CONTINENTAL SCALE VALIDATION OF A DYNAMIC FOREST MODEL WITH REMOTELY SENSED SATELLITE IMAGERY

David A. Lutz, University of Virginia
Herman H. Shugart, University of Virginia

Abstract

The Russian boreal forest is one of the largest terrestrial biomes on the planet and its behavior has significant consequences for global and regional climate, the Russian economy, and the global carbon cycle. In order to make projections about the future of Russian forests, a fine-scale dynamic forest model FAREAST was run and tested throughout the country under a variety of climate and disturbance scenarios. To validate this model output, remotely sensed data derived from satellite imagery was used as a comparison with the results of FAREAST. Baseline model output was matched with an empirically derived forest biomass product for over 28,000 sites at various year-steps of model functioning. After removing immature stands from the analysis, model biomass output and the remotely sensed product were comparable with an r² of 0.872. The results of this study suggest that the FAREAST model is fit to simulate Russian boreal and temperate forests on a continental scale with acceptable accuracy for the scientific community. In addition, a simple investigation into the contribution of agriculture to Russian forest carbon displacement was conducted; croplands throughout Russia were found to reduce the carbon storage capacity of the country by less than 10%.

Introduction

Boreal and temperate forests in the federated territory of Russia store an estimated 42.1 petagrams of Carbon within their biomass (Houghton et al., 2007). Currently, these forests are thought to provide Earth with a large carbon sink (Liski, et al., 2003), important for buffering the effects of increased concentrations of carbon dioxide in the atmosphere. Understanding the dynamic processes within these forests is critical in managing the country’s carbon budget. Afforestation of fallow agricultural lands is a ‘promising’ strategy for sequestering carbon in terrestrial ecosystems (Potter, et al., 2007); in Russia, this strategy may allow the region to store significant amounts of carbon. As well, fire suppression will potentially also store carbon in the short-term (Canadell, et al., 2007). These disturbances, along with insect outbreak, are the major drivers that turn these forests between carbon source to sink (Running, 2008).

Agricultural lands make up a significant component of the Russian Federation’s total territory. Nearly 13% of the country’s total land mass is used for crop production (FAOSTAT 2007); this cropland amounts to 2.16 million square kilometers, or nearly the size of Saudi Arabia. Much of this land was at one time forested and has been converted, thus limiting the total natural carbon storage of the system. In terms of total net flux of carbon, the largest human induced impact that can occur comes from the conversion of undisturbed natural forests into managed cropland. This is due to the direct loss of stored carbon in trees compared to the storage ability of croplands (Houghton and Goodale, 2004) as well as the loss of nearly 30% of the soil carbon stock (Murty et al. 2002). The total natural potential carbon storage for these regions without human disturbance is not known.

Yet what is the total amount of carbon that the boreal and temperate forests of Russia could ever naturally absorb? Given the recent urgings of scientists and politicians to balance anthropogenic output of carbon through mitigation strategies, estimating the maximum storage capacity of our ecosystems should be a critical calculation for the scientific community. The boreal and temperate forests of Russia, which store a significant part of all terrestrial carbon on Earth, are a major piece of the puzzle. However, long-term records lack detail and spatiality to be able to calculate what the original forest cover of Eastern Russia was before it was disturbed and removed for agriculture. As such, we have no ultimate baseline by which to determine vegetation’s ability to capture and store carbon throughout Russia if left undisturbed. The calculation of this ‘Garden of Eden’ scenario may advise scientists and policy makers what potential climate mitigation strategies through reforestation, afforestation, and fire suppression may achieve.

This scientific study attempts to answer the question of the maximum carbon storage of the Russian boreal and temperate forests when disturbance from agriculture is completely eliminated. Following the lead of many forest models to discern pristine forest
structure, species dynamics, and nutrient storage, we use an individual based forest gap model, named FAREAST, to simulate undisturbed conditions for the entire federated territory of Russia. These simulations are checked against current remote-sensing derived forest biomass levels to determine the influence of agriculture on the suppression of carbon in the terrestrial ecosystem. A validation exercise is also performed to serve as a check of model function to ensure simulation accuracy. Results from this exercise provide information about the limitations of climate change mitigation strategy through carbon sequestration methodologies in Russia.

Background Information

Ecological models have been created and used by scientists in order to understand complexities within forested ecosystems. Although simple computing models have been in existence since the late 1960’s, it is in the past 20 years that models have expanded to provide accurate predictions and simulations of forests (Mladenoff, 2004). Individual-based gap models are one of many different types of forest simulation models. Gap models typically are those which simulate individual trees, their growth, mortality, and decomposition into litter in a relatively small area, typically the size of a forest gap (Urban and Shugart, 1992). Generally, these models can be classified as non-spatial, in that they do not include spatial interactions between each small area that they simulate; yet, these models typically contain ecosystem processes such as nutrient cycling and interaction with the local abiotic environment (Scheller and Mladenoff, 2007). Although their size is generally fine scale, they contain large amounts of information at detailed resolution, which is of importance for regional mapping and predictions.

The FAREAST model (Xiaodong and Shugart, 2005) was developed to simulate Eastern Russian and northern Chinese forests. The initial tests to FAREAST were performed on 31 sites in China, of which 24 were simulated with acceptable degrees of accuracy, with an additional several sites simulated correctly after a modification was included to deal with permafrost dynamics of soils (Xiaodong and Shugart, 2005). This modified version, named FURTHEREAST, includes many important European species as well as range maps which constrain forest growth to their appropriate current positions (Shuman et al. in review). Tests of FURTHEREAST have been applied in several scenarios by members of the University of Virginia Center for Regional and Environmental Studies Laboratory (Shuman and Shugart, 2009, Lutz, 2008) and have shown the model to be accurate and acceptable to use to model boreal forests throughout Russia.

Nevertheless, in order to provide accurate projections regarding boreal forest currently displaced by agriculture, a more thorough validation scenario must be completed. Folded into this model application is a rigorous test of the FURTHEREAST model. Country-wide estimates of forest carbon (from biomass) and leaf area index (LAI) derived from remotely sensed data were obtained in collaboration with the CEPF. These data sets were seen to be potential robust tests of model functioning in various locations previously unexamined by FURTHEREAST.

Russian forest ecosystems have been modeled using a variety of different forest model types. Shvidenko et al. created a semi-empirical model based primarily on yield-tables but were limited to dominant tree species only (Shvidenko et al., 2007). Several individual gap models have been applied to locations in Chinese forests containing many similar species as Siberian forests, yet were implemented on small individual sites (Zhang et al., 2009). Yet most studies investigating carbon and agriculture issues regionally in Russia are concentrated on soil modeling or agricultural productivity (e.g. Smith et al., 2007). Modeling the entire continent of Russia’s forest with species level dynamics has not been attempted, and this marks the first validation and application of such an exercise. By expanding a very fine scale model such as FURTHEREAST to a wide near-continental scale, fine-scale information will be available for many areas for the first time. These results can then be used to investigate forest and agriculture questions for the country of Russia.

Validation Exercise

In order to verify the accuracy of FURTHEREAST’s simulations of forest dynamics over such a large geography, a validation exercise was performed. Over 33,500 individual sites were created covering the federated territory of Russia; each site was equally incrementally distributed throughout the area such that the landscape represented by each site composed 22 km². Figure 1 gives a sense of scale of the coverage of these sites. Input files for these sites contained information regarding the topography, soil, and climate of the surrounding area. Elevation and slope information was derived from the Shuttle Radar Topography Mission (Farr et al., 2007); data from this source was resampled using a Nearest-Neighbor methodology to fit the scale of each site (22km²). Information with respect to soil and hydrological properties was derived from the Land Resources of

Lutz
Russia dataset from the International Institute for Applied Systems Analysis and the Russian Academy of Sciences (Stolbovoi and McCallum, 2002).

Weather information for each site was derived from a 60-year record of conditions at stations located throughout the country (Razuvayev, V. N., et al. 1993). These data were interpolated to provide climate information for all generated sites. Subsequent statistical analysis of output from sites selectively excluded from the interpolation scheme suggested no significant change in the simulated forest dynamics when using interpolated weather data compared to original weather station data. This test implied an accurate interpolation procedure.

Following site initialization, the FURTHEREAST model simulated 300 years of forest data at a twelfth hectare scale for 200 individual plots in a similar manner to Shuman and Shugart, 2009. This multitude of simulation plots for each site was necessary to reduce variability within the model output. Forest biomass and leaf area index values for all 44 species at the 300 year time-interval were averaged and integrated into a GIS framework. This data was paired with remotely sensed data from the same location to check model accuracy.

Remotely sensed leaf area indices (LAI) and forest biomass data were derived from SPOT imagery. Both LAI and biomass data from this set was calculated empirically by comparing the satellite imagery to field estimations of the two parameters calculations throughout the Krasnoyarsk region by the Center for Ecology and Productivity of Forests (CEPF) at the Russian Academy of Sciences. This information was post-processed and resampled to a 22 km² scale to be comparable to FURTHEREAST model output.

Before model output was regressed to the remotely sensed satellite data, one modification needed to be made. Since FURTHEREAST has not been adjusted to reflect disturbances such as fire and logging, all simulations represent ‘garden of Eden’ scenarios. Thus, any comparison between these simulations and real world data will be flawed for any validation exercise if the site did not contain mature forest stands (Lutz, 2007). To account for this discrepancy, all young forests needed to be masked out in the comparison. To do this, a threshold LAI value of 4 was chosen; sites containing real-world values lower than this threshold were assumed to be younger than mature phase and eliminated from the comparison; the LAI mask can be found in Figure 2. The result was a direct comparison between only mature phase forests and thus seen as an accurate validation exercise.

FURTHEREAST ‘Garden of Eden’ Application

Following the validation exercise, model biomass data for all sites were compared with the remotely sensed data set, including areas that were not mature phase forests. The difference between model output and ground data from remote sensing was analyzed using Hotspot analysis in ArcGIS 9.2. The Getis-Ord Gi statistics were mapped geographically. These hotspots indicated statistically significant locations of either model over- or under-estimation of biomass.

In order to determine the areas in which agricultural lands were depressing potential biomass, a classified remote sensing dataset was utilized to highlight cropland. The GLC2000 land cover classification data set was used to select areas indicated as agricultural or cropland (Bartholome and Belward, 2005). These sites were then masked and removed from the entire dataset; a map of the masked agricultural lands with respect to the entire federated territory can be found in Figure 3. A point to point regression of biomass for every site was conducted to examine the relationship between a ‘Garden of Eden’ scenario and present day biomass values. A second regression was performed using the dataset with agricultural lands removed. These regression values were then compared to see the total net effect the presence of agriculture had on the country’s terrestrial carbon balance.

Results

Validation Exercise

The results of the validation exercise in which FURTHEREAST model output was compared to satellite-derived biomass information can be found in Figures 4 and 5. Figure 4 shows the results in a comparison graph between the two data sets with an LAI of over 3. The results consisted of 5698 different sites and a linear regression was fit to compare the two data sets. Figure 4 shows the data including sites in which the data derived from remote sensing displayed no biomass values. These zero data points are likely a result of pixel size between the LAI raster data set and the biomass raster data set in which the LAI data incorrectly overlapped an area that had been disturbed. With all data included in the regression, including the zero biomass data, the coefficient of determination is 0.7862. When the 518 outlier data points are removed, the coefficient of determination increases to 0.872. In both cases, the equation for this regression line was forced through zero assuming that a point with no growth possible, for instance a rocky slope, would read as zero by both FAREAST and the satellite data.
Agricultural Exercise

The results of the test to determine the influence of Russian agriculture on carbon storage can be found in Figures 6 and 7. Figure 6 represents the comparison between FAREAST output and biomass estimations from remote sensing data for 26,526 data points across Russia. Figure 7 represents the comparison between these two data sets once sites disturbed by agriculture are removed from the comparison. While the slope of the regressed line remains the same, the coefficient of determination for these comparisons differs. When sites representing agricultural disturbance are considered, the coefficient of determination is 0.4744; when agricultural sites are removed, this value increases to 0.5474.

Analysis and Conclusions

The results of the validation exercise indicate that the FAREAST model correlates well with forests throughout the federated territory of Russia. Previously, the FAREAST model has been tested against forest stands in Northeastern China and the Russian Far East (Xiaodong and Shugart, 2005) as well as managed forest stands scattered throughout the continent (Shuman and Shugart, 2009). This analysis, however, indicates that the FAREAST model not only generates aboveground biomass values that mimic real forests throughout the entire territory of Russia, but also that the model generates these values exceptionally well. Previous gap models have been compared mainly to stand data and are subsequently relegated to a handful of comparisons (e.g. Buggman and Solomon, 2000; Busing et al., 2007). As such, it is difficult for thorough statistical comparisons to be drawn between a model and validation data. This study, however, compares gap model function to nearly 6,000 different field sites.

In addition to the number of validation points with which FAREAST is tested, another interesting point of discussion lies in the relative ability of FAREAST to correctly model biomass dynamics over such an enormous and diverse variety of landscapes. The simplicity of the Russian boreal and temperate forest systems allows the FAREAST gap model to work effectively in many different locations throughout the continent. The addition of range maps to constrain the model such that only species known to exist in a particular site were included during the simulation run. As a result, the model was able to be applied throughout far eastern Russia, Siberia, and European Russia with considerable success with respect to modeling forest biomass. This success indicates that the FAREAST model will be capable of performing many more investigations into boreal forest dynamics and the response of the Russian forest landscape to ecosystem changes and disturbances.

The results of the agricultural disturbance modeling scenario indicate that agricultural lands displace a significant amount of land which would otherwise be forested. Although the slope of the regression line stayed almost exactly the same between the two comparisons (with and without agricultural lands masked out), the coefficient of regression markedly improved once agricultural lands were removed. This indicates that the relationship between modeled biomass and observed biomass is stronger, likely as a result of fewer sites in which the observed biomass was recorded as zero. Generally, the change in the coefficient of regression in this case indicates that disturbance influences about a 0.08 different once croplands are removed from the analysis. In order to investigate precisely the expected aboveground biomass that could be grown on areas that are currently farmed, investigation of individual sites of the model results could yield a robust tonnage value. However, for this initial analysis of the issue, it is significant to say that farming removes around 8% of the potential carbon that could be sequestered within this entire system. A more thorough examination upon the precise areas displaced by farming in this dataset would be able to derive more precise values.

The best-fit linear regression lines for each comparison suggest that the FAREAST model consistently underestimates biomass production for Russian forests. When older mature phase forests are highlighted, the slope of the linear regression equation decreases compared to when all sites are included in the comparison. After 300 years of simulation, the FAREAST model maximizes biomass at 140 t/ha for the most densely forested sites. However, many of these sites were detected by the remotely sensed data to contain upwards of 200 t/ha. The likely reason behind this discrepancy comes from the limitation of model simulation runs to 300 years. Additionally, there is likely some error with respect to remotely sensed data extraction methods which may also skew the regression line.

Conclusively, this marked the first large-scale validation exercise of the FAREAST model throughout the federated territory of Russia. When compared to over 5,000 forested sites, the model successfully simulated biomass dynamics for these sites and compared to satellite derivations of biomass with a coefficient of determination of 0.872. This high correlation indicates a robust model capable of being used for several applications throughout the sites generated in this project. In addition, the initial
investigation into carbon storage in Russian forests currently displaced indicates that croplands remove less than 10% of total storage capacity for Russia in terms of carbon. With a more thorough investigation into individual site dynamics, this value could be refined and adjusted more accurately.

References


Figure 1. Distribution of model test sites across European Russia. Each site accumulates information for an area of 400 km². Red square indicates a single site.

Figure 2. Distribution of forests with a measured leaf area index (LAI) greater than 3. Data compiled by the Center for Ecology and Productivity of Forests.

Figure 3. Distribution of croplands and grasslands of Russia. Classification used from the GLC2000 data set.

Figure 4. Graph displaying results of a direct comparison between remotely sensed biomass and FAREAST model simulations for areas where LAI is greater than 3. The coefficient of determination $r^2$ is 0.7862 and the linear fit equation is $y = 0.6696x$. 
Figure 5. Graph displaying results of a direct comparison between remotely sensed biomass and FAREAST model simulations for areas where LAI is greater than 3 and outliers are removed. The coefficient of determination $r^2$ is 0.872 and the linear fit equation is $y = 0.6696x$.

Figure 6. Graph showing the results of a comparison between remotely sensed biomass and FAREAST model simulations for over 26,500 sites throughout the Russian Federation. The linear fit equation is $y = 0.7147$ and the coefficient of determination $r^2$ is 0.4744.

Figure 7. Graph showing the results of a comparison between remotely sensed biomass and FAREAST model simulations for sites not classified as croplands in the GLC2000 dataset. The linear fit equation is $y = 0.7121$ and the coefficient of determination $r^2$ is 0.5474.