

## Regional adaptation of a dynamic global vegetation model using a remote sensing data derived land cover map of Russia

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2015 Environ. Res. Lett. 10 125007

(<http://iopscience.iop.org/1748-9326/10/12/125007>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 210.77.64.105

This content was downloaded on 13/04/2017 at 06:59

Please note that [terms and conditions apply](#).

You may also be interested in:

[Highlighting continued uncertainty in global land cover maps for the user community](#)

Steffen Fritz, Linda See, Ian McCallum et al.

[Three centuries of dual pressure from land use and climate change on the biosphere](#)

Sebastian Ostberg, Sibyll Schaphoff, Wolfgang Lucht et al.

[3D simulation of boreal forests: structure and dynamics in complex terrain and in a changing climate](#)

Ksenia Brazhnik and Herman H Shugart

[Assessing climate change impacts, benefits of mitigation, and uncertainties on major global forest regions under multiple socioeconomic and emissions scenarios](#)

John B Kim, Erwan Monier, Brent Sohngen et al.

[Potential change in forest types and stand heights in central Siberia in a warming climate](#)

N M Tchebakova, E I Parfenova, M A Korets et al.

[Importance of vegetation dynamics for future terrestrial carbon cycling](#)

Anders Ahlström, Jianyang Xia, Almut Arneth et al.

[Spatial variation in vegetation productivity trends, fire disturbance, and soil carbon across arctic-boreal permafrost ecosystems](#)

Michael M Loranty, Wil Lieberman-Cribbin, Logan T Berner et al.

[Satellite observations of high northern latitude vegetation productivity changes between 1982 and 2008: ecological variability and regional differences](#)

Pieter S A Beck and Scott J Goetz

## Environmental Research Letters



## LETTER

## Regional adaptation of a dynamic global vegetation model using a remote sensing data derived land cover map of Russia

## OPEN ACCESS

## RECEIVED

21 August 2015

## REVISED

28 October 2015

## ACCEPTED FOR PUBLICATION

6 November 2015

## PUBLISHED

8 December 2015

S Khvostikov<sup>1</sup>, S Venevsky<sup>2</sup> and S Bartalev<sup>1,3</sup><sup>1</sup> Space Research Institute of the Russian Academy of Sciences (IKI), 84/32 Profsoyuznaya Str, Moscow, 117997, Russia<sup>2</sup> Center for Earth System Science, Tsinghua University, Haidan District, Beijing, 100084, People's Republic of China<sup>3</sup> Author to whom any correspondence should be addressed.E-mail: [khvostikov@d902.iki.rssi.ru](mailto:khvostikov@d902.iki.rssi.ru), [venevsky@tsinghua.edu.cn](mailto:venevsky@tsinghua.edu.cn) and [bartalev@d902.iki.rssi.ru](mailto:bartalev@d902.iki.rssi.ru)

Keywords: DGVM, land cover, regional adaptation

Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



## Abstract

The dynamic global vegetation model (DGVM) SEVER has been regionally adapted using a remote sensing data-derived land cover map in order to improve the reconstruction conformity of the distribution of vegetation functional types over Russia. The SEVER model was modified to address noticeable divergences between modelling results and the land cover map. The model modification included a light competition method elaboration and the introduction of a tundra class into the model. The rigorous optimisation of key model parameters was performed using a two-step procedure. First, an approximate global optimum was found using the efficient global optimisation (EGO) algorithm, and afterwards a local search in the vicinity of the approximate optimum was performed using the quasi-Newton algorithm BFGS. The regionally adapted model shows a significant improvement of the vegetation distribution reconstruction over Russia with better matching with the satellite-derived land cover map, which was confirmed by both a visual comparison and a formal conformity criterion.

## 1. Introduction

Dynamic global vegetation models (DGVMs) are widely used as components of Earth system models, being important simulators of carbon and water exchange between land and the atmosphere, and providing a vegetation spatial distribution for further estimation of surface albedo and roughness (Cox *et al* 2000, Cramer *et al* 2001, Krinner *et al* 2005). Modelling of future climate-driven vegetation dynamics in the frame of the Arctic Climate Impact Assessment concluded that the northward boreal forest zone shift could reach up to 1000 km in the next 100 years with amplification of global warming because newly forested areas will become darker (albedo feedback) and will absorb more solar radiation (ACIA 2005). It was demonstrated using a global climate model that northward extension of boreal forest may have given a seasonal rise of 4 °C to 1 °C to global warming due to albedo feedback during the Holocene (Foley *et al* 1994). Thus, the ability of DGVMs to give a good representation of the vegetation

distribution at high latitude at large scale is key for better representation of the atmosphere–biosphere feedbacks in the Earth system.

Especially important for Earth system models is a reliable representation of the vegetation distribution in geographical regions with known impacts upon biosphere–atmosphere feedbacks associated not only with albedo variations induced by vegetation zones shifts, but also with possible alterations in large carbon pools, like Northern Eurasia. Northern Eurasia, which to a larger extent coincides with Russia, is a rather sparsely populated area with vast forests serving as a natural stabiliser of the global climate (Groisman and Bartalev 2007). Russia is home to one-fifth of the world's forested area and contains half of the world's coniferous growing stock (Shvidenko and Nilsson 1994). This globally valuable forested area in Russia is, however, under significant threat from on-going and future climate change. So, a rapid change in the mean climate at high latitudes may induce shifts of the zones of climatic suitability of boreal, coniferous, and deciduous forests in Russia, and some tree species may

face the risk of large scale extinction, reducing the stabilization ability of Russian ecosystems. Changes in climate variability may significantly influence the vegetation distribution via permafrost degradation and wildfire intensification in Russia, even if climatic means remain unchanged. While the forecast of future climate-driven vegetation distribution alterations can be performed only using global large scale vegetation dynamic models, there is a strong need to modify and adapt such models for Russia to provide reliable spatial dynamic patterns of ecosystems. The concept of a DGVM requires that a model should be able to represent the recent vegetation distribution starting from the bare soil with the assumption of panspermia (initial availability of seeds for all plants everywhere on the globe). Thus, a DGVM ideally should be able to reproduce the recent vegetation distribution in Russia as close as possible.

Earth observations (EOs) from satellites have already been recognized as the most suitable technique for collection of spatially explicit information on land cover over large territories. The land cover of Russia has been mapped on repeated occasions using remote sensing data within a number of global mapping efforts. The IGBP DISCover (Loveland *et al* 2000), UMD Land Cover (Hansen *et al* 2000), MOD12Q1 (Friedl *et al* 2002), GLC 2000 (Bartholomé and Belward 2005) and GlobCover 2009 (Bontemps *et al* 2011) are among the most well-known examples of the global land cover maps developed based on satellite remote sensing data. Besides the global land cover products, Northern Eurasian land cover has been the focus of some dedicated regional mapping exercises over the last decade. These regional land cover mapping projects have addressed scientific and political demands to improve land cover information for Northern Eurasia. The SPOT-VEGETATION data were used to produce a regional land cover map for Northern Eurasia at a 1 km spatial resolution (Bartalev *et al* 2003), which differentiates 27 classes. A hierarchical approach was developed to map the Northern Eurasian land cover using multi-temporal 500 m surface reflectance and 1 km land surface temperature data from the MODIS instrument (Sulla-Menashe *et al* 2011).

One dedicated mapping effort in particular has been aimed at the development of the TerraNorte Russian land cover (RLC) map as an advanced regional land cover product, with both spatial resolution and thematic accuracy improvements (Bartalev *et al* 2011). The TerraNorte land cover map improvements were achieved as a combined effect of a mapping method based on the locally-adaptive global mapping algorithm (LAGMA) (Bartalev *et al* 2014) being applied to the 231.6 m spatial resolution surface reflectance data acquired by the MODIS instrument Terra. The land cover map legend consists of 22 thematic classes, including 18 various vegetation types, and has been designed to take into account vegetation life forms,

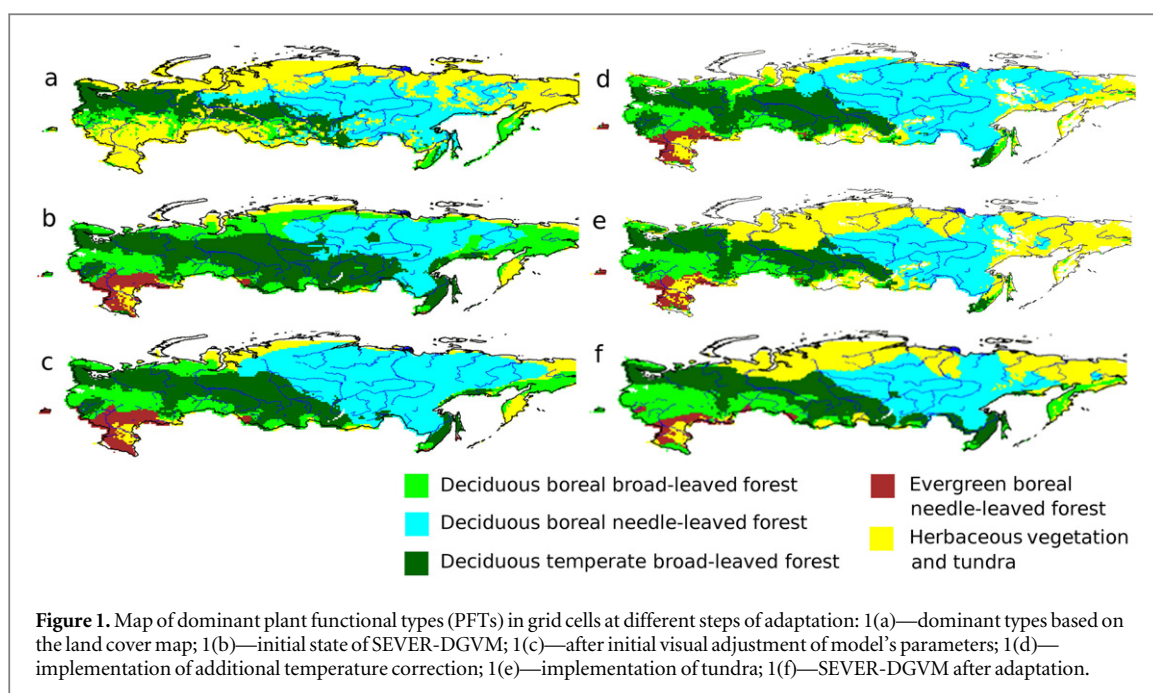
leaf types and phenological dynamics in accordance with the land cover classification system (LCCS) criteria (Di Gregorio 2005).

In some studies, remote sensing data-derived land cover maps were used for visual analysis of models' estimated vegetation distribution quality (Sitch *et al* 2003, Tchebakova *et al* 2009). However, attempts to improve vegetation models using these maps are rare. Parameterisation using formal similarity criteria between models and remote sensing land cover can potentially increase the accuracy of the predicted vegetation distribution.

In this article, the DGVM SEVER (Venevsky and Maksyutov 2007) (Russian: *север*, meaning 'North') was adapted to the territory of Russia using state-of-the-art remote sensing data derived from the land cover map TerraNorte RLC. The DGVM SEVER has developed daily soil temperature and wildfire modules which are important for future climate-driven forecasts of the vegetation distribution and carbon pools in Russia. Adaptation of the SEVER model was done in two steps—first, model routines were modified to address some notable disagreements between the model output and the land cover map. In the second step, formal optimisation of the model parameters was performed with the objective of maximisation of the spatial correlations between the vegetation distributions from the model and from the TerraNorte RLC land cover map. We present here a novel approach for obtaining a correct current distribution of boreal and temperate vegetation zones in high latitude Eurasia within an Earth system modelling framework.

## 2. The SEVER model performance analysis

The SEVER model (Venevsky and Maksyutov 2007) used in our study, is a state-of-the-art process-oriented global vegetation model that imitates processes in 10 plant functional types (PFTs) (which represent plants from tropical to boreal areas) in grid cells at a coarse resolution of 0.5 degrees. Being developed based on the Lund–Potsdam–Jena (LPJ) model (Sitch *et al* 2003), SEVER uses, in particular, daily meteorological data and different wildfire imitation routines. This DGVM imitates the behaviour of an average individual of each plant type in each grid cell throughout the simulation years. Each of these grid cells is processed independently without influence from other cells. The model starts from the 'bare soil' state, placing a minimum quantity of each vegetation type in each grid cell and giving them time to grow and achieve equilibrium (of plant parameters and vegetation distribution), after this main simulation is performed. This DGVM imitates the major processes in vegetation systems: photosynthesis, water balance, reproduction, respiration, light competition, plant mortality and so on. The general scheme of the interactions between the model's processes, described



by separate routines, follows that of the LPJ DGVM (see figure 1 in Sitch *et al* 2003).

The SEVER model uses daily climatic data from the NCEP/NCAR reanalysis dataset (Kalnay *et al* 1996) over a period from 1957 to 2006 interpolated to the model grid resolution (0.5 degrees). The sources of CO<sub>2</sub> concentration and soil type information were the same as those for the LPJ model (Sitch *et al* 2003).

To evaluate the original model quality, the simulated vegetation distribution was compared to the MODIS data-derived TerraNorte land cover map of Russia at 230 m spatial resolution (Bartalev *et al* 2011). Based on the land cover map, the appropriate vegetation types' area fractions were estimated for the model grid cells. Land cover classes were converted in accordance to the SEVER model to four functional plant types, namely C3 herbaceous and the forest types deciduous broadleaved, deciduous needle-leaved and evergreen needle-leaved. As the tundra ecosystem was not present among the SEVER model plant types, this land cover class was assigned to herbaceous vegetation. The land cover types not present in the model, such as peatlands, water, crops, recent burns, urban areas, riparian vegetation, and permanent ice and snow, were ignored in the model quality estimation. Because the model does not account for the existence of ignored types, the area fractions of other types were increased proportionally so that they fill the whole cell and their sum is equal to 1.

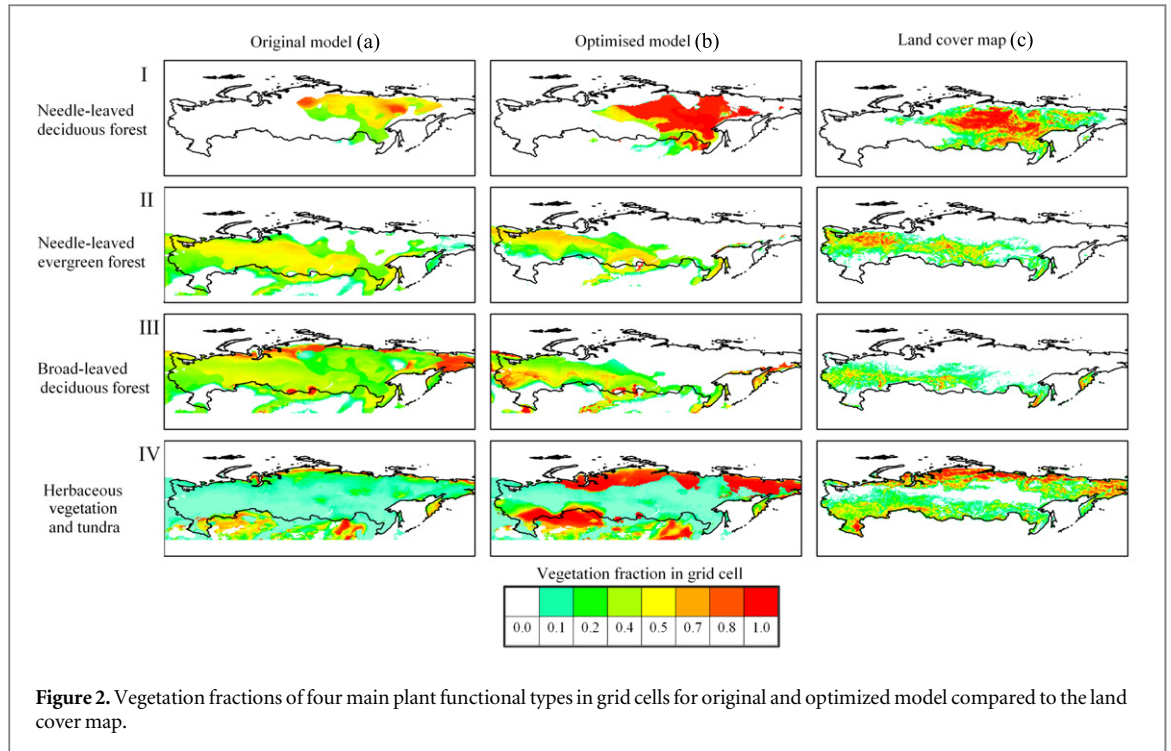
First, the original SEVER DGVM performance was verified through cross-comparison of the vegetation fractions estimated based on the model and derived from the land cover map (figures 1 and 2). While the visual cross-comparison shows that the modelling and land cover mapping results are largely in agreement

(the PFTs are mostly located close to their natural habitat), some major disagreements were also found.

In terms of dominant types (figure 1), the model estimates show major disagreements in northern, south-western and mountainous areas of Russia. According to the land cover map, 26% of Russian high northern latitudes (above 60° N) is covered by tundra (figure 1(a)), but the vegetation fraction estimated based on the model shows only 4% of this area being dominated by herbaceous types (figure 1(b)). Such discrepancies can be explained by the absence of a tundra-associated plant type in the model. In the south-western part of the country, the land cover map shows that 48% of the area is covered by steppes and grasslands, while the model estimates that 90% of this area is covered by forests (figures 1(a), (b)). This difference can be explained by extensive agricultural activities, which are not included in SEVER. Due to altitudinal zonation, vegetation in mountainous areas may show greater variety, with forests in the foothill and alpine grasslands or permanent ice cover closer to peaks. For Russian mountainous areas, the land cover map clearly shows the absence of vegetation or domination of herbaceous at higher elevation (figure 1(a)), but the model's estimates show no difference between those areas and adjacent flatlands (figure 1(b)). This discrepancy is especially noticeable in big mountainous areas like the Ural Mountains, the Middle Siberian Plateau and the Sayan Mountains.

Most of the areas dominated by forests in Russia show agreement regarding dominant types between the model and the TerraNorte land cover map. However, the ratios between different types of forests in some parts of the country significantly differ (figure 2). In the eastern part of Russia, the TerraNorte land cover map shows absolute dominance of larch (100%





of woody types), but the model estimates forests in this region with a composition of 50% larch and 50% evergreen needle-leaved and broad-leaved vegetation (figures 2(I-a), (I-c)). Also, northwestern forests of Russia consist of 80% evergreen needle-leaved and 20% broad-leaved types, and this ratio gradually shifts to 80% broadleaved and 20% needle-leaved in the southern part (figures 2(II-c), (III-c)). At the same time, the vegetation fractions estimated based on the model show a ratio closer to 55% needle-leaved and 45% broad-leaved in the north and a ratio of 45% to 55% in the south (figures 2(II-a), (III-a)). Incorrect ratios between forest types may indicate issues with imitation of competition between needle-leaved and broad-leaved in the model.

All divergences between the model and the land cover map can be summarized as follows:

- (1) Insufficient area of tundra in the Russian north.
- (2) Lack of steppes and grasslands in the southwestern part of Russia.
- (3) Underestimation of the needle-leaved deciduous forest (larch) in the eastern part of Russia.
- (4) Incorrect ratio between the broad-leaved deciduous forest and the needle-leaved evergreen forest in the western part of the country.
- (5) Incorrect vegetation types in mountainous areas.

In addition to the visual analysis, a formal quality assessment was performed using the criteria of the spatial correlations between the plant type fractions estimated by both the original SEVER model and the

land cover map. The spatial correlation ( $\text{corr}(X, Y)_\omega$ ) was estimated for all plant types being weighted by the model's grid cell area as follows:

$$\bar{X}_\omega = \frac{\sum_i \omega_i x_i}{\sum_i \omega_i} \quad (1)$$

$$\text{cov}(X, Y)_\omega = \frac{\sum_i \omega_i (x_i - \bar{X}_\omega)(y_i - \bar{Y}_\omega)}{\sum_i \omega_i} \quad (2)$$

$$\text{corr}(X, Y)_\omega = \frac{\text{cov}(X, Y)_\omega}{\sqrt{\text{cov}(X, X)_\omega \text{cov}(Y, Y)_\omega}} \quad (3)$$

where  $x_i$ ,  $y_i$  are fractions of a vegetation type in a grid cell  $i$  according to the SEVER model for the final year of modelling and the land cover map,  $\omega_i$  is the grid cell area,  $\bar{X}_\omega$  is the average weighted value, and  $\text{cov}(X, Y)_\omega$  and  $\text{corr}(X, Y)_\omega$  are the weighted covariation and correlation between variables. The model quality criterion was introduced as the sum of the squared correlations:

$$Q = \sum_i \text{corr}(X_i^{\text{map}}, X_i^{\text{model}})^2, \quad (4)$$

where  $i$  indicated iteration over all plant types: herbaceous, deciduous broad-leaved, deciduous needle-leaved and evergreen needle-leaved forests. The correlation quality criterion values for the original SEVER model are presented in table 1.

### 3. The SEVER model regional adaptation procedure

Visual qualitative adjustment of the climatic parameters DGVM for Russia improved the model performance according to both visual analysis (figures 1(a)–

**Table 1.** Correlation coefficients between the land cover map and the model's estimated vegetation fractions for the main PFTs at different stages of the SEVER model adaptation.

	Needle-leaved ever-green forest	Needle-leaved decid-u-ous forest	Broad-leaved decid-u-ous forest	Herbaceous	Quality criterion
Original model	0.48	0.39	0.0	0.34	0.50
Visual adjustment	0.51	0.48	0.19	0.33	0.63
Temperature correction	0.38	0.52	0.26	0.25	0.54
Light competition	0.41	0.52	0.30	0.25	0.59
Tundra imitation	0.44	0.63	0.35	0.29	0.80
Optimisation	0.55	0.60	0.42	0.50	1.09

(c) and to the formal quality criterion (table 1). Adaptation of the SEVER model to Russia-specific conditions was performed in two steps, including (1) model routine modification and (2) formal optimisation procedures. Modification of the model routines was performed to address problems in the description of Russian ecosystems, as mentioned in the previous section of the paper. Overall, two new components were implemented in the SEVER model and the input climate data was corrected to reflect altitude dependence. First, the air temperature change rate with elevation was modified in the input data in order to improve the vegetation distribution modelling in mountainous areas. The SEVER model uses NCEP/NCAR Reanalysis meteorological data ( $1.875^\circ \times 1.875^\circ$  lon/lat resolution) interpolated by triangulation to the model's grid cell resolution ( $0.5^\circ \times 0.5^\circ$  lon/lat). It is known that the air temperature decreases with increasing elevation, but this dependency was not taken into account in the interpolation process. To compensate for this shortcoming of the input climate data, the GLOBE digital elevation model (Hastings *et al* 1999) was used to calculate the range between the maximum and minimum elevations within each grid cell. The air temperature in a cell was decreased by a constant value for each 100 meters of altitude difference. The air temperature decrease rate initially was set at  $6^\circ\text{C km}^{-1}$  and was corrected during the formal optimisation procedure performed later.

At first, improvement of light competition was implemented in the model. The effect of competition for light between woody plant types in the original SEVER model was implemented as the proportional mortality of all plant types when the area sum of the projective vegetation cover exceeds the grid area. In our study, the plant height was set to influence the competition for light so that the proportion of surviving plants was determined by the weight coefficient, depending on the height of an average individual

$$\omega_i = A \frac{(h_i - \bar{h})}{\bar{h}} \quad (5)$$

where  $A$  is a heuristically set constant parameter,  $h_i$  is the height of an average individual of the  $i$ -th plant type, and  $\bar{h}$  is the average height of all plant types.

Formula (5) leads to a decrease in the mortality rates from shading for taller trees and an increase in mortality for smaller trees. Coefficient  $A$  was initially set to a value of 10 and was the subject of the formal optimisation procedure (see section 4).

Second, the tundra PFT was implemented in the model. To do this, additional climatic limitations were added for minimum temperature in the year, maximum wind speed in winter and average snow depth. The rationale for these climatic parameters is the observed dependence of krummholz (ecotone between treeless tundra and northern forests, primarily consisting of *Pinus Pumila* stands in Russia) distribution on wind speed (Nagano *et al* 2013), snow depth and minimum winter temperature (Okitsu and Ito 1984)

High wind speed, low temperatures and shallow snow depth are responsible for water stress, frost bite and mechanical abrasion of trees by ice particles. In each grid cell, the average value for each of these parameters over the last 20 years was calculated. A grid cell was considered to be dominated by tundra (all woody types were excluded from the cell) if the following condition was met:

$$f(x_{\max\_wind}) * (1 - f(x_{\min\_temperature})) * (f(x_{\max\_snow})) > 0.5 \quad (6)$$

where

$$f(x_i) = \begin{cases} 0, & x_i < X^{\min} \\ \frac{(x_i - X^{\min})}{(X^{\max} - X^{\min})}, & X^{\min} \leq x_i \leq X^{\max} \\ 1, & X^{\max} < x_i \end{cases} \quad (7)$$

The  $X^{\min}$  and  $X^{\max}$  values for each parameter are presented in table 3 for a total of six climatic limits. These values were obtained using the compass search optimisation algorithm (Kolda *et al* 2003), minimising the number of cells incorrectly classified as tundra/not tundra according to the dominant type of a grid cell in the land cover map. This method located the presence or absence of tundra successfully in 89% of cells above  $60^\circ\text{N}$ .

**Table 2.** Values of model’s parameters at initial stage.

	$T_{min} C^{\circ, a}$	$T_{max} C^{\circ, b}$	$T_{mort, min} C^{\circ, c}$	$T_{w-c, min} C^{\circ, d}$	$X_{min}^{snow} m, e$	$X_{max}^{snow} m, e$	$X_{min}^{temp} C^{\circ, f}$	$X_{max}^{temp} C^{\circ, f}$	$X_{min}^{wind} m s^{-1, g}$	$X_{max}^{wind} m s^{-1, g}$
Boreal broad-leafed deciduous forest	—	-2	23	—	—	—	—	—	—	—
Boreal needle-leafed deciduous forest	—	-2	23	43	—	—	—	—	—	—
Boreal needle-leafed evergreen forest	-32.5	-2	23	—	—	—	—	—	—	—
Temperate broad-leafed deciduous forest	-17	15.5	—	—	—	—	—	—	—	—

- <sup>a</sup> minimum coldest month temperature for survival.
- <sup>b</sup> maximum coldest month temperature for establishment.
- <sup>c</sup> temperature base in the heat damage mortality function.
- <sup>d</sup> larch continental index.
- <sup>e</sup> limits on average snow depth for tundra (see section 3).
- <sup>f</sup> limits on minimum temperature for tundra (see section 3).
- <sup>g</sup> limits on maximum wind speed for tundra (see section 3).

**Table 3.** Values of model’s parameters after optimisation, notations are the same as in table 2.

	$T_{min} C^{\circ}$	$T_{max} C^{\circ}$	$T_{mort, min} C^{\circ}$	$T_{w-c, min} C^{\circ}$	$X_{min}^{snow} m$	$X_{max}^{snow} m$	$X_{min}^{temp} C^{\circ}$	$X_{max}^{temp} C^{\circ}$	$X_{min}^{wind} m s^{-1}$	$X_{max}^{wind} m s^{-1}$
Boreal broad-leafed deciduous forest	<b>-24.4</b>	<b>-12.8</b>	<b>22</b>	—	—	—	—	—	—	—
Boreal needle-leafed deciduous forest	<b>-42.5</b>	-2	23	<b>36.6</b>	—	—	—	—	—	—
Boreal needle-leafed evergreen forest	<b>-27.3</b>	<b>-0.5</b>	<b>24.5</b>	—	<b>0</b>	<b>0.1</b>	<b>-35.4</b>	<b>-12.2</b>	<b>1.88</b>	<b>5.46</b>
Temperate broad-leafed deciduous forest	<b>-18.9</b>	15.5	—	—	—	—	—	—	—	—

**Table 4.** Values of additional model’s parameters at different stages of the SEVER model adaptation and air temperature altitudinal correction.

Parameter name	Initial values	Visual adjustment	Formal optimisation
Decrease of temperature for each km of elevation difference, $C^{\circ} km^{-1}$	0	6	0.011
Coefficient in light competition	1	1	9981.65
Multiplier for larch specific leaf area	1	1	0.6875

### 4. Optimisation

Formal parameter optimisation was performed after implementation of the abovementioned model modifications. Twelve model parameters, mostly defining the climatic limits for different plant types, were chosen for the optimisation. Table 3 shows the optimised climatic limits in bold, their initial values are shown in table 2. The set of those parameters also includes the coefficients for light competition and temperature correction (table 4), as previously discussed.

Global optimisation may require an excessive number of model runs, so execution time reduction was necessary. Thus, in the first step a global optimum was located for a ‘crude’ version of the model. The crude version updates only a small part (5%) of the entire model grid cells, which were chosen by maximisation of the distance between each cell and its

closest neighbours. The model performance is estimated using only this small portion. In preliminary runs of the model, it was observed that the quality criterion values for the crude and original models differed by approximately 10%. However, the correlation coefficient between the quality criterion values for different models was as high as  $R^2 = 0.88$ . This similarity made it possible to use the crude model for an exploration of our parameter space.

The efficient global optimisation (EGO) (Jones *et al* 1998) method was used to find the global optimum for the crude model. This method balances the search for an optimum value with exploration of the parameters’ space by applying the Kriging approximation method. The main idea of an approximation is to replace the original function (model) with some other simple function, which can be evaluated quickly. The Kriging method approximates functions as a sum of polynomial components and stochastic processes.

The main advantage of the Kriging method is that it exactly interpolates values for known values of the real function and that due to its stochastic nature, it can provide an estimate of the variance for this approximation at any point of the parameters' space.

An optimisation approximation function for a set of initial points is created during the EGO and after this, the maximum of the expected improvement (EI) criterion for an approximation is found. The EI criterion is calculated as follows:

$$\text{EI}(\mathbf{x}) = (f_{\min} - f(\mathbf{x}))\Phi\left(\frac{f_{\min} - f(\mathbf{x})}{s(\mathbf{x})}\right) + s\varphi\left(\frac{f_{\min} - f(\mathbf{x})}{s(\mathbf{x})}\right) \quad (8)$$

where  $\mathbf{x}$  is a vector in the parameters' space,  $f_{\min}$  is the minimum function (quality criterion) value found until now,  $f(\mathbf{x})$  is the function value for  $\mathbf{x}$ ,  $s$  is the error (variance) of the approximation at  $\mathbf{x}$ , and  $\Phi$  and  $\varphi$  are the normal cumulative distribution function and the density of the normal distribution. The EI criterion increases if the new value of the function is smaller than the minimum value found in the previous iterations (first term) and the approximation error (described by variance  $s$ ) at the new point is larger (second term).

The newly found parameters were put into the actual model and the updated approximation function was designed. These iterations are repeated until the maximum value of the EI is below a pre-defined threshold criterion, which was set at 0.0001. Use of the EI criteria and the Kriging approximation makes this method relatively fast and effective for global optimisation purposes.

Due to the observed significant difference in the performance criteria of the crude and the original models, a second step of the optimisation was necessary for obtaining at least a local optimum for the original model. This second step was based on using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm with starting parameter values providing the global optimum for the crude model. The BFGS method is a quasi-Newton local optimisation algorithm which uses a function gradient (which this method evaluates directly) and a Hessian (which this method approximates). The main purpose of this algorithm is to change parameter values in the direction of function gradient decrease, additionally using the curvature of the optimized model (Hessian) evaluated using previously obtained function values.

## 5. Results

Correction of the air temperature using the Earth's surface elevation resulted in a significant change of the vegetation distribution in the model output (figure 1(d)). Some mountainous areas can now be clearly distinguished by the domination of herbaceous

types or the absence of vegetation. Introduction of this correction, however, leads to a decrease in both the model quality criterion and the correlations for herbaceous and evergreen needle-leaved plant types (table 1). The temperature correction coefficient was significantly decreased (table 4) as well during optimisation and in effect, imitation of mountain ridges was removed from the model. This may mean that such a simplistic method of temperature correction cannot imitate mountainous areas with their high variability in elevation and complicated land cover.

The new algorithm for light competition changed the ratio between broad-leaved and needle-leaved forest types in the western part of Russia and now it is much closer to values estimated from the TerraNorte land cover map (figures 2(II-b), (II-c), (III-b), (III-c)). This modification also resulted in a small improvement in the correlations of those two types (table 1). After optimisation, the light competition coefficient was set to the maximum possible value (table 4). Such a high value assumes domination in light competition for individuals of PFTs with even a small advantage in height. Thus, the current routine for light competition can be further improved according to our results.

Suggested climatic limitations for tundra in the model simulated a non-significant overestimate of the tundra area, especially in the northern and north-eastern parts of Russia (figure 1(e)). The area of tundra after the modification is slightly larger than that in the TerraNorte land cover map (36% of area above 60° N for the model versus 27% according to the land cover map). This change also resulted in increased correlations for all vegetation types (table 1). This result demonstrates that proper implementation of tundra PFT is needed for a description of the vegetation distribution across Russia.

The optimisation process was performed in the statistical language R (R Core Team 2012) using the EGO method from the package DiceOptim (Roustant *et al* 2012) and the BFGS method from the R standard package *stats*. The language and libraries are distributed under a free software license and the scripts used in the optimisation can be provided as per request. The first step of the optimisation was completed after 1253 evaluations and as a result, the quality criterion increased from 0.80 to 1.06 (by 33%). The second step took 161 original model runs and the quality criterion increased from 1.06 to 1.09. Considering that the crude model is 20 times faster than the original model, the second step was even longer than the first step, but only inclusion of this second step in the optimisation process can guarantee that the new parameter values are at least close to the local optimum of the model.

Formal two-step optimisation resulted in a significant increase in the quality criteria (table 1), and improvements in correlation coefficients between the remote sensing data-derived land cover map and the modelling results for most vegetation types. The area



of tundra decreased to some extent and the area of Kamchatka is filled with the broad-leaved PFT after the optimisation, just like in the land cover map (figure 1(f)).

## 6. Discussion

The vegetation distribution generated by the model after the modification and two step formal optimisation is significantly closer to the TerraNorte land cover map. Significant improvement in the simulated spatial pattern of vegetation types in Northern Eurasia allows a better description of ecological services in the area under conditions of climate change, including potential carbon sequestration, change in accessibility of ecosystems and patterns of human-driven disturbances, change in biodiversity and large scale evapotranspiration and run-off changes. According to our results, the closest fit of the model to the remote sensing data-derived PFT distribution is seen in the Asian part of Russia, where human disturbance activities are mainly in the most southern part near the border. Especially important is a geographically correct representation of areas of boreal needle-leaved deciduous forests (*Larix* forests) and tundra in the Asian part of Russia. Firstly, this confirms that *Larix* forests and tundra are determined mainly by climatic variables. Secondly, specific biophysical features of spatial and temporal dynamics of carbon pools and fluxes of these two major climate-stabilizing PFTs can be further described by the model within their natural geographic areas. These specific features include interactions between vegetation, the active layer of permafrost and aerobic/anaerobic decomposition of soil organic matter for tundra and feedbacks between vegetation, the active layer of permafrost and wildfires for boreal needle-leaf deciduous forests of Russia (Venevsky 2006). This optimisation of the SEVER DGVM can also be further improved and adapted for Russia as available EO data-derived burned area and soil moisture products are implemented into the already-developed two step optimization procedure. This will allow for significant increase in the predictability of future climate-driven changes in the region important for the entire Earth system.

Differences in vegetation distribution between the optimized version of SEVER DGVM and the TerraNorte land cover map seen in the European part of Russia also provide us with important insight. We conclude that areas climatically suitable for boreal and temperate broad-leaved forests are converted to agricultural lands (described as grassland PFTs in SEVER DGVM) by human activities.

Ten of the twelve parameters taken for formal optimization of the SEVER DGVM for Russia are climate limits for the five PFTs of Russia, influencing the physiology of plants. Such a choice diminishes the possible influence of anthropogenically-determined

patterns in vegetation distribution for optimization of the parameter data set. On the other hand, this choice allows a validation of parameters which is hard to find in botanical reviews or to measure in the field.

## Acknowledgments

This work was supported by the Russian Science Foundation under Grant number 14-17-00389. We highly acknowledge comments and corrections of two anonymous reviewers, who contributed in significant improvement of our paper.

## References

- Arctic Climate Impact Assessment (ACIA) 2005 *Impacts of a Warming Arctic: Arctic Climate Impacts Assessment* (Cambridge: Cambridge University Press) p 1042
- Bartalev S, Belward A, Erchov D and Isaev A 2003 A new SPOT4-vegetation derived land cover map of Northern Eurasia *Int. J. Remote Sens.* **24** 1977–82
- Bartalev S, Egorov V, Loupian E and Khvostikov S 2014 A new locally-adaptive classification method LAGMA for large-scale land cover mapping using remote-sensing data *Remote Sens. Lett.* **5** 55–64
- Bartalev S A, Egorov V A, Ershov D V, Isaev A S, Loupyan E A, Plotnikov D E and Uvarov I A 2011 The vegetation mapping over Russia using MODIS spectroradiometer satellite data *Contemporary Earth Remote Sensing from Space* **8** 285–302 (<http://jr.rse.cosmos.ru/article.aspx?id=959>)
- Bartholomé E and Belward A 2005 Glc2000: a new approach to global land cover mapping from earth observation data *Int. J. Remote Sens.* **26** 1959–77
- Bontemps S, Defourny P, Bogaert E V, Arino O, Kalogirou V and Perez J R 2011 *Globcover 2009-Products Description and Validation Report* ESA and UCLouvain
- Cox P M, Betts R A, Jones C D, Spall S A and Totterdell I J 2000 Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model *Nature* **408** 184–7
- Cramer W *et al* 2001 Global response of terrestrial ecosystem structure and function to CO<sub>2</sub> and climate change: results from six dynamic global vegetation models *Global change Biol.* **7** 357–73
- Foley J A, Kutzbach J E, Coe M T and Levis S 1994 Feedbacks between climate and boreal forests during the Holocene epoch *Nature* **371** 52–4
- Friedl M A *et al* 2002 Global land cover mapping from MODIS: algorithms and early results *Remote Sens Environ.* **83** 287–302
- Di Gregorio A 2005 Land cover classification system: classification concepts and user manual (Rome: Food & Agriculture Org)
- Groisman P Y and Bartalev S A 2007 Northern Eurasia earth science partnership initiative (NEESPI), science plan overview *Global Planet. Change* **56** 215–34
- Hansen M, DeFries R, Townshend J R and Sohlberg R 2000 Global land cover classification at 1 km spatial resolution using a classification tree approach *Int. J. Remote Sens.* **21** 1331–64
- Hastings *et al* 1999 *The Global Land One-Kilometer Base Elevation (GLOBE) Digital Elevation Model* version 1. 0 (Boulder, CO: NOAA)
- Jones D R, Schonlau M and Welch W J 1998 Efficient global optimization of expensive black-box functions *J. Global Optim.* **13** 455–92
- Kalnay E *et al* 1996 The NCEP/NCAR 40-year reanalysis project *Bull. Am. Meteorol. Soc.* **77** 437–71
- Kolda T G, Lewis R M and Torczon V 2003 Optimization by direct search: New perspectives on some classical and modern methods *SIAM Rev.* **45** 385–482

- Krinner G, Viovy N, de Noblet-Ducoudré N, Ogée J, Polcher J, Friedlingstein P, Ciais P, Sitch S and Prentice I C 2005 A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system *Global Biogeochem. Cycles* **19** GB1015
- Loveland T, Reed B, Brown J, Ohlen D, Zhu Z, Yang L and Merchant J 2000 Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data *Int. J. Remote Sens.* **21** 1303–30
- Nagano S, Nakano T, Hikosaka K and Maruta E 2013 Pinus pumila photosynthesis is suppressed by water stress in a wind-exposed mountain site *Arct. Antarct. Alp. Res.* **45** 229–37
- Okitsu S and Ito K 1984 Vegetation dynamics of the siberian dwarf pine (pinus pumila regel) in the taisetsu mountain range, hokkaido, japan *Vegetatio* **58** 105–13
- R Core Team 2012 *R: A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing) (<http://R-project.org/>)
- Roustant O *et al* 2012 Dicekriging, diceoptim: two r packages for the analysis of computer experiments by kriging-based metamodeling and optimization *J. Stat. Software* **51** 1–55
- Shvidenko A and Nilsson S 1994 What do we know about the siberian forests? *Ambio* **23** 396–404
- Sitch S *et al* 2003 Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model *Global Change Biol.* **9** 161–85
- Sulla-Menashe D, Friedl M A, Krankina O N, Baccini A, Woodcock C E, Sibley A, Sun G, Kharuk V and Elsakov V 2011 Hierarchical mapping of northern Eurasian land cover using MODIS data *Remote Sens. Environ.* **115** 392–403
- Tchebakova N, Parfenova E and Soja A 2009 The effects of climate, permafrost and fire on vegetation change in Siberia in a changing climate *Environ. Res. Lett.* **4** 045013
- Venevsky S 2006 A method for integrated assessment of vulnerability to climate change in Siberian forests: Example of larch area Mitig. *Adapt. Strategies Global Change* **11** 241–68
- Venevsky S and Maksyutov S 2007 Sever: a modification of the LPJ global dynamic vegetation model for daily time step and parallel computation *Environ. Modelling Software* **22** 104–9