

Bioclimate envelope model predictions for natural resource management: dealing with uncertainty

Michael S. Mbogga, Xianli Wang and Andreas Hamann*

University of Alberta, Department of Renewable Resources, 739 General Services Building, Edmonton, AB, T6G 2H1, Canada

Summary

1. Bioclimate envelope models are widely used to predict the potential distribution of species under climate change, but they are conceptually also suitable to match policies and practices to anticipated or observed climate change, for example through species choice in reforestation. Projections of bioclimate envelope models, however, come with large uncertainties due to different climate change scenarios, modelling methods and other factors.

2. In this paper we present a novel approach to evaluate uncertainty in model-based recommendations for natural resource management. Rather than evaluating variability in modelling results as a whole, we extract a particular statistic of interest from multiple model runs, e.g. species suitability for a particular reforestation site. Then, this statistic is subjected to analysis of variance, aiming to narrow the range of projections that practitioners need to consider.

3. In four case studies for western Canada we evaluate five sources of uncertainty with two to five treatment levels, including modelling methods, interpolation type for climate data, inclusion of topo-edaphic variables, choice of general circulation models, and choice of emission scenarios. As dependent variables, we evaluate changes to tree species habitat and ecosystem distributions under 144 treatment combinations.

4. For these case studies, we find that the inclusion of topo-edaphic variables as predictors reduces projected habitat shifts by a quarter, and general circulation models had major main effects. Our contrasting modelling approaches primarily contributed to uncertainty through interaction terms with climate change predictions, i.e. the methods behaved differently for particular climate change scenarios (e.g. warm & moist scenarios) but similar for others.

5. *Synthesis and applications.* Partitioning of variance components helps with the interpretation of modelling results and reveals how models can most efficiently be improved. Quantifying variance components for main effects and interactions among sources of uncertainty also offers researchers the opportunity to filter out biologically and statistically unreasonable modelling results, providing practitioners with an improved range of predictions for climate-informed natural resource management.

Key-words: climate change, species distribution models, trembling aspen, western Canada

Introduction

Over the last decade, the reality of global climate change has gained wide acceptance among policy makers and natural resource managers, and the demand for modelling and forecasting climate change impacts on the biosphere is growing. Public sector planners would like accurate forecasts of potential land-use changes, threats to biodiversity, or forest health.

In the private sector, decision makers need to know if their natural resource management strategies and long-term business plans are viable in the face of changing environments. One of the most widely discussed issues involves the choice of species and genotypes in reforestation programmes (e.g. Marris 2009; McKenney, Pedlar & O'Neill 2009). Changing practices and policies for large-scale commercial reforestation programmes is a powerful tool to adapt to anticipated climate change, involving little extra cost in addition to current operations.

A useful class of models also referred to as niche models or species distribution models, to guide species choice under

*Correspondence author. E-mail: andreas.hamann@ualberta.ca

observed and anticipated climate change, are bioclimate envelope models. Bioclimate envelope models are simplistic in that they usually do not model demographic or any other ecological processes. Instead, they correlate species census data with environmental predictor variables using a wide range of statistical and machine-learning methods, e.g. reviewed by Guisan & Zimmermann (2000). The limitations and weaknesses of the bioclimate envelope model approach have been thoroughly discussed (Pearson & Dawson 2003; Hampe 2004; Rushton, Ormerod & Kerby 2004; Araujo & Guisan 2006; Guisan *et al.* 2006; Austin 2007; Botkin *et al.* 2007; Thuiller *et al.* 2008). However, some of the most important limitations of bioclimate envelope models do not apply when they are used to match management practices with anticipated climate conditions. Unlike natural species populations, management practices can 'migrate' as rapidly as bioclimate envelope model results suggest. In plantation forestry, seeds are already being moved considerable distances from source to planting locations under normal management, and competition and species interactions can be controlled through spacing of plantations and choice of planting stock.

To guide species choice in reforestation, simple models based on the realized niche space may be preferable to difficult-to-obtain empirical data on species tolerance to climate change. For example, the results of a reciprocal transplant experiment to determine growth across the fundamental niche of lodgepole pine, *Pinus contorta* Douglas ex Loudon, showed that the species may grow well under projected climate warming in many areas as long as there are no moisture limitations (Wang *et al.* 2006b; O'Neill, Hamann & Wang 2008). However, warm and moist growing conditions can also lead to severe needle cast caused by the native fungus *Dothistroma pini* Hulbary (Woods, Coates & Hamann 2005). This biotic interaction therefore reduces the fundamental niche space of lodgepole pine to a more restricted realized niche. A judicious recommendation for reforestation under climate change should therefore exclude warm and wet climate conditions, i.e. a conservative approach to species choice for reforestation should be guided by projections of the realized niche, not the fundamental niche. While this approach may possibly forgo some potential gains in tree growth due to climate change, it is less risky and corresponds to the widely adopted reforestation policy of not planting species outside their observed range. The same principle applies to other applications, such as ecosystem restoration, selection of protected areas, or assisted migration.

Before bioclimate envelope models can be used in practical applications, they need to be validated. Two aspects, model accuracy and robustness to small changes in model parameters, are helpful to evaluate the reliability of predictions (Botkin *et al.* 2007). In a previous study, we evaluated model accuracy (Hamann & Wang 2006), using an independent validation approach by projecting habitat to new geographic regions according to Araujo *et al.* (2005). Here, we focus on the second aspect, uncertainty in model projections due to data quality, modelling approach, and model parameters. A thoroughly investigated source of uncertainty is the choice of modelling method with conclusions ranging from a fair degree of

model consensus to very pessimistic assessments (e.g. Thuiller *et al.* 2004; Hijmans & Graham 2006; Lawler *et al.* 2006; Pearson *et al.* 2006; Araujo & New 2007). A second important aspect is the choice of climate change scenarios (Bakkenes, Eickhout & Alkemade 2006; Beaumont *et al.* 2007; Beaumont, Hughes & Pitman 2008; Iverson *et al.* 2008). Further, the type and quality of predictor variables as well as biological census data has shown considerable effects on modelling results (Beaumont, Hughes & Poulsen 2005; Taverna, Urban & McDonald 2005; Coudun *et al.* 2006; Guisan *et al.* 2007a; Luoto, Virkkala & Heikkinen 2007; Luoto & Heikkinen 2008). To avoid the danger that minor sources of uncertainty are reported and major sources of uncertainty are ignored, as many factors as possible should be considered in sensitivity analysis (Botkin *et al.* 2007). Examples of such efforts include studies by Kadmon, Farber & Danin (2003), Guisan *et al.* (2007b) and Diniz *et al.* (2009).

In this paper we present an approach to evaluating uncertainty in bioclimate envelope model predictions that yields valuable results for practitioners. Rather than evaluating variability in modelling results as a whole, we extract a particular statistic of interest from multiple model runs, e.g. species suitability for a particular reforestation site. Then, this statistic is subjected to analysis of variance, where sources of uncertainty are represented as treatments in a complete factorial design. We evaluate five sources of uncertainty with two to five treatment levels, including modelling methods, interpolation type for climate data, inclusion of topo-edaphic variables, choice of general circulation models, and choice of emission scenarios. As dependent variables we use an ecosystem class variable for a more general evaluation of results from sensitivity analysis. To illustrate a practical application of the modelling results, we also predict suitable habitat of aspen, *Populus tremuloides* Michaux., an important forestry species in western Canada.

Materials and methods

BIOCLIMATE ENVELOPE MODELLING

To assess the effect of modelling method on climate envelope predictions, we included two contrasting modelling techniques, discriminant analysis implemented by PROC DISCRIM of the SAS statistical package (SAS Institute 2008) and classification tree analysis implemented by the RandomForest software package (Breiman 2001) for the R programming environment (R Development Core Team 2008). RandomForest grows multiple classification trees from bootstrap samples of the training data (for this study we used 200 samples) and determines the predicted class by majority vote over all classification trees. Predictions of a class variable with discriminant analysis are based on a reduced set of independent canonical discriminant functions of the original variables to remove multi-collinearity (Hamann & Wang 2006). The approach is similar to using Mahalanobis distances, which uses principal component analysis to remove collinearity (e.g. Farber & Kadmon 2003). Mahalanobis distances to the mean vector of a class are equivalent to Fisher's discriminant functions (Kshirsagar & Arseven 1975).

The key difference between the two methods is that the scaling of the predictor variables matters in discriminant analysis where the classification is ultimately based on a Euclidean distance. In contrast,

the scaling of the predictor variables is irrelevant for classification trees. Log-transformation of a predictor variable, for example, simply results in different threshold values at tree nodes, but the binary decision trees and the predictions remain the same. While these two methods represent contrasting modelling approaches, we cannot evaluate a full range of predictive models in this study because many widely used methods require 'probability of presence' as a dependent variable and cannot predict an ecosystem class variable.

We used mapped ecosystems for western Canada and the USA, rasterized at a resolution of 1 km, as the dependent class variable. From each of approximately 400 mapped ecosystems, 100 grid cells were randomly sampled to be used as training data for classification tree and discriminant analysis. For British Columbia we used the 'Variant' level of the Biogeoclimatic Ecological Classification system version 4 (Meidinger & Pojar 1991). In Alberta, we used the 'Seedzone' level of Natural Regions and Subregions System 2005 release (NRC 2006). 'Ecodistricts' of the National Ecological Framework for Canada were used for Saskatchewan and Manitoba (Selby & Santry 1996), and 'Level 4' delineation of the US Ecoregion System were used for the area west of 100° longitude and north of 42° latitude (EPA 2007).

As in Hamann & Wang (2006), we aggregated predictions of the ecosystem modelling units at a higher hierarchical level of 'Ecoregions' for reporting, and we inferred species distribution maps from known species frequencies for ecosystems. This approach has some disadvantages (e.g. spatial autocorrelations in the ecosystem response variables requires a different approach to model validation), and a community-based modelling approach may restrict individual species response to climate change (Baselga & Araujo 2009). However, the implementation of a community-based modelling approach that we used in this study has been shown to result in individualistic species response and reliable predictions of species range limits and outlying populations far beyond data coverage (Hamann & Wang 2006). The approach also has the practical advantage that the underlying modelling units are used as a framework for natural resource management. Predicted ecosystem or seedzone units directly suggest a set of management practices for anticipated future climates.

PREDICTOR VARIABLES

We included two climate datasets based on thin plate spline interpolation (Rehfeldt 2006) and generated with the Parameter-elevation Regressions of Independent Slopes Model (PRISM) (Daly *et al.* 2008). Both datasets were based on climate normal data observed at weather stations for the 1961–1990 period covering the US and Canada. These interpolated climate surfaces are near identical in areas with good weather station coverage, but diverge significantly in their estimates of climate values for mountainous areas and northern latitudes. The most prominent differences are estimates of seasonal temperatures north of 55° latitude with differences of up to 6 °C, and precipitation estimates in high elevation mountainous regions that can exceed a 50% difference in seasonal precipitation values. From both datasets we calculated or estimated biologically relevant climate variables for modelling according to Wang *et al.* (2006a). These variables were mean annual temperature, mean warmest month temperature (July), mean coldest month temperature (January), continentality (difference between mean January and mean July temperature), mean annual precipitation, mean summer precipitation (May to September), annual heat moisture index, summer heat moisture index, number of forest free days, chilling degree days below 0 °C, growing degree days above 5 °C, and extreme minimum temperature.

The IPCC (2007) recommends that climate change projections for different emission scenarios and from different general circulation models should be treated with equal probability, and ideally a full range of climate projections should be used in predictive biological models to reflect uncertainty in projections. We therefore use the four major SRES emission and population growth scenario families (A1FI, A2, B1, B2) and implementations of these scenarios by five modelling groups (CGCM2, Canada; HADCM3, UK; ECHAM4, Europe; CSIRO2, Australia; and PCM, USA). We only use projections for the 2041–2070 normal periods, hereafter referred to as the 2050s. Interpolated anomalies of climate change projections from various general circulation models were added as deviations from the 1961–1990 normal period to the 1 km resolution baseline climate datasets according to Mbogga, Hamann & Wang (2009), using a software package that we make freely available (<http://www.ualberta.ca/~ahamann/climate.html>).

As a number of ecosystem classes in our study area are primarily defined by bedrock and soil factors, we replicate all model runs including a set of static, topo-edaphic predictor variables in addition to climate variables. We use a relative radiation index as a proxy for exposure due to slope and aspect, and a topographic convergence index as a proxy of water availability. The relative radiation index was generated for our custom digital elevation model according to Pierce, Lookingbill & Urban (2005). This index is an estimate of the amount of solar radiation received as a function of sun angle, slope, aspect and shadowing by adjacent topography. A compound topographic index to describe the effect of soil water accumulation resulting from topography was calculated according to Gessler *et al.* (1995). This index accounts for slope and the upstream contributing area per unit width of the perpendicularly oriented down-slope water flow. In addition to these topographic indices, we use soil descriptors that are available from the International Geosphere-Biosphere Programme at relatively low resolution of 5 arcminutes or approximately 10 km (GSDT 2000). This data was joined to the 1 km master dataset without manipulations except for reprojections and gap-filling. The variables include soil-carbon density (kg/m²), total nitrogen density (g/m²), field capacity (mm), wilting point (mm), profile available water capacity (mm), and bulk density (g/cm³).

SENSITIVITY ANALYSIS

The modelling effort was organized in a factorial experimental design with multiple treatment levels (Table 1) and resulted in 144 projections of approximately 400 ecosystem climate niches for the 2050s (two baseline climate datasets × 2 modelling methods × 2 sets of predictor variables × 5 GCMs and × 4 emission scenarios, minus two GCM-emission scenario combinations that were not available: ECHAM4-A1FI and ECHAM4-B1). The projections were converted into 12 major macroclimatic ecosystem classes (Fig. 1, legend) for display as maps and for analysis. In addition, the ecosystem projections were converted to maps of potential species habitat of aspen by replacing the ecosystem classes with their corresponding species frequencies. Species frequencies for mapped ecosystems were calculated based on forest inventory plot data for Canada according to Hamann *et al.* (2005) and the data coverage was extended to the USA with the Forest Inventory and Analysis database (Bechtold & Patterson 2005).

These projected biome and aspen habitat maps were the basis for queries that we performed on data tables with PROC UNIVARIATE in SAS (SAS Institute 2008). Data tables were arranged so that rows represent 1 km² grid cells of predicted maps, and columns represent multiple model projections. A query consisted of conditional

Table 1. Factorial experimental design to determine which factors and interactions contribute most to the uncertainty in bioclimate envelope model projections

Treatments and treatment levels
(1) Predictor variables (2 levels)
(1a) 12 Climate variables
(1b) 12 Climate variables and 8 topo-edaphic variables
(2) Modelling method (2 levels)
(2a) RandomForest classification tree analysis
(2b) Mahalanobis distance based discriminant analysis
(3) Climate baseline data (2 levels)
(3a) Thin plate smoothing spline interpolation (ANUSPLIN)
(3b) Interpolation with the Parameter-elevation Regression of Independent Slopes Model (PRISM)
(4) General circulation model (5 levels)
(4a–4e) CGCM2, CSIRO2, ECAHM4, HADCM3, PCM
(5) SRES emission scenario (4 levels)
(5a–5d) A1FI, A2, B1, B2

statements that narrowed the total study area to an ecosystem, species, jurisdiction, or management unit of interest. For the selected data rows, we calculated summary statistics, such as the average elevation of an ecosystem or the area of potential habitat for a species (a count of 1 km² grid cells). All summary statistics were subsequently

expressed as the difference from the reference projection for the 1961–1990 climate normal period.

Tables of summary statistics were then merged and transposed to obtain a new data table where treatments (or sources of uncertainty) were represented by five class variables and summary statistics of changes in projected habitat as dependent variables in columns (one for each query). The data were then subjected to an analysis of variance and estimation of variance components with the restricted maximum likelihood method implemented with PROC VARCOMP/REML of the SAS statistical software package (SAS Institute 2008). Additionally, we used box plots for visual representation of variation due to different sources of uncertainty.

Results and Discussion

We first discuss three case studies that we found educational from a scientific perspective, with data queries carried out at the ecosystem level (Fig. 1). Secondly, we discuss how multiple projections may be used to guide species choice for reforestation using projections of suitable habitat for trembling aspen. Note that we do not display or evaluate the US section of the study area. These ecosystems were included in the training data to cover climate niche space equivalent to what is expected under climate change projections in Canada.

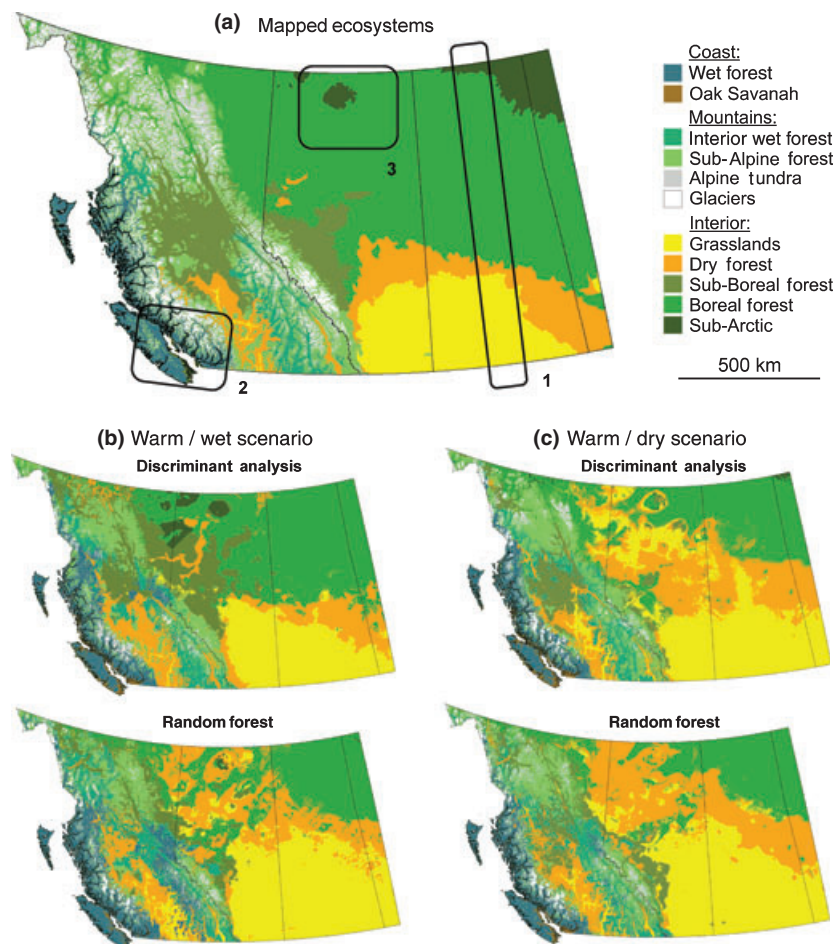
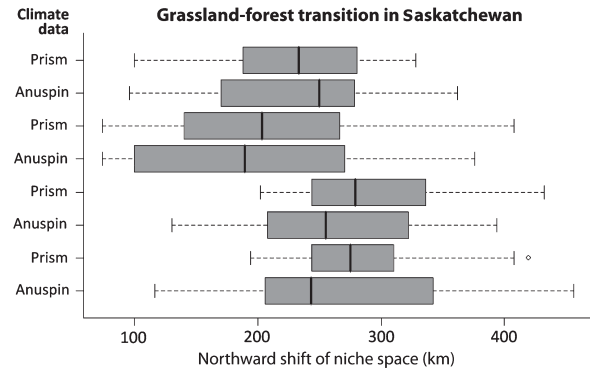


Fig. 1. Current and projected ecosystem climate envelopes for the 2050s. (a) shows the location of queries numbered in the order of discussion in the text, (b) shows different behaviour of modelling methods for a warm and wet scenario (CSIRO-A1) and (c) shows similar model behaviour for a warm and dry scenario (CGCM2-B2).

Fig. 2. Variation in bioclimate envelope modelling results for different datasets and methods (left) and different climate change projections (boxplots). The measured variable is the 90th percentile of latitude of projected grassland ecosystems, reported as northward shift in kilometre relative to the 1960–1990 reference climate projection.

Predictor variables	Modeling method
Climate & topo-edaphic	Random forest
	Discrim. analysis
Climate only	Random forest
	Discrim. analysis



QUERY 1. GRASSLAND–FOREST TRANSITION IN SASKACHEWAN

This first query evaluates the northward shift of the grassland climate envelope between 105 and 107° longitude. The shift under climate change scenarios for the 2050s is expressed in kilometres relative to the 1961–1990 reference projection, and is measured as the location of the 90th percentile of grid cells. By using the location of a percentile rather than the most northern grid cell of the grassland climate envelope we obtain a more robust estimate for the location of its northern boundary. The northward expansion of the grassland climate

Table 2. Variance components corresponding to sources of uncertainty and their interactions. The location of queries are shown in Figs 1 and 6

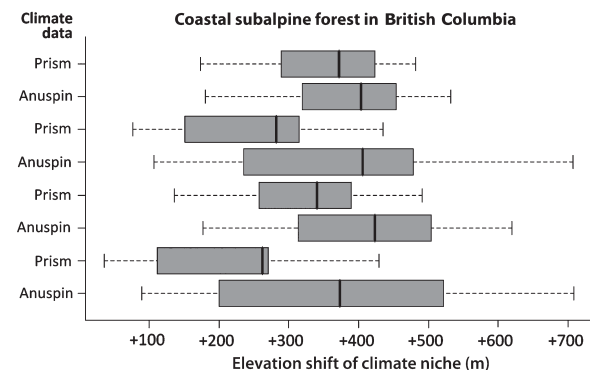
Treatments	Query 1 (SK grasslands)	Query 2 (BC mountains)	Query 3 (AB boreal)	Query 4 (AB FMA)
Main effects				
Predictor variables	15%	0%	20%	1%
Modelling method (MM)	0%	3%	2%	0%
Climate baseline data	1%	11%	7%	0%
General circulation model (GCM)	0%	43%	0%	24%
Emission scenario (SRES)	2%	7%	21%	11%
Interactions				
GCM × MM	42%	11%	31%	15%
GCM × SRES	25%	12%	12%	11%
Other	15%	13%	7%	38%

envelope in Saskatchewan depends largely on whether or not topo-edaphic variables are included as predictor variables and climate change projections indicated by the interquartile range are a large contributor to uncertainty (Fig. 2). A formal analysis of variance reveals another dimension to the modelling results. We already recognized that topo-edaphic variables are an important factor, accounting for approximately 15% of the variance (Table 2), but surprisingly there are no main effects of GCM and SRES emission scenarios. Climate projections only appear in interaction terms, mainly with modelling methods. Under RandomForest, the results for warm and wet scenarios are comparable to a dry scenario, while under discriminant analysis, dry and wet scenarios have very different outcomes (e.g. compare Fig 1b, c).

In this case, we think that the discriminant analysis based approach provides a biologically more plausible result (increased precipitation compensates for increased temperature). RandomForest either used a fairly high precipitation value for the relevant node in the decision tree or did not use precipitation variables to determine the grassland transition zone at all. This is quite plausible because the latitudinal temperature gradient matches the grassland transition zone very closely. We could therefore dismiss RandomForest-based model runs for this particular query. In this way we can narrow plausible results from 144 projections to a smaller number by examining which factors contribute most to the uncertainty in modelling results, and then excluding biologically improbable or statistically questionable results. A smaller number of plausible model projections will usually also result in a narrower range of projections that practitioners need to consider in developing climate change adaptation strategies.

Predictor variables	Modeling method
Climate & topo-edaphic	Random forest
	Discrim. analysis
Climate only	Random forest
	Discrim. analysis

Fig. 3. Variation in bioclimate envelope modelling results for different datasets and methods (left) and different climate change projections (boxplots). The measured variable is the 50th percentile of elevation of the subalpine forest biome, reported as elevation shift in metres relative to the 1960–1990 reference climate projection.



QUERY 2. COASTAL SUBALPINE FORESTS OF SOUTHERN BRITISH COLUMBIA

This second query evaluates the 50th percentile of elevation for the subalpine forest climate envelope for the coast mountains of southern British Columbia (Fig. 3), representing elevational shifts of the climate envelope for this ecosystem. Contrary to the first example, there is no effect due to including topo-edaphic variables. The soils database we use is too low in resolution to provide meaningful information in mountainous areas. However, the high resolution topographic predictor variables CTI and PRR representing exposure and soil moisture due to slope position and aspect did not contribute to variance in modelling results, indicating that they are not essential to characterize the subalpine ecosystem class at this relatively high-level ecosystem summary.

We primarily chose this query because in this area PRISM and ANUSPLIN baseline climate data are quite different, with the PRISM methodology accounting for orographic lift, rain shadows, and slope aspect when estimating climate variables. What we perceive as a much better baseline climate model for this area (PRISM) results in a smaller climate envelope shift and slightly less variable results. However, it is apparent that the quality of baseline climate models for this region is not critical, accounting only for 11% of the total variation, and further improvement of climate data for this region may not be a worthwhile effort.

Another notable observation in this example comes from a comparison with a previous study, where we reported an envelope shift of the Mountain Hemlock Zone of +418 m in elevation (table 3, Hamann & Wang 2006). This is a sufficiently similar query based on a median scenario for British Columbia, but yielding a relatively high value compared with this study. This discrepancy is explained by the fact that a median scenario for British Columbia was not a median scenario for the south coast. Secondly, we choose a median scenario with respect to mean annual temperature and mean annual precipitation, but these may not be the variables that determine the niche space of interest. Thirdly, a median climate change scenario may not always lead to a median modelling result due to the stochastic nature of most predictive models. Therefore, we

want to stress that practitioners would be ill-advised to act on recommendations that are based on a single or a small number of model runs, i.e. the widely used set of a 'median', a 'pessimistic', and an 'optimistic' scenario.

QUERY 3. NORTHERN BOREAL FORESTS OF ALBERTA

The third example evaluates the count of boreal forest raster cells that are predicted to be within a different biome climate envelope by the 2050s (almost always dry forest or grassland). The changes are expressed as per cent loss of boreal forest climate envelope relative to the projection based on 1961–1990 reference climate (Fig. 4, Table 2). We see a repeat of patterns that we have discussed before. Topo-edaphic variables as predictors have an influence, warm and wet scenarios cause an interaction effect in the GCM-method term that is somewhat less pronounced than in the first example, and we see a relatively small baseline climate influence. As in the previous example, the climate baseline datasets differ substantially for this region. The ANUSPLIN estimates for the northern boreal highlands exceed PRISM estimates by 3 °C in mean annual temperature and up to 6 °C in winter temperature, a difference that is larger than projected climate change. Nevertheless, these discrepancies in baseline climate data account for only a minor portion of the total variance in results (Table 2). Again, it appears that bioclimate envelope modelling techniques are surprisingly robust to how ecosystems or species' ranges are climatically characterized with baseline climate.

This query is an example of very high uncertainty in modelling results for the study area. For both subsets, 'climate only' and 'climate and topo-edaphic' as predictor variables, we see a very large range of possible outcomes (about 10%–70% of boreal climate envelope replacement). If we dismiss the results of the RandomForest model runs for the wet scenarios as previously discussed, the overwhelming source of uncertainty is the different climate projections. The variable results may to some degree reflect the biological systems that were subject to this query. Northern boreal ecosystems receive low precipitation (around 300–450 mm mean annual precipitation) and generally have thin, nutrient-poor and acidic soils. Many areas are water-logged coniferous forests or wetlands, such as

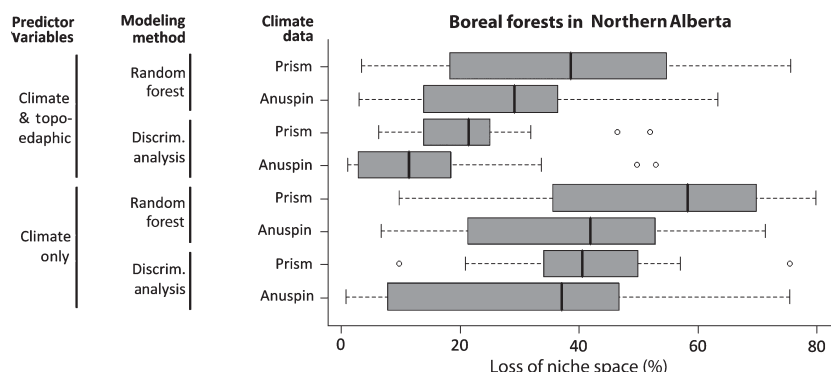


Fig. 4. Variation in bioclimate envelope modelling results for different datasets and methods (left) and different climate change projections (boxplots). The measured variable is the area of projected boreal forest ecosystems, reported as per cent loss relative to the 1960–1990 reference climate projection.

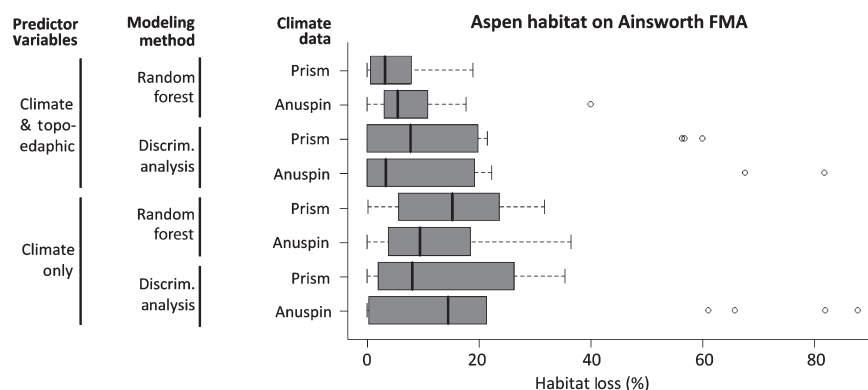


Fig. 5. Variation in bioclimate envelope modelling results for different datasets and methods (left) and different climate change projections (boxplots). The measured variable is the area of projected aspen habitat, reported as per cent loss relative to the 1960–1990 reference climate projection.

sphagnum bogs. There are no obvious biological outcomes if these possibly highly buffered, water-saturated ecosystems are subjected to grassland or dry forest type climates in the future.

QUERY 4. ASPEN HABITAT IN THE AINSWORTH FMA

In many cases, model projections are less variable and easier to interpret than in the previous case study. In the following example we evaluate projections of aspen habitat in the forest management area G16 in Alberta. The area is managed for hardwood supply of a nearby oriented strand board plant, which processes approximately 1 million m³ of hardwood timber annually. We ask if this forest management unit will continue to provide suitable habitat for aspen in the future, and evaluate changes in the count of raster cells with suitable aspen habitat projected for the 2050s

(Fig. 5, Table 2). Projections range from 0–20% loss of habitat and most variation in projections is explained by different climate change scenarios and higher order interactions that are due to erratic behaviour in some model projections for this region (see outliers in Fig. 5). There is no need for filtering these reasonably consistent model results, and we expect only minor changes to hardwood supply from this forest management area by the 2050s, assuming that there are no negative impacts due to maladaptation of local aspen genotypes.

Another way to visualize uncertainty in model projections for aspen habitat over larger geographic areas, are composite maps of all model runs (Fig. 6). Maps of average species frequency indicate where aspen is expected to be a major forest component in the future (Fig. 6c), and counts of presence or absence from all model projections indicate the risk of habitat

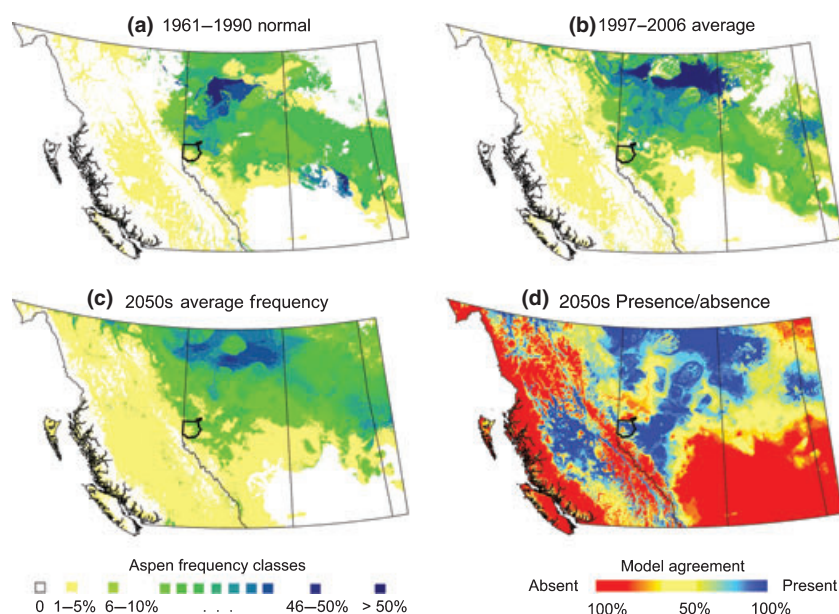


Fig. 6. Predicted aspen frequencies for (a) the 1961–1990 climate normal period, representing the model training data, (b) a recent 10-year average, representing observed climate trends over the last 25 years relative to the climate normal, (c) average aspen frequency projections for the 2050s and (d) the proportion of model runs that predict presence or absence of suitable aspen habitat for the 2050s. The G16 management area for Query 4 is shown as a black outline.

loss (Fig. 6d). These two measures can guide climate-informed forest management. For example, aspen is currently most frequent in the dry mixedwood ecosystem north and northeast of the G16 forest management area of Alberta (Fig. 6a). A majority of model runs, however, project a complete loss of habitat for aspen over much of this area (Fig. 6d). In contrast, moderately high aspen frequencies and low probability of habitat loss are expected along a jet stream driven storm track that originates in the Rocky Mountains and crosses Alberta in a northeast direction. Reforestation or management practices encouraging aspen regeneration should therefore shift to the central mixedwood ecosystems that receive more rainfall.

To further help with confident decisions, we think it is useful to provide model runs based on observed climate trends (Fig. 6b). The 1997–2006 average climate represents an approximately 25-year climate trend relative to the 1961–1990 normal period. We can already observe a shift of high frequency aspen habitat north, and habitat loss along the southern edge of the species distribution in Alberta. This corresponds to drought-related dieback and loss of productivity observed in the parkland ecosystems (Hogg & Bernier 2005; Hogg, Brandt & Michaelian 2008). Thus, the combined information from GCM projections, climate trends that have already materialized, and observed biological response make a strong case for implementing adaptation strategies in the dry mixedwoods and aspen parklands, e.g. reforestation programmes should rely on more drought tolerant species or genotypes in the future, and aspen forestry should concentrate on the moister central mixedwood ecosystems in Alberta.

Conclusions

In this paper we make the case for using bioclimate envelope modelling to match natural resource management practices to anticipated future climates. Because of considerable uncertainty in bioclimate envelope projections, such recommendations should be based on the widest feasible selection of modelling methods, climate change projections, and data sources. If these factors of uncertainty are systematically investigated in a factorial design, large main effects and interaction terms can effectively point to shortcomings in methodology or data quality. This offers an opportunity for the researcher to exclude model runs with biologically or statistically implausible results, and to provide a narrower range of projections that practitioners need to consider in developing climate change adaptation strategies. The task of interpreting a large number of queries is not onerous for the researcher. Even for the varied landscape in this study we only observed a relatively small number of qualitatively different results for variance partitioning. We can therefore also draw some general conclusions about the way in which potential sources of uncertainty contribute to variance in modelling results:

1. Different interpolation techniques for baseline climate did not contribute more than 10% to the uncertainty in modelling results, even though local differences due to interpolation techniques sometimes significantly exceeded climate change projections. It appears that modelling results are surprisingly

robust to how ecosystems or species' ranges are climatically characterized with baseline climate.

2. In several queries, topo-edaphic factors were relevant predictor variables, which had a constraining effect on climate change projections as observed in other studies (e.g. Taverna, Urban & McDonald 2005; Coudun *et al.* 2006; Luoto & Heikkinen 2008). Topo-edaphic variables are often excluded because multi-collinearity among static factors and climate variables can lead to under-estimation of climate change impacts (e.g. Araujo & Guisan 2006). Further, the use of indirect proxies for plant resources (here, CTI and PRR) are not suitable for modelling techniques that rely on a constant statistical relationships over large study areas (Guisan & Zimmermann 2000). However, the two modelling techniques used in this study, account for multi-collinearity and for local interactions of predictor variables. It may therefore be worthwhile to think about whether soil or climate variables are causally related to ecosystem type or species habitat. Again, the objective would be to dismiss a subset of the model projections and narrow the range of predictions that practitioners need to consider.

3. Contrary to other studies (e.g. Hijmans & Graham 2006; Pearson *et al.* 2006), modelling methods were not the largest contributors to uncertainty. However, as we only employ two modelling approaches, we need to be careful in drawing general conclusions. In a recent paper that compares a larger range of methods using a similar variance partitioning approach, Diniz *et al.* (2009) found that modelling methods account for most of the variance while interactions account for approximately 15% in overall species turnover.

4. In this study, general circulation models and their interactions with emission scenarios and modelling methods were the largest contributors to uncertainty. In this situation, a valuable check before implementing adaptation strategies is to analyse locally observed climate trends. We showed that model projections, observed climate trends, and observed biological impacts can make a strong case for changing current management practices. Otherwise, we propose that bioclimate envelope model projections should be used to guide management changes on a moderate scale, e.g. using different species or genotypes for reforestation on 5% of the harvested land base. Over the next decades the success or failure of these changes will provide invaluable empirical data to complement guidance from imperfect models.

Acknowledgements

Advice and help with using RandomForest by Tongli Wang is greatly appreciated. For provision of databases and help with data preparation we thank Todd Schroeder. Alberta Sustainable Development provided additional databases and Laura Khel helped with data preparation. Funding was provided by an NSERC/Industry Collaborative Development Grant CRDPJ 349100-06. We thank Alberta-Pacific Forest Industries, Ainsworth Engineered Canada LP, Daishowa-Marubeni International Ltd., Western Boreal Aspen Corporation, and Weyerhaeuser Company Ltd. for their financial and in-kind support.

References

- Araujo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677–1688.

- Araujo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42–47.
- Araujo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005) Validation of species-climate impact models under climate change. *Global Change Biology*, **11**, 1504–1513.
- Austin, M. (2007) Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological Modelling*, **200**, 1–19.
- Bakkenes, M., Eickhout, B. & Alkemade, R. (2006) Impacts of different climate stabilisation scenarios on plant species in Europe. *Global Environmental Change-Human and Policy Dimensions*, **16**, 19–28.
- Baselga, A. & Araujo, M.B. (2009) Individualistic vs community modelling of species distributions under climate change. *Ecography*, **32**, 55–65.
- Beaumont, L.J., Hughes, L. & Pitman, A.J. (2008) Why is the choice of future climate scenarios for species distribution modelling important? *Ecology Letters*, **11**, 1135–1146.
- Beaumont, L.J., Hughes, L. & Poulsen, M. (2005) Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling*, **186**, 250–269.
- Beaumont, L.J., Pitman, A.J., Poulsen, M. & Hughes, L. (2007) Where will species go? Incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, **13**, 1368–1385.
- Bechtold, W.A. & Patterson, P.L. (2005) *The Enhanced Forest Inventory and Analysis Program – National Sampling Design and Estimation Procedures*. General Technical Report SRS-80, U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC.
- Botkin, D.B., Saxe, H., Araujo, M.B., Betts, R.H., Bradshaw, T., Cedhagen, T.P. *et al.* (2007) Forecasting the effects of global warming on biodiversity. *BioScience*, **57**, 227–236.
- Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5–32.
- Coudun, C., Gegout, J.C., Piedallu, C. & Rameau, J.C. (2006) Soil nutritional factors improve models of plant species distribution: an illustration with *Acer campestre* (L.) in France. *Journal of Biogeography*, **33**, 1750–1763.
- Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J. & Pasteris, P.P. (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology*, **28**, 2031–2064.
- Diniz, J.A.F., Bini, L.M., Rangel, T.F., Loyola, R.D., Hof, C., Nogues-Bravo, D. & Araujo, M.B. (2009) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography*, **32**, 897–906.
- EPA (2007) *Ecoregion Maps and GIS Resources*. U.S. Environmental Protection Agency, Corvallis, OR, Western Ecology Division official website, Available at: <http://www.epa.gov/wed>, accessed 10 May 2008.
- Farber, O. & Kadmon, R. (2003) Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, **160**, 115–130.
- Gessler, P.E., Moore, I.D., McKenzie, N.J. & Ryan, P.J. (1995) Soil-Landscape Modeling and Spatial Prediction of Soil Attributes. *International Journal of Geographical Information Systems*, **9**, 421–432.
- GSDT (2000) Global Soil Data Task Group. Global Gridded Surfaces of Selected Soil Characteristics. IGBP-DIS data set. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, TN, Available at: <http://www.daac.ornl.gov>, accessed 15 June 2007.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J.M.C., Aspinall, R. & Hastie, T. (2006) Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology*, **43**, 386–392.
- Guisan, A., Graham, C.H., Elith, J. & Huettmann, F. (2007a) Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, **13**, 332–340.
- Guisan, A., Zimmermann, N.E., Elith, J., Graham, C.H., Phillips, S. & Peterson, A.T. (2007b) What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecological Monographs*, **77**, 615–630.
- Hamann, A. & Wang, T.L. (2006) Potential effects of climate change on ecosystem and tree species distribution in British Columbia. *Ecology*, **87**, 2773–2786.
- Hamann, A., Smets, P., Yanchuk, A.D. & Aitken, S.N. (2005) An ecogeographic framework for *in situ* conservation of forest trees in British Columbia. *Canadian Journal of Forest Research*, **35**, 2553–2561.
- Hampe, A. (2004) Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography*, **13**, 469–471.
- Hijmans, R.J. & Graham, C.H. (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272–2281.
- Hogg, E.H. & Bernier, P.Y. (2005) Climate change impacts on drought-prone forests in western Canada. *Forestry Chronicle*, **81**, 675–682.
- Hogg, E.H., Brandt, J.P. & Michaelian, M. (2008) Impacts of a regional drought on the productivity, dieback, and biomass of western Canadian aspen forests. *Canadian Journal of Forest Research*, **38**, 1373–1384.
- Iverson, L.R., Prasad, A.M., Matthews, S.N. & Peters, M. (2008) Estimating potential habitat for 134 eastern US tree species under six climate scenarios. *Forest Ecology and Management*, **254**, 390–406.
- Kadmon, R., Farber, O. & Danin, A. (2003) A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, **13**, 853–867.
- Kshirsagar, A.M. & Arseven, E. (1975) A note on the equivalence of two discrimination procedures. *The American Statistician*, **29**, 38–39.
- Lawler, J.J., White, D., Neilson, R.P. & Blaustein, A.R. (2006) Predicting climate-induced range shifts: model differences and model reliability. *Global Change Biology*, **12**, 1568–1584.
- Luoto, M. & Heikkinen, R.K. (2008) Disregarding topographical heterogeneity biases species turnover assessments based on bioclimatic models. *Global Change Biology*, **14**, 483–494.
- Luoto, M., Virkkala, R. & Heikkinen, R.K. (2007) The role of land cover in bioclimatic models depends on spatial resolution. *Global Ecology and Biogeography*, **16**, 34–42.
- Marris, E. (2009) Planting the forest of the future. *Nature*, **459**, 906–908.
- Mbogga, M.S., Hamann, A. & Wang, T. (2009) Historical and projected climate data for natural resource management in western Canada. *Agricultural and Forest Meteorology*, **149**, 881–890.
- McKenney, D., Pedlar, J. & O'Neill, G. (2009) Climate change and forest seed zones: past trends, future prospects and challenges to ponder. *Forestry Chronicle*, **85**, 258–266.
- Meidinger, D.V. & Pojar, J. (1991) *Ecosystems of British Columbia*. Research Branch, Ministry of Forests, Victoria, BC.
- NRC (2006) *Natural regions and Subregions of Alberta*. Natural Regions Committee, Government of Alberta, Alberta Environment, Edmonton, AB.
- O'Neill, G.A., Hamann, A. & Wang, T.L. (2008) Accounting for population variation improves estimates of the impact of climate change on species' growth and distribution. *Journal of Applied Ecology*, **45**, 1040–1049.
- Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.
- Pearson, R.G., Thuiller, W., Araujo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T.P. & Lees, D.C. (2006) Model-based uncertainty in species range prediction. *Journal of Biogeography*, **33**, 1704–1711.
- Pierce, K.B., Lookingbill, T. & Urban, D. (2005) A simple method for estimating potential relative radiation (PRR) for landscape-scale vegetation analysis. *Landscape Ecology*, **20**, 137–147, accessed 1 October 2008.
- R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0, Available at: <http://www.R-project.org>, accessed 1 October 2008.
- Rehfeldt, G.L. (2006) A spline model of climate for the Western United States. Gen. Tech. Rep. RMRS-GTR-165. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO. Available at: <http://forest.moscowfs.wsu.edu/climate>, accessed 15 July 2007.
- Rushton, S.P., Ormerod, S.J. & Kerby, G. (2004) New paradigms for modelling species distributions? *Journal of Applied Ecology*, **41**, 193–200.
- SAS Institute (2008) *SAS/STAT 9.2 User's Guide*. SAS Institute Inc., Cary, NC.
- Selby, C.J. & Santry, M.J. (1996) *A National Ecological Framework for Canada: Data Model, Database and Programs*. Centre for Land and Biological Resources Research, Research Branch, Agriculture and Agri-Food Canada and State of the Environment Directorate, Environment Canada, Ottawa, ON.
- Taverna, K., Urban, D.L. & McDonald, R.I. (2005) Modeling landscape vegetation pattern in response to historic land-use: a hypothesis-driven approach for the North Carolina Piedmont, USA. *Landscape Ecology*, **20**, 689–702.
- Thuiller, W., Araujo, M.B., Pearson, R.G., Whittaker, R.J., Brotons, L. & Lavorel, S. (2004) Uncertainty in predictions of extinction risk. *Nature*, **430**, 34.
- Thuiller, W., Albert, C., Araujo, M.B., Berry, M., Cabeza, A., Guisan, T. *et al.* (2008) Predicting global change impacts on plant species' distributions:

- future challenges. *Perspectives in Plant Ecology Evolution and Systematics*, **9**, 137–152.
- Wang, T., Hamann, A., Spittlehouse, D.L. & Aitken, S.N. (2006a) Development of scale-free climate data for western Canada for use in resource management. *International Journal of Climatology*, **26**, 383–397.
- Wang, T., Hamann, A., Yanchuk, A., O'Neill, G.A. & Aitken, S.N. (2006b) Use of response functions in selecting lodgepole pine populations for future climates. *Global Change Biology*, **12**, 2404–2416.
- Woods, A., Coates, K.D. & Hamann, A. (2005) Is an unprecedented dothistroma needle blight epidemic related to climate change? *BioScience*, **55**, 761–769.

Received 23 October 2009; accepted 6 May 2010

Editor: Jonathan Rhodes