. https://doi.org/10.1111/2041-210x.12261

Myers JH, Simberloff D, Kuris AM, Carey JR (2000) Eradication revisited: dealing with exotic species. Trends Ecol Evol 15(8):316–320

Nealis V (2009) Still invasive after all these years: keeping gypsy moth out of British Columbia. For Chron 85:593–603. https://doi.org/10.5558/tfc85593-4

Nelson, H, Grace, P, McBeath A, Stennes B (2009) Estimating the potential returns from developing a national forest pest strategy: the benefits of developing a proactive approach to managing risk

Nobis MP, Normand S (2014) KISSMig—a simple model for R to account for limited migration in analyses of species distributions. Ecography 37:1282–1287. https://doi.org/10.1111/ecog.00930

Osorio-Olvera L, Lira-Noriega A, Soberón J, Peterson AT, Falconi M, Contreras-Díaz RG, Martínez-Meyer E, Barve V, Barve N (2020) ntbox: an r package with graphical user interface for modelling and evaluating multidimensional ecological niches. Methods Ecol Evol. https://doi.org/10.1111/2041-210x.13452

Padalia H, Srivastava V, Kushwaha SPS (2014) Modeling potential invasion range of alien invasive species, *Hyptis suaveolens* (L.) Poit. in India: comparison of MaxEnt and GARP. Ecol Inf 22:36–43. https://doi.org/10.1016/j.ecoinf.2014.04.002

Paini DR, Mwebaze P, Kuhnert PM, Kriticos DJ (2018) Global establishment threat from a major forest pest via international shipping: *Lymantria dispar*. Sci Rep 8(1):13723

Pejchar L, Mooney H (2009) The impact of invasive alien species on ecosystem services and human well-being. Bioinvasions Glob. https://doi.org/10.1093/acprof:oso/9780199560158.003.0012

Peterson AT, Papeş M, Soberón J (2015) Mechanistic and correlative models of ecological niches. Euro J Ecol 1:28–38. https://doi.org/10.1515/eje-2015-0014

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. https://doi.org/10.1111/j.0906-7590.2007.5203.x

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190:231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecol Appl 19:181–197. https://doi.org/10.1890/07-2153.1

Pines I, Westwood R (2008) A mark-recapture technique for the Dutch elm disease vector the native elm bark beetle, *Hylurgopinus rufipes* (Coleoptera: Scolytidae). Arboric Urban For 34(2):116

Potts JM, Elith J (2006) Comparing species abundance models. Ecol Model 199:153–163. https://doi.org/10.1016/j.ecolmodel.2006.05.025

Qiao H, Soberón J, Peterson AT (2015) No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. Methods Ecol Evol 6(10):1126–1136. https://doi.org/10.1111/2041-210x.12397

Qiao H, Feng X, Escobar LE, Townsend Peterson A, Soberón J, Zhu G, Papeş M (2019) An evaluation of transferability of ecological niche models. Ecography 42(3):521–534. https://doi.org/10.1111/ecog.03986

Rizzo DM, Garbelotto M (2003) Sudden oak death: endangering California and Oregon forest ecosystems. Front Ecol Environ 1:197–204. https://doi.org/10.1890/1540-9295(2003)001%5b0197:sodeca%5d2.0.co;2

Rizzo DM, Garbelotto M, Davidson JM, Slaughter GW, Koike ST (2002) *Phytophthora ramorum* as the cause of extensive mortality of *Quercus* spp. and *Lithocarpus densiflorus* in California. Plant Dis 86(3):205–214

Rizzo DM, Garbelotto M, Hansen EM (2005) *Phytophthora ramorum*: integrative research and management of an emerging pathogen in California and Oregon forests. Ann Rev Phytopathol 43:309–335. https://doi.org/10.1146/annurev.phyto.42.040803.140418

Rodda GH, Jarnevich CS, Reed RN (2011) Challenges in identifying sites climatically matched to the native ranges of animal invaders. PLoS ONE 6(2):e14670

Sawyer A (2007, March) Incidence of Asian longhorned beetle infestation among treated trees in New York. In: Proceedings of the 2006 Emerald Ash Borer and Asian Longhorned Beetle Research and Technology Development Review Meeting. United States Department of Agriculture Forest Service, Forest Health Enterprise Team, Morgantown, WV. FHTET-2007-04, pp 106–107

Sawyer AJ, Panagakos WS, Horner AE, Freeman, KJ (2011) Asian longhorned beetle, over the river and through the woods: habitat-dependent population spread. In: McManus KA, Gottschalk KW (eds) 2010. Proceedings. 21st US Department of Agriculture interagency research forum on invasive species 2010; 2010 January 12–15; Annapolis, MD. Gen. Tech. Rep. NRS-P-75. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station, pp 52–54

Schaefer PW, Strothkamp KG (2014) Mass flights of *Lymantria dispar japonica* and *Lymantria Mathura* (Erebidae: Lymantriinae) to commercial lighting, with notes on female viability and fecundity. J Lep Soc 68:124–129. https://doi.org/10.18473/lepi.v68i2.a5

Senay SD, Worner SP (2019) Multi-scenario species distribution modeling. Insects. https://doi.org/10.3390/insects10030065

Shatz AJ, Rogan J, Sangermano F, Ogneva-Himmelberger Y, Chen H (2013) Characterizing the potential distribution of the invasive Asian longhorned beetle (*Anoplophora glabripennis*) in Worcester County, Massachusetts. Appl Geol 45:259–268. https://doi.org/10.1016/j.apgeog.2013.10.002

Shcheglovitova M, Anderson RP (2013) Estimating optimal complexity for ecological niche models: a jackknife approach for species with small sample sizes. Ecol Model 269:9–17. https://doi.org/10.1016/j.ecolmodel.2013.08.011

Smith MT, Bancroft J, Li G, Gao R, Teale S (2001) Dispersal of *Anoplophora glabripennis* (Cerambycidae). Environ Entomol 30:1036–1040. https://doi.org/10.1603/0046-225x-30.6.1036

Smith MT, Tobin PC, Bancroft J, Li G, Gao R (2004) Dispersal and spatiotemporal dynamics of Asian Longhorned Beetle (Coleoptera: Cerambycidae) in China. Environ Entomol 33:435–442. https://doi.org/10.1603/0046-225x-33.2.435

Socioeconomic Data and Applications Center|SEDAC (n.d.) https://sedac.ciesin.columbia.edu/. Accessed 14 Nov 2019

Srivastava V, Lafond V, Griess VC (2019) Species distribution models (SDM): applications, benefits and challenges in invasive species management. CAB Rev 14(020):1–13

Srivastava V, Griess VC, Keena MA (2020) Assessing the Potential Distribution of Asian Gypsy Moth in Canada: a comparison of two Methodological Approaches. Sci Rep 10(1):1–10

Støa B, Halvorsen R, Mazzoni S, Gusarov VI (2018) Sampling bias in presence-only data used for species distribution modelling: theory and methods for detecting sample bias and its effects on models. Sommerfeltia 38:1–53. https://doi.org/10.2478/som-2018-0001

Stohlgren TJ, Schnase JL (2006) Risk analysis for biological hazards: what we need to know about invasive species. Risk Anal 26(1):163–173

Stolar J, Nielsen SE (2015) Accounting for spatially biased sampling effort in presence-only species distribution modelling. Divers Distrib 21:595–608. https://doi.org/10.1111/ddi.12279

Straw NA, Tilbury C, Fielding NJ, Williams DT, Cull T (2015) Timing and duration of the life cycle of Asian longhorn beetle *Anoplophora glabripennis* (Coleoptera: Cerambycidae) in southern England: life cycle of *A. glabripennis*. Agric For Entomol 17(4):400–411

Straw NA, Fielding NJ, Tilbury C, Williams DT, Cull T (2016) History and development of an isolated outbreak of Asian longhorn beetle *Anoplophora glabripennis* (Coleoptera: Cerambycidae) in southern England: *Anoplophora glabripennis* outbreak history. Agric For Entomol 18(3):280–293

Strobel GA, Lanier GN (1981) Dutch elm disease. Sci Am 245:56–66. https://doi.org/10.1038/scientificamerican0881-56

Syfert MM, Smith MJ, Coomes DA (2013) The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. PLoS ONE 8(2):e55158

The R Project for Statistical Computing. www.r-project.org Accessed 14 Nov 2019

Tibshirani R (1996) Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B (Methodol) 58:267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tomlinson I, Potter C (2010) "Too little, too late"? Science, policy and Dutch elm disease in the UK. J Hist Geogr 36:121–131. https://doi.org/10.1016/j.jhg.2009.07.003

Tooley PW, Browning M, Kyde KL, Berner D (2009) Effect of temperature and moisture period on infection of Rhododendron "Cunningham"s White' by *Phytophthora ramorum*. Phytopathology 99(9):1045–1052

Trotter RT III, Keena MA (2016) A variable-instar climate-driven individual beetle-based phenology model for the invasive Asian longhorned beetle (Coleoptera: Cerambycidae). Environ Entomol 45(6):1360–1370

Trotter RT III, Pepper E, Davis K, Vazquez R (2019) Anisotropic dispersal by the Asian longhorned beetle (*Anoplophora glabripennis*): quantifying spatial risk and eradication effort with limited biological data. Biol Invasions 21(4):1179–1195

Tukey JW (1958) A smooth invertibility theorem. Ann Math Stat 29:581–584. https://doi.org/10.1214/aoms/1177706635

Turgeon JJ, Orr M, Grant C, Wu Y, Gasman B (2015) Decade-old satellite infestation of *Anoplophora glabripennis* Motschulsky (Coleoptera: Cerambycidae) found in Ontario, Canada outside regulated area of founder population. Coleopt Bull 69:674–678. https://doi.org/10.1649/0010-065x-69.4.674

Uden DR, Allen C, Angeler DG, Corral L, Fricke KA (2015) Adaptive invasive species distribution models: a framework for modeling incipient invasions. Biol Invasions 17:2831–2850

Václavík T, Meentemeyer RK (2009) Invasive species distribution modeling (iSDM): are absence data and dispersal constraints needed to predict actual distributions? Ecol Model 220:3248–3258. https://doi.org/10.1016/j.ecolmodel.2009.08.013

Václavík T, Meentemeyer RK (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. Divers Distrib 18(1):73–83. https://doi.org/10.1111/j.1472-4642.2011.00854.x

Venette RC, Kriticos DJ, Magarey RD, Koch FH, Baker RHA, Worner SP, Pedlar J et al (2010) Pest risk maps for invasive alien species: a roadmap for improvement. Bioscience 60:349–362. https://doi.org/10.1525/bio.2010.60.5.5

Vilà M, Hulme PE (2017) Non-native species, ecosystem services, and human well-being. In: Vilà M, Hulme P (eds) Impact of biological invasions on ecosystem services. Springer, Cham. https://doi.org/10.1007/978-3-319-45121-3_1

Wang T, Hamann A, Spittlehouse DL, Murdock TQ (2012) ClimateWNA—high-resolution spatial climate data for

Western North America. J Appl Meteorol Climatol 51:16–29. https://doi.org/10.1175/jamc-d-11-043.1

Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol Appl 21(2):335–342

Warren DL, Wright AN, Seifert SN, Shaffer BH (2014) Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. Divers Distrib 20:334–343. https://doi.org/10.1111/ddi.12160

Webber JF (1990) Relative effectiveness of *Scolytus scolytus*, *S. multistriatus* and *S. kirschi* as vectors of Dutch elm disease. For Pathol 20:184–192. https://doi.org/10.1111/j.1439-0329.1990.tb01129.x

Williams DW, Lee H-P, Kim I-K (2004) Distribution and abundance of *Anoplophora glabripennis* (Coleoptera: Cerambycidae) in natural Acer stands in South Korea. Environ Entomol 33:540–545. https://doi.org/10.1603/0046-225x-33.3.540

Wolfenbarger DO, Jones TH (1943) Intensity of attacks by *Scolytus multistriatus* at distances from dispersion and convergence points. J Econ Entomol 36(3):399–402

Wollerman EH (1979) Dispersion and invasion by *Scolytus multistriatus* in response to pheromone. Environ Entomol 8:1–5. https://doi.org/10.1093/ee/8.1.1

Yang F, Luo Y, Shi J (2017) The influence of geographic population, age, and mating status on the flight activity of the Asian gypsy moth *Lymantria dispar* (Lepidoptera: Erebidae) in China. Appl Entomol Zool 52:265–270. https://doi.org/10.1007/s13355-016-0475-7

Zhu GP, Ye Z, Du J, Zhang DL, Zhen YH, Zheng CG, Zhao L, Li M, Bu WJ (2016) Range wide molecular data and niche modeling revealed the Pleistocene history of a global invader (*Halyomorpha halys*). Sci Rep 6:1–10