Improving the Representation of Fire Disturbance in Dynamic Vegetation Models by Assimilating Satellite Data: a Case Study over the Arctic

E.P Kantzas[1,2], S. Quegan[1] and M. Lomas[1]

[1]{School of Mathematics and Statistics, University of Sheffield, Hicks Building, Hounsfield Rd, Sheffield S37RH, UK}

[2]{Nansen International Environmental and Remote Sensing Centre, Vasilievsky Island, 199034, St. Petersburg, Russia}

Correspondence to: E.P Kantzas (e.kantzas@sheffield.ac.uk)

Abstract

Fire provides an impulsive and stochastic pathway for carbon from the terrestrial biosphere to enter the atmosphere. Despite fire emissions being of similar magnitude to Net Ecosystem Exchange in many biomes, even the most complex Dynamic Vegetation Models (DVMs) embedded in General Circulation Models contain poor representations of fire behaviour and dynamics, such as propagation and distribution of fire sizes. A model-independent methodology is developed which addresses this issue. Its focus is on the Arctic where fire is linked to permafrost dynamics and on occasion can release great amounts of carbon from carbon-rich organic soils. Connected Component Labeling is used to identify individual fire events across Canada and Russia from daily, low-resolution burned area satellite products, and the obtained fire size probability distributions are validated against historical data. This allows the creation of a fire database holding information on area burned and temporal evolution of fires in space and time. A method of assimilating the statistical distribution of fire area into a DVM whilst maintaining its Fire Return Interval is then described. The algorithm imposes a regional scale spatially dependent fire regime on a sub-scale spatially independent model; the fire regime is described by large scale statistical distributions of fire intensity and spatial extent, and the temporal dynamics (fire return intervals) are determined locally. This permits DVMs to estimate many aspects of post-fire dynamics that cannot occur under their current representations of

fire, as is illustrated by considering the modeled evolution of land cover, biomass and Net Ecosystem Exchange after a fire.

1. Introduction

Despite the high uncertainties in estimates of global biomass stocks, analysis of carbon stock studies (Keith et al., 2009) shows that the boreal regions hold considerable biomass per unit area, arising not only from the extent of boreal ecosystems but also from the observed greening of the Arctic (Jia et al., 2003; Xu et al., 2013); indeed, shrub communities have expanded during the 20th century (Sturm et al., 2001), while tundra biomass has increased by almost 20% over the past three decades (Epstein et al., 2012). Furthermore, due to low temperatures and consequent low decomposition rates, enormous stocks of soil carbon exist in the Arctic (Ping et al., 2008; Schepaschenko et al., 2013), some locked in and under permafrost, with the soil organic carbon in the circumpolar permafrost region accounting for approximately 50% of the global soil organic carbon pool (Tarnocai et al., 2009).

Changes and feedbacks in the fluxes of carbon between the land surface and the atmosphere are of utmost importance in the context of global warming. The need to gain quantitative understanding of such processes has encouraged the use of Dynamic Vegetation Models (DVMs), often coupled to atmospheric models. DVMs simulate a host of mechanisms linked to the terrestrial carbon and water cycles, with the aim of reproducing the present status of the terrestrial carbon pools and fluxes and predicting their trends. An influential and dynamic pathway by which terrestrial carbon enters the atmosphere is through burning of vegetation and carbon-rich soils; its implementation in DVMs is investigated in this study.

Key characteristics of fires include their inter-annual variability, size distribution and intensity, which differ between regions of the Arctic (Wooster and Zhang, 2004). Fieldwork and Earth Observation (EO) data show that larger fires contribute most to the total area burned, despite being much rarer. From 1959-1999, out of the 1,000-14,000 forest fires that occur every year in Canadian forests, only 3% exceeded 2 km^2 in area but these accounted for 97% of the total area burned (Stocks et al., 1998). Such fires usually last for several days or even weeks, and extend over large areas and

across biomes; in their central areas nearly all the vegetation is burnt, with reducing degree of burn towards the fire scar edges.

Advances in EO sensors have contributed greatly to the information available on a variety of fire characteristics, beginning with the detection of gas flares from oilfields in images obtained from the AVHRR sensor on board the TIROS-N satellite (Matson and Dozier, 1981). In the 21st century, images acquired from the MODIS instrument are routinely used to identify fire scars at 500 m resolution (Roy et al., 2008) and examine global trends in burned area (Giglio et al., 2010); by measuring thermal anomalies the ATSR instrument can locate active fires and construct time series monitoring their annual evolution (Arino et al., 2012); and measurements of fire radiative power from geostationary and polar orbiting EO sensors allow the amount of biomass lost to fires to be estimated (Wooster et al., 2012; Wooster and Zhang, 2004).

Nonetheless, the representation of fire in most DVMs does not utilize EO information and fails to capture many of the key fire characteristics (Kantzas et al., 2013). A typical DVM will estimate a fraction of area burned for each grid-cell based on climate data (e.g. temperature and precipitation), vegetation characteristics (e.g. plant-specific fire resistance) and other simulated variables (e.g. litter moisture), so the outputs are deterministic and without any random component. Nevertheless, it is well established that the size distribution of forest fires at continental scale follows the law of small numbers and can be simulated stochastically with a Poisson model parameterized with climate data (Jiang et al., 2012; Podur et al., 2010; Wiitala, 1999). This heavily skewed distribution assigns high probability to small fires and lower probability to bigger ones.

Most DVMs are unable to simulate large fires that occupy significant fractions of a model grid-cell (which for a typical DVM resolution of 0.5^0 has dimensions of around 56 km by 28 km at 60^o N). In addition, most DVMs are essentially point-based, with no interaction between neighbouring grid-cells, so cannot simulate the propagation of fire across several grid-cells. Instead, each grid-cell is assigned a small amount of fire each year, with very little inter-annual variability (Kloster et al., 2010; Li et al., 2014; Prentice et al., 2011; Thonicke et al., 2010; Thonicke et al., 2001). As an example, for a typical year over the Arctic, in the LPJ-WM model (Wania et al., 2009a, b), which is a version of the influential LPJ DVM (Sitch et al., 2003) tailored for high-latitudes,

the average fractional area burned per grid-cell is 0.3% with a variance of 0.045%, and it rarely exceeds 1% in any grid-cell. This weakness in fire representation is hidden when only the average fraction of area burned over a long period (whose reciprocal is the Fire Return Interval or FRI) is reported. For example, if a DVM represents fire by burning 0.5% of each grid-cell every year, the FRI will be 200 years, but this completely fails to capture the highly episodic nature of boreal fires, in which severe fire years may give emissions many times greater than the average (van der Werf et al., 2010). In reality, observed FRI data from many small and some rarely occurring big fires generally have a significantly higher variance than that produced by DVMs. Correct simulation of the FRI is insufficient for DVMs to make accurate predictions under changing climate scenarios.

The treatment of fire in DVMs also prevents them from capturing post-disturbance dynamics (e.g. permafrost thawing and carbon fluxes) from large fires which remove a considerable fraction of vegetation and soil carbon (Kantzas et al., 2013), like the unprecedented 2007 Anaktuvuk River fire (Mack et al., 2011) Post-fire carbon fluxes exhibit complicated dynamics (Amiro et al., 2003; Amiro et al., 2006), with consequences for the extent to which vegetation recovery eventually turns a region burned in a large fire from a carbon source into a sink, and how long, if ever, it takes for carbon stocks to return to previous levels under a changing climate (Amiro et al., 2001a).

The amount of litter removed in a fire is a key quantity controlling post-disturbance permafrost degradation (Harden et al., 2006; Yoshikawa et al., 2002) while the water cycle is also affected by large fires, and regional models have showed that changing fire regimes cause changes in evapotranspiration in boreal forests (Bond-Lamberty et al., 2009). Field data show that a substantial loss of canopy will decrease evapotranspiration (Amiro et al., 2006) and canopy interception and consequently increase groundwater recharge (Clark et al., 2012), but vegetation succession would further complicate water dynamics, especially when forests stands are succeeded by grass/shrubs for a number of years or indefinitely (Dore et al., 2010; Dore et al., 2012). The DVMs would also be better coupled with atmospheric models and provide a more realistic gas exchange interface if their simulations were capable of producing large fires, with effects ranging from realistically simulating the carbon and trace gases fluxes of big disturbances (van der Werf et al., 2010) to how smoke affects

cloud formation over the boreal forests (Sassen and Khvorostyanov, 2008) and the Amazon (Koren et al., 2004).

Hence there are pressing reasons to improve the fire representation in the DVMs, but these models are complex, involve highly coupled internal processes, operate on a grid-cell basis, and are often embedded in climate models. In addition, significant resources have been spent to calibrate fire processes so that the FRI compares well (in some cases) with data (Prentice et al., 2011; Thonicke et al., 2010). Hence it is desirable to keep model restructuring to a minimum and preserve its estimate of FRI, while ensuring that fire characteristics, such as structure and size distribution, are consistent with observational data.

The first step towards this goal is to obtain realistic statistical information on fire at spatial scales appropriate to the models, i.e. 0.5°-1°, for example the number of fires per year, their size distribution and their spatial characteristics. Currently, historical information on wildfires in the Arctic, such as their number, area burned and boundaries, is compiled in databases by fire agencies in Canada (Canadian National Fire Database) and Alaska (Alaska Interagency Coordination Centre); these consist mostly of ground observations supplemented by EO data. Due to the remoteness of much of the boreal zone, there are large data gaps, and similar data do not exist for the much larger area of the Russian Arctic. In Section 2.1 we show how readily-available image analysis tools, specifically Connected Component Labeling, can be employed to identify individual fires in EO burned area data and extract the information needed to characterize fires in the Arctic statistically. We then exploit this information in Section 2.2 to develop a model-independent methodology for creating fires with a realistic size distribution in DVMs while maintaining their FRI and involving little model restructuring. In Section 3 we verify both methodological approaches and demonstrate some of the consequences for post-fire dynamics, while in Section 4 we discuss the limitations of the approach and possible ways to address them.

2. Methodology

2.1. Connected Component Labeling

Connected component labeling (CCL) (Gonzalez et al., 2003) or “blob detection” in the context of image processing is a method where unique clusters in a binary image are identified based on the connectivity of their sides and/or edges. In two

dimensions, two categories exist: 4-connected and 8-connected. In 4-connected labeling, each pixel with coordinates (x, y) can be connected to those pixels with which it shares an edge, i.e. the pixels with coordinates ($x \pm 1$, y) and (x, $y \pm 1$). In 8-connected labeling, pixels with a common vertex are also included, so there are extra possible connections to the pixels at positions ($x \pm 1$, $y \pm 1$). Thus in a binary image, CCL would *label* or *cluster* connected blobs of 1s against a background of 0s. Two dimensional CCL has numerous applications in image analysis, and has been used for clustering pixels in fire scars in single images (Koltunov et al., 2012; Morisette et al., 2005). CCL can also be applied in three dimensions, where the third dimension can be time, and we exploit this capability to determine the growth of fire scars in sequences of daily EO images of burned area. For each image, pixels identified as burned are assigned the value 1 and the rest are given the value 0. Additionally each image and its pixels are labeled by the associated day of the year, t, to yield a 3-dimensional dataset (x, y, t) on which we apply the CCL algorithm.

Three-dimensional CCL has 6, 18 and 26-connected categories, defined respectively at a given voxel by those voxels having a common face, plus those with a common edge, plus those with a common vertex; the first 2 categories are depicted in Fig. 1. Assuming the accuracy of the underlying daily burned area images, CCL should be able to track the progress of a particular fire from its ignition, through its temporal and spatial propagation to its extinction, by following the connections between burned pixels. As fire scars are continuous in both space and time, individual fires will be labeled and subsequently categorized based on burned area. It should be noted that the labelling of an individual fire may depend on the spatial resolution of the EO sensor, since what is seen as a single burned pixel at lower resolution may in fact be resolved as several fires when imaged at higher resolution. However, this does not cause problems when assimilating the data, as the model used will have the same spatial resolution as the fire database created.

In principle, the 6-connected variety of CCL should be sufficient to capture fire spread as a fire could not propagate diagonally in space without affecting the adjacent pixels. For example, a fire at (x, y, t) propagating to ($x + 1$, $y + 1$, t) would most likely affect ($x + 1$, y, t) and/or (x, $y + 1$, t). However, in some cases the fire may propagate diagonally with no detectable effect on the adjacent pixels, for example because of sensor detection sensitivity, so 6-connected CCL would detect it as two independent

fires instead of a single event. It is also possible that two fires occurred on the same day in diagonal grid-cells with independent ignition sources, whether natural or anthropogenic. Weather conditions conducive to lightning can cover large areas and lead to lightning igniting more than one fire in a wide front, so whether these fires are independent smaller fires or a single larger one is a matter of interpretation. Hence we applied both the 6 and 18-connected CCL algorithm and compared the results with available data on fire statistics in order to determine which was more appropriate.

Following the approach of *Jiang el al.* (2012), the individual fires obtained by CCL-6 and CCL-18 were assigned to five categories according to fire size: (1) 2 km^2 to 10 km^2, (2) 10 km^2 to 30 km^2, (3) 30 km^2 to 100 km^2, (4) 100 km^2 to 500 km^2, (5) greater than 500 km^2; the aggregate of (1) to (5) is defined as a sixth category. We then applied two non-parametric statistical tests to test the null hypothesis that the fire sizes obtained from CCL and from a reference dataset detailed below represent samples from the same distribution. The two-sample Kolmogorov-Smirnov (KS) test uses a statistic that quantifies the distance between the cumulative distribution functions of the two samples; small values of this statistic indicate that the samples originate from the same distribution. The two-sample Mann-Whitney-Wilcoxon (MWW) test examines whether two independent samples originate from distributions with equal medians. Both tests were performed at a 90% confidence interval with results shown in Fig. 2.

2.1.1. Datasets

The CCL algorithm was applied to the latest version (v.4.0) of the influential Global Fire Emissions Database (GFED4) (Giglio et al., 2013). This is based on the algorithm of Giglio et al. (2009) and provides two products: burned area (GFED-BA), which gives the area burned within each grid-cell, and fire emissions, which gives fire-induced emissions of various chemical species, such as CO_2, CH_4 and NO_x (van der Werf et al., 2010). For the period used in this study, from the mid-2000s to the present day, the GFED-BA is derived daily from the MODIS MCD64A1 500 m burned area product (Roy et al., 2008), which is based on changes in reflectance in the visible channels of MODIS, but the GFED-BA also takes into account information on active fire counts (Giglio et al., 2009). It is not offered at the MODIS 500 m resolution but instead is aggregated to a resolution of 0.25^o to facilitate interfacing the

fire data to biochemical and atmospheric models which run at such resolutions (Castellanos et al., 2014; Kaiser et al., 2012; Valentini et al., 2014).

The Canadian Large Fire Database (CLFD) (Stocks et al., 2002) offers the best tool to test the outputs from the CCL analysis. It reports on forest fires greater than 2 km^2 in extent occurring in Canada from 1959-1999, including their date, location and size, together with metadata such as cause of ignition, when available. The CLFD has been used extensively in various contexts, such as investigating temporal trends in burned area (Krezek-Hanes et al., 2011), evaluating fire emissions (Amiro et al., 2004; Amiro et al., 2001b) and modeling fire frequency (Jiang et al., 2012).

In order to evaluate the CCL algorithm against the CLFD, the 0.25º GFED4 grid-cells that contain forest in Canada must first be identified. Instead of utilizing a land cover product, which would add unnecessary uncertainty, we built the forest map by combining the CLFD and GFED4 data. To do this, we first applied the CCL algorithm to the GFED4 data and assigned the value 1 to a grid-cell if it also contained a fire record in the CLFD. Clearly this would omit forest grid-cells where the CLFD did record any fire over its 40-year period, so to generate a forest mask the set of identified pixels was morphologically closed. This assigns 1s to grid-cells in close proximity to or surrounded by grid-cells already assigned the value 1. All other pixels were considered as non-forest and assigned the value 0.

We also applied the CCL algorithm over Russia, but here used the GlobCover 2000 land cover map (Arino et al., 2008) to distinguish forest from non-forest. The area of forest-related classes within each grid-cell was aggregated and if this exceeded 50% the grid-cell was assigned as forest, otherwise as non-forest.

2.2. Assimilating CCL fire products into a Dynamic Vegetation Model

Applying CCL to the daily GFED4 burned area images from 2001-2012 allows the creation of a database of individual fire events that includes their geographical location, daily propagation, fire size and geometry, i.e. how many grid-cells were affected and the fraction of each that was burned. We now give details of a methodology that assimilates this information to produce a realistic fire regime in a DVM whilst maintaining its internally simulated FRI. The algorithm can be applied to any sub-grid of pixels whose aggregate geographical representation is considered to

have a spatially independent fire regime in terms of size and intensity. Here it is applied separately to Canada and Russia, which are considered to have different fire regimes.

1) A DVM calculates the annual fraction of area burned in year *y*, **BA**(*lat*, *long*, *y*) for each grid-cell, where *lat* and *long* denote latitude-longitude. As described in the Introduction, in most current DVMs only 0.1% - 5% of each grid-cell burns annually. Each year we accumulate this fractional burned area into a new cumulative array, **BAC**, which gives the total fractional area per grid-cell burned after y years, and is defined as:

$$\mathbf{BAC}(lat, long, y) = \sum_{n=1}^{y} \mathbf{BA}(lat, long, n). \quad (1)$$

2) For each year we integrate **BA**(*lat*, *long*, *y*) over its spatial dimensions to give the aggregated fraction of area burned in year *y*,

$$int\_f(y) = \iint_{lat/lon} \mathbf{BA}(lat, long, y). \quad (2)$$

As an example, in LPJ-WM the value of *int_f* for a representative year is approximately 28.0 for Canada and 47.0 for Russia. Since the numbers of LPJ 0.5° grid-cells in the two countries are approximately 8000 and 12500 respectively, the model burns an average fraction of 0.35% per grid-cell for Canada and 0.375% for Russia; in both cases, northern Arctic grid-cells significantly reduce the overall average fraction burned.

3) Using CCL-6, we created a database **[*CCL-6*]** which, as explained above, labels all grid-cells belonging to a single fire and records the fraction burned in each of these grid-cells. For each fire, we sum these fractional areas. For the majority of the fires, e.g. a small fire over a single grid cell, the summation will yield a percentage close to 0.1%, but for larger fires that spread over multiple grid-cells, e.g. 15 grid cells with an average 10% burn, the summation can exceed 100%. We then average these summations for all fires in the database to give μf_{fire}. For Canada, μf_{fire} is 1.23% and for Russia it is 0.76%.

4) We define the total number of fires in a specific year y to be

$$no_{fires}(y) = {int\_f(y)}\Big/{\mu f_{fire}}, \quad (3)$$

which amounts to approximately about 2000 fires for Canada and 6000 for Russia, depending on year.

5) We then randomly select with replacement from the **[*CCL-6*]** database a number of fires occurring in year *y* equal to $no_{fires}(y)$.The total fraction of area burned will therefore be a normally distributed random variable with mean

$$\mu_{fire} * no_{fires}(y) \quad (4)$$

and variance

$$no_{fires}(y) * variance([\boldsymbol{CCL-6}]), \quad (5)$$

where variance(**[*CCL-6*]**) is 1.02% and 2.69% for Canada and Russia respectively. This process would cause the total fraction of area burned for that year to be a random variable, but we wish to fix it to *int_f*(*y*) which is the fraction that the model wants to burn in year *y* (Step 2), so we normalize the size of each fire so that

$$int\_f(y) = \mu_{fire} * no_{fires}(y). \quad (6)$$

6) Each fire selected from the **[*CCL-6*]** database is then overlaid on a randomly selected subset of **BAC**(*lat*, *long*, *y*) with the same spatial dimensions as the fire, e.g. if the selected fire is extended over 3x1 grid-cells then a 3x1 grid-cell area will be randomly selected from **BAC**. If each grid-cell in the **BAC**(*lat*, *long*, *y*) subset has an accumulated fractional area burned greater than or equal to that of the corresponding-grid-cell in the selected fire, then the fire will be accepted, i.e., considered to occur, and the fraction of each affected grid-cell as given by the **[*CCL-6*]** overlay will be subtracted from **BAC**(*lat*, *long*, *y*). Otherwise, a new subset will be selected at random from **BAC** until a subset capable of accommodating the fire is found.

7) The chance of finding a suitable location for a particular fire event decreases with increasing fire intensity and extent, and such a location may not exist. In the rare cases when this occurs the fire is forced to fit the location that came closest to accommodating the fire and the deficit in **BAC** is taken from other pixels to maintain the regional average of FRI. These cases occur on average once every 2 years and the **BAC** deficit is approximately 50% of a grid-cell.

8) Since DVM calculations of FRI differ between regions according to climate and vegetation, the subsets of **BA** with higher values will also have higher

values of **BAC** since the fractional burned area will accumulate there faster. Hence these regions will be able to accept more fires and the random process of selecting grid-cells will converge to produce a FRI equal to the reciprocal of **BA**.

This methodology requires an initial run of the DVM to produce **BA** for each year. These values are then fed into the above procedure to define the fires that are accepted in the **BAC** array for that specific year. The grid-cells which experience burn and the fraction burned are stored. The model is then rerun but with area burned read from the outputs of the algorithm. Even though this requires two runs, the initial run to acquire **BA** is not always needed. As long as the FRI of the model does not change significantly, either the fires produced by a previous application of the algorithm can be used or the algorithm can be run again with **BA** obtained from a previous run of the model. In the latter case, and since the process is stochastic, a different set of fires will be produced but the FRI will not change.

The LPJ-WM DVM used in this study calculates a daily fire probability for each grid cell as a function of temperature and litter moisture (Thonicke et al., 2001). The fire probability is then summed over the course of a year, from which the length of the fire season and fraction of area burned per grid cell is derived; the values of the latter populate the **BA** array. As DVMs are designed for global simulations, the built-in fire probability function needs to produce FRIs that vary considerably (10-1000 years) depending on the ecosystem. The sensitivity of fire probability to driving variables is thus optimized to the geographic variability of climate, which is considerably higher than the temporal variability over a grid cell; consequently the burned area of a grid-cell remains largely constant. Our approach corrects this unrealistic behavior by converting the fire simulations of DVMs into stochastic processes, the random component of which provides temporal variability with the aid of data assimilation.

3. Results

3.1. Applying Connected Component Labeling to the Canadian Large Fire Database

The best agreement was achieved between the CLFD and CCL-6 on Canadian forests. Here, Categories 2, 3, 4 and 5 all passed the KS test while categories 1, 2, 3 and 5 passed the MWW test. Category 1 failed the KS test and Category 4 the MWW test.

The broad category 6 passed the MWW but not the KS test. When applied to Canadian forests, CCL-18 detected 15% fewer fires than CCL-6 because the increased number of connecting points in CCL-18 merged fires that CCL-6 characterized as distinct. Nevertheless, the frequency distributions remained largely unchanged and consequently the results of both statistical tests were identical for every category.

CCL identified fewer Canadian non-forest fires than forest fires as most of the non-forest cover is in the smaller expanse of the Great Plains in the south and in the Arctic north, where climate causes a much smaller fire occurrence frequency. The smaller number of fires in non-forest grid-cells, in combination with the division into 5 categories and subdivision into 15 bins per category, reduces the size of the sample, causing higher sample variance and a less smooth histogram than for forests (Fig. 2). Nevertheless, categories 2 to 5 passed both statistical tests, but the 6^{th} aggregated fire category did not pass any of the tests. As seen in Fig. 2, this is because in Canada non-forest fires produced by CCL have smaller sizes than for forest, which is also the case when CCL is applied over Russia.

As no extensive, fire-related ground data are available for Russia, we compared the results of the CCL algorithm over Russia against the CLFD. For forests, the MWW test was passed for categories 1, 4 and 5 and the KS test for categories 3, 4 and 5. Neither test was passed for the overall category 6 for forest or non-forest. As seen in Fig. 2, this is because of the larger fraction of small forest fires compared to Canada. Nevertheless, as noted earlier, the bigger fires contribute disproportionately to the annual area burned and consequently are the most important to incorporate correctly in the DVMs. Indeed, in the CLFD, fires over 30 km^2 accounted for 30.3% of the total number of fires but contributed 91.2% of the area burned; similar results were obtained with CCL-6 for Canadian forests (30.7% and 92.5%) and Russian forests (19.8% and 89.6%) despite the non-overlapping time periods of the analysis.

Even though the fire size sample distributions for Canada and Russia were similar, the statistical tests show that they did not originate from the same distribution. This may be associated with the known differences in intensity between Canadian and Russian forest fires (Harden et al., 2000; Wooster and Zhang, 2004). However, it could also be a sampling artefact arising from the small number of years for which there are data in GFED. For example, if the data available for Russia had covered more years with

large fires, the distribution of fire sizes would be shifted to the right and would more closely match the distribution of the Canadian fires. Hence the lack of a database analogous to the CLFD prevents safe conclusions to be drawn regarding the validity of CCL results over this region.

The statistical tests show that the CCL algorithm produces a histogram of forest fire sizes closely matching that from the CLFD, and it also produces a similar probability function for Canadian non-forest, especially for the categories containing larger fires. This agreement occurs despite the CLFD recording fires from 1959-1999 while the GFED4 starts in 2001. This could indicate that, despite fluctuations in the number of fires and area burned each year, their size distribution remains essentially unchanged, an assumption implicit in the statistical tests performed.

To simplify the assimilation of the CCL database into a DVM we pooled forests and non-forest fires as identified by CCL-6 together but maintained the distinction between fires occurring in Canada and Russia.

3.2. Fire disturbance simulations with assimilated CCL fire products

To test whether the FRI is conserved between the initial and rerun version of the model where CCL fires are utilized, we calculated **BA** from 1000 years of spin-up and 112 years of transient runs (1901-2012) of LPJ-WM driven by CRU 3.0 climatology (Mitchell and Jones, 2005); we then ran the algorithm described in Section 2.2 for the full 1112 years and produced a set of fires for each year for both Canada and Russia. The FRI obtained using the new algorithm, referred to hereafter as a CCL run, closely matched the FRI obtained from the original run, demonstrating that fire can be included in a DVM in a way that retains the model structure and FRI, but is also consistent with the size distribution of burned area observations (Fig. 3). Even though the CCL run adds random spatial variability to the FRI, the average magnitude of FRI remains largely unaffected over sub-regions of both Canada and Russia.

We investigated whether this variability in FRI is caused by the short spin-up time of the DVM (1000 years) compared to the long FRI for the region (100-1000 years), which may not allow enough time for the FRI to converge to the original model value. However, even CCL runs with over 4000 years of spin-up failed to produce the original spatially smooth FRI. Only after excluding the very large fires (Categories 4 & 5 in Fig. 2) from the CCL algorithm was the spatial variability reduced: an almost

exact match to the original FRI was then achieved. This seems to indicate that the spatial variability arises from the limited number of large fires found by CCL, both because they are rare and because GFED4 is derived from data covering only a decade. The conditions imposed by the algorithm make it hard for them to be accepted under comparison with the accumulated burned area array **BAC** (Section 2.2, Step 6). As a result, these larger fires are frequently allowed to burn the same subsets of grid-cells, which hinders production of a smooth FRI across the region. Nevertheless, as Fig. 3 shows, the FRI produced by the CCL run does capture the FRI of the original run in the sub-regions of the Arctic. Possible ways to reduce the variability are discussed in Section 4.

3.3. Post-Fire Dynamics

Of greater importance is that the CCL run produces fire size characteristics consistent with those derived from EO data. Fig. 4 demonstrates this by comparing the fraction of burned area over a year of the transient run (2007) obtained from an original run of LPJ-WM, a CCL run for the same year and GFED4. As noted in the Introduction, the original LPJ-WM representation of fire (top) causes a very small fraction of most of the grid-cells to burn, and the area burned (either per grid-cell or total) remains largely unchanged in different years; such behaviour is common to many DVMs. In contrast, using the CCL methodology (centre) gives rise to fires whose sizes cover the entire range of burned areas, as shown in Fig. 2. Additionally, since the algorithm accumulates grid-cell fire potential and then consumes it when one or more grid-cells are chosen to accommodate a fire, regions are prevented from unrealistic behavior in which big fires are separated by only a short time span. The long FRI in the Arctic means that several decades need to pass after a big disturbance before a grid cell has accumulated enough fire potential in order to experience another fire.

To investigate the benefits of this new ability to simulate large fires in DVMs, a fire occurring in northwest Canada in simulation year 1910 is examined in Fig. 5. This particular fire spread over 16 grid-cells; the fraction burned was approximately 80% in the two central grid-cells and fell off towards the edges of the fire scar. Post-fire competition amongst species in the CCL run gives rise to evolution of vegetation cover that is consistent with field data (Dore et al., 2012). The fire occurred in a forest region surrounded by herbaceous vegetation (Fig. 5 top, Fire Year -1). As expected, in the year of the fire and that following (Fire Year 0 & +1), the dominant cover

switches to bare ground. By year +4, plant competition processes lead to the vegetation becoming a mixture of grass and trees, with trees, as saplings, becoming the dominant species by year +5. In contrast, biomass requires much more time to recover. In Fire Year 0 and the years immediately after (Fig. 5, middle), the biomass of the forest is similar in magnitude to that of the neighbouring grass grid-cells. As the forest regenerates, biomass slowly recovers to pre-fire levels while the fire scar remains visible in the model calculations even 50 years after the fire. Carbon fluxes are expressed through the annual Net Ecosystem Exchange (NEE), which is the net flux of carbon to the atmosphere from all possible pathways (Fig. 5, bottom). In Fire Year 0, fire emissions turn the grid-cells into strong sources whose NEE is about an order of magnitude off the scale used. Even though vegetation begins to recover, the fire scar is initially (Fire Year +2) not a strong carbon sink, since the cover is mostly grasses and saplings with limited carbon uptake rates. However, by Fire Year +10 it has developed into a marked sink, even though the surrounding region for that particular year happens to be a source. This indicates the value of this new approach for simulating carbon dynamics in the Arctic boreal zone, since these effects cannot occur in the original version of the DVM. However, as discussed in Section 4, it needs further extension before being able to account for effects such as large scale inter-annual variability.

4. Discussion

The new methodology for simulating fire disturbance in DVMs described in this paper significantly improves on current model approaches by providing a representation of fire sizes that is consistent with large scale satellite observations and requires minimal modifications to model structure. Having realistic fire sizes offers scope for investigating both post-fire carbon dynamics and effects on the water cycle caused by large fires; this is of particular interest in the Arctic due to the presence of permafrost, carbon-rich soils and an expected increase in fire activity caused by changing climate (Balshi et al., 2009a; Balshi et al., 2009b; Krawchuk et al., 2009; Stocks et al., 1998). However, our approach has certain limitations, which depend on the DVM in which the methodology is applied, the region under study and the methodology itself.

Taking advantage of this new capability requires DVMs with sufficiently rich process representations; indeed, the lack of such a capability has meant there has been no motivation for the DVMs to embody the process coupling that is set in train when severe fires occur. This is particularly true as regards the connections between fire, land cover and permafrost. It is also the reason why our post-fire analysis was qualitative and restricted to examining whether a DVM has the capability to simulate the expected ecosystem response following a fire; without all the necessary linked processes in place, it is premature to attempt a quantitative comparison of carbon and water fluxes between field data and simulations. For example, even though LPJ-WM considers permafrost, the upper boundary value for soil heat transfer is the daily air temperature provided by the climatology driver. Hence, even after a large fire that removes most of the canopy, thermal conduction is unaffected, and no account is taken of heating of the soil by incoming radiation. These shortcomings can be alleviated in an ad hoc fashion by using an extinction equation parameterized by Leaf Area Index to characterize temperature during canopy recovery but, as shown by Kantzas et al. (2013), what is really required is a more sophisticated radiative/heat transfer process. The JULES model (Best et al., 2011), for example, does consider radiative transfer through the canopy and has a recently-added permafrost representation which considers the thermal properties of organic soils (Chadburn et al., 2015), but JULES does not contain a fire component.

The CCL algorithm captures all the temporal characteristics of a fire event (dates of ignition and extinction and temporal evolution), but this information cannot be assimilated in LPJ-WM because this DVM only calculates fire effects annually. This significantly weakens the ability of LPJ-WM (and other DVMs with annual fire accounting) to model, for example, post-fire permafrost dynamics, biomass burning and litter/soil emissions, since the timing of a fire relative to summer defines its effect on carbon pools and soil heat transfer. Nevertheless, and since the algorithm described here is model-independent, a DVM with a daily fire step could be used, such as the Community Land Model (CLM) (Kloster et al., 2010), although CLM does not currently consider carbon soils or specific Arctic plant functional types like LPJ-WM. However, the temporal characteristics of fires obtained by CCL could be assimilated into a DVM by making the probability distribution of fire occurrence conditional on

day of the year or season, thus allowing post-fire dynamics to be studied at higher temporal resolutions.

The long FRI in the Arctic means that the 12 years of data in the GFED4 daily product is insufficient for adequate sampling of the rarer large fires. This can distort the local occurrence statistics and give rise to spatial variability in the simulated FRI, even though at larger scales the CCL runs agree well with the FRI produced by the original model. The restricted number of larger fires means that the algorithm tries to accommodate the same large fires over regular time intervals, which slightly alters the FRI produced by the original model run. It seems likely that this spatial variability would be reduced in regions and sub-regions with shorter FRI where acquiring statistically representative data is less problematic. Available fire data would then offer a more representative picture of the local fire regime. Alternatively, in such regions, at each position an empirical distribution can be fitted to the histogram of fire sizes identified by CCL, and this probability density function could be used for sampling; this resolves problems associated with unoccupied bins in the histogram.

Burned area data from GFED4 over the Arctic reveals that in a given year fires tend to cluster spatially (Fig.4, bottom), presumably because of fuel availability and conditions that locally favour fire ignition and propagation, such as high temperatures and winds, low precipitation and abundance of natural or anthropogenic ignition sources. In contrast, the fires simulated by assimilating the CCL algorithm in LPJ-WM (Fig. 4, centre), even though they have the correct size distribution, do not appear in clusters and give a smaller IAV in burned area than GFED4 data. This is because the assimilation preserves the original annual burned area produced by the model; since the area burned in LPJ-WM is nearly the same area every year, the IAV for a region is therefore forced to be small. The reason why simulated fires do not appear in clusters is because the location of each one is decided during assimilation based on random allocation of its point of ignition and the FRI of the region; even though the FRI produced by a DVM for each grid-cell depends on local climate conditions, it does so on long time scales and is relatively insensitive to inter-annual variations which, for example, could cause multiple fires to ignite in close proximity.

Refining the algorithm so that it simulates fire activity in accordance with the IAV and clustering exhibited by GFED4 is a daunting task, especially as lightning, which

is not considered in most DVMs, is the main ignition source at these latitudes (Stocks et al., 2002) and is projected to increase in frequency (Romps et al., 2014). Furthermore, even though it is desirable, it is not necessary for a DVM to capture the IAV and spatial variability in annual fire locations in order to make medium to long term predictions on the effects of fire activity on net carbon and water fluxes. As long as the FRI produced by the DVM has the correct magnitude and captures the trend in fire activity in accordance with climate change, and fire size is linked to a complete suite of post-fire processes, then the model is capable of accounting for the effects of fire activity on an ecosystem.

Nevertheless, there are possible ways to modify the spatial distribution and IAV of fires so they are closer to what is seen in GFED4. For example, even though LPJ-WM produces annual fires whose areas rarely exceed 3% of a grid-cell, it uses climate data and soil/litter properties to derive the fraction of area burned; this value increases or decreases in a given year, albeit slightly, depending on how favourable the conditions are for fire. This annual fluctuation could be used as an indicator of the magnitude of fire activity. Instead of selecting a fire from the CCL pool and randomly assigning it to grid-cells that can accommodate it (Step 6, Sect. 2.2), grid-cells that experience an increase in area burned during consecutive years, or in the current year and a running average of the previous ones, could be prioritized to accommodate a fire against grid-cells that experience a decrease. This would allow the model not only to generate fires with a realistic size distribution, as in this study, but also their location would be linked to regional climatic conditions, thus further improving the fire representation. The accuracy of this approach depends not only on the ability of the model to identify annual fire hotspots based on climate but also on the random component of fire activity. Regarding IAV, as mentioned in Step 5 of the methodology, all fires are normalized so the area burned in a given year of the CCL implementation matches the area burned in the original model run. This normalization was performed to preserve model cohesion, but the algorithm can operate with any user-defined IAV whilst maintaining FRI. Unfortunately, at these latitudes we do not yet have enough information to characterize statistically the temporal behaviour of burned area (either its distribution or its second order correlation properties). However, the continuing availability of suitable space-based sensors will progressively fill this knowledge gap.

Despite the limitations described above, the assimilation methodology described here gives DVMs hitherto unavailable capabilities to study post-fire behaviour under the large climatic changes projected to occur in the Arctic. As long as a DVM has the necessary processes to simulate post-fire dynamics (e.g. canopy radiative transfer, vegetation succession, permafrost-related processes and parameterization) and is correctly calibrated against field data, model runs driven by climate scenarios can now offer insights into the role of fire by answering questions such as: (1) Will permafrost recover after a big fire when the atmospheric temperature is rising, especially in regions where it is discontinuous, and what will be the effect of the projected increase in precipitation? (2) How will post-fire vegetation succession be affected at ecosystem boundaries under the greening effect in the Arctic? (3) How will evapotranspiration be affected under increases in fire activity and precipitation? (4) How will the magnitude of fire emissions vary over sub-regions, and can changes in fire activity change the sign of the land-atmosphere net carbon exchange?

Acknowledgements

This study was supported by the EU FP7 project EURUCAS (grant no. 295068), European-Russian Centre for cooperation in the Arctic and sub-Arctic environmental and climate research. E.P Kantzas acknowledges the support of Nansen Centre, St Petersburg, Russia.

References

Amiro, B. D., Logan, K. A., Wotton, B. M., Flannigan, M. D., Todd, J. B., Stocks, B. J., and Martell, D. L.: Fire weather index system components for large fires in the Canadian boreal forest, Int J Wildland Fire, 13, 391-400, 2004.

Amiro, B. D., MacPherson, J. I., Desjardins, R. L., Chen, J. M., and Liu, J.: Post-fire carbon dioxide fluxes in the western Canadian boreal forest: evidence from towers, aircraft and remote sensing, Agr Forest Meteorol, 115, 91-107, 2003.

Amiro, B. D., Orchansky, A. L., Barr, A. G., Black, T. A., Chambers, S. D., Chapin, F. S., Gouldenf, M. L., Litvakg, M., Liu, H. P., McCaughey, J. H., McMillan, A., and Randerson, J. T.: The effect of post-fire stand age on the boreal forest energy balance, Agr Forest Meteorol, 140, 41-50, 2006.

Amiro, B. D., Stocks, B. J., Alexander, M. E., Flannigan, M. D., and Wotton, B. M.: Fire, climate change, carbon and fuel management in the Canadian boreal forest, Int J Wildland Fire, 10, 405-413, 2001a.

Amiro, B. D., Todd, J. B., Wotton, B. M., Logan, K. A., Flannigan, M. D., Stocks, B. J., Mason, J. A., Martell, D. L., and Hirsch, K. G.: Direct carbon emissions from Canadian forest fires, 1959-1999, Canadian Journal of Forest Research, 31, 512-525, 2001b.

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., and Weber, J. L.: GLOBCOVER The most detailed portrait of Earth, Esa Bull-Eur Space, 24-31, 2008.

Arino, O., Casadio, S., and Serpe, D.: Global night-time fire season timing and fire count trends using the ATSR instrument series, Remote Sens Environ, 116, 226-238, 2012.
Balshi, M. S., McGuire, A. D., Duffy, P., Flannigan, M., Kicklighter, D. W., and Melillo, J.: Vulnerability of carbon storage in North American boreal forests to wildfires during the 21st century, Global Change Biol, 15, 1491-1510, 2009a.
Balshi, M. S., McGuirez, A. D., Duffy, P., Flannigan, M., Walsh, J., and Melillo, J.: Assessing the response of area burned to changing climate in western boreal North America using a Multivariate Adaptive Regression Splines (MARS) approach, Global Change Biol, 15, 578-600, 2009b.
Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Menard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description - Part 1: Energy and water fluxes, Geosci Model Dev, 4, 677-699, 2011.
Bond-Lamberty, B., Peckham, S. D., Gower, S. T., and Ewers, B. E.: Effects of fire on regional evapotranspiration in the central Canadian boreal forest, Global Change Biol, 15, 1242-1254, 2009.
Castellanos, P., Boersma, K. F., and van der Werf, G. R.: Satellite observations indicate substantial spatiotemporal variability in biomass burning NOx emission factors for South America, Atmos Chem Phys, 14, 3929-3943, 2014.
Chadburn, S., Burke, E., Essery, R., Boike, J., Langer, M., Heikenfeld, M., Cox, P., and Friedlingstein, P.: An improved representation of physical permafrost dynamics in the JULES land-surface model, Geosci. Model Dev., 8, 1493-1508, 2015.
Clark, K. L., Skowronski, N., Gallagher, M., Renninger, H., and Schafer, K.: Effects of invasive insects and fire on forest energy exchange and evapotranspiration in the New Jersey pinelands, Agr Forest Meteorol, 166, 50-61, 2012.
Dore, S., Kolb, T. E., Montes-Helu, M., Eckert, S. E., Sullivan, B. W., Hungate, B. A., Kaye, J. P., Hart, S. C., Koch, G. W., and Finkral, A.: Carbon and water fluxes from ponderosa pine forests disturbed by wildfire and thinning, Ecol Appl, 20, 663-683, 2010.
Dore, S., Montes-Helu, M., Hart, S. C., Hungate, B. A., Koch, G. W., Moon, J. B., Finkral, A. J., and Kolb, T. E.: Recovery of ponderosa pine ecosystem carbon and water fluxes from thinning and stand-replacing fire, Global Change Biol, 18, 3171-3185, 2012.
Epstein, H. E., Raynolds, M. K., Walker, D. A., Bhatt, U. S., Tucker, C. J., and Pinzon, J. E.: Dynamics of aboveground phytomass of the circumpolar Arctic tundra during the past three decades, Environ Res Lett, 7, 2012.
Giglio, L., Loboda, T., Roy, D. P., Quayle, B., and Justice, C. O.: An active-fire based burned area mapping algorithm for the MODIS sensor, Remote Sens Environ, 113, 408-420, 2009.
Giglio, L., Randerson, J. T., and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), Journal of Geophysical Research: Biogeosciences, 118, 317-328, 2013.
Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D. C., and DeFries, R. S.: Assessing variability and long-term trends in burned area by merging multiple satellite fire products, Biogeosciences, 7, 1171-1186, 2010.
Gonzalez, R. C., Woods, R. E., and Eddins, S. L.: Digital Image Processing Using MATLAB, Prentice-Hall, Inc., 2003.
Harden, J. W., Manies, K. L., Turetsky, M. R., and Neff, J. C.: Effects of wildfire and permafrost on soil organic matter and soil climate in interior Alaska, Global Change Biol, 12, 2391-2403, 2006.
Harden, J. W., Trumbore, S. E., Stocks, B. J., Hirsch, A., Gower, S. T., O'Neill, K. P., and Kasischke, E. S.: The role of fire in the boreal carbon budget, Global Change Biol, 6, 174-184, 2000.

Jia, G. S. J., Epstein, H. E., and Walker, D. A.: Greening of arctic Alaska, 1981-2001, Geophys Res Lett, 30, 2003.
Jiang, Y., Zhuang, Q., and Mandallaz, D.: Modeling Large Fire Frequency and Burned Area in Canadian Terrestrial Ecosystems with Poisson Models, Environmental Modeling & Assessment, 2012.
Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J. J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527-554, 2012.
Kantzas, E., Lomas, M., and Quegan, S.: Fire at high latitudes: Data-model comparisons and their consequences, Global Biogeochem Cy, 27, 677-691, 2013.
Keith, H., Mackey, B. G., and Lindenmayer, D. B.: Re-evaluation of forest biomass carbon stocks and lessons from the world's most carbon-dense forests, P Natl Acad Sci USA, 106, 11635-11640, 2009.
Kloster, S., Mahowald, N. M., Randerson, J. T., Thornton, P. E., Hoffman, F. M., Levis, S., Lawrence, P. J., Feddema, J. J., Oleson, K. W., and Lawrence, D. M.: Fire dynamics during the 20th century simulated by the Community Land Model, Biogeosciences, 7, 1877-1902, 2010.
Koltunov, A., Ustin, S. L., and Prins, E. M.: On timeliness and accuracy of wildfire detection by the GOES WF-ABBA algorithm over California during the 2006 fire season, Remote Sens Environ, 127, 194-209, 2012.
Koren, I., Kaufman, Y. J., Remer, L. A., and Martins, J. V.: Measurement of the effect of Amazon smoke on inhibition of cloud formation, Science, 303, 1342-1345, 2004.
Krawchuk, M. A., Moritz, M. A., Parisien, M. A., Van Dorn, J., and Hayhoe, K.: Global Pyrogeography: the Current and Future Distribution of Wildfire, Plos One, 4, 2009.
Krezek-Hanes, C. C., Ahern, F., Cantin, A., and Flannigan, M. D.: Trends in large fires in Canada, 1959-2007, Canadian Councils of Resource Ministers, Ottawa, 2011.
Li, F., Bond-Lamberty, B., and Levis, S.: Quantifying the role of fire in the Earth system - Part 2: Impact on the net carbon balance of global terrestrial ecosystems for the 20th century, Biogeosciences, 11, 1345-1360, 2014.
Mack, M. C., Bret-Harte, M. S., Hollingsworth, T. N., Jandt, R. R., Schuur, E. A. G., Shaver, G. R., and Verbyla, D. L.: Carbon loss from an unprecedented Arctic tundra wildfire, Nature, 475, 489-492, 2011.
Matson, M. and Dozier, J.: Identification of Subresolution High-Temperature Sources Using a Thermal Ir Sensor, Photogramm Eng Rem S, 47, 1311-1318, 1981.
Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int J Climatol, 25, 693-712, 2005.
Morisette, J. T., Giglio, L., Csiszar, I., Setzer, A., Schroeder, W., Morton, D., and Justice, C. O.: Validation of MODIS Active Fire Detection Products Derived from Two Algorithms, Earth Interact, 9, 1-25, 2005.
Ping, C. L., Michaelson, G. J., Jorgenson, M. T., Kimble, J. M., Epstein, H., Romanovsky, V. E., and Walker, D. A.: High stocks of soil organic carbon in the North American Arctic region, Nat Geosci, 1, 615-619, 2008.
Podur, J. J., Martell, D. L., and Stanford, D.: A compound Poisson model for the annual area burned by forest fires in the province of Ontario, Environmetrics, 21, 457-469, 2010.
Prentice, I. C., Kelley, D. I., Foster, P. N., Friedlingstein, P., Harrison, S. P., and Bartlein, P. J.: Modeling fire and the terrestrial carbon balance, Global Biogeochem Cy, 25, 2011.
Romps, D. M., Seeley, J. T., Vollaro, D., and Molinari, J.: Projected increase in lightning strikes in the United States due to global warming, Science, 346, 851-854, 2014.
Roy, D. P., Boschetti, L., Justice, C. O., and Ju, J.: The collection 5 MODIS burned area product - Global evaluation by comparison with the MODIS active fire product, Remote Sens Environ, 112, 3690-3707, 2008.

Sassen, K. and Khvorostyanov, V. I.: Cloud effects from boreal forest fire smoke: evidence for ice nucleation from polarization lidar data and cloud model simulations, Environ Res Lett, 3, 2008.
Schepaschenko, D. G., Mukhortova, L. V., Shvidenko, A. Z., and Vedrova, E. F.: The Pool of Organic Carbon in the Soils of Russia, Eurasian Soil Sci+, 46, 107-116, 2013.
Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Global Change Biol, 9, 161-185, 2003.
Stocks, B. J., Fosberg, M. A., Lynham, T. J., Mearns, L., Wotton, B. M., Yang, Q., Jin, J. Z., Lawrence, K., Hartley, G. R., Mason, J. A., and McKenney, D. W.: Climate change and forest fire potential in Russian and Canadian boreal forests, Climatic Change, 38, 1-13, 1998.
Stocks, B. J., Mason, J. A., Todd, J. B., Bosch, E. M., Wotton, B. M., Amiro, B. D., Flannigan, M. D., Hirsch, K. G., Logan, K. A., Martell, D. L., and Skinner, W. R.: Large forest fires in Canada, 1959-1997, J Geophys Res-Atmos, 108, 2002.
Sturm, M., Racine, C., and Tape, K.: Climate change - Increasing shrub abundance in the Arctic, Nature, 411, 546-547, 2001.
Tarnocai, C., Canadell, J. G., Schuur, E. A. G., Kuhry, P., Mazhitova, G., and Zimov, S.: Soil organic carbon pools in the northern circumpolar permafrost region, Global Biogeochem Cy, 23, 2009.
Thonicke, K., Spessa, A., Prentice, I. C., Harrison, S. P., Dong, L., and Carmona-Moreno, C.: The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: results from a process-based model, Biogeosciences, 7, 1991-2011, 2010.
Thonicke, K., Venevsky, S., Sitch, S., and Cramer, W.: The role of fire disturbance for global vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model, Global Ecol Biogeogr, 10, 661-677, 2001.
Valentini, R., Arneth, A., Bombelli, A., Castaldi, S., Gatti, R. C., Chevallier, F., Ciais, P., Grieco, E., Hartmann, J., Henry, M., Houghton, R. A., Jung, M., Kutsch, W. L., Malhi, Y., Mayorga, E., Merbold, L., Murray-Tortarolo, G., Papale, D., Peylin, P., Poulter, B., Raymond, P. A., Santini, M., Sitch, S., Laurin, G. V., van der Werf, G. R., Williams, C. A., and Scholes, R. J.: A full greenhouse gases budget of Africa: synthesis, uncertainties, and vulnerabilities, Biogeosciences, 11, 381-407, 2014.
van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997-2009), Atmos Chem Phys, 10, 11707-11735, 2010.
Wania, R., Ross, I., and Prentice, I. C.: Integrating peatlands and permafrost into a dynamic global vegetation model: 1. Evaluation and sensitivity of physical land surface processes, Global Biogeochem Cy, 23, 2009a.
Wania, R., Ross, I., and Prentice, I. C.: Integrating peatlands and permafrost into a dynamic global vegetation model: 2. Evaluation and sensitivity of vegetation and carbon cycle processes, Global Biogeochem Cy, 23, 2009b.
Wiitala, M. R.: Assessing the risk of cumulative burned acreage using the Poisson probability model, Gen. Tech. Rep. PSW-GTR-173, 1999.
Wooster, M. J., Xu, W., and Nightingale, T.: Sentinel-3 SLSTR active fire detection and FRP product: Pre-launch algorithm development and performance evaluation using MODIS and ASTER datasets, Remote Sens Environ, 120, 236-254, 2012.
Wooster, M. J. and Zhang, Y. H.: Boreal forest fires burn less intensely in Russia than in North America, Geophys Res Lett, 31, 2004.
Xu, L., Myneni, R. B., Chapin, F. S., Callaghan, T. V., Pinzon, J. E., Tucker, C. J., Zhu, Z., Bi, J., Ciais, P., Tommervik, H., Euskirchen, E. S., Forbes, B. C., Piao, S. L., Anderson, B. T., Ganguly,

S., Nemani, R. R., Goetz, S. J., Beck, P. S. A., Bunn, A. G., Cao, C., and Stroeve, J. C.: Temperature and vegetation seasonality diminishment over northern lands, Nature Climate Change, 3, 581-586, 2013.
Yoshikawa, K., Bolton, W. R., Romanovsky, V. E., Fukuda, M., and Hinzman, L. D.: Impacts of wildfire on the permafrost in the boreal forests of Interior Alaska, J Geophys Res-Atmos, 108, 2002.

Figures

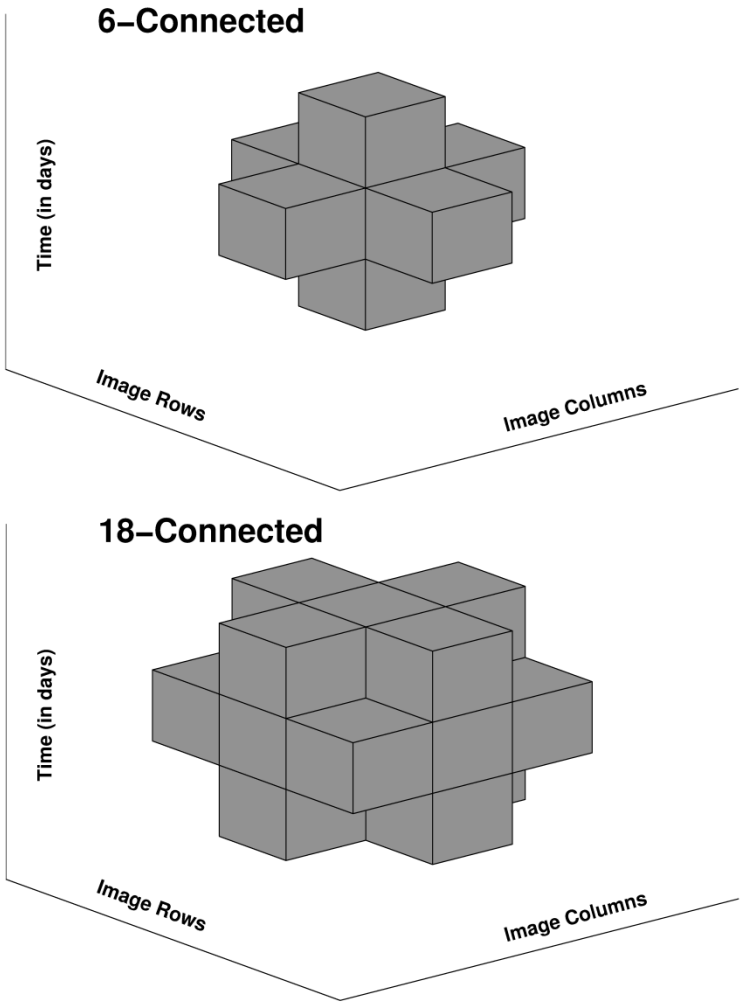

Figure 1: 6- and 18-connected pixel connectivity in 3-dimensional CCL analysis; the axes are image row, image column and day of the year.

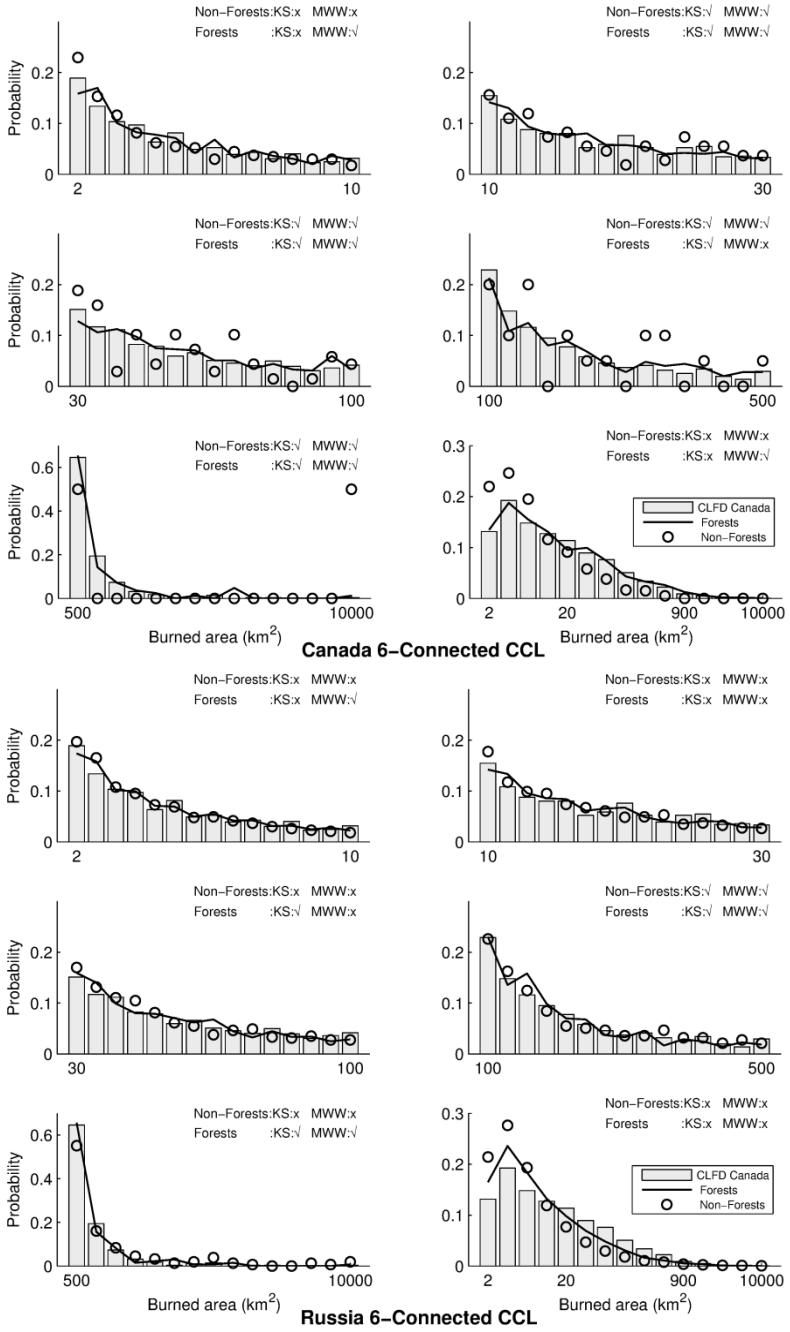

Figure 2: Histograms of area burned in each fire size category obtained from the CLFD and the application of 6-connected CCL to the GFED burned area daily product. The CLFD results only describe forest fires in Canada, while the CCL-6 results are given separately for forests and non-forests in Canada (top) and Russia (bottom). The limits of the x-axes in each figure give the range of burned area studied in each category; the x-axis in the bottom right figure for each region uses a logarithmic scale. A tick or cross shows whether the CCL-derived distributions for forests and non-forests pass or fail the Kolmogorov-Smirnov (KS) and Mann-Whitney-Wilcoxon (MWW) test when compared to their size-respective CLFD distributions.

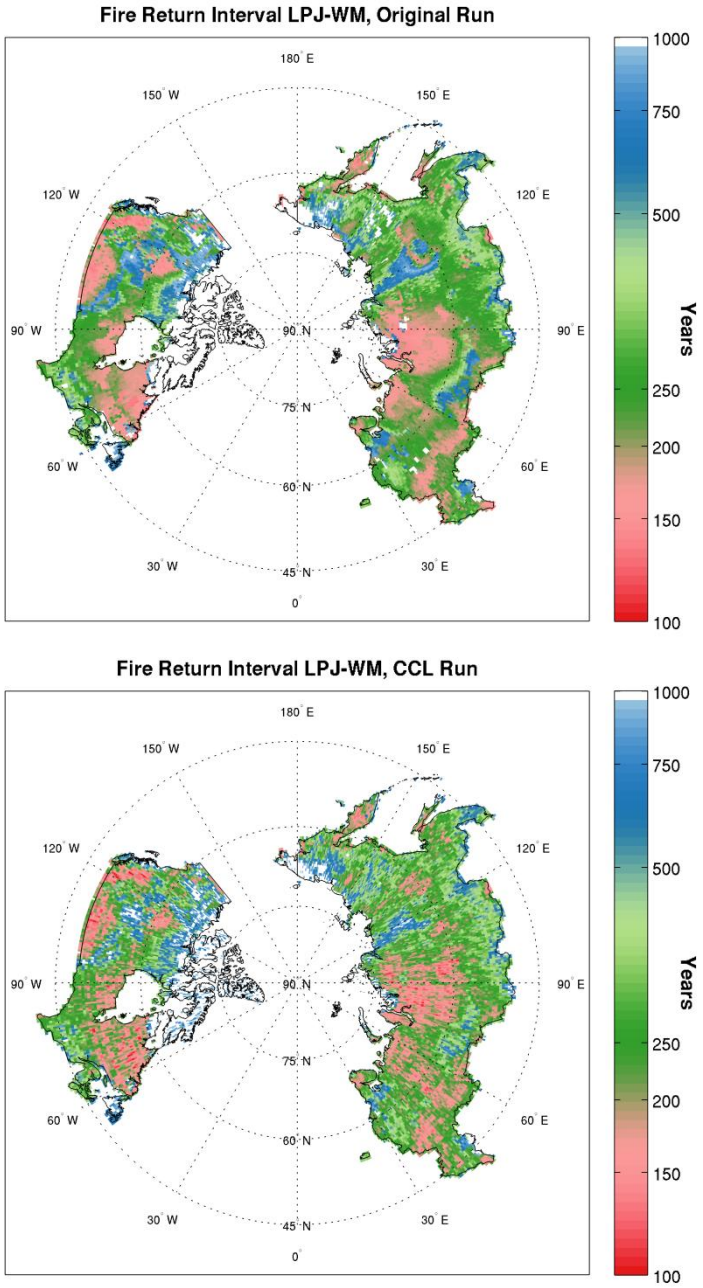

Figure 3: (top) Fire Return Interval produced by an original LPJ-WM run over a 1000 years spin-up combined with a transient run (1901-2012) for Canada and Russia. (bottom) FRI produced by a LPJ-WM run over the same period with the CCL methodology.

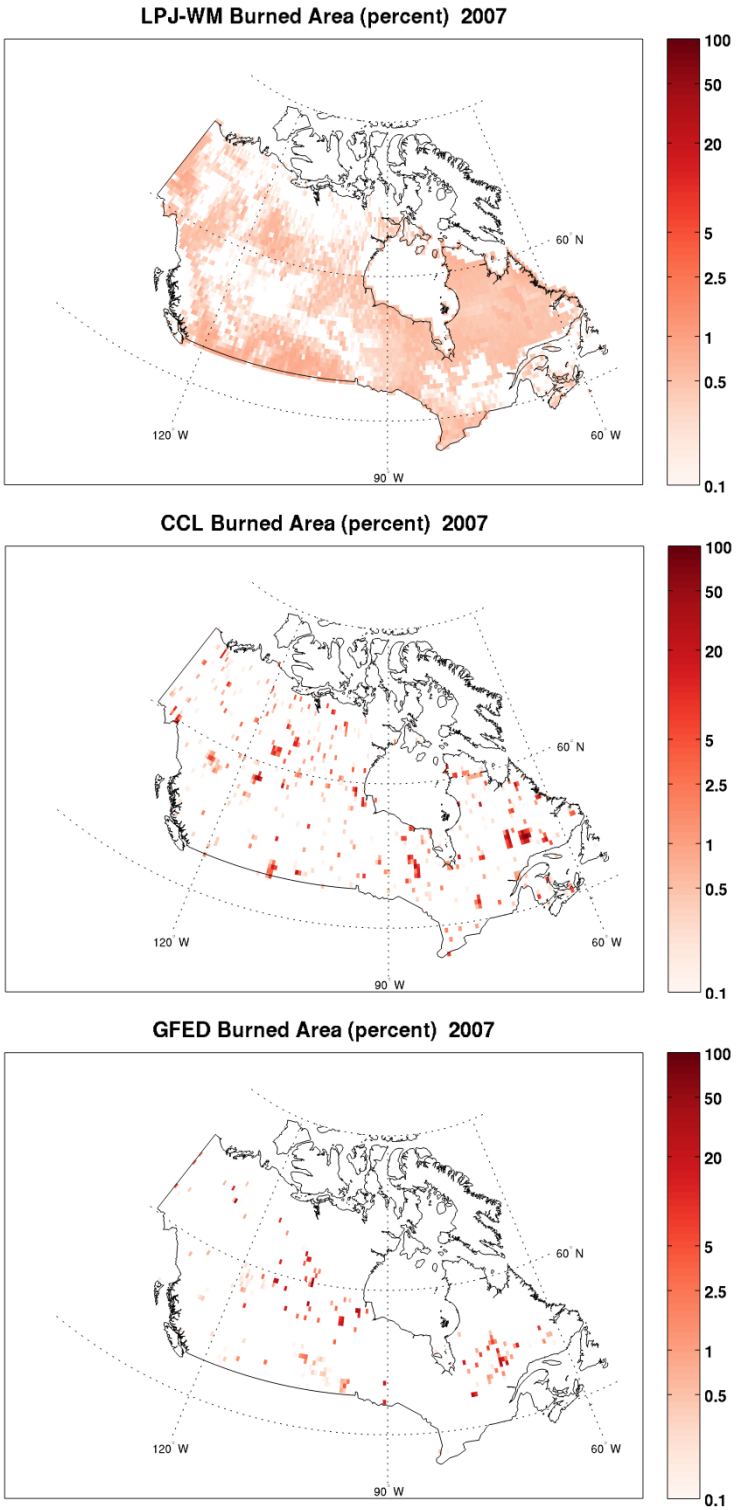

Figure 4: Fractional burned area per grid-cell (%) for 2007 in an original LPJ-WM run (top), a CCL run (centre) and GFED (bottom). Note that, since the fire is stochastic in the CCL run, different runs would produce different fires for the same year but the overall fraction of area burned would remain equal to that of the original run.

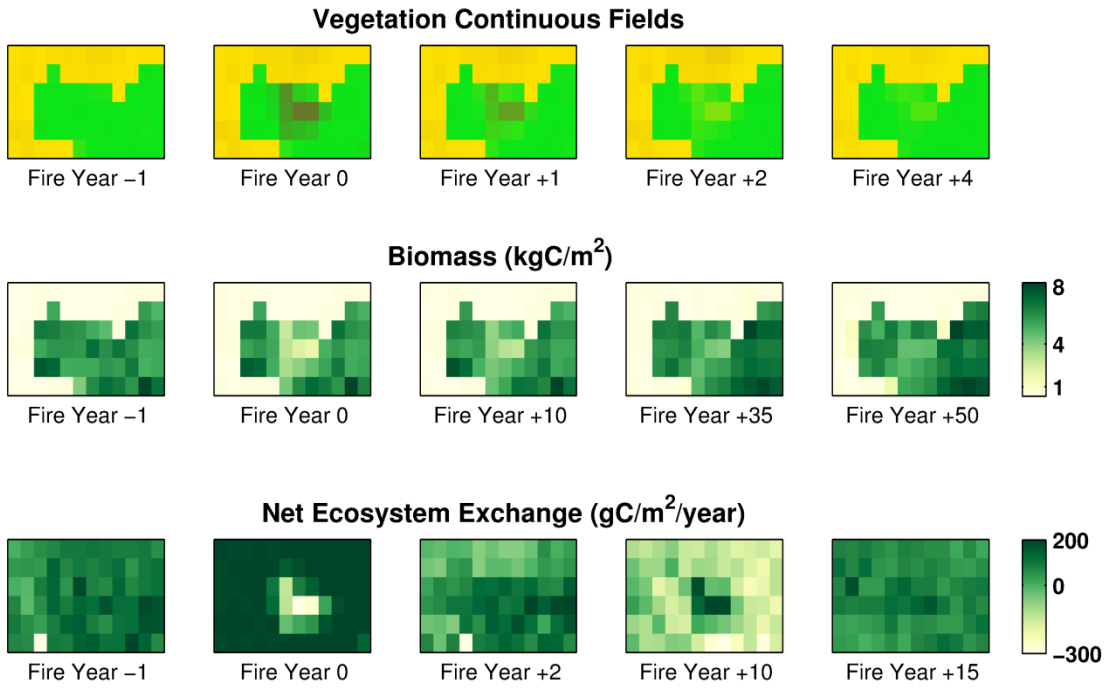

Figure 5: Post-fire evolution of the carbon stocks and fluxes after the fire disturbance indicated in Fig. 4 (centre). Top: vegetation cover in Vegetation Continuous Fields format (Green = Trees, Yellow = Grass, Brown = Bare Ground); middle: biomass density in kg of carbon m^{-2}; bottom: Net Ecosystem Exchange in g of carbon m^{-2} y^{-1}.