Projects and Initiatives addressing Environmental Impact issues in Northern Mongolia and the Lake Baikal Region (Friedrich-Schiller-University Jena, Germany)

Karsten Frotscher and Christiane C. Schmullius
Friedrich-Schiller-University Jena, Department of Geography, Löbdergraben 10, D-07743 Jena, Germany, karsten.frotscher@uni-jena.de

Abstract
Fast-growing economies and worldwide growing consumer demands have a considerable impact on natural resources and thus on the way Earth Science community addresses the acquisition, storage and analyses of spatial data. Reliable and up-to-date information on land surface characteristics and changes are therefore required by decision makers in order to fulfil several international and national treaties and for its own policy. Satellite-based Earth Observation (EO) serve thereby as an independent and unbiased framework to analyse landscape structure and related environmental processes as well as impacts at multiple scales. Advanced image processing techniques such as for Synthetic Aperture Radar (SAR) or time series, embedded in a hierarchical system of an object-oriented GIS framework help to handle the complex character of the problem. One of the essential components of information distribution is the development of spatial data infrastructures by using open standards as recommended by the Open Geospatial Consortium (OGC™).

Keywords: Earth Observation, Data Infrastructure, Forest Inventory, Watershed Hydrology, Landscape Assessment, Object-oriented, SAR processing, XML

1 Introduction

China's exports of wood products have been fast growing over the last decade and hence their imports of timber (White et al., 2006). Forest resources of Mongolia and the Lake Baikal region are affected by this development and since then information provision for forest protection with respect to natural and human-induced disturbances is needed. Alongside, the region is rich in many types of mineral resources which are of newly interest to both, international investors and local communities. Artisanal mining in form of illegal gold placer mining have become a major problem for water quality through the uncontrolled use of mercury. Under such circumstances an issue that must be addressed is the monitoring of the region and the subsequent analyses pertaining to environmental impacts.

One of the major challenges for an integrative environmental development is the integrated management of water and forest resources and socio-economic developments in order to secure a sufficient availability of fresh water and timber. Thus, for an integrated management methodical designs are necessary which refer to the complexity of the resources to be managed and the difficulty to monitor them. A recent, sophisticated approach to understand landscapes consists in modelling their structures as a fuzzy system composed of complicated, dynamic, stochastic processes. This methodology supports specifically landscape-ecosystem based approaches in which ecosystems of different scales are regarded as primary units for quantification and modelling (Seppelt & Voinov, 2002). Because of the fuzzy character of the problem, the integration of a detailed landscape description in form of a multi-layer GIS, remotely sensed data and regional environmental models is suggested for the projects and initiatives outlined in this paper.

Earth observation (or remote sensing) platforms are the primary data source from which landscape patterns can be assessed (Herzog et al., 2001; Blaschke et al., 2001). Without *a* *priori* information about these patterns, observations made by remote sensing sensors supply an independent and unbiased framework to analyse the land cover at multiple scales (Marceau & Hay, 1999; Hay et al., 2002). Two important scale-specific characteristics need to be accounted for: first the spatial, spectral and temporal resolution of each image pixel and secondly the image characteristics themselves, i.e. geographical area, combined band-widths and temporal duration. Within these data sets, only objects with "real-world relevance" may serve as suggested units over a range of scales. Recent studies describe how "image-objects", nested in a hierarchical system of a multi-scale GIS framework are being detected and described by a Fractal Net Evolution Approach (FNEA) to reduce this problem (Baatz et al., 2000; Schiewe, 2003; Hay et al., 2003). The projects outlined in this paper follow an approach that considers spatial, spectral and temporal resolution demands by combining a variety of EO products and is thus in agreement with the suggested key issues by latest landscape research: "data acquisition" and "scaling" (Wu & Hobbs, 2002). We mainly focus on object-oriented and SAR image processing and limitations analysing remote sensing data in rough topography and under forest cover as supplement to algorithms that worked well in flat regions. We provide techniques for their improved application in mixed forest classification. The investigation aims to support the establishment of spatial tools that will contribute directly to water management and forest inventory activities. The philosophy of incorporating concepts and knowledge keeps the system open for new advanced thematic and EO data, that simply can import and extent existing databases. Moreover, the user with knowledge in forestry can align classification rules and therefore revert to a couple of area-wide and high grade processed optical and SAR- EO data sets. For our resulting data infrastructures, community involvement is the most challenging objective and remains an ongoing process.

2 Projects, objectives and study area

We resume to projects which address environmental impacts like forest degradation or water pollution in Northern Mongolia and the Lake Baikal region. The multi-scale character of such investigations implies the acquisition, storage and analyses of spatial data, like satellite-based EO data. The projects introduced here are carried out within international consortiums.
For the Technical Cooperation Programme Mongolia we used the full capacity of area-wide and mostly free available Remote Sensing data to support the development of **(1)** *Participatory Forestry*. The approach provides an object-oriented GIS environment containing intersected spectral, topographical and topical data for the preparation of forest maps. Optimised post-processing and sensor synergies enabled classifications of prevailing forest types, disturbances, timber density and timber volume. Moreover technology consultancy and the provision of advanced training on the use of Image Processing/GIS encouraged capacity building at the Mongolian Ministry for Nature and Environment (MNE). The project was founded by UN's 'Food and Agriculture Organization' (TCP/MON/2903).

The **(2)** *Irkutsk Regional Information System for Environmental Protection – 'IRIS'* will assess the current status and dynamics of the Irkutsk Region's (South-eastern Siberia, watershed area of Lake Baikal) forestry environment, influenced by man-made changes and anthropogenic impact arising from pollution sources and other negative anthropogenic drivers located in the region and in adjacent areas. It will investigate the responsiveness and vulnerability of forestry environment within the Region under different scenarios of industrial development and nature-preserving measures. The output of the project is the adaptation of the existing GIS layers, completion and transfer into operative testing and exploitation a simplified version of the Regional Information System that serve as a prototype for other regions of Northern Eurasia. The project also includes the preparation of the detailed prospective studies and explorations aiming at the development of the efficient simulation and management tool for practical use by regional governance and nature-protection service(s). The tool will help to manage the risks associated with man-made changes and anthropogenic stress affecting the forest ecosystem of the region under investigation, as well as other regions of Northern Eurasia. The project is founded by the European Commission (INCO-CT2006-015110).

The **(3)** *'GMES (Global Monitoring for Environment and Security) Service Element Forest Monitoring'* provides another powerful tool for effective forest monitoring and inventory at regional scale for the Irkutsk Oblast. Reliable and up-to-date information on forest characteristics and changes are required by the State Forest Service of Irkutsk General Survey of Natural Resources (FS of GSNR) in order to fulfil several international and national treaties, for its own forest policy as well as to perform the task of delivering data to the federal level. Specific information needs arise in the area of continuous inventory of forest and forest fund land as a basis for perspective and operative planning, an assessment of changes and trends such as ARD (Afforestation, Reforestation, Deforestation) resilience of forests, etc. and information provision of forest protection with respect to natural and human-induced disturbances such as fire, insect outbreaks or illegal logging. Without the proper implementation of a multi-sensor EO-concept, the FS of GSNR is not able to satisfactory solve its major tasks. The project is founded by the European Space Agency.

To answer the question "How do land cover/ land use changes affect hydrology and thus the detrital discharge of sediments and dissolved particles into Lake Baikal?" we need a better understanding of sediment fluxes as controlled by present day hydrological conditions. The **(4)** *SELENGA initiative* provides a generic methodology using EO data to overcome information gaps for evaluating processes as controlled by the Selenga River and other large and inaccessible watersheds. These are relevant for assessing the hydrological balance in the catchment area and understanding related sediment mass and pollutant transfer to the Lake Baikal. The main scientific innovation results from a combination of EO- based land surface monitoring, an assessment of landscape patterns, and a water-bound sediment transport model. The proposed approach serves as the basis for regionalized land use scenarios under changing climate conditions and their impact to the watershed. Furthermore, hot spot detection and rating of illegal gold placer mining areas will be carried out by using new advanced high resolution sensor systems. The initiative will be undertaken in collaboration with the Centre for Environmental Research Leipzig and the GeoForschungszentrum Potsdam, Germany.

Based on open standard technologies the **(5)** *Siberian Earth System Science Cluster – 'SIB-ESS-C'* will be developed as our comprehensive spatial data infrastructure for remote sensing product generation, data dissemination and scientific data analysis. SIB-ESS-C is the follow-on activity to the EU funded SIBERIA-II project (Multi-Sensor Concepts for Greenhouse Gas Accounting of Northern Eurasia, EVG2-2001-00008). The tools and systems which have been employed for it include a selected yet spectrally and temporally diverse set of 15 Earth observation instruments on 8 satellites, detailed GIS databases and some of the worlds most advanced Dynamic Global Vegetation Models (the Lund-Potsdam-Jena LPJ-DGVM and the Sheffield-DGVM) to account for greenhouse gas fluxes between land and atmosphere. Following the principle of interoperability SIB-ESS-C is planned to become part of a distributed network of similar systems where not only data is being distributed and shared, but also applications (e.g. analysis functionalities, processing modules) are being offered and used throughout the network. The project is founded by the University of Jena.

3 Data

To achieve the above mentioned objectives, three points are critical for the integration of EO data. First, a distinction between spatial resolution and image extent in the EO data is essential. Generally, the higher the spatial resolution of the pixels the smaller the total area covered by the image. Second, the lower spatial resolution, the higher temporal availability of images, so far. EO products such as time series from Terra/Aqua MODIS can serve as baseline information for higher resolution images at a more or less particular point in time. Third, ground truth data like forest inventories are essential for the classification validation. Such data are often available as analog-to-digitized maps. Additionally provided forest parameter attribute tables from ongoing inventories have no established spatial relationships to both, the maps and the EO images which would a prerequisite for automated classification and parameter retrieval validation. An object-oriented approach that might offer a solution to this methodological problem is described later on.

To ensure a continuous, area-wide approach we have chosen spatially low (250-1000 meters per pixel) but temporally high resolution data from ENVISAT-MERIS and Terra/Aqua MODIS, spatially medium resolution data (25 – 50 meters per pixel) from Landsat-5 and –7 as well as from the Radar sensor systems ERS-1,-2, ENVISAT ASAR and ALOS. Until today the data were acquired both at no costs from internet databases and external project agreements or at low cost reproduction fees. Expenses were therefore be minimised. For the future we are looking forward to major data contingents on spatial high resolution data (1 – 15 meters per pixel) from the new Radar sensor system TerraSAR-X. The University of Jena is a principle investigator on the use of TerraSAR-X data in forest environments.

4 Methodologies

The use of satellite-based EO data and derived information for water and forest applications is commonly hindered by mismatches between information user and information provider. The underlying cause is that remote sensing sensors are the front-end stage and complex data manipulation is needed in the process flow in order to meet the user's requirements for the final thematic content. The initial satellite vision system focuses on a specific spectral and spatial domain at a particular point in time. The transformation of such records into reliable information is controlled by the capabilities of image processing and expertise about the ground. Data of different EO sensors in different spectral, spatial and temporal dimensions need to be made compatible and synergistically available to be in line with user's demands. Altogether this limits the use of EO data but nevertheless EO platforms are the primary data source for up-to-date information, whereas derived landscape units (or objects) are regarded as primary units for further landscape quantification and modelling. Thus, our general concepts set high value on advanced database integration of multi-resolution EO data. From the multitude of image manipulations to be undertaken (atmospheric correction, topographic normalization, unsupervised classification) we will focus on object-oriented and SAR image processing. For concept and knowledge building, the involvement of forester and hydrologist community is the most challenging objective and remains an ongoing process.

4.1 Data infrastructures

For the SELENGA initiative (see figure 1), final data infrastructures will provide the object-related spatial, spectral, temporal and thematic content in conventional vector and grid formats and in a data description meta-language such as the Extensible Markup Language (XML). The core application is the derivation of land surface characteristics from spatial medium resolution SAR and optical EO data by taking topographic properties into account. Land surface baseline information will be generated from several time series products of the Terra/Aqua MODIS sensor. The GIS-Client summarises tools for a hierarchical landscape discretisation starting with river network topology, land cover classification, object creation and GIS- and time series data intersection. These objects find entrance into the Open Source Relational Data Base (RDB) that finally hosts landscape units from different levels of spatial discretisation, featured with different parameters over time. The units will be warehoused in different formats like vector, grid or metadata. An XML-metadata file describes one parameter over one object/unit for a specific time span such as the statistical mean of surface temperature for a specific sub-catchment for the days 145 to 153 in year 2005. The reason for using XML is it's familiarity to RDF (Resource Description Framework) or OWL (Ontology Web Language) and thus it's capability in the development of concepts and knowledge which is requisite for the better understanding of multidimensional environmental processes.

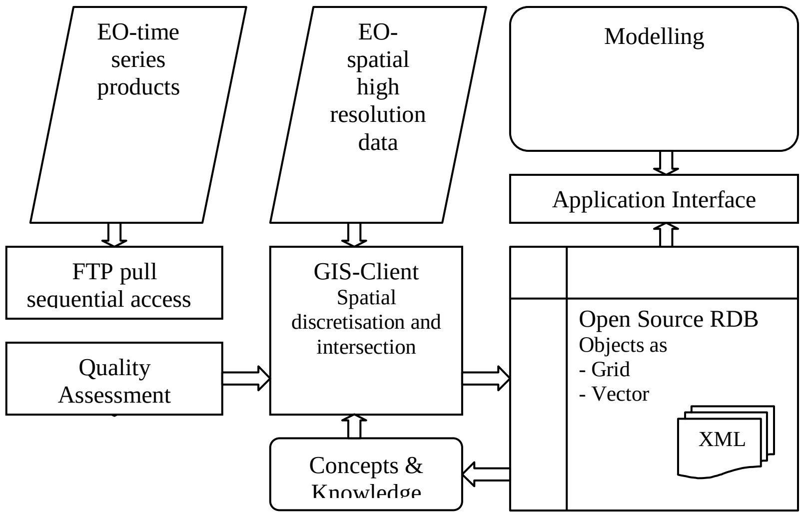

Figure 1: Workflow and data infrastructure for a Selenga watershed evaluation.

4.2 Object-oriented image processing

We applied an object-oriented image processing approach which allows the textual and spatial intersection of the continuous EO raster data (see figure 2) to "sculpture" objects based on the corresponding spatial thematic vector data sets. Multiresolution segmentation is a bottom up region-merging technique starting with one-pixel objects. In numerous subsequent steps, smaller image objects are merged into bigger ones. For the segmentation we used the layers with the highest feasible spatial und textural resolution. The forest inventory classes serve as super objects (compartment external border) and sub-objects were created from image grey level channels within the super object boundaries. As result we obtain vector polygon files and corresponding attribute tables containing intersected forest inventory and image grey level statistics. Manual object fusion depends on local knowledge but is recommended to reduce complexity. The strategy for using eCognition object-oriented image classification can briefly outline as (1) build up of a knowledge base for the interpretation process, (2) image object generation, (3) class hierarchy generation and (4) class description by representative sample objects and a nearest neighbour classifier or the formulation of knowledge using one-dimensional membership functions.

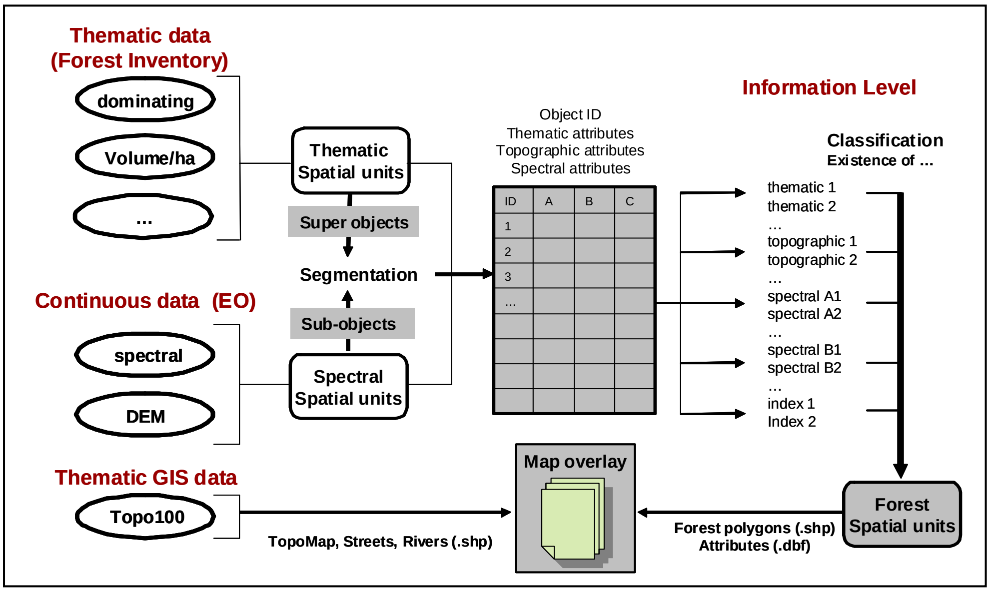

Figure 2: An object-oriented approach within a GIS-Client allows the textual and spatial intersection of the continuous EO raster data with thematic forest inventory vector data.

4.3 SAR data processing

SAR data application in hilly and mountainous areas is limited due to inherent image geometry. According to the cross- and along-track components of radar images, a strong dependence of radar backscatter to mountain slopes is given. Thus, for forest mapping, radiometric and geometric corrections are necessary. Prerequisite is the availability of precise elevation data. Forest test sites in the region can be characterized by rough topography with steep slopes. For forestry, the use of low incidence angles enhances the sensitivity to biomass, whereas the use of high incidence angles enhances the discrimination of forest types through interaction with forest structure. The use of cross polarisation improves the discrimination between volume scattering (vegetation) and surface scattering (soil), in our case the forest/non-forest discrimination and the retrieval of low biomass values (forest regeneration, re-growth). In case of forest area/type applications AP mode with VV/VH polarisation for swaths IS4-6 is recommended (ASAR Science Advisory Group, 1998). For processing we used GAMMA SAR Geocoding and Image Registration Software (GEO) which is a collection of tools for SAR geocoding and multi-source image registration. Geocoding is the coordinate transformation between the coordinates of the imaging system, in our case the range-Doppler coordinate of the ASAR, and orthonormal map coordinates from SRTM-DEM for terrain correction. The processing steps can conclude as follows: (1) Generation of Multi-Look intensity images (sigma nought) in slant range geometry including radiometrically normalization for the antenna patterns, (2) coregistration in slant range geometry, (3) geocoding slant and ground range to DEM geometry, (4) geometrical normalization by local incidence-angle correction, (5) GammaMAP filtering and (6) calculation of dB-values for better visualisation.

4.4 Concept and knowledge building

For concept and knowledge building we applied General Linear Models (GLM) and here the one-way analysis of variance (ANOVA) to test specific image and inventory object attributes for significant differences between means by comparing variances. The GLM is a generalization of the linear regression model, such that effects can be tested for continuous predictor variables, as well as for effects for categorical predictor variables. In one-way ANOVA designs, the effect of a single categorical grouping variable like tree type or volume per hectare on one or more continuous dependent variables (e.g. objects mean NDVI or Radar intensity) can be evaluated. We are able to compare the variance due to the between-groups (spectral values between different channels) variability with that due to the within-group variability (spectral values within one channel). Under the null hypothesis (that there are no mean differences between groups), the variance estimated from the within-group variability should be about the same as the variance estimated from between-groups variability. This latter least square means effect from between-groups variability is then tested on statistical significance. Furthermore, Terra/Aqua MODIS 16-day vegetation indices are intersected with the forest inventory units (super-objects) and sub-units (sub-objects) for a better understanding of annual and inter-annual variations of the phenological turnover.

5 Summary

In this paper projects and initiatives addressing environmental impact issues in Northern Mongolia and the Lake Baikal Region are presented. Some of the concepts of spatial data infrastructures and remote sensing image processing techniques currently being established at the University of Jena are introduced. The major goal is to efficiently share Earth Observation data and domain-specific information within earth science community and thus using standards published by the World Wide Web Consortium (W3C®), the Open Geospatial Consortium (OGC™) or the International Organization for Standardization (ISO). On the long-term, decision makers and earth science communities will highly profit from the evolving field of semantic webs, where domain-specific knowledge has been rigorously categorized in an also machine readable format.

Forest area maps for the Irkutsk region are generated using medium-resolution, cross-polarized Envisat ASAR precision images acquired at large incidence angles (swath 7). For quality control and product verification archived Landsat TM5 and Landsat ETM 7 as well as ground reference data will be used. The following products will be provided: “Forest Area Map 2005/2006”, “Forest Area Change Map”, “ARD Area Map”, “Burned Area Map”, “Clear Cut Map”. The service includes a documentation of the applied methodology as well as the assessment of product quality and accuracy. The acceptability threshold of the thematic mapping accuracy is 90% for forest areas and 85% for forest area changes, respectively.
For *tree type classification* linear regression models (ANOVA) were applied to test grey level statistics of segmented image objects with forest inventory parameters for significance. The results imply that spatially medium resolution EO data such as from Landsat work well in the case of pure forest stands. For mixed forests, inner- and inter-channel variances of object spectral as well as object texture values are not reliable for the discrimination of prevailing tree types. However, inner- and inter-channel variability is given and therefore show the potential of such data for using their object-related spectral and texture statistics in the nearest neighbour classification feature space. Additionally, rules like the dependence of Siberian Pine with elevation, the occurrence of Birch after fires or the inner-annual trend of NDVI for Larch-dominated areas were implemented by one-dimensional membership functions. For the classification, only few rules can be validated since forest inventory statistics often are not congruent with forest inventory boundaries. Manual object fusion is therefore highly recommended to reduce complexity where forest inventory boundaries include non-forest

land cover. The task depends on local knowledge in which the proposed approach offers a solution to this methodological problem.
A methodology to retrieve *timber volume* using SAR data has been developed as part of the SIBERIA-I project. The algorithm has been proven to work well for Siberian forests in flat regions (Wagner et al. 2003; Eriksson et al. 2003). Due to the rough topography in the Mongolian test sites and partly absent coverage from ERS-1/2 and JERS, this product can only be provided for the relatively smooth undulated western parts of Selenge-Mandal test site. Due to insufficient quality and/or test site coverage, the data were not considered in the general object-oriented analyses concept, but on demand can easy implemented in the database environment.

With an adequate infrastructure, Earth Observation data can serve as the base for analysing landscape structure and related socio-economic and environmental processes at multiple scales. The discussed projects and approaches generate and compound domain-specific knowledge on both sides, the information providers and the information users. Besides investigating long-term regional environmental problems, Earth Observation data sets are of major relevance for providing on-demand information such as for large and partly inaccessible regions like the transboundary basin of the Selenga river. EO data sets are normally less precise than field survey data available at state or regional level. But, however, the integration of the latter data from different sources (when river basins extend into a number of administrations) is also surrounded by considerable uncertainty. In contrast, satellite-based EO data serve as an independent and unbiased framework for less precise but standardized information. With different image processing methodologies using optical and SAR data we were able to demonstrate the potential of EO data for investigating meso- and macro-scale environmental impact problems in Northern Mongolia and the Lake Baikal Region. The methodology and analysis discussed in this paper can be applied to other (transboundary) regions, using free of charge EO data. New sensors like TerraSAR-X (SAR), EnMAP (hyperspectral) or RapidEye (optical) will significantly improve the integration of EO data to support the aims of environmental protection.

References

ASAR Science Advisory Group (1998): ASAR Science and Applications. In Harris (Ed.), European Space Agency 1998, ESA SP-1225.

Baatz, M. and A. Schäpe (2000): Multiresolution segmentation – an optimisation approach for high quality multi-scale image segmentation. Angewandte Geographische Informationsverarbeitung XI, Vorträge zum AGIT-Symposium Salzburg, Strobl, J., Blaschke, T., Griesebener, G. (Eds.), Herbert Wichmann, Heidelberg, 12-23.

Blaschke, T., Conradi, M., Lang, S. (2001): Multi-scale image analysis for ecological monitoring of heterogeneous, small structured landscapes. *Proceedings of SPIE 4545*, 35-44.

Eriksson, L. E. B., M. Santoro, A. Wiesmann, and C. Schmullius (2003):
Multi-Temporal JERS Repeat-Pass Coherence for Growing Stock Volume Estimation of Siberian Forest. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 7, pp. 1561-1570.

Hay, G.J., Dubé, P., Bouchard, A. and D.J. Marceau (2002): A scale-space primer for exploring and quantifying complex landscapes. *Ecological Modelling 153*, 27-49

Hay, G.J., Blaschke, T., Marceau, D.J. and A. Bouchard (2003): A comparison of three image-object methods for the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry & Remote Sensing 57*, 327-345.

Herzog, F., Lausch, A., Müller, E., Thulke, H.-H., Steinhardt, U. and S. Lehmann (2001): Landscape metrics for the assessment of landscape destruction and rehabilitation. *Environmental Management 27/1*, 91-107.

Marceau, D.J. and G.J. Hay (1999): Contributions of remote sensing to the scale issue. *Canadian Journal of Remote Sensing 25 (4)*, 357-366.

Schiewe, J. (2003): Integration of multi-sensor data for landscape modeling using a region-based approach. *ISPRS Journal of Photogrammetry & Remote Sensing 57*, 371-379.

Seppelt, R. and A. Voinov, (2002): Optimization methodology for land use patterns using spatially explicit landscape models. *Ecological Modeling*. 151 (2/3), 125–142.

Wagner, W., A. Luckman, J. Vietmeier, K. Tansey, H. Balzter, C. Schmullius, M. Davidson, D. Gaveau, M. Gluck, T. LeToan, S. Quegan, A. Shvidenko, A. Wiesmann, JJ. Yu (2003): Large-scale mapping of boreal forest in SIBERIA using ERS tandem coherence and JERS backscatter data. *Remote Sensing of Environment*, 85, pp. 125-144.

White, A., X. Sun, K. Canby, J. Xu, C. Barr, E. Katsigris, G. Bull, C. Cossalter, and S. Nilsson (2006): China and the Global Market for Forest Products: Trends, Implications, and Steps to Transform the Trade to Benefit Forests and Livelihoods, Washington D.C.: Forest Trends, Center for International Forestry Research, and Chinese Center for Agricultural Policy.

Wu, J. & R. Hobbs (2002): Key issues and research priorities in landscape ecology. *Landscape Ecology, 17 (4)*, 355-365.