Satellite-based peatland mapping: Potential of the MODIS sensor

Dirk Pflugmacher a,⁎, Olga N. Krankina a, Warren B. Cohen b

a Department of Forest Science, Oregon State University, 321 Richardson Hall, Corvallis, OR 97331, United States
b Forestry Sciences Laboratory, Pacific Northwest Research Station, USDA Forest Service, 3200 SW Jefferson Way, Corvallis, OR 97331, United States

Received 27 May 2005; accepted 19 July 2006
Available online 3 November 2006

Abstract

Peatlands play a major role in the global carbon cycle but are largely overlooked in current large-scale vegetation mapping efforts. In this study, we investigated the potential of the Moderate Resolution Imaging Spectroradiometer (MODIS) to capture the extent and distribution of peatlands in the St. Petersburg region of Russia by analyzing the relationships between peatland cover fractions derived from reference maps and ∼ 1-km resolution MODIS Nadir BRDF-Adjusted Reflectance (NBAR) data from year 2002.

First, we characterized and mapped 50 peatlands from forest inventory and peat deposit inventory data. The peatlands represent three major nutritional types (oligotrophic, mesotrophic, eutrophic) and different sizes (0.6–7800 ha). In addition, parts of 6 peatlands were mined for peat and these were mapped separately. The reference maps provided information on peatland cover for 1105 MODIS pixels. We performed regression analysis on 50% of the pixels and reserved the remainder for model validation. Canonical correlation analysis on the MODIS reflectance bands and the peatland cover fractions produced a multi-spectral peatland cover index (PCI), which served as the predictor in a reduced major axis (RMA) regression model. The results suggest a high potential for mapping peatlands with MODIS. The RMA regression models explained much of the variance in the PCI ($r^2 = 0.74$ for mined and $r^2 = 0.81$ for unmined peatlands). Model validation showed high correlation between observed versus predicted peatland cover (mined: $r = 0.87$; unmined: $r = 0.92$). We used the models to derive peatland cover estimates for the St. Petersburg region and compared the results to current MODIS land cover maps.

© 2006 Elsevier B.V. All rights reserved.

Keywords: peatlands; mapping; satellite methods; regression analysis; carbon cycle; Russian Federation

1. Introduction

Peatlands constitute one of the most widespread wetland types in the world. The most significant regions in terms of absolute extent of peatlands are in Europe, including the former Soviet Union, and North America, in particular above 45° N (Charman, 2002). Despite their importance in the global carbon and hydrological cycle, and their significance as wildlife habitat, the global distribution and extent of peatlands remains uncertain (Maltby and Immirzi, 1993).

Peatlands are often referred to as organic wetlands, because of their characteristic layer of peat, which is plant detritus accumulated under anaerobic, waterlogged conditions. Peat represents a major pool of organic carbon that amounts to about one-third to a half of the global soil carbon, which is almost the equivalent to the global atmospheric carbon pool (Charman, 2002). Total carbon stored in northern peatlands alone has been estimated to be about 455 Pg C (Gorham, 1991) with a
current uptake rate in extant northern peatlands of 0.07 Pg C/yr (Clymo et al., 1998). At the same time, peatlands represent a major source of methane (CH\textsubscript{4}) (Matthews and Fung, 1987) and dissolved organic carbon (DOC) (Freeman et al., 2004). Since atmospheric greenhouse gases are linked with the global climate system, global information on size and distribution of peatlands is of fundamental importance to climate change research. Currently, no clear consensus exists about the net effects of and feedbacks on future climate (IPCC, 2001).

For the last decades, global climate and biogeochemistry models largely depended on data derived from preexisting maps and atlases. Some of the most commonly used global data sets on wetland distribution were compiled by Olson and Watts (1982), Matthews and Fung (1987), and Aselmann and Crutzen (1989; later revised by Stillwell-Soller et al., 1995). These data sources rely on extensive historic sit surveys and represent the best available information at the time. Nevertheless, their accuracy has not been rigorously assessed. As peatlands change over time in the process of their natural evolution and under the impacts of natural disturbances (e.g. fires), and human activities (e.g. peat mining, drainage and conversion to agricultural and forest land), the old maps likely become even less accurate.

Peatlands, particularly in the boreal region, tend to lack tree cover and represent distinct vegetation types (Botch and Masing, 1983; Gorham, 1991) that can be identified on satellite imagery. Since ground information on peatlands is often limited or even lacking in remote regions (Sheng et al., 2004), satellite remote sensing could provide a valuable tool for monitoring peatlands, especially in the northernmost latitudes. Several local and regional studies have mapped peatlands with high and medium spatial resolution sensors such as Landsat TM (30 m) and SPOT HRV (20 m) (Markon and Derksen, 1994; Poulin et al., 2002). However, mapping vast regions such as Northern Eurasia at fine spatial resolutions does not appear practical because of the lack of cloud-free imagery for many areas (DeFries et al., 1997; Krankina et al., 2004b). Further, there are considerable costs and logistical difficulties involved with handling such high data volumes, limiting the repeatability of such studies.

In comparison, the medium to coarse resolution sensor MODIS provides consistent and frequent observations of global land cover and land cover change processes at essentially no cost to the user (Townshend and Justice, 2002). Since the first global satellite based land cover map produced by DeFries and Townshend (1994) with data from the advanced high-resolution radiometer (AVHRR), substantial advancements have been made towards the development of comprehensive global vegetation and land cover data sets (Friedl et al., 2002; Justice and Townshend, 2002; Hansen et al., 2003). Nevertheless, peatlands are largely overlooked in large-scale land cover mapping efforts.

This study tested the capability of the MODIS sensor to map peatland cover proportions in a taiga landscape of the East-European plain. Results presented here may provide useful information for future mapping algorithms and therefore may promote the development of global land cover maps where peatlands are adequately represented.

2. Study area

The St. Petersburg region in Russia (Fig. 1) was selected as a test site, because of its location in one of the most significant peatland regions of the world. The abundance of peatlands in the St. Petersburg region is representative for northwestern Russia and is also close to the overall peatland proportion of the whole Russian Federation (8%) (Kobak et al., 1998). The study area occupies about 8 million hectare of flat terrain that rests on ancient sea sediments covered by a layer of moraine deposits. The natural vegetation belongs to the southern taiga type, and the climate is cool maritime. In terms of land cover, the study area is quite diverse. The region includes a major metropolitan area, St. Petersburg (5 million inhabitants) and a large agricultural region south and west from the city. Fifty-three percent of the region is covered with closed canopy forest, and repeated logging is a major disturbance factor, as is urban expansion and agricultural change (Krankina et al., 2004a).

The dominant peatland type in the region is the ‘raised string bog’ (Botch and Masing, 1983). Raised bogs have a dome-shaped surface built up of Sphagnum peat. In contrast to minerotrophic fens, raised bogs receive all their water and nutrients from the atmosphere (ombrotrophic). Therefore, they tend to be acid and low in nutrient availability (oligotrophic). According to Botch et al. (in preparation) oligotrophic bogs account for about 75% of the total peat volume in the St. Petersburg region, while transitional peat from mesotrophic peatlands and low-lying peat from fens or eutrophic peatlands comprise about 14% and 11%, respectively. In some areas peat is mined for use as fuel or soil conditioner. Mining removes the upper layers of peat, leaving bare peat surfaces that are often converted to agricultural or forested land. The consequences include rapid oxidation of the extracted peat into carbon dioxide and discharge of
particulate carbon, DOC, nutrients and heavy metal (Charman, 2002).

3. Data

3.1. MODIS

The MODIS instrument provides near-daily repeated coverage of the earth’s surface with 36 spectral bands and a swath width of approximately 2330 km. Seven bands are specifically designed for land remote sensing with a spatial resolution of 250 m (band 1–2) and 500 m (band 3–7) (Table 2). In addition, MODIS science teams are producing a suite of higher level radiation products (e.g. systematically atmospherically corrected surface reflectance, surface temperature and emissivity, bidirectional reflectance distribution function (BRDF) and albedo) and vegetation and land cover products (e.g. land cover and cover change, vegetation indices, leaf area index and fraction of photosynthetically active radiation, thermal anomalies and fire) (Justice and Townshend, 2002).

For this study we selected the Nadir BRDF-Adjusted Reflectance (NBAR) data set with a spatial resolution of ~1 km (MOD43B4, collection 4). MODIS NBAR is based on multi-date, cloud-cleared and atmospherically corrected surface reflectance (MOD09) acquired over a 16-day period. NBAR values are computed for each of the seven spectral bands and normalized to the mean solar zenith angle of each 16-day period. Since view angle effects and cloud and aerosol contamination have been minimized, MOD43B4 delivers more stable and consistent data that is suitable for global land surface modeling (Schaff et al., 2002). This is of particular value for high latitudes where the availability of high quality imagery is constrained by persistent cloud cover and extreme variations in sun-target-sensor geometry. To assess the performance of the BRDF/Albedo algorithm, MODIS NBAR data are supplied with extensive quality assurance information for each pixel. We examined the band-averaged quality control flags from all imagery acquired within the active growing season (May–August) of the year 2002. NBAR data from Julian Day 145 (acquired between 05/25/02–06/09/02) showed best quality (processed, good quality) for nearly all (99%) of the pixels belonging to the sample population and was therefore selected for further analysis.

MODIS NBAR also serves as primary input for the MODIS land cover product MOD12Q1. MOD12Q1 provides maps of global land cover at a spatial resolution of 1 km using several vegetation classification systems (Friedl et al., 2002). Among these, the IGBP (International Geosphere–Biosphere Programme) classification is currently the only one that includes peatlands (class 11: permanent wetlands). The MODIS IGBP land cover map analyzed in this study is dated to year 2001 (collection 4).

3.2. Reference data

In the St. Petersburg region of Russia all lands under state forest management are surveyed approximately
every 10 years by the Northwest State Forest Inventory Enterprise. In this study we used inventory data and maps from 1992–93. Forest inventory maps are based on detailed topographic maps and aerial photographs, and peatlands are classified as bogs and mires with less than 40% tree cover. We georeferenced the inventory maps to an orthorectified Landsat TM scene (UTM-Zone 36, WGS84) and onscreen-digitized 50 peatlands into a polygon data set. The Landsat scene had a spatial resolution of 30 m and was acquired on May 29th 1992 at path 184 and row 18 on the World Reference System 2 (Fig. 1).

Reference peatlands were selected to represent three major nutritional types (oligotrophic, mesotrophic, eutrophic), different sizes and development states (mined, unmined). Total peatland areas vary from less than a hectare to about 7800 ha with a median of 287 ha; 16 peatlands are smaller than 100 ha (1 km²), 14 peatlands are larger than 1000 ha (10 km²). Mined areas occur in 6 peatlands and make up about 10% of the reference data. About 90% of the unmined reference peatland area is oligotrophic, 8% is mesotrophic and 2% is eutrophic. Sphagnum moss, along with sphagnum–sedge associations, is the dominant vegetation type on over 90% of the reference peatlands. More eutrophic areas support sparse tree cover, comprised mainly of Scots pine (Pinus sylvestris L.) but birch (Betula pendula L.) and willow species (Salix spp.) also occur.

To spatially match the high-resolution reference data with the MODIS imagery, we reprojected the digitized peatland polygons to the MODIS sinusoidal projection (Snyder, 1987) and nearest neighbor resampled them to a grid with a cell size of 18.5 m. Then we computed fractional peatland cover for each 1-km MODIS NBAR grid cell as the proportion of reference grid cell area covered by peatlands. This process yielded 1105 1-km grid cells with peatland cover ranging from 0% to 100%. Finally, we assigned every other pixel to a model development set, and retained the remainder for model validation.

4. Methods

4.1. Approach

The broad swath of the MODIS instrument allows for frequent observation of the earth’s surface over broad geographic regions, facilitating regional and global analyses of land cover and change processes. Such temporal and spatial coverage, however, comes at the expense of spatial grain size: pixels 1 km in size typically include several types of land cover, especially in heterogeneous landscapes. Discrete land cover classification systems require that a single label be assigned to each cell, often based on the class occupying the majority area within the cell. Because peatlands often occur in patches that are much smaller than a 1-km MODIS NBAR pixel, such an approach would result in an underestimation of the extent of peatlands. This approach also obscures subtle changes in land cover that do not involve an extreme conversion from one cover type to another (DeFries et al., 1995; DeFries et al., 2000). Thus, a discrete land cover classification is likely not appropriate for accurate mapping of peatlands with 1-km grain-size imagery.

Therefore, we used a method to map sub-pixel proportions of peatlands in each pixel. We patterned our approach on successful efforts to map sub-pixel proportions of forest cover (Zhu and Evans, 1994; DeFries et al., 1997; Cohen et al., 2001; Hansen et al., 2003), land cover (Atkinson et al., 1997) and other biophysical attributes (DeFries et al., 1995; Cohen et al., 2003). Typically, statistical models are used to relate the spectral measurements in a pixel to the proportion of the desired vegetation property in that pixel. Here, we developed empirical regression models linking percentages of peatland cover (see Section 3.2) with MODIS NBAR reflectance.

4.2. Analysis

Previous studies suggest that reduced major axis (RMA) regression provides a superior tool to fit regression models based on remotely sensed data, compared to the more widely used ordinary least squares (OLS) regression (Cohen et al., 2003). RMA regression is an

![Fig. 2. Fractional cover of unmined peatlands in MODIS pixels: reference data (mapped by forest inventory) versus predicted (modeled).](image-url)
orthogonal regression method, which minimizes the sum of squared orthogonal distances from the data points to the fitted line. Unlike OLS, RMA makes no assumptions about dependencies between the \( X \) and \( Y \) variables and does not presume the \( X \) variable to be obtained without error. While the formula for the RMA regression model (Eq. (1)) and the intercept \( \beta_0 \) (Eq. (2)) is identical to the OLS regression, the RMA slope \( \beta_1 \) is defined as the ratio of sample standard deviation of the \( X' \)s and sample standard deviation of the \( Y' \)s (Eq. (3)),

\[
Y = \beta_0 + \beta_1 X + \varepsilon \\
\beta_0 = \bar{Y} - \beta_1 \times \bar{X} \\
\beta_1 = \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}}
\]

where \( X \) is the predictor variable, \( Y \) is the variable to be predicted and \( \bar{X} \) and \( \bar{Y} \) are the sample means of \( X \) and \( Y \), respectively.

A simple application of RMA requires a single \( X \) and a single \( Y \) variable. Consequently, we reduced the dimensionality of the 7 MODIS bands. Spectral vegetation indices such as the normalized difference vegetation index (NDVI), which is based on a ratio of the red and near-infrared band, have been widely used to model vegetation properties. By using only two bands, such indices contain only a fraction of the available spectral information. Improved mapping is often possible by incorporating additional spectral bands (Cohen and Goward, 2004). To take advantage of the full spectral range we linearly combined multiple MODIS bands into a single index by means of canonical correlation analysis (Cohen et al., 2003).

Canonical correlation analysis (CCA) is a generalized form of multiple regression that permits the examination of linear relationships between two sets of variables (Tabachnick and Fidell, 1989). The CCA yields a pair of linear combinations between the two sets of variables such that the correlation between the two combinations is maximized. In this study CCA provided a set of coefficients for the MODIS spectral bands that aligned them with the variation in the observed peatland cover fractions. We applied CCA to fractional peatland cover as a single variable and the MODIS bands as a set of variables. We selected only MODIS bands that significantly predicted mined and unmined peatland coverage by means of forward stepwise regression analysis.
(p<0.01). The standardized canonical coefficients (z transformed with a mean of 0 and a standard deviation of 1) signify each band’s contribution to the final predictive model: the larger the absolute value, the greater is the unique positive or negative effect of the band (Table 2). We refer to the resulting canonical variable as the peatland cover index (PCI). The PCI served as predictor variable and the observed fractional peatland cover as dependent variable in the RMA regression.

4.3. Error assessment

In regression methodology, accuracy of the model is typically evaluated with the coefficient of determination ($r^2$) and the root mean squared error (RMSE). Both approaches only describe the variance described by the model for the sample population, and do not provide a true measure of accuracy of the predictions. Here, we built our statistical models using half of our reference data (see Section 3.2), and then applied the models to the reserved model validation set. When the model is applied to the validation data, it yields a set of predicted values, which are plotted against the observed fractional peatland cover. The regression between predicted versus observed values results in a new coefficient of determination. We calculated overall bias as the mean of the predicted values minus the mean of the observed values. A mean overprediction would result in a positive bias, and vice versa. We further calculated variance ratios as the standard deviation of the predicted values divided by the standard deviation of the observed values. A variance ratio of greater than one means that the prediction standard deviation was greater than the observed standard deviation, and vice versa.

Another common expression of accuracy is the classification error matrix. Error matrices allow comparison of the reference data with the model predictions on a category-by-category basis. For this purpose, we aggregated the data into 4 classes of peatland proportion (0–25%, 26–50%, 51–75%, 76–100%) and constrained predictions by truncation to between 0 and 100%.

5. Results

Our results (Figs. 2 and 3) show a strong relationship between MODIS NBAR data and observed fractional peatland cover. The reduced major axis (RMA) regression models explained much of the variance in the peatland cover index (PCI) for unmined ($r^2=0.81$) and mined ($r^2=0.74$) peatlands. While the unmined PCI model includes all seven MODIS bands, the mined model is based only on five bands (Table 1). In the latter case bands 3 and 6 were discarded by forward stepwise regression ($p<0.01$). Fractional unmined peatland cover was most strongly correlated with the red band 1, while mined peatlands showed highest correlation with the shortwave infrared band 7. This relationship is reflected in large standardized canonical coefficients indicating the contribution of these bands to their respective PCI variable (Table 2).

Results of the PCI model validation against an independent data set show that predicted fractional peatland cover estimates are highly correlated with the observed peatland cover (mined: $r=0.87$; unmined: $r=0.92$), with essentially no bias (Figs. 2 and 3). The variance ratio of nearly 1 suggests overall preservation of observed variance in the predictions (Table 3). Both unmined and mined peatland predictions have acceptable root mean squared errors (RMSE) of 16% and 9%, respectively. All three nutritional types seem to be well represented by the model. Pixels grouped by dominant nutritional type exhibit similar mean prediction errors (eutrophic=0.11, mesotrophic=0.15, oligotrophic=0.16). The error matrices in Tables 4 and 5 confirm the models are capable of predicting unmined and mined peatland cover with a good overall accuracy (75% and 94% respectively) at the 25% interval level, which translates into a minimum mapping unit of about 22 ha.

When applied to MODIS NBAR data for the entire St. Petersburg region, our models predict that unmined and mined peatlands make up about 9% and 1% of the land.

### Table 4

Validation error matrix for fractional unmined peatland cover

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
<th>Commission</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–25%</td>
<td>199</td>
<td>30</td>
</tr>
<tr>
<td>26–50%</td>
<td>15</td>
<td>41</td>
</tr>
<tr>
<td>51–75%</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>76–100%</td>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td>Omission</td>
<td>92%</td>
<td>38%</td>
</tr>
</tbody>
</table>

### Table 5

Validation error matrix for fractional mined peatland cover

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Observed</th>
<th>Commission</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–25%</td>
<td>511</td>
<td>3</td>
</tr>
<tr>
<td>26–50%</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>51–75%</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>76–100%</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Omission</td>
<td>98%</td>
<td>36%</td>
</tr>
</tbody>
</table>
area, respectively. These regional estimates are based on a minimum fractional mapping unit of 25% peatland cover per pixel (~22 ha) and do not include pixels that are mapped as water, urban area or cropland by the MODIS IGBP (2001, collection 4) land cover product (Figs. 4 and 5). Our area estimates agree with those from Oetter et al. (in review) who reported a proportion of peatlands of 10% for the same region using high-resolution satellite data.

The comparison with our model results shows that the current MODIS IGBP land cover product (2001, collection 4) fails to represent peatlands adequately. In the St. Petersburg region MODIS IGBP considerably underestimated the peatland area. The MODIS product identified only 0.2% of the land area in the St. Petersburg region as permanent wetlands. Pixels where our models projected a total (unmined plus mined) peatland proportion over 75% were most often classified by MODIS IGBP as evergreen needleleaf forests (44%) followed by open shrubland (14%), woody savanna (14%), savanna (10%) and mixed forests (9%). Conversely, our models estimated a third of the MODIS wetland pixels to have less than 25% peatland coverage.

6. Discussion and conclusion

The observed relationships between surface reflectance measured by the satellite sensor MODIS and percent peatland cover information from forest inventory data indicate good potentials for MODIS-based peatland mapping. The predictions of percent peatland cover based on these empirical relationships suggest that a
map of percent peatland cover could be derived with adequate accuracy for the St. Petersburg region. To develop global or continental maps, i.e. for Northern Eurasia, further research including additional reference sites is necessary. Prediction accuracy could be improved with multi-temporal data sets. By incorporating time series of imagery, seasonal variations in land surface and vegetation properties (e.g. vegetation phenology) can add considerable information to the classification process. Further analyses could also test whether and how additional MODIS data sets such as the vegetation indices, surface temperature and vegetation continuous fields can aid in the determination of percent peatland cover.

As expected, the regression models for mined and unmined peatlands utilized different combinations of spectral bands, indicating the capability of MODIS to distinguish between the two development states. However, visual examination of high-resolution satellite data (Landsat TM) during analysis suggested that mined areas may be confused with agricultural lands, whereas peatlands that have once been mined and later abandoned may resemble pristine peatlands after natural processes of regeneration occurred. Further research will be needed to address classification errors associated with other land cover types. It is clear that land surface information derived from medium to coarse resolution sensors cannot be as spatially explicit as those derived from high-resolution sensors like Landsat. However, recent research has increasingly focused on continuous modeling of sub-pixel land cover fractions, and much progress has been made to improve the accuracy of such continuous field products. The failure of the MODIS IGBP land cover product to adequately map peatlands in the St. Petersburg region may point to three limitations of the product. First, the MODIS IGBP product may overestimate woody cover proportions. Forest cover of the reference peatlands used in this analysis was limited by forest inventory classification to 40%. In the IGBP classification scheme, forests are defined as lands dominated by woody vegetation with a percent cover of at least 60% and height exceeding 2 m. Overestimation of woody cover was also observed by Cohen et al. (2006) at an evergreen needleleaf site with low tree cover proportion and at an arctic tundra site in Alaska. Second, ambiguities or generalities in land cover class definitions may not allow to distinguish between cover types that do not follow certain vegetation classification logics alone, but are characterized by additional, in this case hydro-morphological attributes. Therefore, wetlands and peatlands often cannot be classified correctly. Finally, the MODIS IGBP product is generated using an empirical supervised classification strategy that relies on a global database of training sites. Since the quality of the classification results is strongly influenced by the quality of the training site database, a lack of reference sites in Northern Eurasia would result in a land cover map that is less accurate in that region.

In summary, this study indicates promising opportunities for peatland mapping based on MODIS data. The improvements are especially important in remote boreal regions where available data is clearly inadequate. Since the ecological role of peatlands is vastly different from vegetation types on mineral soils, the poor representation of peatlands may lead to significant misinterpretations of the role of terrestrial ecosystems in the global carbon cycle. More precise and up-to-date information on the extent of change in the carbon balance resulting from peatland alteration is needed (Maltby and Immirzi, 1993) to better understand possible responses and feedbacks to climate change or other environmental factors. With MODIS data from multiple years, changes in the distribution of peatlands and the extent of peat mining and regeneration could be quantified. MODIS-derived maps would provide a valuable means to facilitate and improve regional and global estimates of the amount of carbon stored in peatlands.

Acknowledgements

This research was conducted at Oregon State University and at USDA Forest Service PNW Research Station with support from the Land Cover/Land-Use Change Program of the National Aeronautics and Space Administration (grant number NAG5-11250 and NNG04GB81G). We greatly appreciate the reviews and thoughtful comments of Robert Kennedy and an anonymous reviewer.

References


Technical note TN-416+STR. National Center for Atmospheric Research, Boulder, Colorado, USA.


