

International Conference on Computational Science, ICCS 2013

Estimating basal area of spruce and fir in post-fire residual stands in Central Siberia using Quickbird, feature selection, and Random Forests

Martin Jung^{a,*}, Susanne Tautenhahn^a, Christian Wirth^b, Jens Kattge^a

^aMax Planck Institute for Biogeochemistry, Hans-Knöll-Str. 10, 07745 Jena, Germany

^bUniversity of Leipzig, Johannisallee 21-23, 04103 Leipzig, Germany

Abstract

The occurrence and spatial arrangement of post-fire alive residual stands affect the recolonization of trees and animals of burned areas in boreal ecosystems. Because the analysis of residual stands in the field is prohibitively expensive we lack understanding on how residual stands are distributed and why. Here, we explore the use of high-resolution Quickbird satellite imagery in conjunction with in-situ measurements and machine learning techniques to map basal area of spruce and fir for two fire areas in Central Siberia, and analyze the distribution of residual stands with respect to topography.

First, an advanced feature selection algorithm which combines a genetic algorithm with guided local search is wrapped around the Random Forests regression technique, to identify suitable variable subsets out of a large number of candidate variables that were derived from Quickbird data. Second, we train and apply Random Forests using the derived variable subsets to the two fire areas to generate spatially explicit estimates of basal area for spruce and fir. Third, we analyze species specific differences and the relationship between basal area and topography using a high resolution digital elevation model from ASTER data.

Our results show that the main gradients of species specific basal area can be reproduced using Quickbird data but stress the importance of variable selection. We find associations of residual stands with topography - depressions and channels exhibit larger prevalence of residual stands than ridges or plateaus, the latter being more often subject to severe fires. We further found that the relationship between basal area and elevation tends to be reversed inside the burned area in comparison to the surrounding unburned forest. Our results suggest that local topography may control the sensitivity of ecological processes to a changing fire regime with more severe fires, and highlight the synergistic use of high resolution satellite remote sensing and machine learning methods for fire ecological applications.

Keywords: residual stands; Quickbird; basal area; feature selection; topography; Siberia; boreal

1. Introduction

Fires are a key factor in boreal ecosystems that concern nature conservation and management practices [1], the natural regeneration [2], as well as carbon dynamics [3]. Fires in the boreal forest are predicted to move towards a higher frequency, size and severity [4, 5]. Most research on fire related effects on ecosystems concentrated in

*Corresponding author. Tel.: +49-3641-57-6261
E-mail address: mjung@bgc-jena.mpg.de.

North America, while comparatively little is known for Siberia. In Central Siberia forest fires are abundant with about 1.1×10^7 ha burning each year [6].

Fires generally do not burn the entire landscape - residual alive stands, either distributed as isolated trees or patches that survived the fire are an important but often overlooked factor for the post-fire development of the area. Surviving residual stands provide important seed sources for the revegetation [7, 8]. Seed dispersal limitation could lead to changes in species composition [9] and biomass recovery [10] – which could lead to influences of residual stands on carbon fluxes [11]. Residual stands further serve as refuges for animals to recolonize the area when it becomes revegetated [12].

To understand effects of an intensifying fire regime on post fire processes and population dynamics especially knowledge of the spatial variability of burned areas and their arrangement of residual stands is needed [13]. Such analysis is extremely expensive and time consuming to conduct entirely in the field, which is particularly true for Central Siberia because the burned areas there are hardly accessible without helicopter flights. Some studies used high resolution aerial photography to analyze residual stands and counted survived trees [14, 15, 16, 17] but such data are not available for Central Siberia, and a more informative measure like basal area (fractional area covered by cross-sections of trees) would be desirable. Although highlighted as a research gap [13], we are not aware of a study that explores commercial high resolution satellite data (e.g. Quickbird) to map the basal area of different tree species of residual stands.

In this paper we aim to produce and analyze species specific maps of basal area for two fire areas in Central Siberia. To do so we integrate in-situ measurements from field campaigns, high resolution Quickbird data, and machine learning methods for feature selection and regression. Our specific objectives are:

- to assess the feasibility of producing species specific basal area maps for trees that survived the fire for burned area using Quickbird data and machine learning methods
- to identify species specific differences of the spatial distribution of pre-fire basal area
- to assess the spatial distribution and structure of residual stands with respect to topographic variations

2. Data and Methods

2.1. Data

Field data. We studied two burned areas located near to Zotino (about 5 to 10 km east of the Yennissey river) in the *Abies sibirica* (fir) and *Picea obovata* (spruce) dominated dark taiga zone of the Central Siberian Plateau (Figure 1). The fire area F4 of 8.4 km² burned in 1990, the other one F6 of 23.4 km² burned in 1999, while at least parts burned already before in 1965 (ages determined dendrochronologically). During two extensive field expeditions in 2007 and 2008 we installed in total 157 plots (71 on F4, 86 on F6) with GPS coordinates along transects through the burned areas as well as in the unburned forest edge and measured the diameter at breast height *DBH* of fir and spruce trees with *DBH* > 10cm. Here, only trees that survived the fire are accounted for (no regeneration) because we are primarily interested in their role as potential seed sources (see [9]). The plot sizes A_i were commonly 14 x 14 meters but were adapted to the local density of individuals i of different size classes and species (details are available in [9]).

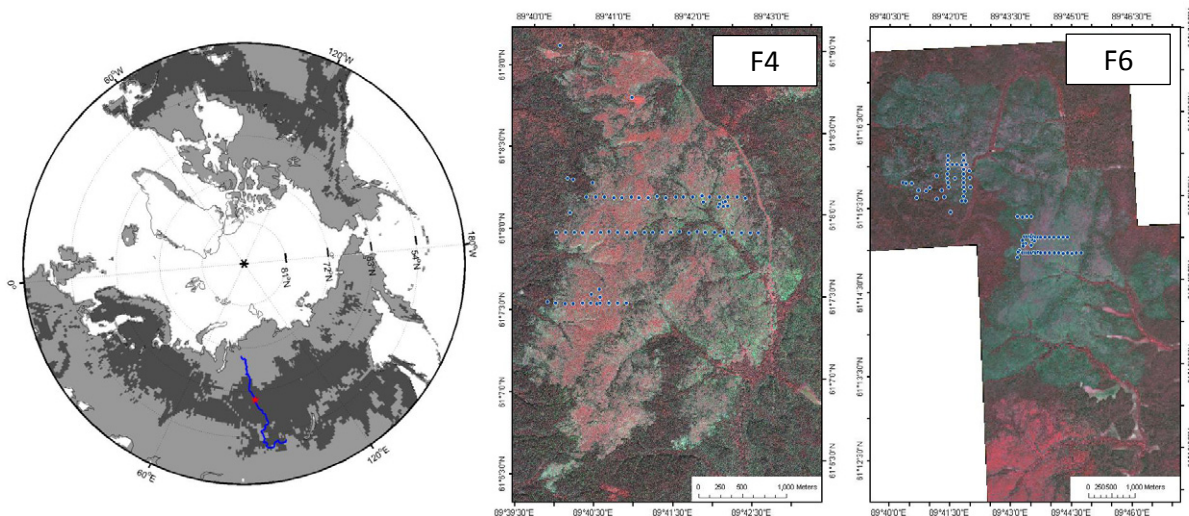


Fig. 1. Study site and plot location drawn on false color Quickbird RGB composites.

For each plot and species, we calculated the basal area for pre-fire seed trees of species s as:

$$BA = \sum_{i=1}^N 0.25 \cdot DBH_i^2 / A_i \cdot \pi \cdot p_i, \tag{1}$$

where p_i is the probability of tree i being recruited pre-fire (obtained from a statistical model based on extensive dendrochronological and biometric data from the fire areas [9]).

Quickbird data. We acquired Quickbird satellite data from two different acquisition dates for F4 and F6 because both fire areas were not covered by one single Quickbird scene. The Quickbird data were taken on 24th June 2003 and 30th July 2003 for both F4 and F6. Quickbird provides a ultra-high spatial resolution images with pixel sizes of 0.6 m in the panchromatic channel and of 2.4 m multispectral. Quickbird measures in four spectral ranges: 450-520 nm (blue), 520-600 nm (green), 630-690 nm (red), and 760-900 nm (near infrared, NIR). The Quickbird products were radiometrically calibrated using ENVI 4.5.

We calculated the normalized difference vegetation index (NDVI, [18]), enhanced vegetation index (EVI, [19]), modified enhanced vegetation index that omits the blue band (EVI2, [20]), and the modified soil adjusted vegetation index (MSAVI, ref):

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{2}$$

$$EVI = 2.5 \cdot \frac{NIR - RED}{NIR + 6 \cdot RED - 7.5 \cdot BLUE + 1} \tag{3}$$

$$EVI2 = 2.5 \cdot \frac{NIR - RED}{NIR + 2.4 \cdot RED + 1} \tag{4}$$

$$MSAVI = \frac{2 \cdot NIR + 1 - \sqrt{(2 \cdot NIR + 1)^2 - 8 \cdot (NIR - RED)}}{2} \tag{5}$$

Digital elevation model. We used a digital elevation model (ASTER GDEM version2) to analyze how residual stands vary with topography. It was acquired from <http://gdex.cr.usgs.gov/gdex/>, and comes with a spatial resolution of 1 arc second. We reprojected the DEM to UTM, Zone 45 North to conform with the Quickbird based basal area maps, and resampled it to a spatial resolution of 28m resolution.

2.2. Methods

2.2.1. Machine learning methods

Random Forests regression. We use the non-parametric regression algorithm Random Forests [21] to find a model that predicts basal area from the explanatory variables. Random Forests is an ensemble of $nTrees$ regression trees, where the individual trees were trained based on a bootstrap sample [22] of the training data set. The individual regression trees are fully grown without pruning until the leaf nodes inhabit at least $S_{splitMin}$ examples. Random Forests makes further use of the Random Subspace Method [23], where only a random subset of predictors $nS_{splitVar}$ is searched for the best split at a given node. One advantage of Random Forests is that an unbiased estimate of the performance can be obtained from the training using the out-of-bag examples of each tree in the ensemble. Here, each tree predicts the examples that were not used in constructing the tree, and averaging all available predictions per example provides the cross-validated prediction of the response variable Y_{cv} .

Feature selection algorithm. We use the Guided Hybrid Genetic Algorithm (GHGA, [24]) to identify suitable variable subsets for the prediction of basal area. GHGA is a feature selection algorithm that combines a Genetic Algorithm with guided variable elimination and is tailored to the problem where (1) the number of variables required for a good fit is not known, (2) the evaluation of the cost function is expensive because it involves frequent trainings of a machine learning algorithm, and (3) one cannot assume that uninformative variables do not harm (i.e. in sparse domains with a large ratio of number of variables to training examples). GHGA is wrapped around a base learning algorithm, here Random Forests, and repeatedly calls it with different feature subsets to evaluate the performance. GHGA uses a dynamic cost function j , which is the sum of the number of included variables n_{c^1} for feature subset c , and a penalty term p for the suboptimal performance measured by the modeling efficiency m [25]. The penalty term depends further on a tolerance parameter ϵ , and the largest modeling efficiency found so far $\max(M_A)$.

$$j(c) = n_{c^1} + p(m(c), \max(M_A), \epsilon), \quad (6)$$

$$p(m(c), \max(M_A), \epsilon) = e^{\left[\frac{\max(M_A) - m(c)}{\epsilon}\right] \frac{\ln(2)}{\epsilon}} - 1 = 2^{\frac{\max(M_A) - m(c)}{\epsilon}} - 1 \quad (7)$$

$$m(c) = 1 - \frac{\text{var}(Y - Y_{CV}(c))}{\text{var}(Y)} \quad (8)$$

where Y is the observed response variable, and Y_{CV} is the cross-validated modeled response variable. The behavior of the cost function is such that the penalty for suboptimal performance is small as long as it stays within the tolerance ϵ , and rapidly increases when the performance is poor. Hence, the cost function value is primarily controlled by the penalty term when the performance associated with the current feature subset is poor (outside the tolerance margin: $\max(M_A) - m(c) > \epsilon$) and otherwise by the number of included features. Since $\max(M_A)$ changes over the course of the search, cost function values also change and are recomputed in such cases (dynamic cost function).

2.2.2. Estimating basal area

General approach. We first compute various statistical measures from Quickbird data for all plot locations to yield a large set of potentially informative predictor variables. Then we use a feature selection algorithm to identify suitable variable subsets that optimally predict pre-fire basal area for spruce and fir. The identified variable subsets are then used to compute spatially explicit basal area estimates using the Quickbird data only. We use a digital elevation model to visualize and analyze relationships between topography and basal area of the burned areas.

In our approach we do not extract explicitly different vegetation types or what grew before or after the fire from the satellite data in the first place; this is rather implicit in the way the response variables are formed and it is the task of the machine learning algorithms to find appropriate relations with the predictor variables.

Candidate predictors. For each plot location, we extracted Quickbird data for a 14x14m window for the available scenes and calculated mean, variance, coefficient of variation (standard deviation / mean), and percentiles (1st, 5th, 10th, 25th, 75th, 90th, 95th, 99th) for the panchromatic channel, all multispectral channels, and all vegetation indexes. The mean value and various measures that describe the distribution of the values within the 14x14m window should help to pick up both, differences in foliage, as well as differences in stand structure. In addition, shadow effects that can be useful for biomass estimation [26] are captured implicitly. We further considered acquisition date and fire area as categorical predictors such that we yield a set of 101 variables.

Feature selection and training. GHGA was run separately for spruce and fir using the cost function and default settings [24]. The termination criterion for GHGA was that it stops if no improvements were found within the last 500 function evaluations (trainings). We used an ensemble of 120 trees for Random Forests, and set $S_{splitMin} = 1$ (determined empirically). We controlled the random number generator such that in each training using a different feature subset, exactly the same bootstrap samples are taken to minimize subtle differences in performance due to the stochastic nature of Random Forests. The final training of Random Forests uses the feature subset with the smallest cost function value returned by GHGA.

Mapping basal area. The required variables returned by GHGA for spruce and fir were computed for all Quickbird scenes using a 14x14m moving window, and the trained models for spruce and fir were applied. The resulting basal area maps for each Quickbird scene were mosaicked using ENVI 4.5 to archive a full coverage of the burned areas F4 and F6 respectively, while we used a histogram matching algorithm to remove residual differences due to acquisition dates. The basal area mosaics were resampled to 28m using bi-linear interpolation for subsequent analysis that involves a digital elevation model at that resolution.

3. Results and discussion

3.1. Feature selection, training, and validation

The feature selection tested 3802 and 4302 different combinations of variables for spruce and fir respectively. The results show clearly the importance of feature selection (Figure 2) - for spruce the model containing all variables (101) is substantially worse (modeling efficiency of 0.28) in comparison to the best variable subset with 6 variables (modeling efficiency of 0.55). This is typical for a training data set with a large ratio of candidate variables to examples [27].

We infer variable importance of the final variable by computing two metrics (Figure 2): the change in modeling efficiency if a variable is dropped from the set ('leave out'), and the change in mean squared error if a variable is randomly permuted (i.e. destroying the information content). The first metric measures the relevance or added value of a variable, while the second metric measures the sensitivity of the model to the variable. For spruce we find fire area number (the categorical variable that distinguishes the two fire areas) as most relevant, while the model seems most sensitive to the 10th percentile of the red reflectance. For fir, the permutation and leave out test indicate the largest importance of 75th percentile of the blue and 1st percentile of the panchromatic channel. The information by the 1st percentile of the panchromatic channel might reflect shadow fractions that yields information on biomass [26]. For both species we note that the selected variables are mainly percentiles that are rather low or high while the median or mean was never selected. This indicates the usefulness of structural information rather than average spectra for predicting basal area from high resolution remote sensing data.

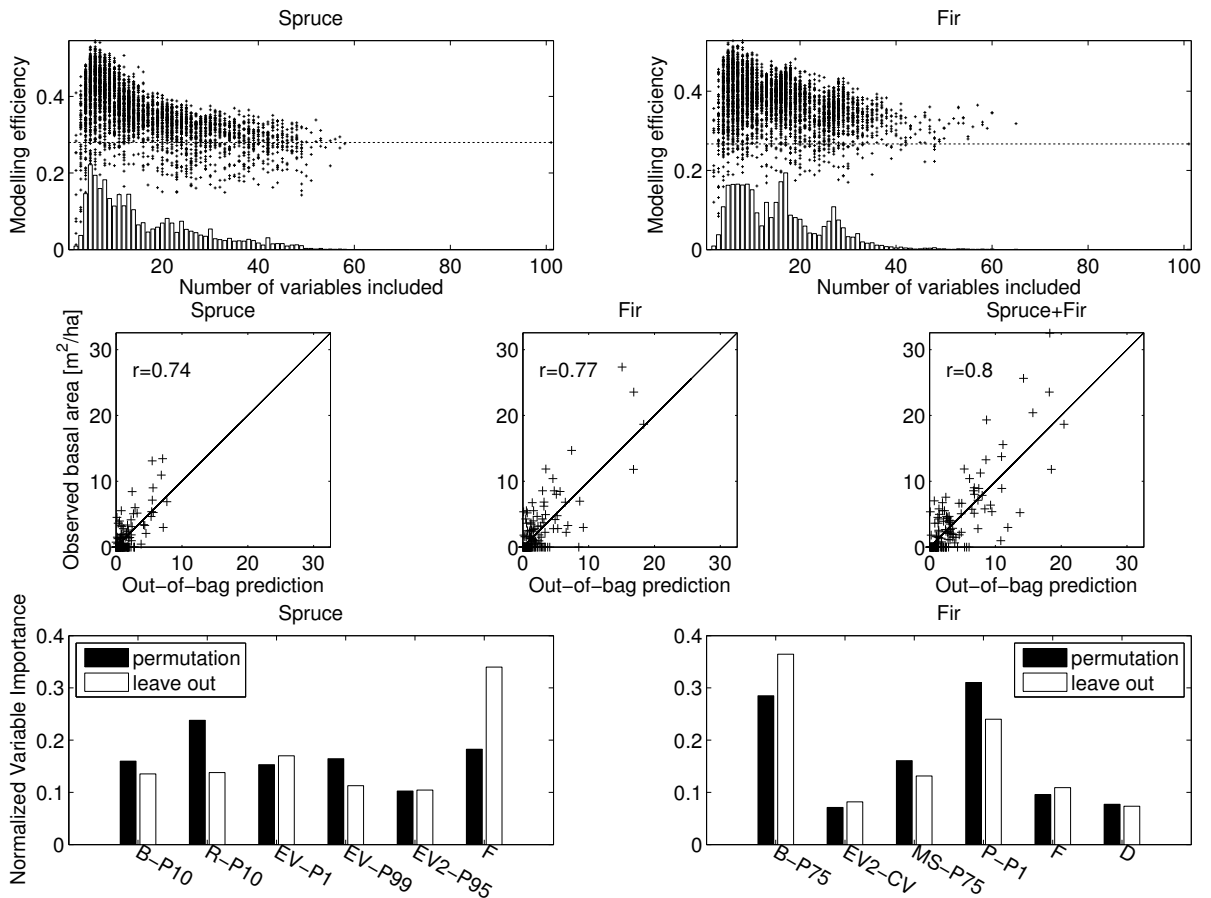


Fig. 2. Results from feature selection, variable importance estimation, and model evaluation. Top left: Modeling efficiency as a function of the number of included variables for all tested combinations for spruce. The bars represent the frequency (scaled to match y axis) of tested combinations per number of variables. Top right: same as previous for fir. Middle row: Cross-validated predictions vs observed basal area for spruce, fir, and the sum of spruce and fir. Lower left: Estimates of variable importance for spruce. B-P10: 10th percentile of blue reflectances; R-P10: 10th percentile of red reflectances; EV-P1: 1st percentile of EVI; EV-P99: 99th percentile of EVI; F: fire area number. Lower right: Estimates of variable importance for fir. B-P75: 75th percentile of blue reflectances; EV2-CV: coefficient of variation of EVI2; MS-P75: 75th percentile of MSAVI; P-P1: 1st percentile of panchromatic channel; F: fire area number; D: acquisition date index.

The cross-validation of the trained models show correlations of 0.74, 0.77, and 0.8 with the observations for spruce, fir, and the sum of the two. This result is acceptable given that we aim at estimating basal area for individual tree species for diverse spots rather than total biomass including deciduous pioneer trees (which are ubiquitous) and the understorey. Some error is also likely introduced by insufficiently precise GPS coordinates that are difficult to obtain in Central Siberia. Additional information such as multi-temporal complete coverage of the fire areas during key phenological stages of deciduous trees and the understorey (spring, summer, autumn) as well as higher order textural measures of the spatial arrangement could have helped to improve the results but was not available or not explored.

3.2. Spatial estimates of species specific basal area and its topographic control

The species specific basal area maps reveal differences in the distribution of basal area for spruce and fir (Figure 3). In the unburned region of F4 the basal area of spruce shows a much more homogeneous distribution than the basal area of fir which is much lower in the eastern part in comparison to the northern and south western part of the region. Within the fire area the fraction of spruce is higher around the residual stands than in the parts with little post-fire basal area (high fire severity). This might either suggest that spruce is slightly more resistant

to fire with low severity than fir, or might indicate different moisture preferences of spruce and fir. The former and the latter would represent direct and indirect topographic influences respectively. Satellite images obtained before the fire would be necessary to clarify the relative contributions of direct and indirect topographic influences on fir and spruce basal area after the fire.

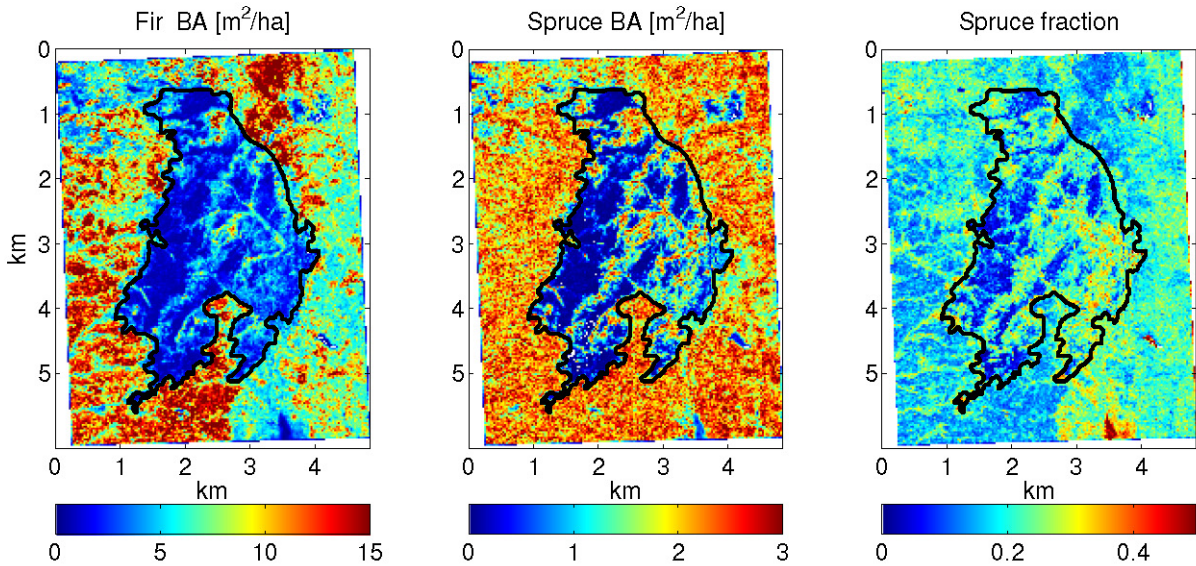


Fig. 3. Basal area maps of spruce and fir and the spruce fraction (which is calculated by $BA_{spruce}/(BA_{spruce} + BA_{fir})$) for F4.

The visual interpretation of the distribution of residual stands with topographic variations reveals two main patterns (Figure 4) that conform with the impressions from the field campaign. firstly, residual stands are nearly absent in some parts that are characterized by higher altitudes with ridges or small plateaus. Secondly, residual stands often occur moving along depressions or channels at lower altitudes. To further analyze these impressions we produced maps of moving 1x1km window correlations between basal area and elevation and find that the correlations are predominantly negative within the burned areas and predominantly positive outside the burned areas (Figure 5).

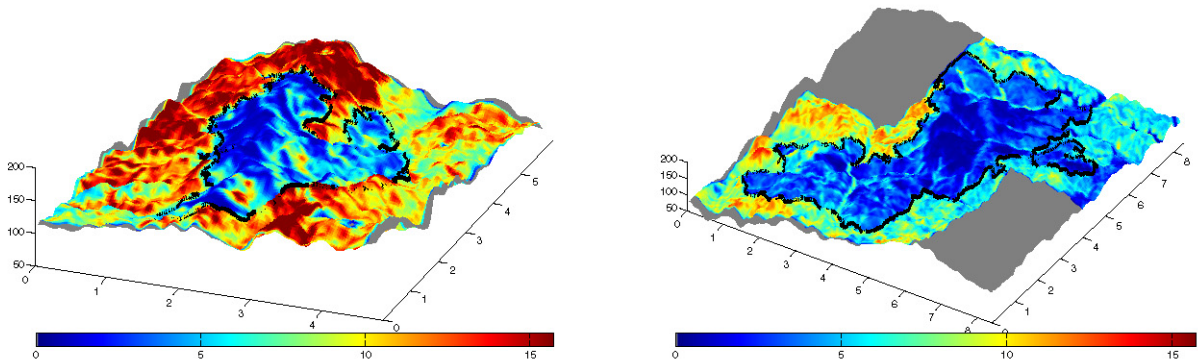


Fig. 4. Basal area maps (spruce + fir) of F4 (left) and F6 (right) projected on a digital elevation model. Spatial distance are in km; elevation is in m; basal area (color code) is in m^2/ha .

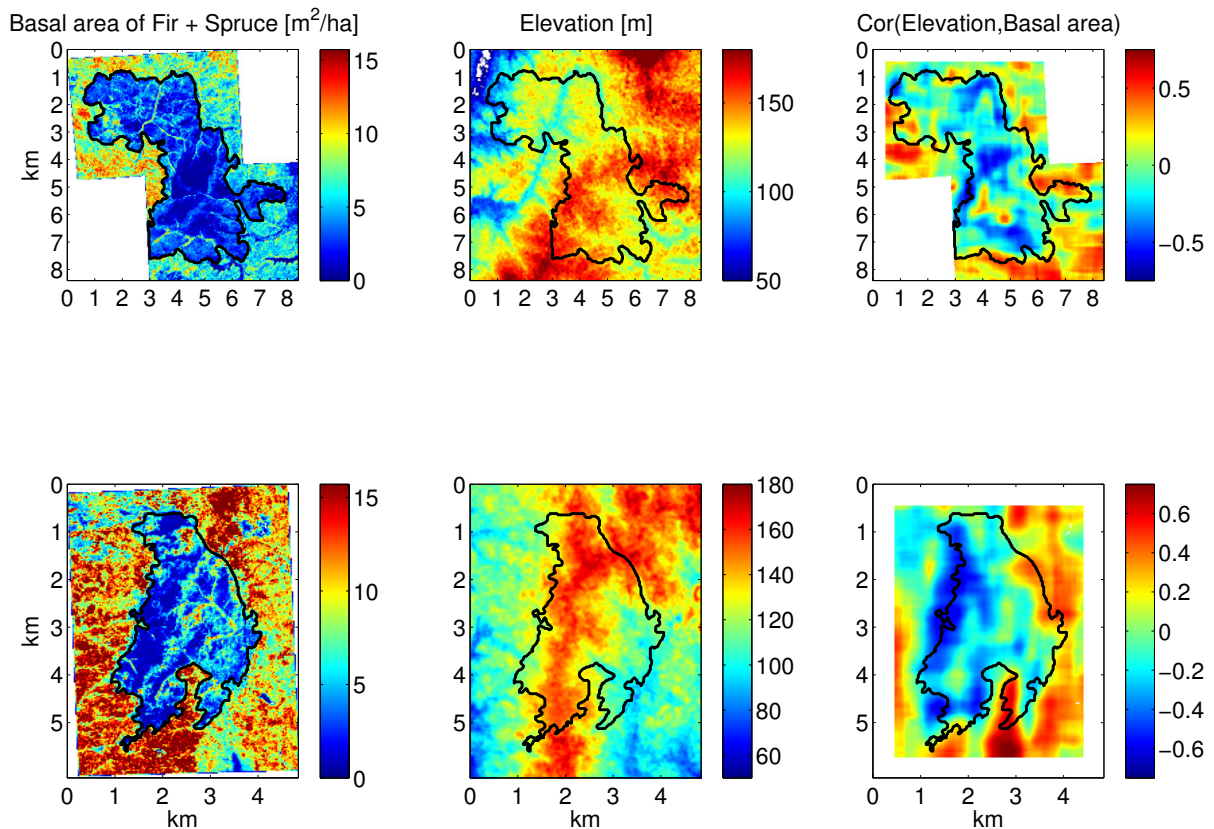


Fig. 5. Basal area [m^2/ha] of fir + spruce, elevation, and the moving window correlation between elevation and basal area (moving window of 1 km²) for the two burned areas F6 (top) and F4 (bottom).

Increasing basal area with elevation in the unburned parts suggests better growth conditions here, presumably because of less wet soils in comparison to the depressions and channels that often had gley and pseudogley soils (not shown). Fire spread due to topography is lower on ridges and plateaus, which together with enhanced fuel load and drier conditions likely caused severe fires there. The wetter conditions in the depressions and channels along with less fuel availability likely caused less severe fires there with more residual stands, which is consistent with previous findings that drainage conditions are the primary control of direct fire emissions in boreal ecosystems [3].

The topographic control of the distribution of residual stands may have some important implications for the carbon dynamics and the successional pathways of fire areas in Central Siberia. The enhanced prevalence of residual stands in the depressions at low altitudes serves as a buffering mechanism for carbon emissions, which is likely a challenging feature for global fire models and might cause overestimated fire emissions.

The coniferous species in Central Siberia need to recolonize the burned area by seeding into the area, but exhibit rather low dispersal distances. Therefore, the distribution of residual stands is crucial for the regeneration of the boreal forest, at least for fire areas of a few square kilometers or larger where the edge of the unburned forest is far away for most places in the burned area. Our findings that residual stands tend to be less frequent or absent on plateaus and ridges where fire severity was large might explain the in the field observed prevalence of deciduous forests (Danillo Mollicone, pers. comm.) on plateaus and ridges. Deciduous trees (primarily birch) are pioneers that typically dominate approximately the first 100-150 years after fire in Central Siberian dark taiga [28]. In contrast to spruce and fir, birch is much less limited by dispersal. Due to the high fire severity insufficient seed sources might be available for the conifers to recolonize the plateaus and ridges, possibly leading to deciduous forests that are stable over a longer time period in comparison to more lowland regions with abundant seed sources.

Recently, it was postulated that the successional pathways in Central Siberia might shift towards more deciduous

ous species dominated trajectories if fires become more severe and leave fewer coniferous seed sources within the burned area [9]. Such a scenario would have further implications to regional and possibly global climate via biogeochemical and especially biophysical (esp. albedo) feedbacks. Residual stands could be classified into 'event specific' refugia that are due to fire characteristics and the stochastic nature of fire spread, and 'inherent' refugia that largely originate from the topography. Our results indicates the occurrence of such 'inherent' refugia such that the local topography affects the sensitivity to such a potential biome shift, with regions of a larger topographic complexity being less sensitive than regions with lower topographic complexity.

4. Conclusions

We used in-situ measurements in conjunction with high resolution satellite imagery from Quickbird and machine learning approaches to map basal area for spruce and fir of two burned areas in Central Siberia. Our approach demonstrated that estimating species specific basal area in diverse ecosystems is feasible. Multi-temporal images, advanced textural measures and potentially other data sources such as LIDAR would likely improve our results.

We found pronounced relationships of the distribution of residual stands with topography. Residual stands were often located in channels and depressions and often absent on better drained uplands due to the direct and indirect topographic controls on fire spread and severity. Topographically induced variations of residual stands implies different landscape scale sensitivities of ecological processes like recolonization after fire to (changing) fire regimes. An extensive analysis based on many fire areas of different ages, sizes, and structures would be desirable to assess the relevance of this effect for the entire Central Siberian dark taiga.

Acknowledgements

We thank Anatoly Prokushkin, Danilo Mollicone, Marina Bryukhanova, Alexey Sharper, and Sergey Verkhovets for their great help in preparing, and organising the field expeditions. We appreciate the hard work under harsh conditions of Anja Kahl, Jan Hertwig, Waldemar Ziegler, Ulrich Pruschitzki, Norman Gentsch, Luisa Hiese, Sergey Titov, Vladimir Kislytsyn, Kolya Savaronsky, and Roman Bachman in the field. Henrik Hartmann, Angela Günther, Corinna Höhl assisted the dendroecological analysis. The project was funded by the Max-Planck Society.

References

- [1] J. Niemelä, Management in relation to disturbance in the boreal forest, *Forest Ecology and Management* 115 (1999) 127–134.
- [2] E. Shorohova, T. Kuuluvainen, A. Kangur, K. Jõgiste, Natural stand structures, disturbance regimes and successional dynamics in the Eurasian boreal forests: a review with special reference to Russian studies, *Annals of Forest Science* 66 (2) (2009) 201–201. doi:10.1051/forest/2008083.
- [3] J. Harden, S. Trumbore, The role of fire in the boreal carbon budget, *Global Change Biology* 6 (2000) 174–184.
- [4] F. Chapin, A. McGuire, Arctic and boreal ecosystems of western North America as components of the climate system, *Global Change Biology* 6 (2000) 211–223.
- [5] A. J. Soja, N. M. Tchepakova, N. H. French, M. D. Flannigan, H. H. Shugart, B. J. Stocks, A. I. Sukhinin, E. Parfenova, F. S. Chapin, P. W. Stackhouse, Climate-induced boreal forest change: Predictions versus current observations, *Global and Planetary Change* 56 (3–4) (2007) 274–296. doi:10.1016/j.gloplacha.2006.07.028.
- [6] Y. Kajii, Boreal forest fires in Siberia in 1998: Estimation of area burned and emissions of pollutants by advanced very high resolution radiometer satellite data, *Journal of Geophysical Research* 107 (D24) (2002) 4745. doi:10.1029/2001JD001078.
- [7] S. Tautenhahn, H. Heilmeyer, M. Jung, A. Kahl, J. Kattge, A. Moffat, C. Wirth, Beyond distance-invariant survival in inverse recruitment modeling: A case study in Siberian *Pinus sylvestris* forests, *Ecological Modelling* 233 (2012) 90–103. doi:10.1016/j.ecolmodel.2012.03.009.
- [8] D. Greene, E. Johnson, Tree recruitment from burn edges, *Canadian Journal of Forest Research* 30 (8) (2000) 1264–1274. doi:10.1139/cjfr-30-8-1264.
- [9] S. Tautenhahn, M. Jung, A. Prokushkin, J. Lichstein, S. Bohlman, J. Kattge, A. Kahl, C. Wirth, Seed dispersal limitation controls successional pathways in Central Siberia: Implications for a biome shift under an intensifying fire regime.
- [10] M. C. Mack, K. K. Treseder, K. L. Manies, J. W. Harden, E. a. G. Schuur, J. G. Vogel, J. T. Randerson, F. S. Chapin, Recovery of Aboveground Plant Biomass and Productivity After Fire in Mesic and Dry Black Spruce Forests of Interior Alaska, *Ecosystems* 11 (2) (2008) 209–225. doi:10.1007/s10021-007-9117-9.
- [11] B. Bond-Lamberty, C. Wang, S. T. Gower, Annual carbon flux from woody debris for a boreal black spruce fire chronosequence, *Journal of Geophysical Research* 108 (D23) (2003) 8220. doi:10.1029/2001JD000839. URL <http://www.agu.org/pubs/crossref/2002/2001JD000839.shtml>

- [12] J. Bendell, Effects of fire on birds and mammals., in: T. T. Kozlowski, C. E. Ahlgren (Eds.), *Fire and ecosystems*, physiologi Edition, Academic Press, New York, 1974, pp. 73–138.
- [13] L. B. Lentile, Z. a. Holden, A. M. S. Smith, M. J. Falkowski, A. T. Hudak, P. Morgan, S. a. Lewis, P. E. Gessler, N. C. Benson, Remote sensing techniques to assess active fire characteristics and post-fire effects, *International Journal of Wildland Fire* 15 (3) (2006) 319. doi:10.1071/WF05097.
- [14] K. Eberhart, P. Woodard, Distribution of residual vegetation associated with large fires in Alberta, *Canadian Journal of Forest Research* 17 (1987) 1207–1212.
- [15] S. C. Delong, D. Tanner, Managing the pattern of forest harvest: lessons from wildfire, *Biodiversity and conservation* 1205 (1996) 1191–1205.
- [16] K. McCleary, Using Low-Level Aerial Photography to Measure Structural Features, Tech. Rep. 2, Alberta Foothills Disturbance Ecology Methodology Series (2005).
- [17] A. H. Perera, B. D. Dalziel, L. J. Buse, R. G. Routledge, Spatial variability of stand-scale residuals in Ontarios boreal forest fires, *Canadian Journal of Forest Research* 39 (5) (2009) 945–961. doi:10.1139/X09-024.
- [18] C. Tucker, Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sensing of Environment* 8 (2) (1979) 127 – 150. doi:10.1016/0034-4257(79)90013-0.
- [19] A. Huete, K. Didan, T. Miura, E. Rodriguez, X. Gao, L. Ferreira, Overview of the radiometric and biophysical performance of the modis vegetation indices, *Remote Sensing of Environment* 83 (2002) 195 – 213. doi:10.1016/S0034-4257(02)00096-2.
- [20] Z. Jiang, A. Huete, K. Didan, T. Miura, Development of a two-band enhanced vegetation index without a blue band, *Remote Sensing of Environment* 112 (10) (2008) 3833 – 3845. doi:10.1016/j.rse.2008.06.006.
- [21] L. Breiman, Random forests, *Machine Learning* 45 (2001) 5–32.
- [22] B. Efron, Bootstrap methods: Another look at the jackknife, *Annals of Statistics* 7 (1979) 1–26.
- [23] T. Ho, The random subspace method for constructing decision forests, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (8) (1998) 832 – 844. doi:10.1109/34.709601.
- [24] M. Jung, J. Zscheischler, A guided hybrid genetic algorithm for feature selection with expensive cost functions, *Procedia Computer Science*.
- [25] J. Nash, J. Sutcliffe, River flow forecasting through conceptual models part I: A discussion of principles, *Journal of Hydrology* 10 (3) (1970) 282 – 290.
- [26] A. Leboeuf, A. Beaudoin, R. Fournier, L. Guindon, J. Luther, M. Lambert, A shadow fraction method for mapping biomass of northern boreal black spruce forests using QuickBird imagery, *Remote Sensing of Environment* 110 (4) (2007) 488–500. doi:10.1016/j.rse.2006.05.025.
- [27] A. Jain, D. Zongker, Feature selection: evaluation, application, and small sample performance, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (1997) 153–158.
- [28] E. Schulze, C. Wirth, D. Mollicone, W. Ziegler, Succession after stand replacing disturbances by fire, wind throw, and insects in the dark taiga of central siberia, *Oecologia* 146 (2005) 77–88.