Drivers of fire in the boreal forests: Data constrained design of a prognostic model of burned area for use in dynamic global vegetation models

Cyril Crevoisier,1,2 Elena Shevliakova,3 Manuel Gloor,4 Christian Wirth,5 and Steve Pacala3

Received 21 December 2006; revised 6 August 2007; accepted 19 September 2007; published 28 December 2007.

Boreal regions are an important component of the global carbon cycle because they host large stocks of aboveground and belowground carbon. Since boreal forest evolution is closely related to fire regimes, shifts in climate are likely to induce changes in ecosystems, potentially leading to a large release of carbon and other trace gases to the atmosphere. Prediction of the effect of this potential climate feedback on the Earth system is therefore important and requires the modeling of fire as a climate driven process in dynamic global vegetation models (DGVMs). Here, we develop a new data-based prognostic model, for use in DGVMs, to estimate monthly burned area from four climate (precipitation, temperature, soil water content and relative humidity) and one human-related (road density) predictors for boreal forest. The burned area model is a function of current climatic conditions and is thus responsive to climate change. Model parameters are estimated using a Markov Chain Monte Carlo method applied to on ground observations from the Canadian Large Fire Database. The model is validated against independent observations from three boreal regions: Canada, Alaska and Siberia. Provided realistic climate predictors, the model is able to reproduce the seasonality, intensity and interannual variability of burned area, as well as the location of fire events. In particular, the model simulates well the timing of burning events, with two thirds of the events predicted for the correct month and almost all the rest being predicted 1 month before or after the observed event. The predicted annual burned area is in the range of various current estimates. The estimated annual relative error (standard deviation) is twelve percent in a grid cell, which makes the model suitable to study quantitatively the evolution of burned area with climate.


1. Introduction

Boreal forest is the largest forested area on earth, accounting for more than 25 percent of the world’s forests [Intergovernmental Panel on Climate Change, 2000]. The boreal region also hosts large belowground carbon pools, partially in a frozen state and in form of peat, which, together, account for about 1000 GtC [Gorham, 1991; Zimov et al., 1999]. The boreal region witnessed the largest warming trend on the globe over the last decades [Giorgi, 2006], which may induce substantial changes in the functioning of land ecosystems. Current General Circulation Model simulations indicate that this warming trend will continue in the future decades [Intergovernmental Panel on Climate Change, 2001]. In addition, changes in the regional and temporal patterns and intensity of precipitation are predicted, which may lead to an increase of extreme climate events (droughts and floods). As fire frequency depends strongly on climate and on extreme climate events, an increase of fire occurrence and related carbon emissions to the atmosphere is likely in the future. Such an increased trend is supported by observations made over the last two decades [Stocks et al., 2002; Kajii et al., 2002; Mouillot et al., 2006; Westerling et al., 2006]. However, Amiro et al. [2004] found few trends in weather indices computed from observed fires in Canada. Most of the studies based on various fire indices, which depend on climate variables, have shown an increase of potential fire danger over northern Eurasia in a warming and drying climate [e.g., Stocks et al., 1998; Groisman et al., 2004].
Predicting the effect of climate change on boreal forest through changes in the fire regimes and in their related feedback on climate is therefore needed to accurately predict both future boreal ecosystem composition and climate.

[3] In Earth System Models, carbon release to the atmosphere is commonly computed as the product of burned area, fuel loads and combustion efficiency, integrated over the spatiotemporal resolution of interest [Seiler and Crutzen, 1980; Hao et al., 1990; Pereira et al., 1999; van der Werf et al., 2003; Arora and Boer, 2005; Randerson et al., 2005; Mouillot et al., 2006]. Burned area is generally considered to be the most uncertain of these three parameters [van der Werf et al., 2003; Mouillot and Field, 2005] and is derived either from observations or from model simulations. A set of global or regional observation-derived estimates of burned area are available, mostly from satellite observations, and these have been used to drive fire modules of vegetation models such as the Carnegie-Ames-Stanford Approach (CASA) [van der Werf et al., 2003, 2004; Randerson et al., 2005]. This approach allows one, in principle, to study specific fire events and to link emissions to observations of trace gases in the atmosphere. However, it relies centrally on estimates of burned area derived from satellite observations, which are still uncertain [e.g., Boschetti et al., 2004], especially in the boreal regions. Moreover, this approach is limited to the time period of the observations, which prevents any simulation of future evolution. A model describing the effect of fire on vegetation at a resolution compatible with that of dynamic global vegetation models (DGVM) is thus needed in order to simulate future evolution of fire, and to study the impact of fires on vegetation mortality and the flux of carbon from the terrestrial biosphere to the atmosphere in climate models.

[4] Modeling fire at the coarse spatial (~100 km) and temporal (monthly to annual) scale of climate models requires considering only the major regional characteristics of fire. In addition, the relationship linking fire and its drivers must be sufficiently general to be applicable to every grid cell in the region considered. In particular, DGVMs typically represent plant diversity by a few taxonomic categories, such as plant functional types (PFTs) which are distinguished according to their plant physiological, phonological and physiognomic characteristics. For instance, LM3V, the new vegetation model of the Geophysical Fluid Dynamics Laboratory (GFDL), considers five vegetation types: C3 and C4 grasses, temperate deciduous trees, tropical trees and cold evergreen trees (E. Shvetliakova et al., Carbon cycling under 300 years of land-use changes in the Dynamic Land Model LM3V, submitted to Global Change Biology, 2007, hereinafter referred to as Shvetliakova et al., submitted manuscript, 2007). The biogeography parameterization utilizes total biomass in combination with prevailing climatic conditions to determine the distributions of vegetation types. In this model, the boreal forest comprises cold evergreen trees and temperate deciduous trees. Fire models, which are to be used in DGVMs, thus must capture the overall characteristics of the region, rather than distinguish every species. Finally, the fire model must rely on a reduced set of climate and human-related variables commonly simulated by climate models.

[5] A few models have been designed to model fire in a mechanistic way. The fire model of the Lund-Potsdam-Jena (LPJ) DGVM [Thonicke et al., 2001; Venevsky et al., 2002], also used in the Organizing Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) model [Krinner et al., 2005], estimates the probability of fire as a sigmoid function of the topsoil layer moisture content. It has been developed using data from three limited zones in central Portugal, southern California and northern Australia. Accordingly, its predictive skill is limited in the boreal region [Thonicke et al., 2001]. The fire module of the Canadian Terrestrial Ecosystem Model (CTEM) [Arora and Boer, 2005] estimates the probability of fire occurrence depending on fuel availability (given by CTEM), readiness of fuel to burn (given by a sigmoid function of the root zone soil moisture status, similar to the one used by Thonicke et al. [2001]), and presence of a source ignition (anthropogenic or natural). Thus only one climate variable is used. Model predictions have been compared to on-ground observations made at six locations across the globe, one being located in the boreal forest (northern Canada), for which human influence is assumed to be negligible and where the predicted seasonality of the fire season is shorter than observed. In a different approach, Cardoso et al. [2003] have developed a fire model for Amazonia, which estimates the number of fire pixels in a given grid cell, depending on various climate (total precipitation, minimum precipitation, cloud cover) and human-related (distance to the nearest road and deforestation) variables. They used it to study the evolution of the tropical forest in the context of deforestation and their model is not directly applicable to boreal forest.

[6] A model estimating the amount of burned area and specifically designed for the boreal region, is therefore needed to account for this important region in climate simulations and to link fire regime and its effects on vegetation dynamics. Several extensive data sets of burned area estimates exist for Canada, Alaska and, to a lesser extent, Siberia. They provide an opportunity to design and validate a strongly data-constrained fire model by identifying statistical correlation between burned area patterns and a set of climate and human-related variables.

[7] In this paper, we develop a strongly data-constrained prognostic fire model for the boreal region, which estimates Potential Burned Area (PBA) in a typical grid cell of 2° × 2.5°, on a monthly basis, using four climate (precipitation, temperature, soil water content and relative humidity) and one human-related (road density) variables as predictors. We made the choice of not including fuel availability in the fire model, hence the term potential. This variable will be included once the model is implemented as a fire module of GFDL/LM3V, which will provide information on fuel load and species. Section 2 presents the various data sets of on-ground and remote sensing data that are used in this study for the design and validation of the PBA model. The coupling of fire data to various climate and human-related variables is developed in section 3. Section 4 presents the design of the PBA model using on-ground observation in Canada. Its performance in boreal forests of Canada, Alaska and Siberia are tested against observations in section 5. Conclusions are found in section 6.

2. Data

2.1. Fire

[8] Two independent sources of information on fire are available and used in this study. First, in situ observations
give the most valuable information when an efficient network is in place. For Canada, extensive fire history data are archived in the Large Fire Database (LFDB) [Stocks et al., 2002] (data available at http://fire.cfs.nrcan.gc.ca/research/climate_change/lfdb/lfdb_download_e.htm). The LFDB reports information on all fires larger than 200 ha that occurred in Canada during the 1959–1999 period. It includes fire location, detection date, final burned area and, for some fires, cause and suppression action. The LFDB represents only 3.1% of Canadian fires, but account for about 97% of the total area burned [Stocks, 1991]. For Alaska, another LFDB has been created from fire records held by the Alaska Fire Service. It contains the digitized boundary of Alaskan fires greater than 200 ha recorded for the period 1950 to present [Kasischke et al., 2002] (data available at http://agdc.usgs.gov/data/blm/fire/index.html). In general fires reported after 1970 have more accurate and complete information. Errors affect both LFDBs [Amiro et al., 2001; French et al., 2004]. For instance, unburned regions within a fire are often unmapped, leading to an overestimate of total burned area by a factor of 15%. On the other hand, missing fire records and nonrecording of fires smaller than 200 ha could result in an underestimation of 9–13%. As opposed to North American boreal forests, official data for Siberia cover only protected forests, which represent about 60% of the total forested area [Sukhinin et al., 2004]. This percentage varies from year to year and has continuously decreased in the last years because of financial and logistic constraints [Goldammer and Stocks, 2000]. Thus, together with the fact that official statistics have been purposely manipulated in the past, this explains why these statistics strongly underestimate the burned area in Siberia [Conard et al., 2002; Sukhinin et al., 2004].

[v] A second source of information on fire comes from remote sensing, which allows access to regions where in situ observations are rare or nonexistent, and provides observations on a large scale. The time coverage is however limited to the lifetime of the satellite missions. Two kinds of products are usually derived from space observation: fire hot spot detection and burned area maps. Several instruments give access to these products: AVHRR; ATSR [e.g., Simon et al., 2004]; SPOT-VEGETATION [e.g., Tansey et al., 2004]; and MODIS [Justice et al., 2002]. Although there is a good agreement between the different remote sensing data sets in terms of spatial distribution, there is large disagreement among their estimates of burned area [Boschetti et al., 2004], and they are not consistent with in situ observations. Moreover, these problems are most severe in the boreal regions [Ichoku et al., 2003] for a number of reasons. In addition to the usual limitations of remote sensing observations of fires (cloud cover, saturation of spectral channels, sensor spatial resolution) [e.g., Cahoon et al., 2000], boreal observations are especially affected by light contamination and predominance of low solar angles at high latitudes [Simon et al., 2004]. Moreover, burned area in boreal forest tends to be overestimated because old burned scars masked by snow tend to be misinterpreted with fresh burns when first uncovered by snowmelt [Simon et al., 2004]. Two products have been specifically designed for mapping fires in boreal forests in Canada for the period 1994–1998 [Fraser et al., 2000] and in Siberia for the period 1996–2002 [Soja et al., 2004a, 2004b; Sukhinin et al., 2004] (available at http://glcf.umiacs.umd.edu/data/burned/) from AVHRR observations. They have the advantage of not being designed for a global application and thus take into account specific characteristic of the boreal regions. They both combine hot spot detections and burned area detection in hybrid algorithms.

2.2. Climate

[10] The set of climate fields used in the PBA model includes atmospheric temperature, wind, relative humidity, soil water content (SWC), and snow. In the following, SWC is defined as the sum of root-zone water store, excluding wetlands, divided by the maximum water store the grid cell soil can contain [Milly and Shmakin, 2002]. Because the model is designed for use in a climate model, it uses only fields simulated by the model. Temperature and wind are from an experiment with the GFDL AM2 model at a resolution of 2° × 2.5° longitude by latitude [Anderson and the GFDL Global Atmospheric Model Development Team, 2004]. Relative humidity, fraction of soil water content, and snow are computed in the stand-alone dynamic vegetation model LM3V (Shevlakova et al., submitted manuscript, 2007) forced by the output from the same experiment with AM2 and observed precipitation [Nijssen et al., 2001]. The land surface component of LM3V operates on fast timescales with a 30 min time resolution. Fast timescale processes include canopy biophysics, ecosystem CO2 exchange, soil/snow thermodynamics and water balance, and radiation exchange. The treatment of soil and snow thermodynamics and water balance is based on the Land Dynamics (LaD) model [Milly and Shmakin, 2002]. The land surface model tracks three lumped water pools: snowpack, root-zone soil water, and groundwater. Soil heat conduction and temperature are tracked in multiple layers, and snowpack is assumed to be isothermal and in thermal equilibrium with the soil surface.

2.3. Roads

[11] Fires in boreal forest may be caused by human activities. This is particularly true for Siberia, where more than 85% of the fires have been related to anthropogenic sources [Shvidenko and Nilsson, 2000; Mollicone et al., 2006]. In Canada, lightning-ignited fires account for about 35% of the fires and 85% of the area burned. In the far north, where both the population and the fire control are sparse [Stocks et al., 2002], lightning-ignited fires become dominant. These large fires represent only a small number of fires but account for most of the area burned because they are remote (difficult to access) and allowed to burn. Most of the anthropogenic fires start close to roads [Korovin, 1996]. Distance to the nearest road has thus been used as input in some fire models [e.g., Cardoso et al., 2003]. In the following, road density (including railways) from the database of Hearns et al. [2002] is used as a proxy for human influence.

2.4. Training and Validation Data Sets

[12] In this study, the basic unit is a grid cell of 2° × 2.5° on a monthly time step. All data are aggregated to this spatiotemporal resolution. Each grid cell (1 month, 2° × 2.5°) is referred to as “one situation.” The use of such a coarse resolution presents several advantages. First, most of the existing climate models use a similar or coarser spatial...
resolution. Second, it is easier to model the main features of fires at a coarse resolution rather than at a finer resolution, with the local events being harder to catch than the general patterns [e.g., Cardoso et al., 2003]. Third, it reduces the error of mismatch between the different data sets. The temporal resolution also matches the most common resolution used in climate models to output results. A monthly resolution may miss some particular events affecting a grid cell on a short time period (days). However, at such a spatial resolution, short-term weather conditions, which strongly affect fire behavior, are expected to be smoothed out.

The time period used in this study is 1983–1999, which extends from the first year of precipitation data availability from Nijssen et al. [2001] to the last year covered by the Canadian LFDB [Stocks et al., 2002]. The situations occurring in the period 1983–1989 chosen from the Canadian LFDB make the training data set used to estimate the model parameters. The validation data set is made up of the remaining part of the Canadian LFDB, of the Alaskan LFDB [Kasischke et al., 2002], and of AVHRR estimates of burned area for Siberia from Sukhinin et al. [2004]. Therefore model parameters are estimated exclusively on the basis of on-ground data from Canada.

3. Influence of Climate and Human-Related Variables on Fire

The potential importance of different climate and human-related variables in explaining fire behavior may be inferred by comparing frequency distributions of these variables in all grid cells located in the Canadian boreal forest versus frequency distributions of the same variable only in the burnt grid cells for the fire season. These distributions are plotted in Figure 1. The relative frequency distribution of burnt grid cells is calculated by normalizing the frequency distribution for burnt cells by the inverse of the all grid cell frequency distribution. For most variables, the two distributions differ substantially, indicating a strong correlation between the variables and fire occurrence. As expected, drier and warmer situations coincide with increased fire occurrence compared to more humid and colder conditions.

Soil water content (Figure 1f) and precipitation (Figure 1b) give different information on fire. Using both of them is similar to using minimum and total precipitation, as in Cardoso et al. [2003], the soil water content being an integrator of past precipitation. However, soil water content integrates precipitation and evaporation. This is important because evaporation is expected to increase in some parts of the boreal forest, because of rising air temperature, and this may offset increases in precipitation [Groisman et al., 2004; McGuire et al., 2006].

The distribution of wind speed based on all grid cells is similar to the distribution based on burnt grid cells (Figure 1c). This is due to the coarse spatial and temporal resolutions used here. At finer resolutions, wind does drive fire spread and wind speed is still used in calculating fire weather index [Flannigan and Harrington, 1988].

As expected, snow plays a substantial role in predicting fire. Figure 1g shows that most of the fires occur when snow is melted. This is in agreement with Westerling et al.
Table 1. Maximum Likelihood Estimates of the Parameters and Their Standard Deviation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate ± Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scaling α</td>
<td>1.48 ± 0.31</td>
</tr>
<tr>
<td>Precipitation β</td>
<td>-3.95 ± 0.09</td>
</tr>
<tr>
<td>γ</td>
<td>2.77 ± 0.18</td>
</tr>
<tr>
<td>Temperature β</td>
<td>0.89 ± 0.07</td>
</tr>
<tr>
<td>γ</td>
<td>-2.41 ± 0.19</td>
</tr>
<tr>
<td>Soil water content β</td>
<td>2.23 ± 0.07</td>
</tr>
<tr>
<td>γ</td>
<td>2.61 ± 0.15</td>
</tr>
<tr>
<td>Relative humidity β</td>
<td>-9.23 ± 0.11</td>
</tr>
<tr>
<td>γ</td>
<td>6.46 ± 0.23</td>
</tr>
<tr>
<td>Road density β</td>
<td>-3.16 ± 0.11</td>
</tr>
<tr>
<td>γ</td>
<td>2.11 ± 0.24</td>
</tr>
</tbody>
</table>

[2006] who showed that snow controls the length of the fire season, and thus the amount of burned area in 1 year (a).

[18] Finally, road density plays a discriminating role: above a given threshold, no burning is expected to happen. Below, fire will be likely to happen provided that favorable climate conditions exist. This result is in agreement with several studies showing that fires are more likely to initiate close to roads [Korovin, 1996]. However, when the road density is too high, fires are expected to be stopped more easily because of higher fire suppression actions and because roads act as fire breaks.

4. Design of the Burned Area Model

4.1. Choice of the Functional Form

[19] We chose logit functions to describe the influence of each variable on burned area because they are consistent with the shapes observed on Figure 1. Our formulation models burned area as the product of five sigmoid functions, one for each of our predictive variables, similar to the sigmoid function used to model soil moisture in Thonicke et al. [2001] and Arora and Boer [2005]. The monthly burned area \( BA \) in a 2° × 2.5° grid cell, expressed in terms of a fraction of a grid cell, is given by

\[
BA = \alpha \times H(\text{snow}) \times \prod_{i=1}^{5} \frac{1}{1 + \exp[-(\beta x_i + \gamma_i)]} \tag{1}
\]

where \( x_i \) stands for the 5 reduced variables temperature, relative humidity, precipitation, soil water content and road density, and

\[
H(\text{snow}) = \begin{cases} 
0 & \text{if } \text{snow} > 0 \\
1 & \text{if } \text{snow} = 0
\end{cases} \tag{2}
\]

\( H(\text{snow}) \) reflects the importance of snow in governing the fire season and thus BA (see section 3). In this model, there are 11 parameters to be estimated: 2 parameters (\( \beta \), \( \gamma \)) per variable, and the scaling parameter \( \alpha \). The independent variables are normalized to have zero mean and unit standard deviation:

\[
x_i = \frac{X_i - \mu_i}{\sigma_i} \tag{3}
\]

where \( X_i \) is the variable, \( \mu_i \) and \( \sigma_i \) are, respectively, the mean and the standard deviation of the variable \( X_i \) computed over the period 1983–1999.

[20] The 11 parameters are estimated using a Markov Chain Monte Carlo (MCMC) method based on the Metropolis simulated annealing algorithm [Metropolis et al., 1953] and applied to the burned area training data set derived from the Canadian LFDB (see section 2.4). The normal log likelihood is used as an objective function to compare the merits of various sets of parameters

\[
\ell(\theta) = \sum_{i=1}^{n} -\ln(\sigma) - \frac{(BA_i - BA_i^{\text{obs}})^2}{2\sigma^2} \tag{4}
\]

where \( \theta \) represents the set of 11 parameters, \( BA_i^{\text{obs}} \) is the observed burned area for situation \( i \), \( BA_i \) is the modeled burned area for situation \( i \), and \( n \) is the number of situations in the training data set. The MCMC method tries to minimize the error between observed and modeled burned area by optimizing the set of 11 parameters. Maximum likelihood estimates of the parameters, along with their standard deviations, as given by the MCMC method are shown in Table 1.

4.2. Relative Importance of the Variables

[21] The log likelihood function given by equation (4) measures the goodness of fit of the model to the data and thus may be used to measure the relative contribution of each variable to the estimation of burned area. To assess this contribution, each variable is removed one at a time, and a new estimation procedure is launched. The decrease in the log likelihood function computed using equation (4) gives the influence of each variable to the estimation of burned area. The relative importance of each variables to the most important one derived from this study is plotted on Figure 2 (an arbitrary value of 100 has been assigned to the most important variable).

[22] The two most important variables in the estimation of burned area are precipitation and temperature. They are followed by soil water content, relative humidity and road density. Temperature and precipitation have previously been identified as drivers of fire and are used by all operational agencies to estimate fire danger. However, it should be noted that all the variables used in the PBA model have similar importance in the estimation. As expected, road density is an important factor influencing the amount of burned area and should be taken into account in any fire models. However, as compared to climate variables, its importance is less than in regions where humans heavily perturb the ecosystems. For instance, in Amazonia, where the highest deforestation rates in the world are found [Laurance et al., 2001], Cardoso et al. [2003] found that distance from paved roads had the highest relative importance in modeling fire, before any climate variables.

[23] Figure 2 also shows the importance of wind speed (relative to precipitation) that has been obtained by
performing the same study while including wind in the functional form given by equation (1). At the spatial and temporal resolutions considered here, wind is of one order less important than the previous five variables, confirming the conclusion drawn from Figure 1 (see section 3).

5. Evaluation of the Model Across the Whole Boreal Region

[24] In this section, the PBA model is evaluated against observation-derived data for three boreal regions: Canada over the period 1990–1999 (the period 1983–1989 has been used to estimate the model parameters), Alaska (1983–1999) and Siberia (1996–2000). To be able to compare the simulated burned area with real data, it is necessary to use realistic climate fields. As described in section 2, we use precipitation fields from Nijssen et al. [2001], which limits the evaluation to the time period 1983–2000.

[25] In order to evaluate the fire model, fire history must be taken into account. In the absence of any information on fuel load (that will be available once the model is coupled to a dynamic vegetation model), we use a rather simple formulation that has the mere purpose of taking into account the blocking period that follows fire. The surface available to burn is not the entire surface of the grid cell, but the surface of the grid cell minus the surface burned in the past and that has not yet recovered. Therefore the burned area estimated for 1 month of year t, noted \( B_A_t \), will be given by

\[
\frac{S - \sum_{j=1}^{r} B_A_{t-j}}{S} = B_A_{t}^{\text{mod}} \times \frac{S - \sum_{j=1}^{r} B_A_{t-j}}{S} \quad (5)
\]

where \( S \) is the surface of the grid cell for which burned area is computed, \( B_A_{t}^{\text{mod}} \) is the output of the model for year t given by equation (1), and \( T \) is the number of years for which past burned area is taken into account. For Canada and Alaska, for years before 1983, \( B_A_t \) is derived from the Canadian and Alaskan LFDBs. \( T \) is chosen to be equal to 25 a, which is approximately the number of years needed to build up sufficient fuel loads (about 0.5 kgC) to carry a fire [Nalder and Wein, 1999; Wirth et al., 2002]. It is also the time between the first year of the Canadian LFDB and the first year of burned area estimation from the PBA model. For Siberia, no historical data is available and thus fire history is not taken into account.

5.1. Evaluation of the Model in Canada

[26] The model is first applied to Canada. Figures 3a and 3b show the monthly evolution of burned area as derived from the LFDB [Stocks et al., 2002] and as simulated by the PBA model for the years 1983–1989 (training period) and 1990–1999. The model is able to reproduce the interannual variability of burned area. In particular, most of the severe fire years that occurred in the region (1995 and 1998) are well simulated by the model. With the largest exception of 1994 (see below).

[27] Figures 3c and 3d show the spatial distribution of the annual mean of percent burned area observed and simulated in Canada for the period 1990–1999. The highest percent burned area is found in the northern boreal zone, and the model is able to reproduce the zone of high burning in the Taiga and Boreal Shield West ecozones, as defined by Stocks et al. [2002] (regions extending from south of the northwestern territories to north of Manitoba). However, the model does not reproduce the high percent burned area observed at the border of Quebec and Ontario (southern Taiga Shield East ecoregion in the work by Stocks et al. [2002]).

[28] Figure 4 shows the performance of the model in terms of seasonality, for the period 1990–1999. The PBA model accurately simulates a fire season running from April through October, with a maximum in June–August, the highest values being found in June (Figure 4a). Figure 4b shows the histogram of the difference between observed and
Figure 3. Monthly time series of burned area in Canadian boreal forest (a) as observed in situ and (b) as predicted by the model in ha. Percent annual burned area distribution for 1990–1999 (c) observed and (d) predicted by the model.
predicted month of highest burning for the whole grid cells in Canada affected by fire over the period 1990–1999. The month of highest burning estimated by the model agrees with observations in 67% of the situations; for the remaining situation, the model tends to predict the highest burning the following month (19%, mainly August instead of July) or the month before (9%).

Figure 5 shows the histogram of the relative difference between observed and predicted annual burned area for all the grid cells affected by fire over the period 1990–1999. The mean of the distribution is 0.02 with a standard deviation of ±0.12. The high positive values observed on Figure 5 (around 0.21) mainly come from the year 1994 for which the model predicts less burned area than what was observed.

Figure 5. Histogram of the difference of annual burned area in each grid cell as derived from the Canadian LFDB and as predicted by the model for the period 1990–1999, expressed as a fraction of a grid cell.
Figure 6. For the period May–August 1995, monthly evolution of (a) observed burned area as given by the Canada LFDB, (b) same as Figure 6a but simulated by the PBA model, (c) temperature anomalies, and (d) precipitation anomalies.
Figure 7. Same as Figure 6 but for the period May–August 1994.
from the model follows these climate fields: in May, fires occur mainly in the western part of Canada; in June, as observed in situ, the model simulates two spots of high burnings, one in Ontario, and the other one in the western part of Canada, following the two zones where hot and dry conditions are taking place (see Figures 6c and 6d); in July, as in the observations, fires take place in central Canada; finally, in August, the model catches the favorable conditions for fire in the east, but simulates too much burning in the southeast, following the strong negative anomaly in precipitation (drought) associated with the positive temperature anomaly located in this region.

Monthly distributions of burned area, temperature anomalies and precipitation anomalies are plotted in Figure 7 for the period May–August 1994, for which the model clearly underestimate the amount of burned area (Figure 3b). It may be seen that the low burning simulated by the PBA model mainly comes from the distribution of the two climate fields, which show small anomalies in temperature and very high positive anomalies in precipitation (wet year). A comparison of the Nijssen et al. [2001] data set with the observation-derived precipitation fields from Xie and Arkin [1996] shows a large discrepancy in the amount and distribution of precipitation in Canada for 1994 between the two data sets, whereas they agree well for 1995. In particular, the region north of the Great Lakes, where huge fires were observed in June (Figure 7a), is wetter in the Nijssen et al. [2001] data set. On the average, the precipitation anomalies derived from Nijssen et al. [2001] are 30% higher than those derived from Xie and Arkin [1996] in 1994. Therefore, while the PBA model predicts well the large burned area in 1995 (when precipitation fields from Nijssen et al. [2001] agree with those of Xie and Arkin [1996]), it underestimates burned area in 1994 as a result of biases in the precipitation field.

5.2. Evaluation of the Model in Alaska

The annual evolution of burned area as derived from the Alaska LFDB [Kasischke et al., 2002] and as simulated by the PBA model are plotted in Figure 8. Both time series are in good agreement with regards to interannual variation and

![Figure 8. Annual variation of burned area in Alaska as given by the Alaska LFDB (dashed line) and as simulated by the PBA model (solid line).]
amount of burned area, with, once more, the exception of 1994. The geographic distribution simulated by the PBA model is also in agreement with the Alaskan LFDB observations (Figure 9). Similar to Canada (Figure 5), the difference between observed and predicted Alaskan annual burned area, expressed in terms of fraction of the grid cell, for all the grid cells affected by fire over the period 1983–1999 is unbiased with a standard deviation of ±0.12.

5.3. Evaluation of the Model in Siberia

[33] The annual variation of burned area in Siberian boreal forest given by the Russian Forest Fund [Shvidenko and Goldammer, 2001], derived from AVHRR [Soja et al., 2004a, 2004b; Sukhinin et al., 2004], and estimated by our fire model are plotted in Figure 10. The model reproduces fairly well the interannual variation of burned area as observed by both in situ data and remote sensing data, with two severe fire events occurring in 1998 and 2000. The amplitude is in the range of existing estimates. As expected, the amplitude is larger than the one given by official statistics, which are known to be underestimated (see section 2.1). On the other hand, it is lower than the amplitude derived from AVHRR.

[34] The location of fires retrieved by the model is in agreement with remote sensing observation as can be seen in Figure 11, which shows the burned area as estimated by the model (Figure 11, right) and as observed by AVHRR [Sukhinin et al., 2004] (Figure 11, left). In 1997, fires occurred around 90°E while, in 1998, large fires occurred in the far east. In 1999, fire extended over a large region in middle Siberia. These patterns are retrieved by the model. However, it can be seen that the model tends to predict more burning in two regions, where few fires are seen: the first region is located above 60°N and west of 120°E; and the second is located above 64°N and east of 140°E. Two reasons might explain this. First, the past evolution of fire is not taken into account for Siberia and some burning may have occurred in this part of Siberia in previous years. Second, these regions are predominantly covered by sparse and prostrate larch trees [Nikolov and Helmisaari, 1992]. Hence the fuel available is very limited. Now, the model does not have any information about vegetation and only predicts the potential amount of burned area, depending on climate and road density. Therefore the model tends to overestimate the amount of burned area in these regions. This is particularly true in 1997 and explains the high value of the estimated annual burned area plotted in Figure 9.

[35] For the severe fire year of 1998, Figure 12 shows the season of highest burning estimated by the PBA model: springtime (from April through June), summertime (July and August) and fall (September and October). It can be directly compared to Kajii et al. [2002, Figure 1], which shows the accumulated hot spots due to fire obtained by AVHRR image in 1998 during the same three seasons. In both observations and estimations, fires primarily occurred in eastern and western Siberia in spring; they then moved eastward during the summer; the highest burning occurred in the far east from July to October.

6. Conclusion

[36] A prognostic model predicting the amount of burned area, in a grid cell of 2° × 2.5°, on a monthly basis, specifically for the boreal forest, has been designed for use in DGVMs. Its parameters have been estimated using a large set of in situ observations made in Canada [Stocks et al., 2002]. Predictor variables are one human-related variable (road density) and four climate variables (precipitation, temperature, soil water content and relative humidity), which are routinely produced by most dynamic vegetation models.
Figure 11. Annual area burned ($10^3$ ha) for 1996, 1997 and 1998, as observed by AVHRR [Sukhinin et al., 2004] and as simulated by the model, on a $2^\circ \times 2.5^\circ$ grid.
The estimations are therefore functions of current climatic conditions and are thus responsive to climate change. Model formulation is general enough to be included in any dynamic vegetation model in order to include the interaction between vegetation and fire.

The model has been validated against large sets of in situ observations for Canada and Alaska (LFDBs [Stocks et al., 2002; Kasischke et al., 2002]) and of burned area estimates from AVHRR for Siberia [Sukhinin et al., 2004]. Provided realistic climate predictors, the model successfully reproduces the spatiotemporal evolution of burned area in the boreal forest, including seasonality and interannual variability. In particular, the model simulates the correct month of highest burning, with two thirds of the situations predicted the right month and the rest being predicted 1 month before or after. The predicted burned area is in the range of various current estimations. The estimated annual relative error (standard deviation) is twelve percent in a grid cell, which makes it suitable to study quantitatively the evolution of burned area with climate.

Using a log likelihood metric to measure the fit of model to data, monthly precipitation and temperature have proven to be the two main drivers of the amount of area burned in a grid cell (2° × 2.5°, 1 month), followed by soil water content, relative humidity and road density. Temperature and precipitation have previously been identified as drivers of fire in the boreal region and are used by all operational agencies to estimate fire danger. As expected, the importance of road density is less in boreal forest than in regions heavily influenced by human, like Amazonia. However, our study shows that neglecting such a human-related factor has a negative effect on the estimation of burned area.

The model tends to overestimate burned area in some regions. This may be explained by various factors. First, the model does not account properly for the past evolution of fire. Second, the climate drivers are not derived from observations but from simulations, which may make them biased in some regions. Third, no information on fire suppression policies has been taken into account in the model, whereas such policies can dramatically impact fire behavior, as has been observed in the US [Mouillot and Field, 2005]. However, a significant part of the boreal region is remote and both the population and fire control are sparse. Fourth, no information on species is used. This could induce several biases, since conditions associated with large fires may vary between ecozones [Amiro et al., 2004]. For instance, some mismatch between observations and simulations might come from the inclusion of Cordilleran (British Columbia) and Maritime (Nova Scotia and New Brunswick) forests in the study of Canadian fires. However, the amount of area burned in those two regions accounts for less than 6% of the total area burned annually in the boreal region, as suggested by Figures 3c and 3d. In addition, in its current stage the model does not account for the fact that, in some regions of Siberia, the postfire communities are dominated by deciduous forests (Betula spec. and Populus spec.), with a comparatively low flammability for approximately 80 a [Johnson, 1992; Schulze et al., 2005]. This may lead to an overestimation of annual area burned. Moreover, coniferous species endemic to the North American and Siberian boreal regions exhibit vastly different fire adaptation strategies with significant implica-
tion for the fire regime. For instance, Scots pine and various larch species capable of surviving fires and favoring recurring surface fires are common in Eurasia. In North America, serotinous species, that depend on hot but less frequent crown fires for successful regeneration, dominate [Wirth, 2005]. These issues will be solved when the PBA model is coupled to the dynamic vegetation model LM3V.

[40] The model could be improved by including cloud-to-ground lightning activity as a precursor. Indeed, even if the severity of fires, and hence the surface burned, are mostly dependent on climate, lightning is a major source of ignition, mainly in North America. In particular, lightning frequency is expected to increase with warmer climate [Williams, 2004], which might have an impact on future fire ignition at high latitudes [e.g., Price and Rind, 1994]. In the context of simulation of future climate, this would however require to simulate accurately lightning activities and its relation with climate.

[41] Acknowledgments. This research was funded by the Cooperative Institute for Climate Science under award NA17RJ2612 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce, and by the Carbon Mitigation Initiative, with support provided by the Ford Motor Company. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce. The authors wish to thank Edward Hyer for his help in getting the AVHRR data; Amber Soja, Ted Schuur and Andy Turner for fruitful discussion; Tsering W. Shawa, GIS librarian of Princeton University, for his help in exploiting GIS data; and the three anonymous reviewers for their helpful comments.

References
Kasischke, E. S., D. Williams, and D. Barry (2002), Analysis of the patterns of large fires in the boreal forest region of Alaska, Int. J. Wildland Fire, 11, 1–144.
Nikolov, N., and H. Helmerssati (1992), Silvics of the circumpolar boreal forest tree species, in A Systems Analysis of the Global Boreal Forest,


