Sensitivity analysis of the tree distribution model PHENOFIT to climatic input characteristics: implications for climate impact assessment

XAVIER MORIN and ISABELLE CHUINE
Centre d’Ecologie Fonctionnelle et Evolutive, Equipe BIOFLUX, CNRS, 1919 route de Mende, 34293 Montpellier cedex 5, France

Abstract
Species distributions are already affected by climate change. Forecasting their long-term evolution requires models with thoroughly assessed validation. Our aim here is to demonstrate that the sensitivity of such models to climate input characteristics may complicate their validation and introduce uncertainties in their predictions. In this study, we conducted a sensitivity analysis of a process-based tree distribution model PHENOFIT to climate input characteristics. This analysis was conducted for two North American trees which differ greatly in their distribution and eight different types of climate input for the historic period which differ in their spatial (local or gridded data) and temporal (daily vs. monthly) resolution as well as their type (locally recorded, extrapolated or simulated by General Circulation Models). We show that the climate data resolution (spatial and temporal) and their type, highly affect the model predictions. The sensitivity analysis also revealed, the importance, for global climate change impact assessment, of (i) the daily variability of temperatures in modeling the biological processes shaping species distribution, (ii) climate data at high latitudes and elevations and (iii) climate data with high spatial resolution.

Keywords: climate change, general circulation model, plant distribution, process-based model, sensitivity analysis

Introduction
Concern about the current climate change is increasing, probably because ecological (Walther et al., 2001; Parmesan & Yohe, 2003; Root et al., 2003), as well as socio-economic (Michaelis, 1994; Keeney & Mc Daniels, 2001; Yohe & Schlesinger, 2002) impacts are more and more perceptible. Climate change is significantly affecting species development timing (Menzel & Fabian, 1999; Walther et al., 2001; Parmesan & Yohe, 2003), physiology (Keeling et al., 1996; Myneni et al., 1997; Cannell et al., 1998; Hughes, 2000), competition (Hughes, 2000) and geographical ranges (Parmesan et al., 1999; Walther et al., 2002; Parmesan & Yohe, 2003) with major consequences on biodiversity, silviculture and agriculture (Nicholis, 1997; Jingyun et al., 2001). The development of accurate predictive models is, therefore, urgent to anticipate harmful consequences and alert stake-holders. Predicting changes in plant distribution and phenology is also a key requirement to forecast accurately future climate because of the vegetation’s feedback on the atmosphere (Betts et al., 1997; de Noblet, 2000).

Climate’s impact on species’ distribution has been for long attested (Holdridge, 1947; Budyko, 1974) and used to reconstitute paleoclimates from information on species past distributions (Guiot, 1994). This relationship has been a basis for numerous models in ecology, and especially in biogeography (Austin, 1985, 1999, 2002; Stephenson, 1990; Prentice et al., 1992; Brzeziecki et al., 1993; Neilson, 1995; Kleidon & Mooney, 2000; Chuine & Beaubien, 2001; Dullinger et al., 2004; Thuiller et al., 2003). A few models have already been used to predict the impact of climate change on species and ecosystems distribution and functioning using General Circulation Models (GCMs) predictions for the 21st century (Huntley et al., 1995; Iverson et al., 1999; Shafer et al., 2001; Bakkenes et al., 2002; Beaumont & Hughes, 2002; Berry et al., 2002; Erasmus et al., 2002; Lehmann...
et al., 2002; Pearson et al., 2002; Peterson et al., 2002; Pearson & Dawson, 2003; Raxworthy et al., 2003; Thuiller, 2003, 2004; Thomas et al., 2004; Thuiller et al., 2004).

Two main kinds of models have been developed to study species distribution and predict species distribution change: models based on the correspondence between species observed distributions and environmental variables (i.e. niche-based models), and models based on the mechanisms involved in the delimitation of species distribution (i.e. process-based models). Providing accurate predictions for the future with such models requires first a robust validation and second a quantification of the different possible sources of variation in their predictions (Rykiel, 1996; Aspinall, 2002; Fleishman et al., 2003).

In the field of ecological modeling, model validation is of primary importance and is most discussed (Loehle, 1983; Oreskes et al., 1994; Rykiel, 1996). Several studies concluded that crossvalidation, (i.e. validation with an independent data set, is the most robust and straightforward validation method for predictive models (Lebreton et al., 1992) and especially for biogeography models (Guisan & Zimmermann, 2000). Validation is easy for process-based models as observed distributions are not used to fit the model parameters and can be used to validate the model. It does not correspond to a crossvalidation senso-stricto as the data used to fit the model parameters and the data used to test the model are of different kind. Crossvalidation is on the contrary very difficult for niche-based models as the observed distribution is used to fit the model parameters. The use of half of the distribution chosen randomly to fit the model and the other half to validate it does not provide a robust validation as both data sets could be strongly auto-correlated. However, the use of one continuous part of the distribution to fit and the rest to validate does not provide an accurate model fit either. Validation is, therefore, an important problem for niche-based models, which is currently driving most attention (Thuiller, 2003, 2004). Nevertheless, few biological models aimed to provide predictions under global change scenarios, either niche-based or process-based are crossvalidated.

Quantifying the different sources of variation in the model predictions has, on the contrary, very rarely been addressed. Recently, Thuiller (2003, 2004) showed that intrinsic sources of variation in niche-based models could be responsible for much higher variance in the model projections than the climate scenarios. In the present study, we focus on the extrinsic sources of variation that may affect model predictions, (i.e. variation in the model’s inputs). Apart from the accuracy of the information on the species distribution that alters niche-based models fit and process-based models validation, a first source of variation in the model predictions for the future is the climate scenarios. Several scenarios actually need to be considered to provide a range of possible outcomes. A second source of variation in the model predictions is the GCMs simulated climate. This source of variation can be partly taken into account using the GCM specific anomalies, (i.e. the differences between simulated and observed historic climate), to correct the simulated future climate. As these anomalies may not be constant through time, the use of raw data should be preferred when the simulated historic climate does not differ substantially from the observed one. Just like for scenarios, several GCMs should also be used to assess climate change impacts. A third source of variation in the model predictions is the climatic variables used in the modeling. This source of variation has to our knowledge never been questioned. If sensitivity analyses are sometimes undertaken on the model parameters, they rarely concerned the input variables (Bachelet et al., 1998).

This last source of variation especially concerns process-based models (vegetation functioning or biogeography), as an increasing number of them use daily climatic data to simulate key biological processes (Neilson, 1995; Kleidon & Mooney, 2000; Chuine & Beaubien, 2001). Gridded daily climate data do not exist at a global scale with a high spatial resolution and for the 20th century, and are only recently available for GCM runs at a low spatial resolution. For these reasons, modelers have been using monthly means, either generating, when required, the daily variability or adapting the models to run with monthly means.

Here we aim to (i) determine the impact of the type (locally recorded, interpolated or simulated), as well as the spatial and temporal resolution of climate data on the predictions of the process-based model, PHENOFT (Chuine & Beaubien, 2001) and (ii) identify the most relevant type of data set to simulate species’ distributions at a continental scale (i.e. our model’s scale). For that purpose, we conducted a sensitivity analysis of the model to the climate input characteristics over 1950–2000. The model originally uses daily climatic data and generates tree species distributions. It has been previously validated for two species, quaking aspen (Populus tremuloides Michx) and sugar maple (Acer saccharum Marsh), using daily climate data from 92 weather stations located all over North America. The comparison of pairs of simulations of sugar maple and quaking aspen distributions obtained with different types of climatic data allowed quantifying the five following effects on the model predictions: temporal resolution of temperature and water balance, extra-
PHENOFIT consists in fitting phenological models depending on the climatic data available. Calibration of equation (Thornthwaite, 1948) on a monthly basis, Taylor, 1972) on a daily basis, or using Thornthwaite–Taylor equation and specific coefficients (Priestley evapotranspiration (PET) is calculated using the Priestley–Taylor equation and specific coefficients (Priestley & Taylor, 1972) on a daily basis, according to the historic period. For each location, climatic data used were daily mean and minimum temperature, precipitation, dew point temperature (temperature to which, while keeping unaltered the current barometric conditions, air becomes saturated of water steam) and solar radiation. Time series available varied between 8 years (Canadian and Mexican stations) to 50 years (United States stations).

Material and methods

PHENOFIT

PHENOFIT is a process-based model that predicts tree species distributions. It relies on the principle that the adaptation of a tree species to the environmental conditions strongly depends on the synchronization of its timing of development to the seasonal variations of climate. The model outputs a probability of presence of an adult individual of the studied species after several years (Fig. 1). This probability is estimated by the fitness of the individual, calculated as the product of its probability to survive until the next reproductive season and to produce viable seeds before the end of the annual cycle (reproductive success). The potential evapotranspiration (PET) is calculated using the Priestley–Taylor equation and specific coefficients (Priestley & Taylor, 1972) on a daily basis, or using Thornthwaite equation (Thornthwaite, 1948) on a monthly basis, depending on the climatic data available. Calibration of PHENOFIT consists in fitting phenological model parameters with phenological observations from several native populations. Other ecological parameters of the model are provided for each species from experimental results provided in the literature. The model is cross-validated by comparing the predicted probability of presence to observed presence. For further details concerning the model see Chuine & Beaubien (2001).

Climatic data

Three types of climate data sets were used:

1. Meteorological stations data: We used daily climatic data from 92 meteorological stations overall North America (Three in Northern Mexico, 70 in the USA – including Alaska – and 19 in Canada) from the NOAA database (www.ncdc.noaa.gov). For each location, climatic data used were daily mean and minimum temperature, precipitation, dew point temperature (temperature to which, while keeping unaltered the current barometric conditions, air becomes saturated of water steam) and solar radiation. Time series available varied between 8 years (Canadian and Mexican stations) to 50 years (United States stations).

2. CRU data (Climatic Research Unit, University of East Anglia, United Kingdom): The CRU TS 2.0 data set provides monthly means of mean and minimum temperature and precipitation, from 1901 to 1995, for a 0.5° × 0.5° grid resolution. These data have been obtained by interpolation of observed climatic data of more than 20000 weather stations all over the world (New et al., 2000). We used the 51 most recent years of this database (1950–2000).

3. GCM data: We used the raw data of three different GCMs (LMD, HADCM3 and ARPEGE) corresponding to the historic period.
LMD data (Laboratoire de Météorologie Dynamique, CNRS, France): The LMD data consist of daily mean and minimum temperature and daily precipitation simulated for a 10-year period, (i.e. with 1990s atmospheric CO₂ concentration). The grid resolution varies with latitude (narrower near the equator and wider near the poles), and is 2.3° × 3.5° on average over North America.

HADCM3 data (Hadley Center for Climate Predictions and Research, United Kingdom): The HADCM3 data consist of daily mean and minimum temperatures and daily precipitation, simulated for a 21-year period (1980–2000). The spatial resolution is 2.5° × 3.75° and is constant.

ARPEGE data (Météo-France, France): The ARPEGE data consist of daily mean and minimum temperatures and daily precipitation, simulated for a 21-year period (1980–2000). The spatial resolution is 2.79° × 2.8125° and is constant.

The data from the meteorological stations were our reference to quantify the different effects. The model simulations were thus conducted with the data from the 92 grid cells of CRU and GCMs corresponding to the 92 meteorological stations. As the GCMs grid cells were much larger than the CRU grid cells, we disaggregated the GCM temperatures at the CRU resolution with an elevation adjustment.

Characteristics of the climatic data sets are summarized in Table 1.

The simulations

The different simulations were conducted as follows:

Simulation 1 (S1): Daily mean and minimum temperature, precipitation, snow, dew point, radiation recorded at the 92 meteorological stations were used. These data allow for the Priestley–Taylor’s PET calculation (Priestley & Taylor, 1972) and provide a daily PET.

Simulation 2 (S2): Daily mean and minimum temperature and monthly precipitation recorded at the 92 stations were used. The PET is calculated with Thornthwaite’s equation (Thornthwaite, 1948) to provide a monthly PET.

All subsequent simulations (S3–S6) used monthly mean precipitation and PET.

Simulation 3 (S3): Observed monthly mean temperatures of the 92 meteorological stations (temperature constant for each day of a month), and corresponding monthly mean precipitation were used.

Simulation 4 (S4): Daily mean and minimum temperatures generated from the monthly means of the 92 meteorological stations, and corresponding monthly mean precipitation were used.

Simulation 5 (S5): Daily mean and minimum temperatures generated from extrapolated monthly means (CRU database), and corresponding monthly mean precipitation were used.

Simulations 6a, 6b and 6c (S6a, S6b, S6c): Raw daily mean and minimum temperatures, and precipitation simulated by the LMD (S6a), HADCM3 (S6b) and ARPEGE (S6c) GCMs for the historic period were used.

Simulations are summarized in Table 2.

The comparison of five pairs of these simulations allowed for the discrimination of the following five effects:

1. monthly (S2) vs. daily (S1) water balance;
2. monthly (S3) vs. daily (S2) mean temperatures;
3. generated (from observed monthly means) (S4) vs. observed (S2) daily temperatures;
4. extrapolated (CRU's data) (S5) vs. observed (S4) monthly mean temperatures;
5. raw GCMs (S6a, S6b, S6c) vs. observed (S2) climate.

Effects tested are summarized in Table 3.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Characteristics of the climatic data sets used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data set</td>
<td>Spatial resolution</td>
</tr>
<tr>
<td>Weather Stations</td>
<td>Local</td>
</tr>
<tr>
<td>CRU</td>
<td>0.5° × 0.5°</td>
</tr>
<tr>
<td>LMD</td>
<td>2.3° × 3.5°</td>
</tr>
<tr>
<td>HADCM3</td>
<td>2.5° × 3.75°</td>
</tr>
<tr>
<td>ARPEGE</td>
<td>2.79° × 2.8125°</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Table 2</th>
<th>Summarized description of the simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulations</td>
<td>Data set</td>
</tr>
<tr>
<td>S1</td>
<td>Weather stations</td>
</tr>
<tr>
<td>S2</td>
<td>Weather stations</td>
</tr>
<tr>
<td>S3</td>
<td>Weather stations</td>
</tr>
<tr>
<td>S4</td>
<td>Weather stations</td>
</tr>
<tr>
<td>S5</td>
<td>CRU</td>
</tr>
<tr>
<td>S6a</td>
<td>Raw LMD</td>
</tr>
<tr>
<td>S6b</td>
<td>Raw HADCM3</td>
</tr>
<tr>
<td>S6c</td>
<td>Raw ARPEGE</td>
</tr>
</tbody>
</table>

*From the monthly means.
Maps and simulated distribution maps is misleading levels of agreement between observed distribution (Leemans, 1992). Indeed the use of $r^2$ requires presence/absence values. As PHENOFIT outputs continuous values between 0 and 1, following the classical methodology (Thuiller, 2003), we used a cut-off threshold to assign the presence or the absence of the species, and chose the threshold that minimized the mean square error. As imposing a cut-off threshold is a source of uncertainty, we also used a threshold-independent index (i.e. the area under the relative operating characteristics curve (AUC)). The greater the AUC the higher the agreement (for more details, see Fielding & Bell, 1997). AUC is ranged from 0.5 to 1, and 0.9 for instance, means very good agreements (Swets, 1988).

However, because the AUC provide a global estimation of the model accuracy, do not depict local discrepancies, and only apply to presence/absence binary data, we also used the simulations’s likelihood calculated as in Chuine & Beaubien (2001).

The different simulations were ranked according to their log-likelihood $L$, $\kappa$ and AUCs. Each of the five effects was quantified by the log likelihood ratio (LLR) of the corresponding paired simulations (i.e. the difference between their log-likelihood). Local differences between paired simulations at each of the 92 locations (LLR) were mapped and interpolated using the inverse distance weighted interpolation method (calculation on the 12 nearest neighbors) of ESRI ArcMap™ 8.3.8.d. Interpolations have been undertaken to help visualizing geographical patterns, but the numerical results presented in the text and in the tables are based on the 92 points of comparison.

Table 3  Summarized description the studied effects

<table>
<thead>
<tr>
<th>Record number</th>
<th>Compared simulations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S2–S1</td>
<td>PET calculation time step</td>
</tr>
<tr>
<td>2</td>
<td>S3–S2</td>
<td>Temperature daily variability</td>
</tr>
<tr>
<td>3</td>
<td>S4–S2</td>
<td>Temperature daily variability</td>
</tr>
<tr>
<td>4</td>
<td>S5–S4</td>
<td>Climate extrapolation (CRU)</td>
</tr>
<tr>
<td>5</td>
<td>S6a–S2</td>
<td>Climate simulation (GCM)</td>
</tr>
<tr>
<td></td>
<td>S6b–S2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S6c–S2</td>
<td></td>
</tr>
</tbody>
</table>

PET, potential evapotranspiration; CRU, Climatic Research Unit; GCM, General Circulation Model.

Method for generating daily temperatures from monthly means

We followed the classical and simple method of generation of daily temperature from monthly means used by several weather generators, (e.g. CLIGEN (Nicks et al., 1995)). Daily values of a given variable of month $i$ of year $j$ were obtained by a random draw from the normal distribution $N(\mu_{i,j}, \sigma_{i,j})$ with $\mu_{i,j}$ the monthly mean of the variable. The standard error $\sigma_{i,j}$ was randomly drawn from the normal distribution $N'(\mu', \sigma')$, where $\mu'$ and $\sigma'$ are, respectively, the mean and the standard error of the monthly standard errors. Random draws were obtained with the Marsaglia et al. (1990) procedure.

Two sets of daily series were generated for simulations 4 (S4a and S4b) and 5 (S5a and S5b) to account for stochasticity effect. Simulations S4 and S5 results are the means of S4a and S4b; and S5a and S5b, respectively.

Species observed distribution

The species observed distributions consisted of digital maps from the US Department of Agriculture Forest Service (http://climchange.cr.usgs.gov/data/atlas/little). These maps were compiled by Little and Critchfield (Critchfield & Little, 1966; Little, 1971, 1976, 1977) based on 10-year field observations.

Predictions’ accuracy and simulations comparison

A usual, yet controversial, statistics for comparing simulated distribution maps to observed distribution maps is the $\kappa$ index (Landis & Koch, 1977; Monserud & Leemans, 1992). Indeed the use of $\kappa$ for quantifying levels of agreement between observed distribution maps and simulated distribution maps is misleading because it breaks the condition of statistical independence between the two distributions. Consequently, the empirical scaling of $\kappa$ (0.74 excellent agreement; 0.60–0.74 very good; 0.40–0.59 fair; <0.40 poor), which is arbitrary, is also misleading (Landis & Koch, 1977). In the special case of process-based models (as it is in the present study), the condition of statistical independence is fulfilled. However, lacking a proper framework for the use of $\kappa$ in our case, we used it to rank the simulations, but not to quantify the effects.

$\kappa$ requires presence/absence values. As PHENOFIT outputs continuous values between 0 and 1, following the classical methodology (Thuiller, 2003), we used a cut-off threshold to assign the presence or the absence of the species, and chose the threshold that minimized the mean square error. As imposing a cut-off threshold is a source of uncertainty, we also used a threshold-independent index (i.e. the area under the relative operating characteristics curve (AUC)). The greater the AUC the higher the agreement (for more details, see Fielding & Bell, 1997). AUC is ranged from 0.5 to 1, and 0.9 for instance, means very good agreements (Swets, 1988).

However, because the AUC provide a global estimation of the model accuracy, do not depict local discrepancies, and only apply to presence/absence binary data, we also used the simulation’s likelihood calculated as in Chuine & Beaubien (2001).

The different simulations were ranked according to their log-likelihood $L$, $\kappa$ and AUCs. Each of the five effects was quantified by the log likelihood ratio (LLR) of the corresponding paired simulations (i.e. the difference between their log-likelihood). Local differences between paired simulations at each of the 92 locations (LLR) were mapped and interpolated using the inverse distance weighted interpolation method (calculation on the 12 nearest neighbors) of ESRI ArcMap™ 8.3.8.d. Interpolations have been undertaken to help visualizing geographical patterns, but the numerical results presented in the text and in the tables are based on the 92 points of comparison.

Results

The log-likelihood $L$, $\kappa$ and AUC of each simulation are shown in Table 4. The LLRs of the paired simulations quantifying the five effects are shown in Table 5, and local differences are shown in Fig. 2.

Differences between the simulations were usually confined within the species distributions or close to their limits except for the GCM simulations (Fig. 2), the estimated fitness outside the present distributions usually remaining low. For sugar maple, the simulations ranked approximately in the same order whatever
the statistics used, rank switch between $L$ and $\kappa$ or AUC (i.e. S1 and S2 or S6a and S6c) concerned simulations with very similar $L$ and $\kappa$ or AUC. For quaking aspen, except simulation S6b (HADCM3 data) that was ranked first by $\kappa$ and AUC but fifth by the likelihood, simulations were ranked approximately the same order by the different statistics.

Water balance calculation (Effect 1, S2 vs. S1)

$L$, $\kappa$ and AUC values show that the use of daily vs. monthly water balance does not affect much the model predictions (Table 4), which are slightly better with a daily water balance except for sugar maple according to $L$. The use of a monthly water balance actually provides slightly better predictions within the distribution of sugar maple, whereas the use of a daily water balance provides better results in the Great Plains (west of the distribution) where water constraints are stronger (Fig. 2, A1). This result is in agreement with the species sensitivity to water stress, which is very weak for quaking aspen and a bit stronger for sugar maple. The use of monthly water balance is thus on average as relevant as a daily water balance at this scale.

According to this result, S2 was taken as the reference simulation and each effect was expressed as a percentage of S2 likelihood.

Daily vs. monthly temperature (Effect 2, S3 vs. S1)

The use of monthly mean temperatures instead of daily temperatures decreased the accuracy of the predictions sharply for both species, but especially for quaking aspen (Table 4). The use of monthly temperatures decreased the simulation likelihood by 50.2% of S2 likelihood for sugar maple and 69.6% for quaking aspen (Table 5). Local differences between the two runs were mainly marked overall the distribution of sugar maple (Fig. 2, A2), and at high latitudes and elevations (Rocky Mountains) of the distribution of quaking aspen (Fig. 2, B2).

Observed vs. generated daily temperatures (Effect 3, S2 vs. S4)

The daily temperature generation effect differed greatly between the two species (Table 5). For quaking aspen the generation of the daily variability of temperatures reduced the likelihood of the simulation by 46.1% of S2 likelihood, discrepancy was mainly localized at high latitudes (Fig. 2, B3). The decrease in likelihood was much less for sugar maple (−28.3% of S2 likelihood), and discrepancy was localized within the observed distribution (Fig. 2, A3).

Observed monthly means vs. extrapolated monthly means (Effect 4, S5 vs. S4)

The use of extrapolated monthly means (CRU data) instead of observed monthly means (used to generate daily data), affected very differently the simulated distribution of the two species. Indeed, it decreased the

Table 4  Log-likelihood ($L$), $\kappa$ and AUC of each simulation (S1–S6c) for sugar maple and quaking aspen

<table>
<thead>
<tr>
<th></th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6a</th>
<th>S6b</th>
<th>S6c</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sugar maple</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L$</td>
<td>−16.75</td>
<td>−16.02</td>
<td>−24.05</td>
<td>−20.54</td>
<td>−27.25</td>
<td>−60.46</td>
<td>−48.68</td>
<td>−67.36</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.83</td>
<td>0.82</td>
<td>0.55</td>
<td>0.64</td>
<td>0.50</td>
<td>0.17</td>
<td>0.40</td>
<td>0.31</td>
</tr>
<tr>
<td>AUC</td>
<td>0.99</td>
<td>0.98</td>
<td>0.94</td>
<td>0.97</td>
<td>0.94</td>
<td>0.78</td>
<td>0.86</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Quaking aspen</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L$</td>
<td>−44.23</td>
<td>−44.71</td>
<td>−75.83</td>
<td>−65.31</td>
<td>−50.99</td>
<td>−136.41</td>
<td>−65.76</td>
<td>−110.30</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.49</td>
<td>0.49</td>
<td>0.31</td>
<td>0.27</td>
<td>0.39</td>
<td>0.20</td>
<td>0.50</td>
<td>0.39</td>
</tr>
<tr>
<td>AUC</td>
<td>0.86</td>
<td>0.85</td>
<td>0.77</td>
<td>0.75</td>
<td>0.85</td>
<td>0.70</td>
<td>0.88</td>
<td>0.78</td>
</tr>
</tbody>
</table>

AUC, area under the relative operating characteristics curve.

Table 5  Log-likelihood ratio (LLR) of the paired simulations for sugar maple and quaking aspen expressed as a percentage of $L_2$

<table>
<thead>
<tr>
<th>Effect</th>
<th>LLR</th>
<th>Sugar maple (%)</th>
<th>Quaking aspen (%)</th>
<th>Average (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$L_2$−$L_1$</td>
<td>4.6</td>
<td>−1.1</td>
<td>1.7</td>
</tr>
<tr>
<td>2</td>
<td>$L_3$−$L_2$</td>
<td>−50.2</td>
<td>−69.6</td>
<td>−59.9</td>
</tr>
<tr>
<td>3</td>
<td>$L_4$−$L_2$</td>
<td>−28.3</td>
<td>−46.1</td>
<td>−37.2</td>
</tr>
<tr>
<td>4</td>
<td>$L_5$−$L_4$</td>
<td>−41.9</td>
<td>32.0</td>
<td>−4.9</td>
</tr>
<tr>
<td>5</td>
<td>$L_{5a}$−$L_2$</td>
<td>−277.5</td>
<td>−205.1</td>
<td>−241.3</td>
</tr>
<tr>
<td>5</td>
<td>$L_{5c}$−$L_2$</td>
<td>−203.9</td>
<td>−47.1</td>
<td>−125.5</td>
</tr>
<tr>
<td>5</td>
<td>$L_{5c}$−$L_2$</td>
<td>−320.6</td>
<td>−146.7</td>
<td>−233.7</td>
</tr>
</tbody>
</table>

1–5 refers to the effects numbered in the Material and methods and Table 3.
likelihood of the simulation for sugar maple (−40.1% of S2 likelihood), whereas it increased the likelihood for quaking aspen (+32.4% of S2 likelihood, Table 5). As for effect 3, discrepancy was mainly at high latitude and elevation for quaking aspen (Fig. 2, B4) and within the species distribution for sugar maple (Fig. 2, A4). At temperate latitudes, the two types of climatic data provided similar results. These results suggest that extrapolated monthly means from the CRU database are relevant for global scale vegetation modeling, especially for the boreal latitudes.

**Observed vs. GCM daily data (Effect 5, S6a vs. S2, S6b vs. S2; S6c vs. S2)**

Results varied a lot among GCMs, and the likelihood of the simulations decreased highly (−200.2% of S2 likelihood on average). However HADCM3 data provided the most accurate predictions among the three GCMs. LMD and ARPEGE data provided similar predictions compared with HADCM3. In both cases (LMD and ARPEGE), discrepancy was localized within and outside the distribution of sugar maple (Fig. 2, A5a, A5c), and within the distribution of quaking aspen (Fig. 2, B5a, B5c). On the contrary, with HADCM3, discrepancy was mostly localized within the species distributions and mainly at high latitudes for quaking aspen (Fig. 2, B5b).

**Discussion**

The process-based model PHENOFIT showed high sensitivity to climatic inputs, especially within the species distributions and at the boundaries. This means that the simulated absence outside the present range is much less affected by climate inputs than the simulated presence within and nearby the present range. This is because of the fact that transforming a simulated absence into a presence requires all fitness component indices to become different from zero, whereas transforming a simulated presence into an absence requires just one fitness component index to become zero. Changes in climate inputs are thus more likely to affect the model predictions within the species range than outside the range.
If we rank the five effects from the strongest to the weakest according to the likelihood ratio, we end up with (1) effect 5, GCM simulated vs. observed daily temperatures (−200.2%), average over simulations S6a, S6b and S6c, (2) effect 2, monthly means vs. daily temperatures (−59.9%), (3) effect 3, generated (from monthly means) vs. observed daily temperatures (−7.2%), (4) effect 4, extrapolated vs. observed monthly means (−4.9%), (5) effect 1, monthly vs. daily water balance and PET (+1.7%) (Table 5). Modeling species distribution with PHENOFIT using daily temperatures generated (effect 3) from extrapolated monthly means (CRU) (effect 4), which has been and still is a common situation, reduces the simulation likelihood by about 70.2% of S2 likelihood for A. saccharum and 14.0% for P. tremuloides (Table 5, sum of independents effects 3 and 4) compared with using recorded daily temperatures.

**PET calculation at global scale**

The differences observed between simulations using monthly PET and daily PET at a continental scale are very small, even for the most water sensitive species, sugar maple. At least at a continental scale, the use of PHENOFIT with monthly PET thus appears as relevant as with daily PET, which is fortunate as daily PET calculation is much more difficult and requires much more climate variables than monthly PET calculation.

**High latitude and elevation climate**

Differences between recorded vs. generated daily temperatures are especially localized at high latitude and high elevation sites as shown by the simulation of quaking aspen distribution (Fig. 2, B3), which is the only one of the two species to occur in these regions. These differences can be explained by the high daily climatic variability that characterizes these regions (Table 6), which makes realistic daily data more difficult to generate from monthly means. Although our results concern only two species, it seems that the generation of temperature daily variability (from monthly means) could affect much more the modeling of boreal species distribution than temperate species distribution. This result is particularly important as boreal regions are expected to show the greatest climatic change by the end of the century (Pastor & Post, 1988; Prentice et al., 1991; IPCC, 2000), and boreal and alpine species and ecosystems are expected to respond quickly to global warming (Prentice et al., 1993; Thorhallsdottir, 1998).

**Importance of the temperature’s daily variability in biological processes**

PHENOFIT calculates annual variables such as dates of leaf unfolding, flowering or fruit maturation, frost damage index. If these variables could be accurately estimated using monthly means instead of daily climatic data, a nonnegligible gain of time in the modeling and computer memory allocation to data storage could be spared. However, whereas monthly means of precipitation can be used instead of daily precipitation without a substantial loss of accuracy (effect 1), this is not the case of temperatures. This is because the processes involved in species niche breadth simulated by the model, such as phenology and frost resistance, are sensitive to temperature extremes and temperature threshold effects.

So far, few ecological models have been using daily climatic data, firstly for technical reasons (as it increases the size of the data set and the model runtime) and secondly because daily data do not exist at the CRUs grid resolution at global scale, but this may change in a near future.

**Use of GCMs in ecological modeling**

The accuracy of the predictions using the GCMs’ data (S6) varied greatly among the GCMs. The LMD simulations (S6a, Fig. 2, A5a, B5a) provided inaccurate predictions throughout the continent for sugar maple, and at temperate latitudes (around 40°) for quaking aspen but accurate predictions at high latitude and elevation for the latter. This was because of an important overestimation of the temperature throughout the year (up to 5°C in summer) above 32° latitude. Warmer temperatures within sugar maple distribution reduced its fitness because of a lack of chilling requirement for budburst, whereas outside the distribution, warmer temperatures increased the simulated fitness by decreased frost damage and increased fruit maturation success. For quaking aspen, warmer temperatures decreased the simulated fitness, because of a lack of chilling requirements, except at high latitudes (above 55°) and elevations where chilling

---

**Table 6** Averaged standard deviation of the daily mean temperatures of February, July and November over the 1994-2001 period for five northern latitudes stations (more than 53° N) (N); five temperate latitudes stations (between 53° and 40° N) (T); five southern latitudes stations (less than 40° N) (S)

<table>
<thead>
<tr>
<th>Stations</th>
<th>February</th>
<th>July</th>
<th>November</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>7.76</td>
<td>2.98</td>
<td>5.98</td>
</tr>
<tr>
<td>T</td>
<td>4.87</td>
<td>2.69</td>
<td>3.69</td>
</tr>
<tr>
<td>S</td>
<td>4.04</td>
<td>1.88</td>
<td>3.38</td>
</tr>
</tbody>
</table>

© 2005 Blackwell Publishing Ltd, Global Change Biology, 11, 1493–1503
requirements were still fulfilled and frost damage decreased.

The HADCM3 simulations provided the most accurate predictions overall for the GCMs either for sugar maple or quaking aspen. The region showing the greatest deviance from S2 is northeastern North America for sugar maple where temperatures are abnormally warm in winter and cold in summer decreasing fitness because of a lack of chilling requirements in winter and higher risk of frost damage in late summer. The simulated distribution of quaking aspen was mainly affected in northwestern North America where temperatures were abnormally cold, increasing frost damage.

The ARPEGE simulations (S6c, Fig. 2, A5c, and B5c) provided inaccurate predictions throughout the continent for sugar maple and in western United States for quaking aspen. This result was because of an over-estimation of winter temperatures in northern latitudes (above 50° in the West and 35° in the East), and an underestimation in southern latitudes compared with observed climate. Colder temperatures reduced sugar maple simulated fitness over its present distribution because of increased frost damage and a later flowering; whereas warmer temperatures increased the species probability of presence outside its distribution because of decreased frost damage and increased fruit maturation success. Warmer temperatures in northern latitudes increased quaking aspen’s probability of presence over most of its distribution because of decreased frost damage and increased fruit maturation success, whereas colder temperatures in western United States increased the probability of presence outside the observed distribution.

These results stress two points. Firstly, the impact of climate change on species distribution will not be easily assessed because climate change varies through space and time and the climate-adapted traits limiting species fitness vary within the species distribution. Secondly, to predict species distribution change under climate change scenarios with PHENOFIT, it would be recommended to use specific anomalies for LMD and ARPEGE, and raw data for HADCM3.

Simulations using GCMs’ data would probably be more accurate if data were available at a higher spatial resolution and for longer time spans (the GCMs time series were much shorter than the weather stations or CRU data, see Table 1). Spatial disaggregation of temperatures according to elevation cannot indeed replace temperatures generated at a high spatial resolution by the GCMs, and short climate series are not adequate to model long lived organisms’ distribution. Regional Circulation Models, which can provide high spatial and temporal resolution data, will thus have a major role to play in the assessment of global climate change’s impact, especially on alpine and boreal ecosystems.

Acknowledgements

The authors thank D. Viner (Climate Research Unit, University of East Anglia, United Kingdom), J. Polcher (Laboratoire de Météorologie Dynamique, CNRS, France), and J.-F. Royer and A. Rascol (Météo-France) for providing us the climate data. Observed distribution maps of quaking aspen and sugar maple are from the USDA Forest Service. The authors are grateful to J. Roy and N. Viévy for their constructive comments, B. Hautdidier for mapping help, R. Pradel for statistical help, and also three anonymous referees whose constructive remarks highly increased the quality of this paper. Support was provided to X. Morin by a Bourse de Docteur Ingénieur du Centre National de la Recherche Scientifique.

References


Cannell MGR, Thornley JHM, Mobbs DC et al. (1998) UK conifer forests may be growing faster in response to increased
N deposition, atmospheric CO₂ and temperature. *Forestry*, 71, 277–296.


