VU Research Portal

Challenges in using land use and land cover data for global change studies

Verburg, P.H.; Neumann, K.; Nol, L.

published in
Global Change Biology
2011

DOI (link to publisher)
10.1111/j.1365-2486.2010.02307.x

document version
Publisher's PDF, also known as Version of record

Link to publication in VU Research Portal

citation for published version (APA)
Verburg, P. H., Neumann, K., & Nol, L. (2011). Challenges in using land use and land cover data for global change studies. *Global Change Biology*, *17*(2), 974-989. https://doi.org/10.1111/j.1365-2486.2010.02307.x

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 25. Sep. 2021

Global Change Biology

Global Change Biology (2011) **17**, 974–989, doi: 10.1111/j.1365-2486.2010.02307.x

REVIEW

Challenges in using land use and land cover data for global change studies

PETER H. VERBURG*, KATHLEEN NEUMANN† and LINDA NOL‡

Institute for Environmental Studies, VU University Amsterdam, de Boelelaan 1085, 1081 HV Amsterdam, The Netherlands, †Land Dynamics Group, Wageningen University, PO Box 47, 6700 AA, Wageningen, The Netherlands, ‡CAH Dronten University of Applied Science, De Drieslag 1, 8251 JZ Dronten, The Netherlands

Abstract

Land use and land cover data play a central role in climate change assessments. These data originate from different sources and inventory techniques. Each source of land use/cover data has its own domain of applicability and quality standards. Often data are selected without explicitly considering the suitability of the data for the specific application, the bias originating from data inventory and aggregation, and the effects of the uncertainty in the data on the results of the assessment. Uncertainties due to data selection and handling can be in the same order of magnitude as uncertainties related to the representation of the processes under investigation. While acknowledging the differences in data sources and the causes of inconsistencies, several methods have been developed to optimally extract information from the data and document the uncertainties. These methods include data integration, improved validation techniques and harmonization of classification systems. Based on the data needs of global change studies and the data availability, recommendations are formulated aimed at optimal use of current data and focused efforts for additional data collection. These include: improved documentation using classification systems for land use/cover data; careful selection of data given the specific application and the use of appropriate scaling and aggregation methods. In addition, the data availability may be improved by the combination of different data sources to optimize information content while collection of additional data must focus on validation of available data sets and improved coverage of regions and land cover types with a high level of uncertainty. Specific attention in data collection should be given to the representation of land management (systems) and mosaic landscapes.

Keywords: climate change, emission inventory, integrated assessment, land cover, land use, spatial data

Received 19 April 2010; revised version received 26 July 2010 and accepted 31 July 2010

Introduction

Land use and land cover change are one of the prime driving forces of changes in the Earth system and climate in particular. Agricultural land uses are estimated to contribute to changes in atmospheric concentrations of greenhouse gases (GHG): non-CO_2 GHG account for 10–12% of the total global anthropogenic emissions. At the same time, agricultural lands generate large CO_2 fluxes both to and from the atmosphere, but the net flux is small. The expansion of cropland and pastures at the cost of forests results in an increase in atmospheric CO_2. This decreases the sink capacity of the global terrestrial biosphere, and thereby may amplify the atmospheric CO_2 rise. Apart from managed lands, large sources and sinks of GHGs are found in natural land cover types: wetlands are a well-known source of methane emissions (Hines *et al*., 2008; Kayranli *et al*., 2010) and much research focuses on the consequences of global change on emissions from arctic and boreal ecosystems (Chapin *et al*., 2000).

Besides the impact of land cover change on GHG emissions, more subtle changes in land management and land use practices can have important consequences. Differences in forest management are likely to influence carbon emissions or sequestration in different ways (Schulp *et al*., 2008b) while agricultural management effects carbon sequestration and GHG emissions (Dendoncker *et al*., 2004). Methane emissions from rice fields are highly dependent on water management (Verburg & Denier van der Gon, 2001) and rice variety choice (Denier van der Gon *et al*., 2002).

Land use and land cover change not only impacts GHG emissions but also land surface properties of relevance to climate. At the global scale, changes in land surface properties associated with changes in vegetation can have impacts on continental and global atmospheric circulation, with possible large impacts on regional and continental climate (Pielke *et al*., 1998; Chapin *et al*., 2000). Impacts of

Correspondence: P. H. Verburg, e-mail: Peter.verburg@ivm.vu.nl

© 2010 Blackwell Publishing Ltd

land use/cover change upon climate include remote impacts upon local circulation regimes often labeled as land use driven 'teleconnections' because effects are communicated to regions distant from the actual changes in surface characteristics (Eastman *et al*., 2001).

The impact of land use and land cover change on climate change is mostly studied from the perspective of land use/cover change as a driver of climate change (Schulp *et al*., 2008a). However, in many cases the interrelations are more complex. Land use/cover change can also be induced by climatic changes. Droughts have strong effects on vegetation and may increase the risk of forest fires. Climate change may also make areas more or less suitable for certain land use management practices leading to changes in land use decisions. Multidirectional impacts may be linked through feedbacks that strengthen or attenuate the interaction between land use/cover change and climate change (Verburg, 2006). Such feedbacks make it difficult to distinguish impacts from drivers (Bossel, 1999). Large-scale deforestation changes climate conditions and, hence, influences vegetation patterns and occurrence of forest fires. These may affect land requirements and reclamation potential (Nepstad *et al*., 2001; Foley *et al*., 2003; Carvalho *et al*., 2004).

A feedback that deserves special attention is the adaptation of land management in response to climate change. Adaptation aims at reducing the potential negative consequences of climate change through adapting management practices in such a way that they anticipate the changed climatic conditions. Adaptation may include a wide variety of measures (Aerts, 2005; EEA, 2008). These include spatial planning measures, for example to restrict new residential locations in areas with high flooding risk, the use of improved crop varieties that are better adapted to the changed climatic conditions, resource management and diversification (Polsky & Easterling III, 2001). Finally, the emerging incentives for reducing emissions from deforestation and forest degradation (REDD) require high quality monitoring of forest dynamics directly related to land use change (Eva *et al*., 2010).

Given the very central role of land use/cover change almost all assessments related to climate change somehow make use of data on land use/cover. These data may originate from different sources and inventory techniques. Each source of land use/cover data has its own domain of applicability and quality standards. In many studies, no specific attention is given to the quality and content of these data. This paper reviews the use of land use/cover data in climate studies and aims at identifying critical issues related to the use of land use/cover data in climate assessments. Based on an inventory of data sources, the domain of applicability and the constraints of particular data sets are discussed ('Inventory of data types on land use/cover'). An inventory of large uncertainties and inconsistencies in land use/cover data that may have critical implications for climate change assessments is provided in 'Consistency issues of land use and land cover data' whereas 'Methods and opportunities to deal with data inconsistencies' discusses several methods and techniques that may reduce such errors and inconsistencies. The implications for global change assessments and integrated assessment models are described in 'Implications for global change analysis and modelling'. The paper concludes with a list of the main findings and recommendations.

Inventory of data types on land use/cover

Sources for land use/cover data

The most common source of data on land cover is remote sensing. Remotely sensed data include information gathered by aerial photography and satellites. Solar radiation (or in case of radar actively emitted radiation) is reflected from the surface of the earth – from soil, water, vegetation and building – to sensors that measure the intensity of different frequencies. Each type of surface reflects or absorbs different frequencies. Based on these measurements, it is possible to make inferences about the earth surface. The reflection is related to the cover of the earth surface which not always reveals the intended use of the land. However, land cover in combination with the spatial structures and additional attributes may allow, to some extent, to derive land use. Park & Stenstrom (2008) subdivide urban land cover derived from remote sensing data into several land use classes based on the dominant use of the area. They indicate that high accuracies in classifying different types of urban land use can be achieved.

For the interpretation of remote sensing imagery, a wide range of methods exists ranging from fully automated, to semiautomated, to pure interpretation (Richards & Jia, 2006). Fully automated approaches, i.e. direct processing of image data using models or other techniques, have the advantage of fast processing allowing high temporal resolution and replication which is especially useful in monitoring rapid deforestation (Asner *et al*., 2009). Supervised techniques and visual interpretation allow the integration of expert-knowledge, field observations and pattern recognition (Sirén & Brondizio, 2009), but also introduce an element of subjectivity making the interpretations dependent on the observer (Foody, 2002). This may cause problems in replication and therefore change detection.

Another commonly used source of land use/cover data are surveys and census data. Many countries and international agencies collect statistical information on land use. Agricultural land use types are often reported as part of an agricultural census. Besides cropping areas such information often also includes management information such as irrigation, fertilizer application rates and crop yields. Similarly, forestry statistics may provide information about management practices. Therefore, census information is highly suitable to provide land use information which can never be collected through remote sensing. However, most census information is focused on economic sectors and information on (semi-)natural land uses is often not included. In addition, data are often aggregated to the level of administrative units while the original data are not available as result of privacy legislation (Sabor *et al.*, 2007).

Besides remote sensing and census information other, but less frequently used, sources of land use/cover data include maps based on field surveys, participatory maps (Rambaldi *et al.*, 2007) and cadastral information (Aspinall, 2004). An overview of the characteristics of different sources of land use/cover data is given in Table 1.

Applicability domains

It is important to notice that the different sources of data have clearly different possible applications. High-resolution remote sensing data, for example from IKONOS and Quickbird, provide detailed land cover information and the possibility to infer land use information. However, acquisition and interpretation of such data is only feasible for relatively small areas and may be very costly. Therefore, global land cover data sets based on remote sensing data are using medium to coarse resolution sensors, e.g. in case of the GLC2000 dataset (http://www-tem.jrc.it/glc2000/) use was made of SPOT images (1 km resolution) while the recent GLOBCOVER data (http://ionia1.esrin.esa.int/index.asp) are based on MERIS images from ENVISAT with a spatial resolution of 300 m. The differences between sensors and the spatial resolution have an impact on the representation of landscapes and the thematic information that can be derived. In many land cover data sets, each pixel is classified as one specific land cover type. When land cover types mainly occur in small patches the occurrence of these land cover types may be underestimated because they will not dominate the reflectance characteristics of the pixel. Moody & Woodcock (1994) show that large proportion errors can arise as landscapes are represented at increasingly coarse scales. Therefore, the appropriate spatial resolution of the sensor may depend on the landscape under analysis (Ozdogan & Woodcock, 2006). Various methods are developed to correct, to some extent, for this aggregation problem either through statistical techniques (Moody & Woodcock,

Table 1 Overview of the spatial, temporal and thematic properties of various sources of land-use and land cover data

Data source	Spatial resolution	Spatial extent	Temporal resolution	Temporal extent	Thematic properties
Remote sensing/ Aerial photography	Dependent on sensor (remote sensing mostly between 0.6 m and 1 km)	Dependent on sensor. Coverage is limited in case of clouds (not for radar)	Frequent depending on sensor/satellite	Depending on launching and life time of sensor. Few remote sensing data are available before 1970s except for aerial photographs	Land cover classes. Classification is based on sensor characteristics and user preferences
Census/ survey data	Administrative units	Often national level	Infrequent depending on census, often less than every 10 years	Country specific depending on statistical system	Focus on economic sectors (mostly agriculture and forestry)
Land-use maps based on field survey	Dependent on scale of mapping (often between 1 : 25 000 and 1 : 1 million	Varying	Often made for 1 year only	–	Varying and fixed within a specific map
Participatory maps	Dependent on scale of mapping	Often restricted to territory of one or more communities	For one moment only	Participatory back casting possible	Depending on purpose of mapping
Cadastral information	Precise information at property level	Dependent on cadastral system	Continuously updated	Often available for long time period	Limited to tenure conditions with limited information about land use, especially in urban environments

1994; Marceau & Hay, 1999) or through unmixing of the spectral information in the pixel during the interpretation and reporting of the fractional coverage of each land cover within the pixel.

Whereas the discussion on spatial resolution effect is mostly concerned with raster data a similar effect can be found in polygon representations of land use data. An important generalization mechanism inherent to polygon maps is the minimum mapping unit (MMU), i.e. the minimum size that a land unit must exceed in order to be represented in the map. Land units smaller than the MMU are simply not represented. This may result in an overrepresentation of the land cover/use types that occur in larger units. An exception is the case of mosaic polygons, which have a compound label representing a mosaic of patches from different classes, all smaller than the MMU. In this kind of polygons, the percentage cover from each class may be reported, but information regarding their actual distribution within the polygon is missing (Castilla & Hay, 2007).

Besides spatial scale, temporal scale is also important. Survey and census data tend to be infrequent and sampling schemes and definitions change between surveys. Remote sensing data offer the best possibility to monitor changes in land cover through time. However, time series based on remote sensing tend to suffer from inconsistencies due to improved resolution of sensors and changes in classification schemes. Moreover, accuracies of classification are often around 80% while change in land cover over the period analyzed is often small. Therefore, temporal consistency needs be ensured by focusing on change in the spectral information rather than providing two individual interpretations for the different years (Hansen *et al*., 2008; Berberoglu & Akin, 2009). Furthermore, cloud cover may cause problems in large parts of the humid tropics during a considerable part of the year (Ju & Roy, 2008). Radar images may overcome these problems to some extent (DeFries, 2008; Freitas *et al*., 2008).

Consistency issues of land use and land cover data

Similar studies conducted in different regions and at different times often use different sources of data, which hinders a comparison between them. In addition, researchers may choose from different sources of land use/cover data that contain different information. Comparison and integration of different data sources is hampered by different issues including:

- difference between land cover and land use;
- temporal consistency;
- spatial consistency and scaling bias;
- thematic differences and inconsistencies.

The following paragraphs will discuss the underlying reasons for consistency problems in more detail and provide a number of illustrations.

Temporal consistency issues

For monitoring and analysis of change in land use/cover, it is essential to have consistent data over a longer period of time. Preferably, these are derived from exactly the same data source with applying the same processing techniques. However, often this is not possible. Because technology, science and policy objectives are continuously changing, repeated natural resource inventories rarely employ the same methods as in previous surveys (Wadsworth *et al*., 2008). The production of spatial data sets from remote sensing has often been driven by short-term funding and specific information requirements by the funding agencies. As a result, a wide variety of data sets exist that were generated using different atmospheric classification algorithms, land cover classes, training sites, map projections and spatial resolutions. Remote sensing sensors have changed in time and interpretation techniques have become more advanced. Furthermore, many interpretations of remote sensing data include a ‘supervised’ component in which classification decisions and generalizations are made by an expert. This is not necessarily the same person as the one that has processed the data for an earlier year.

Also survey methods and inventory techniques may change in time, including classification systems, legend classes and spatial detail. When analyzing time series of data, it is important to distinguish changes caused by inconsistencies between the data sets and the ‘real’ changes. In practice, these are often difficult to separate and it is likely that the land use/cover change reports from such exercises over- or underestimates the ‘real’ change.

Spatial consistency issues

Spatial inconsistencies between data sources may be related to positional errors caused by georeferencing of remote sensing data or the level of generalization of a vector map (Castilla & Hay, 2007). If georeferencing is appropriately done such inconsistencies are small as compared with inconsistencies as result of differences in spatial scale and aggregation. Nol *et al*. (2008) compare a series of land cover data from different sources available for the fen meadow landscape in Western Netherlands (Fig. 1). The landscape is dominated by grassland with a large number of small, linear, water bodies (ditches). The less detailed data do not represent the linear figures of the landscape and grassland cover dominates in the coarse scale data sets. These

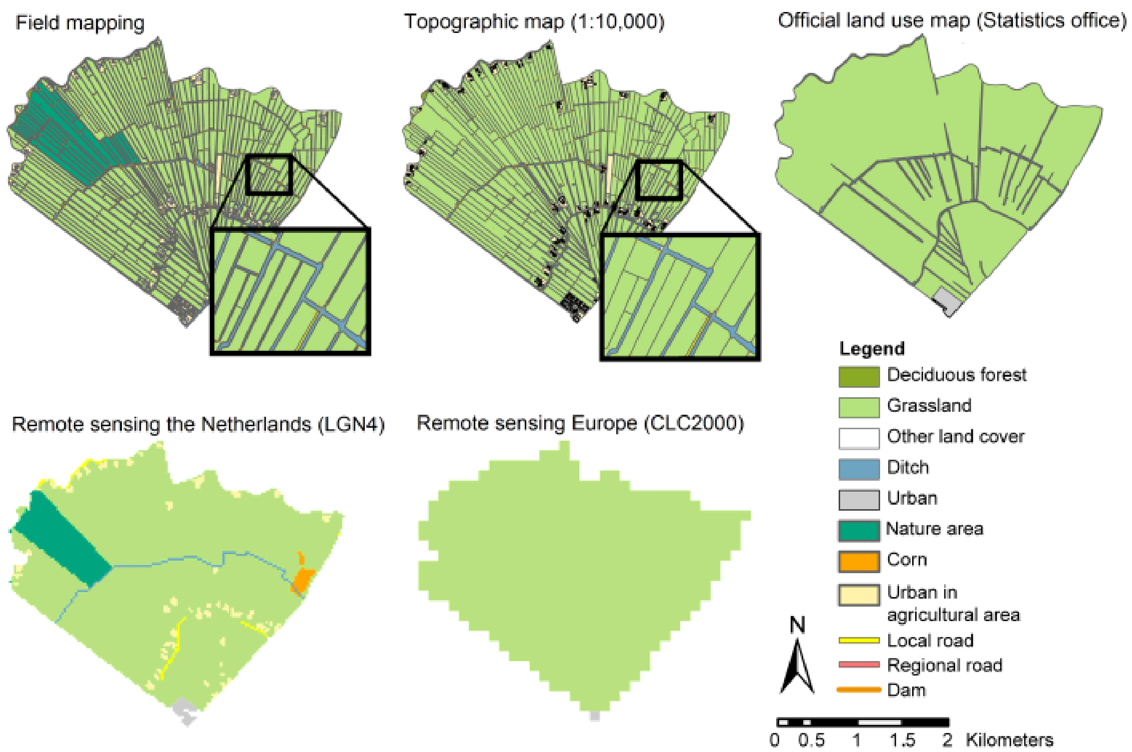

Fig. 1 Representations of polder Zegveld using different land cover databases.

differences have large consequences for the land use statistics that can be derived from the data. In the field mapping grassland only covers 77% of the area and water covers 11%. The CLC2000 map classifies 100% of the area as grassland. Such differences have large impact on the estimation of GHG emissions given the large differences in emissions of nitrous oxide and methane from grassland and water bodies.

Comparable findings were reported by Fassnacht *et al*. (2006) who found the class 'broadleaf trees,' which forms narrow linear features along rivers in the landscape studied, to be particularly susceptible to changes in resolution. Ozdogan & Woodcock (2006) also noted that large landscape elements can support large pixels, but when the landscape elements of interest are small, fine resolution is needed to correctly estimate surface areas. In many landscapes, linear or other small land cover types have important impacts on the functioning of the hydrological and climatic system (Ellis *et al*. 2009).

Spatial inconsistencies may also originate from improper aggregation of the data to the level of analysis. The level of aggregation and the aggregation method are extremely important determinants of the resulting land use distributions and spatial configurations. Dendoncker *et al*. (2008) compared differences in landscape composition and configuration between three aggregation methods and three spatial resolutions. Differences between aggregation results were found to be at least as large as the differences between the results of widely diverging scenarios of land use change. This study, as well as a similar study by Schmit *et al*. (2006) demonstrated that both the rasterizing method and the level of aggregation can become important sources of uncertainty. The spatial bias introduced by aggregation is strongly dependent on the landscape type and level of analysis. In most studies related to climate change aggregation methods are arbitrarily chosen as a result of the software used and hardly ever reported.

Thematic differences and inconsistencies

Different data sources have different capacities in capturing specific land use/cover types given the characteristics of the observation technique. Each data source will therefore lead to specific categorical uncertainties in the final land use/cover data. Categorical uncertainty is commonly assessed using a contingency table of agreement between predicted and observed values, typically referred to as the confusion matrix (Castilla & Hay, 2007). Comparison and assessment of different data sources for the same area has shown that some land use/cover types, independent of the data source, are easier to delineate than other land use/cover types. For an area in Germany, Bach *et al*. (2006) compared the performance of three different data sets on land use/cover with accurate field information. The results showed that the congruency of the land use classes forest, urban and traffic areas are higher than the congruency of the land use classes of the open land (arable land, pastures and meadows, fallow). A similar conclusion was reached by an analysis of global land cover maps at 1 km spatial resolution by Herold *et al.* (2008). The authors harmonized the thematic legends of four available global land cover maps (IGBP DISCover, UMD, MODIS 1 km and GLC2000), which have arisen from different initiatives and are based on different remote sensing data and used different methodologies. Analysis of the agreement among the global land cover maps and existing validation information highlights general patterns of agreement, inconsistencies and uncertainties. The thematic classes of evergreen broadleaf trees, snow and ice and barren show high accuracy and good agreement among the data sets. The classes of mixed tree types show high inconsistency. Overall, the results show a limited ability of the four global products to discriminate mixed classes characterized by a mosaic of trees, shrubs and herbaceous vegetation. There is a strong relationship between class accuracy, spatial agreement among the data sets and the heterogeneity of landscapes.

In another comparison between different data sets for Europe, Neumann *et al.* (2007) identify that comparison of different data sets using a similar legend is hampered by differences in the thematic content of the classes between maps. Interpreters have often used dissimilar definitions of land use/cover classes and used different aggregation schemes. Differences in thematic content are also a direct result of the specific purpose for which a data set was constructed.

A wide range of definitions are used for the same land cover class between different data sets and inventories. Wadsworth *et al.* (2008) shows [based on an inventory of Lund (2004)] that official definitions of forest include the full range of canopy cover between 10% and 80%. Estimates of the Earth's land surface covered by rangelands vary even more and range from 18% to 80% of the earth surface (Lund, 2007). This wide range of variation is due to differences in definitions used but also due to different estimates of the land surface (inclusive or exclusive ice covered surface, etc.) and data source. Large inconsistencies in the definition of rangelands are also a result of the lack of an international organization responsible for the assessment and reporting on the world's rangelands as there is for the periodic global forest assessments by the Food and Agricultural Organization of the United Nations (FAO) (Lund, 2007; Grainger, 2008). The classification of rangelands is especially troublesome in case of sparse tree cover, such as in savannahs. If separate definitions of forest and rangeland are used that are not mutually exclusive there is a risk of double-counting or underestimating the respective areas with major implications for the assessments based on these numbers.

Land cover vs. land use

An important issue underlying differences in area reported is the distinction between land use and land cover. Several studies have identified the relation of land cover and land use as one of the major challenges for monitoring, modelling and communicating land change (Comber, 2008; Verburg *et al.*, 2009). Land cover addresses the layer of soils and biomass, including natural vegetation, crops and human structures that cover the land surface. Land use in contrast refers to the purposes for which humans exploit the land cover (Fresco, 1994; McConnell & Moran, 2001). Land use is not always easily observable, although, in many cases, land use may be inferred from observable activities (e.g., grazing) or structural elements in the landscape (e.g., the presence of logging roads). An example of the potential effects of the differences between the land cover and land use definition is the documentation of land abandonment. While agricultural statistics indicate in Europe a strong decreases of agricultural areas these are, in many cases, not observed in data derived from remote sensing. One of the reasons for not being able to observe the ongoing process of agricultural abandonment by remote sensing is the use of 'abandoned' grasslands for other functions, for example horse keeping. Although the land cover in these areas remains the same the changing land use has large implications for the functioning of the land and the rural economy. Because in many countries hobby-horses are not included in agricultural statistics the extent and areas used for this type of land use are largely unknown. Many authors have also reported large declines of agricultural areas in Europe's mountain areas (MacDonald *et al.*, 2000; Etienne *et al.*, 2003; Tasser *et al.*, 2007). These mountain areas are facing two related trajectories of change: part of the meadows are more intensively used, while other parts have been converted to pasture or have been abandoned (Mottet *et al.*, 2006). These changes in intensity and the actual use of the grasslands, either for pasture or hay-making is not observable by remote-sensing leading to a discrepancy with agricultural statistics.

Methods and opportunities to deal with data inconsistencies

The differences and inconsistencies between data set have been acknowledged by several researchers. Different methods are proposed to deal with these inconsistencies and make optimal use of the available data. This section will discuss a number of methods for data integration and data management that acknowledge the complementarities and inconsistencies of the different land use/cover data.

Integrating remote sensing and census data

Many land cover/use data sets combine (sub-)national census data with remote sensing information (Table 2). Ramankutty *et al*. (2008) assume that the inventory data represent the 'true' areas of agricultural land (except for identified outliers), while satellite data are used to spatially disaggregate these inventory data within each administrative unit. This way the strength of the ground observations is combined with the high spatial detail of remote sensing data. However, the use of inventories from various, national level censuses causes a risk for the use of inconsistent definitions of the agricultural land use classes between countries leading to a globally inconsistent map.

Integration techniques have been used in the preparation of data sets for individual countries or continents, e.g. Hurtt *et al*. (2001) combined land use statistics and remote sensing data to generate a land use map for the United States, Cardille & Foley (2003) combined remote sensing and census data to estimate land cover change in the Brazilian Amazon and Pelorosso *et al*. (2009) used a similar combination to detect land use changes for a small area in Central Italy. While many of the data sets listed in Table 2 focus on agricultural land, a number of other studies have focused on preparing data sets for urban areas. Although urban areas only cover relatively small areas at a global scale they may have a large impact on (local) climate. For such data sets census data are combined with remote sensing data, often complemented with imagery of nighttime lights as an indicator of urban activity (Schneider *et al*., 2009).

For climate change studies, the history of land cover change is also interesting. Several attempts have been made to reconstruct historic land use based on sparse data of human population, potential vegetation, old topographic maps and model assumptions. Based on such data, authors prepared global land cover reconstructions (Ramankutty & Foley, 1999; Klein Goldewijk, 2001; Wang *et al*., 2006; Pongratz *et al*., 2008). Ruddiman & Ellis (2009) have used explorations of land use change during the Holocene to explain historic trends in atmospheric CO_2 concentration and have thus been able to better understand the role of land use change on historic climate changes.

Techniques for integration of different remote sensing data

Combining different existing land cover data sets into one consistent database has the risk of inconsistent use of information and unknown sources of uncertainty. Alternative methods have been presented that use primary data from different sensors through equalizing thematic and spatial content of the data or through change detection techniques. Other methods explore the information contained in data inconsistencies to detect land cover changes. Petit & Lambin (2001) present a methodology to integrate multisource remote sensing data into a homogeneous time series of land cover maps for change detection. The authors developed a method to increase the comparability between

Table 2 Overview of the main characteristics of a number of harmonized, global data sets for land cover based on the combination of different individual land cover products

Method	Reference	Thematic coverage	Spatial resolution (min)	Time period
Remote sensing and (sub)national inventory data	Ramankutty & Foley (1998)	Croplands	5	1992
Remote sensing and (sub)national inventory data	Monfreda *et al*. (2008); Ramankutty *et al*. (2008)	Croplands, grasslands, 175 crop types	5	Circa 2000
National level census data and available thematic spatial datasets	Erb *et al*. (2007)	Cropland, grazing, forestry, urban, transportation	5	2000
FAO national statistics, IGBP DIScover, Global Land Cover 2000 (GLC2000)	Goldewijk *et al*. (2007)	Cropland and grasslands	5	1990–2000
Satellite imagery, ecological modeling, country surveys, existing maps of potential land cover and layers of the major anthropogenic land covers	Sterling & Ducharne (2008)	Cropland, built-up land, grazing land, wetlands, irrigated land, inundated land	5	1990–2000

land cover maps coming from panchromatic aerial photographs and SPOT XS (multispectral) data by equalizing their levels of thematic content and spatial detail. By controlling successively the parameters that influence the level of map generalization, the equalization of the thematic content and of the level of spatial detail of the two land cover maps can be significantly improved.

Postclassification comparison is the most common method to detect land cover changes between two images, i.e. the change is identified based on a comparison of the interpreted land cover maps of time 1 and time 2. However, the accuracy of this change detection technique is only as good as the result of the multiplication of the accuracies of each individual classification. Studies have identified image differencing as being the most accurate change detection technique (Macleod & Congalton, 1998; Ridd & Liu, 1998; Petit *et al*., 2001). The difference between two images is calculated by finding the difference between each pixel in each image, and generating an image based on the result. This method avoids the detection of change as result of differences in interpretation of the individual images.

For detection of deforestation over larger areas, Hansen *et al*. (2008, 2009) have proposed the use of a combination of easily available images with coarse spatial resolution and high-resolution images for sample areas used to calibrate the information of the coarse resolution images. By using this combination of different image types, the method allows for fast monitoring of tropical deforestation across large areas such as the Brazilian Amazon and Indonesia.

Wadsworth *et al*. (2008) present a method utilizing aspects of quantified conceptual overlaps and semantic-statistical approaches (Comber *et al*. 2005). The method is applied to reconcile three independent land cover maps of Siberia, which differ in the number and types of classes, spatial resolution, acquisition date, sensor used and purpose. A map of inconsistency scores is presented that identifies areas of most likely land cover change based on the maximum inconsistency between the maps. The method of quantified conceptual overlaps was used to identify regions where further investigations on the causes of the observed inconsistencies seem warranted. The method highlights the value of assessing change between inconsistent spatial data sets, provided that the inconsistency is adequately considered.

A common approach to address the problem of classification uncertainty in remote sensing and GIS data is fuzzy logic (Woodcock & Gopal, 2000; Robinson, 2003; Fritz & See, 2005). Jung *et al*. (2006) present a straightforward method that merges existing land cover data sets into a desired classification legend for a specific application. This process follows the idea of convergence of evidence and generates a 'best-estimate' data set using fuzzy agreement. The authors apply the method to develop a new joint 1 km global land cover product (SYNMAP) with improved characteristics for land cover parameterization of the carbon cycle models. The overall advantage of the SYNMAP legend is that all classes are properly defined in terms of plant functional type mixtures, which can be remotely sensed and include the definitions of leaf type and longevity for each class with a tree component. See & Fritz (2006) have used fuzzy logic to incorporate expert knowledge for comparing the GLC2000 and the MODIS land cover product. To capture classification uncertainties of both data sources the authors mapped spatial disagreement by using a combination of fuzzy logic and expert knowledge. Fuzzy membership matrices were generated based on knowledge of classification experts to indicate the different levels of difficulty in classifying different land cover types and to map spatial disagreement. The areas of highest disagreement were validated using additional land cover information and a hybrid land cover map was generated by fusion of the GLC2000 and the MODIS land cover product.

For local scale applications, Alfieri *et al*. (2007) propose a reclassification method to better assign the various parameters needed for land surface model simulations in a case study in south-eastern Kansas, USA. Besides land use/land cover data sets normalized difference vegetation index (NDVI) measurements, elevation and slope are used to provide a more accurate data set.

Standardization and harmonization of land cover data

Instead of dealing with limited compatibility and comparability of existing land cover data sets and their thematic legends others have aimed at improving the flexibility and usability of land cover data to avoid a translation of different classification systems and to allow for comparability of different data sets (Neumann *et al*., 2007). International initiatives, such as the Group on Earth Observation (GEO), the Global Terrestrial Observing System (GTOS), and the Global Observation of Forest and Land Cover Dynamics (GOFC–GOLD), foster the establishment of international standards and protocols with respect to standardized development and harmonization of land cover data (Herold, 2006). These initiatives are largely driven by needs of international conventions (GCOS, 2004; GEOSS, 2005). An example of an already established strategic and methodological framework for harmonization of land cover classification systems is the UN Land Cover Classification System (LCCS) (Jansen & Gregorio, 2002; Gregorio, 2005). The LCCS applies a flexible but standardized set of classifiers and thresholds and is currently evolving as an internationally agreed standard for land cover characterization (Herold, 2006). Comparable land cover classification systems and understanding of semantic differences between data sets is essential for comparative accuracy analyses of different land cover data sets. Harmonization and validation of land cover data are therefore often parallel processes. Analogous efforts for land use data are still missing since international consensus on the definition and classification of land use has not yet been reached (Jansen, 2006).

Improved validation of land use and land cover data

Validation can help the user to select the most appropriate land cover or land use map for a specific purpose based on its correspondence with field observations or other data of relevance to the specific study for which the land use/cover data are prepared. Classical validations use field observations to judge the suitability of land cover data. Mostly these compare ground observations of land cover with the information in the land cover database and the percentage of correctly represented observations is measured. Field observations require an enormous investment, especially in case of the validation of global data sets; consequently most global data sets have not been validated (Herold, 2006). Therefore, Iwao *et al*. (2006) propose the use of Degree Confluence Project (DCP) information as a new method for validating land cover maps. The DCP is a volunteer-based project that aims to collect onsite information from all the degree confluences (intersections of integer level latitude and longitude gridlines) in the world. The paper of Iwao *et al*. (2006) assesses the reliability and effectiveness of DCP-derived data in validating land cover maps. DCP-derived validation information (at 749 confluences) was used to evaluate existing land cover maps for Eurasia (GLC2000, MOD12, UMD and GLCC). The agreements between the DCP-derived validation information and the land cover maps were between 50% and 58%. If the confluence and its surrounding 1 km square area contain more than one type of land cover, it is not always possible to determine whether the DCP-derived validation information faithfully represents the land cover of a 1 km square surrounding the confluence, leading to potential inconsistencies. The authors indicate that agreement may be improved by combining DCP-derived validation information with (visual) interpretation of high-resolution imagery (Iwao *et al*., 2006).

For some regions validation efforts of global land cover data are even showing lower levels of agreement. Frey & Smith (2007) use a large set of ground-based observations to see if global and regional land cover products are suitable for climate and ecosystem assessments. The authors used a collection of 2161 geolocated, irregularly spaced field observations of land cover throughout West Siberia to validate a number of currently available global land cover characteristics databases. The study indicated that overall agreement with ground observations of land cover is between 11% and 21% only. Permanent wetlands and waterbodies are underestimated in all databases. These results raise into question the efficacy of incorporating currently available land cover products into terrestrial ecosystem models in northern wetland environments.

In the absence of a consistent global database of field observations, Sterling & Ducharne (2008) validate their global land cover maps by comparing the areas of land cover with a wide range of estimates from the literature. The comparison, based on estimates of anthropogenic land cover show a wide variation in the percentage of the earth surface covered by anthropogenic land cover indicating the high degree of uncertainty in global maps. Based on these observations, the authors highlight the need for improvements in land cover mapping. Especially, the area and distribution of grazing land needs to be determined more accurately given its large influence on land surface processes. Similar attention is needed to map global tree plantations given the large importance and dynamics of timber plantations in North America and oil-palm plantations in Southeast Asia.

Implications for global change analysis and modelling

Land cover data and emission inventories

The role of land cover data in calculating emissions is especially important for the reporting requirements for GHG inventories to the Intergovernmental Panel on Climate Change (IPCC) commission. To make these country specific reporting as much comparable as possible, guidelines and best practice documents were developed (see: http://www.ipcc-nggip.iges.or.jp/public/gpglulucf/gpglulucf_files/Chp2/Chp2_Land_Areas.pdf for the full guidelines with respect to the estimation of land cover areas). These guidelines address many of the issues raised in this paper regarding proper use of the data with respect to harmonization of legends, scaling issues and other uncertainties. However, these guidelines actually contain little information on how to estimate land areas (changes) and leave it to the individual countries to select the land cover data used from a variety of sources (agricultural census surveys, forest inventories, and remote sensing data). Although this approach enables countries to make best use of existing data, inconsistencies between data sources will ultimately lead to inconsistent estimates of emissions that are not easily corrected.

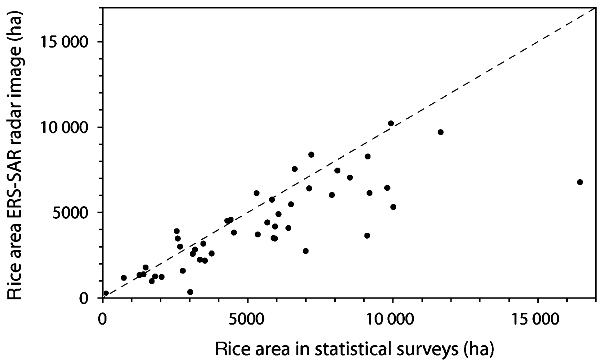

Fig. 2 Rice area per municipality based on interpretation of Radar images and as reported in statistical surveys (based on Verburg *et al.*, 2006).

The choice of data used in emission inventories may have large consequences for the resulting emission estimates. Verburg *et al.* (2006) made an assessment of the uncertainties involved in regional estimates of methane emissions from rice fields in one of the main rice growing regions of the Philippines. The rice area was determined by supervised classification of a radar image (ERS-SAR). A comparison between the rice areas derived from the ERS-SAR image and the rice area given in the statistical surveys at the level of municipalities shows large deviations between the two sources for a number of municipalities (Fig. 2). In general, the rice grown area as identified by ERS-SAR is smaller than the area reported in statistics (almost 30% for the total study area). Causes for this difference include inconsistency in statistical sources due to definitions and sampling methods and interpretation and classification problems of the ERS-SAR image. Many rice fields are relatively small and irrigation canals and dykes are abundant. The small size parcels may lead to an underestimation of the rice area in the ERS-SAR interpretation. It is clear that the regional emission of methane calculated with the rice area of the ERS-SAR is considerably lower than the emission calculated with the statistical data. In comparison to other uncertainties in emission estimates for this region, such as model accuracy and upscaling procedure, the uncertainty in land cover data is highest.

Earlier in this paper the influence of the spatial resolution of the land cover maps on the areas of, respectively, grassland and water bodies in the Dutch Fen Meadow landscape was discussed. Nol *et al.* (2008) assessed the impact of the spatial resolution of the land cover data for emissions of nitrous oxide from this region. Based on the estimated surface areas, agricultural N_2O emissions were estimated using different inventory techniques. All four databases of land cover overestimated the grassland area when compared to the field map. This caused a considerable overestimation of agricultural N_2O emissions, ranging from 9% for more detailed databases to 11% for the coarsest database. The effect of poor land cover representation was larger for an inventory method based on a process model than for inventory methods based on simple emission factors. Although the effect of errors in land cover representations were, in this study, small compared with the effect of uncertainties in emission factors, these effects are systematic (i.e., cause bias) and do not cancel out by spatial upscaling. Moreover, bias in land cover representations can be quantified or reduced by careful selection of the land cover database (Nol *et al.*, 2008).

Quaife *et al.* (2008) show how the use of different land cover maps influences calculated large-scale, bottom-up estimates of terrestrial carbon fluxes. The authors compare calculated fluxes based on globally available moderate resolution satellite-derived land cover maps with fluxes calculated using a reference high resolution (25 m) land cover map specific to Great Britain (the Land Cover Map 2000). The authors demonstrate that uncertainty is introduced into carbon flux calculations by (1) incorrect or uncertain assignment of land cover classes to plant functional types (PFTs), (2) information loss at coarser resolutions, and (3) difficulty in discriminating some vegetation types from satellite data. Differences in land use data account for differences between −15.8% and 8.8% in calculated Gross Primary Production.

Considerations with respect to integrated assessment modelling

Many models used in global climate studies [e.g., ORCHIDEE (Krinner *et al.*, 2005) and LPJmL (Bondeau *et al.*, 2007)] use PFTs as a basis for simulating vegetation dynamics. Sterling & Ducharne (2008) indicate that the reclassification process of land cover type to plant functional types and especially assigning a percentage of bare soil to the land cover types is an important step in land surface modelling. The percentage of bare soil drives the major land surface fluxes and properties, determining, among others, LAI and albedo. Wang *et al.* (2006) present a table to convert land cover classes to Plant Functional Types using available information to make the PFTs as consistent with the land cover description as possible but also mention that this translation could be another source of uncertainty given the wide range of PFTs possible in one land cover type.

Many recent integrated assessments of environmental change use a series of coupled models instead of one single assessment model. In such approaches often a specific model or module dealing with land change is found. This module converts the demands for commodities and services into change in land cover areas. These land cover changes are allocated within a spatial representation of the landscape being input to the emission and climate models (e.g., Rounsevell *et al.*, 2005; Sohl *et al.*, 2007; Verburg *et al.*, 2008). The difficulty in translating demands for commodities in land cover claims is complicated by the different data sources used in different disciplinary traditions. Demands for commodities are often determined by (multi-)sectoral economic models and are expressed in the units of the goods or service under consideration, such as agricultural production (Meijl *et al.*, 2006; Lotze-campen *et al.*, 2008). The conversion of change in the demand for commodities into land cover change is not always straightforward. In case of agricultural commodities, farming system characteristics such as multiple-cropping, intercropping and other management practices need to be accounted for.

Expansion of arable area is only one possible way of fulfilling an increasing demand for agricultural commodities. In many cases intensification by means of increasing inputs, efficiency or cropping intensity are a more likely means of fulfilling the demand (Neumann *et al.*, 2010). Similar considerations apply to forestry. Increasing wood demands do not necessarily lead to deforestation but in many cases to forest degradation or changed management practices which are difficult to detect using remote sensing (Lambin, 1999). Most land change models focusing on deforestation are only capable of addressing complete deforestation and ignore forest degradation.

Integrated assessment modelling and coupling of economic and geographic models means explicitly dealing with inconsistencies between data sources and the difference between land use and land cover. If not properly accounted for such differences in underlying data, integrated assessment modelling runs the risk of propagating errors due to inconsistent use of data across the different models or modules.

Conclusions and recommendations

Based on the review of literature, a number of recommendations are formulated to improve the practice of using land use and land cover data in climate change studies.

Explore and document data inconsistencies and uncertainties

Uncertainty assessment is an integral part of many climate change studies (Diniz *et al.*, 2009; Brown, 2010; Hastings *et al.*, 2010). In most analyses, the uncertainty and error in the land use/cover data used as input to the assessments is not addressed. Because part of the errors and uncertainties in the data may be structured towards under- or overestimating a specific land cover type due to observation or aggregation bias this may lead to a one-directional error in the climate change assessments. It is especially this type of errors and uncertainties that should be given explicit attention in climate change assessments. Scientists should use their awareness of these inconsistencies and uncertainties in land use/cover data to study the implications of these uncertainties for the results of their studies and properly communicate those implications.

Differences between data sets representing land use or land cover are often seen as inconsistencies or errors. However, in many cases such differences are the result of different representations of the data in terms of classification and temporal and spatial scale. An analysis of the underlying reasons for inconsistencies and errors will help to make more appropriate use of the data in climate change assessments. Inconsistencies between data may, in fact, indicate that complementary information is available which, if properly used, may benefit the overall assessment.

Many global change assessments regard land use/cover data as just one of the many input variables for which easily available data are used without explicit consideration of alternative data sources. This often results in a mismatch between the characteristics of the data and the use of the data in further calculations. Models calibrated with point or field level data will, most likely, not perform optimally with coarse scale data (Heuvelink, 1998). Therefore, a careful selection of the land use/cover data and a documentation of this selection for a specific application is mandatory.

The risks involved in allowing countries to make best use of available data for their emission inventories could be reduced by requiring a better documentation of the potential bias involved in selecting a specific data source for the inventory. Enhanced guidelines could assist in identifying potential bias in inventories to make estimates more comparable between countries.

Many climate change assessments focus on improving model performance and representation, e.g. by moving from Tier 2 to Tier 3 methods in emission inventories. While the Tier 2 approach applies (country-specific) emission factors and activity data to link land use/cover areas to emissions the Tier 3 approach uses higher-order methods including models with a closer link to soil and vegetation processes (IPCC, 1997). Given the uncertainties involved in land use/cover data and the key role of these data in climate change assessments a better balance of research efforts between improvement of input data and model representation should be achieved.

Provide full documentation of land use/cover classifications used

For a proper selection of data on land use/cover and their intended application it is essential to have a clear and extended documentation of the land use or land cover classes. Especially in translating the land cover information to PFTs, common in land surface modelling, this information is essential. Given the wide variety of definitions for the same land cover type, it is not wise to use data without a proper documentation of the categories. In spite of many efforts to create uniform descriptions of classes a wide variety of definitions exists. This variety also reflects the diverse intended applications of the collected data and, in many cases, the categories are defined to best suit the application the data are collected for.

A proper documentation of land use/cover data can best be made in a uniform system. The UN LCCS (http://www.glcn.org/) provides guidelines and software for this and is now accepted by many users as a proper means for documenting and exchanging documentation on land cover classifications.

Select data based on the needs for a specific application

Land cover and land use information is often confused and not explicitly distinguished. Most information currently used in climate change assessments is based on remote sensing and primarily contains land cover information. However, for many climate change assessments management aspects of the land cover are essential. Land use may be very different on the same land cover. Grasslands may be natural, intensively managed or extensively used with highly different implications for e.g. carbon sequestration. Recent efforts to create data sets of land management are an important step forward including global and regional inventories of irrigated areas (Siebert *et al*., 2005; Wriedt *et al*., 2009; Portmann *et al*., 2010), global data on crop yields and fertilizer use (Monfreda *et al*., 2008; Potter *et al*., 2010) and methods to derive parcel size and structure from remote sensing images (Kuemmerle *et al*., 2009).

It is common belief that land use/cover data with high spatial resolution are the 'best' data. This is not always the case. Thematic resolution or fit of the land cover classes with the model description of the processes under investigation may be a more important criterion for data quality for a specific application. Selection of land use/cover data should therefore primarily be based on the match of the data with the model or assessment approach.

Several studies have indicated the poor fit of global data on land use/cover to specific regions. The classification algorithms and categories considered in global studies are not targeted towards the specific regional conditions. Therefore, the use of global land use/cover data for regional climate change assessments should be avoided. Regional specific data often better represent the regional situation and are therefore better suited for this level of analysis.

Select scaling and aggregation methods based on data and landscape characteristics

Most studies require edits to the data prior to analysis. Such edits include thematic aggregations, geographic projections or spatial aggregations. Such changes to the data may influence the characteristics of the data and the information contained in the data. Proper documentation of the data processing is essential but often lacking.

Spatial aggregations may be done by different methods with different effects on the data. Some aggregation methods lead to an overall loss of information while other aggregation methods can structurally change the representation of specific classes in the data depending on the prevalence and structure of the classes in the landscape. Different landscapes require different aggregation techniques. Methods are available to correct unintended bias due to aggregation. Careful documentation of the aggregation methods and its effects on the data is essential and researchers may benefit from methods for knowledge exchange developed in geographic information science (Fonseca *et al*., 2000; Kuhn, 2001).

Combine different data sources when complementary information is available

Several attempts of integrating different, complementary data sources have provided high quality results as well as assessments of the uncertainty of the different data sources. By combining the strengths of different data more robust and reliable data can be constructed. In many countries and regions, multiple data sources that are not yet explored as common sources of land cover data are available and could be used to improve remote sensing-derived land cover estimates or census-based land use inventories. Examples of such data are for the European Union (EU) the Farm Accountancy Data Network (FADN) which was established to monitor the effects of the EU common agricultural policy (CAP). Derived from national surveys, the FADN contain harmonized micro-economic data, i.e. the bookkeeping principles are the same in all countries. The methodology applied aims to provide representative data along three dimensions: region, economic size and type of farming. This type of data can potentially improve insights in the land use practices that land cover data cannot provide. Because of privacy issues the location of the sample points is not disclosed, which largely restricts the applicability of these data. Another

© 2010 Blackwell Publishing Ltd,

useful data set at the European level are for obtaining additional land use/cover data is the LUCAS '*Land Use/Cover Area frame statistical Survey*' pilot project which was launched by Eurostat (European statistical bureau) in 2001. In contrast to mapping approaches this project uses area frame sampling to collect data. Based on the visual observation of sample geo-referenced points, area estimates are computed and used as a valid generalization without studying the entire area under investigation. The approach has also the important advantage of not involving/disturbing the land owners and the farmers. At European level, currently around 250 000 points have been sampled providing an excellent coverage of ground observations. A third example of alternative existing data sets that may provide useful information on land use are parcel registration systems and databases describing cropping practices and agricultural land use at the field scale. These data are collected to compute nutrient balances as requested by law or to apply for subsidies. Although such data are not always publicly available several authors have successfully applied such databases for research purposes (Schmit *et al*., 2006; Lucas *et al*., 2007). A change in the data policies of the responsible institutions for collecting these data adapting to the needs of the scientists dealing with global change research would benefit the quality of global change assessments.

Focus the collection of new data to the most uncertain land use/cover types and regions

In addition to making most efficient use of existing data new data may be collected to improve the representation of land use/cover in climate assessments. For newly collected data, documentation of methods and uncertainty (including validation) should receive sufficient attention. Documentation of the classification of data should follow internationally agreed guidelines. For global scale applications, new land cover data should focus on the main uncertainties in current data. Most important is the representation of grassland areas. Global estimates are deviating largely between current sources. Also the distinction between forest and forest plantations deserves explicit attention. Well organized campaigns aimed at a combination of remote sensing with ground observations and inventory data may improve the representation of grasslands and forest plantations in global data sets.

Explicitly represent mosaics and/or land use systems in land classifications

Most land cover data attempt to classify each pixel by the dominant land cover. For high spatial resolution data, this may provide a good representation of reality. For medium to coarse resolution data, this representation may lead to problems since many landscapes contain a mosaic of land cover types. Although the classification into one dominant land cover type is convenient in many models the characteristic conditions of these mosaic landscapes are disregarded. In terms of processes relevant to climate change mosaic landscapes can have specific characteristics that may not be represented correctly by the dominant land cover types. A recent study indicated that agricultural areas and agro-forestry systems worldwide contain large amounts of trees that may not be ignored given their important role in climate change processes (Zomer *et al*., 2009). Therefore, the explicit representation of these mosaics as separate classes in land cover data should be considered.

Besides land cover information land use information is often essential for climate change assessments. The same land use may, depending on the local context, represent different levels of interaction of humans with the environment. Instead of representing the landscape in terms of land cover or land use a representation in terms of land use systems that integrate the land cover and management aspects may be considered. Ellis & Ramankutty (2008) have attempted to map land use systems, in their terminology called anthropogenic biomes (or anthromes). These biomes share a common level of interactions between humans and the environment, examples include 'dense settlements,' 'pastoral villages' and 'populated rainfed croplands,' Each of these biomes consists of a heterogeneous landscape mosaic combining a variety of different land covers. Through some of this heterogeneity might be explained by the relatively coarse resolution of the analysis, a more fundamental explanation is that human–environment interactions lead to different mosaics due to natural variation in terrain, human enhancement of the natural heterogeneity by concentrating activities at the most productive locations and heterogeneity caused directly by the specific activity types of the considered biome. The use of an anthropogenic biome map instead of a conventional biome or land cover map has major advantages given the better representation of the human-environment interactions and its intensity that cannot directly be observed from land cover data. The collection of data that assist the representation of land use systems may be further elaborated by using available information on livestock densities (van de Steeg *et al*., 2010), production efficiencies (Verburg *et al*., 2000), crop types (Monfreda *et al*., 2008), cropping periods (Sacks *et al*., 2010), remoteness and land management (e.g., fertilizer application).

Acknowledgements

The authors would like to thank Hsin Huang for useful comments on an early draft of this paper.

References

Aerts J (2005) Adaptation for river basins: connecting policy goals to the water resources system. *Water Science and Technology*, **51**, 121–131.

Alfieri JG, Niyogi D, Lemone MA, Chen F, Fall S (2007) A simple reclassification method for correcting uncertainity in land use/land cover data sets used with land surface models *Pure and Applied Geophysics*, **164**, 1789–1809.

Asner GP, Knapp DE, Balaji A, Paez-Acosta G (2009) Automated mapping of tropical deforestation and forest degradation: CLASlite. *Journal of Applied Remote Sensing*, **3**, 033543, doi: 10.1117/1.3223675.

Aspinall R (2004) Modelling land use change with generalized linear models – a multi-model analysis of change between 1860 and 2000 in Gallatin Valley, Montana. *Journal of Environmental Management*, **72**, 91–103.

Bach M, Breuer L, Frede HG, Huisman JA, Otte A, Waldhardt R (2006) Accuracy and congruency of three different digital land-use maps. *Landscape and Urban Planning*, **78**, 289–299.

Berberoglu S, Akin A (2009) Assessing different remote sensing techniques to detect land use/cover changes in the eastern Mediterranean. *International Journal of Applied Earth Observation and Geoinformation*, **11**, 46–53.

Bondeau A, Smith PC, Zaehle S *et al.* (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology*, **13**, 679–706.

Bossel H (1999) *Indicators for Sustainable Development: Theory, Method, Applications*. International Institute for Sustainable Development, Winnipeg, Manitoba.

Brown JD (2010) Prospects for the open treatment of uncertainty in environmental research. *Progress in Physical Geography*, **34**, 75–100.

Cardille J, Foley JA (2003) Agricultural land-use change in Brazilian Amazonia between 1980 and 1995: evidence from integrated satellite and census data. *Remote Sensing of Environment*, **87**, 551–562.

Carvalho G, Moutinho P, Nepstad D, Mattos L, Santilli M (2004) An Amazon perspective on the forest–climate connection: opportunity for climate mitigation, conservation and development? *Environment, Development and Sustainability*, **6**, 163–174.

Castilla G, Hay GJ (2007) Uncertainties in land use data. *Hydrology and Earth System Sciences*, **11**, 1857–1868.

Chapin FS, Mcguire AD, Randerson J *et al.* (2000) Arctic and boreal ecosystems of western north America as components of the climate system. *Global Change Biology*, **6**, 211–223.

Comber A, Fisher P, Wadsworth R (2005) Comparing the consistency of expert land cover knowledge. *International Journal of Applied Earth Observation and Geoinformation*, **7**, 189–201.

Comber AJ (2008) The separation of land cover from land use using data primitives. *Journal of Land Use Science*, **3**, 215–229.

DeFries R (2008) Terrestrial vegetation in the coupled human-earth system: contributions of remote sensing. *Annual Review of Environment and Resources*, **33**, 369–390.

Dendoncker N, Schmit C, Rounsevell M (2008) Exploring spatial data uncertainties in land-use change scenarios. *International Journal of Geographical Information Science*, **22**, 1013–1030.

Dendoncker N, Van Wesemael B, Rounsevell MDA, Roelandt C, Lettens S (2004) Belgium's CO_2 mitigation potential under improved cropland management. *Agriculture, Ecosystems and Environment*, **103**, 101–116.

Denier van der Gon HAC, Kropff MJ, Van Breemen N *et al.* (2002) Optimizing grain yields reduces CH_4 emissions from rice paddy fields. *Proceedings of the National Academy of Sciences*, **99**, 12021–12024.

Diniz JAF, Bini LM, Rangel TF, Loyola RD, Hof C, Nogues-Bravo D, Araujo MB (2009) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography*, **32**, 897–906.

Eastman JL, Coughenour MB, Pielke RA (2001) The regional effects of CO_2 and landscape change using a coupled plant and meteorological model. *Global Change Biology*, **7**, 797–815.

EEA (2008) *Impacts of Europe's changing climate – 2008 indicator-based assessment*. Report 4/2008. European Environmental Agency, Copenhagen.

Ellis EC, Neerchal N, Peng K *et al.* (2009) Estimating long-term changes in China's village landscapes. *Ecosystems*, **12**, 279–297.

Ellis EC, Ramankutty N (2008) Putting people in the map: anthropogenic biomes of the world. *Frontiers in Ecology and the Environment*, **6**, 439–447.

Erb K-H, Gaube V, Krausmann F, Plutzar C, Bondeau A, Haberl H (2007) A comprehensive global 5 min resolution land-use data set for the year 2000 consistent with national census data. *Journal of Land Use Science*, **2**, 191–224.

Etienne M, Le Page C, Cohen M (2003) A step-by-step approach to building land management scenarios based on multiple viewpoints on multi-agent system simulations. *Journal of Artificial Societies and Social Simulation*, **6**, Available at http://jasss.soc.surrey.ac.uk/6/2/2.html.

Eva H, Carboni S, Achard F, Stach N, Durieux L, Faure JF, Mollicone D (2010) Monitoring forest areas from continental to territorial levels using a sample of medium spatial resolution satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, **65**, 191–197.

Fassnacht KS, Cohen WB, Spies TA (2006) Key issues in making and using satellite-based maps in ecology: a primer. *Forest Ecology and Management*, **222**, 167–181.

Foley JA, Costa MH, Delire C, Ramankutty N, Snyder PK (2003) Green Surprise? How terrestrial ecosystems could affect earth's climate. *Frontiers in Ecology and Environment*, **1**, 38–44.

Fonseca FT, Egenhofer MJ, Davis CA, Borges KAV (2000) Ontologies and knowledge sharing in urban GIS. Computers. *Environment and Urban Systems*, **24**, 251–272.

Foody GM (2002) Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, **80**, 185–201.

Freitas CD, Soler LD, Anna SJSS, Dutra LV, dos Santos JR, Mura JC, Correia AH (2008) Land use and land cover mapping in the Brazilian Amazon using polarimetric airborne P-band SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, **46**, 2956–2970.

Fresco LO (1994) Imaginable futures, a contribution to thinking about land use planning (eds Fresco LO, Stroosnijder L, Bouma J, van Keulen H), pp. 1–8. John Wiley and Sons, Chichester.

Frey KE, Smith LC (2007) How well do we know northern land cover? Comparison of four global vegetation and wetland products with a new ground-truth database for West Siberia. *Global Biogeochemical Cycles*, **21**, GB1016, doi: 10.1029/2006GB002706.

Fritz S, See L (2005) Comparison of land cover maps using fuzzy agreement. *International Journal of Geographical Information Science*, **19**, 787–807.

GCOS (2004) *Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC*. WMO, Geneva.

GEOSS (2005) The Global Earth Observation System of Systems (GEOSS) 10-Year Implementation Plan and Reference Document, 11pp. Available at http://earth-observations.org.

Goldewijk KK, Van Drecht G, Bouwman AF (2007) Mapping contemporary global cropland and grassland distributions on a 5 times 5 minute resolution. *Journal of Land Use Science*, **2**, 167–190.

Grainger A (2008) Difficulties in tracking the long-term global trend in tropical forest area. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 818–823.

Gregorio AD (2005) *Land Cover Classification System Classification Concepts and User Manual Software Version (2)*. FAO, Rome.

Hansen MC, Roy DP, Lindquist E, Adusei B, Justice CO, Altstatt A (2008) A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin. *Remote Sensing of Environment*, **112**, 2495–2513.

Hansen MC, Stehman SV, Potapov PV, Arunarwati B, Stolle F, Pittman K (2009) Quantifying changes in the rates of forest clearing in Indonesia from 1990 to 2005 using remotely sensed data sets. *Environmental Research Letters*, **4**, 034001, doi: 10.1088/1748-9326/4/3/034001.

Hastings AF, Wattenbach M, Eugster W, Li CS, Buchmann N, Smith P (2010) Uncertainty propagation in soil greenhouse gas emission models: an experiment using the DNDC model and at the Oensingen cropland site. *Agriculture Ecosystems and Environment*, **136**, 97–110.

Herold M, Mayaux P, Woodcock CE, Baccini A, Schmullius C (2008) Some challenges in global land cover mapping: an assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, **112**, 2538–2556.

Herold MM (2006) A joint initiative for harmonization and validation of land cover datasets. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 1719–1727.

Heuvelink GBM (1998) Uncertainty analysis in environmental modelling under a change of spatial scale. *Nutrient Cycling in Agroecosystems*, **50**, 255–264.

Hines ME, Duddleston KN, Rooney-Varga JN, Fields D, Chanton JP (2008) Uncoupling of acetate degradation from methane formation in Alaskan wetlands: connections to vegetation distribution. *Global Biogeochemical Cycles*, **22**, GB2017, doi: 10.1029/2006GB002903.

Hurtt GC, Rosentrater L, Frolking S, Moore B III (2001) Linking remote-sensing estimates of land cover and census statistics on land use to produce maps of land use of the conterminous united states. *Global Biogeochemical Cycles*, **15**, 673–685.

IPCC (1997) *Intergovernmental Panel on Climate Change Revised 1996 Guidelines for National Greenhouse Gas Inventories*. Workbook, OECD, Paris.

Iwao K, Nishida K, Kinoshita T, Yamagata Y (2006) Validating land cover maps with degree confluence project information. *Geophysical Research Letters*, **33**, L23404, doi: 10.1029/2006GL027768.

Jansen LJM (2006) Harmonization of land use class sets to facilitate compatibility and comparability of data across space and time. *Journal of Land Use Science*, **1**, 127–156.

Jansen LJM, Gregorio AD (2002) Parametric land cover and land-use classifications as tools for environmental change detection. *Agriculture, Ecosystems and Environment*, **91**, 89–100.

Ju JC, Roy DP (2008) The availability of cloud-free Landsat ETM plus data over the conterminous united states and globally. *Remote Sensing of Environment*, **112**, 1196–1211.

Jung M, Henkel K, Herold M, Churkina G (2006) Exploiting synergies of global land cover products for carbon cycle modeling. *Remote Sensing of Environment*, **101**, 534–553.

Kayranli B, Scholz M, Mustafa A, Hedmark A (2010) Carbon storage and fluxes within freshwater wetlands: a critical review. *Wetlands*, **30**, 111–124.

Klein Goldewijk K (2001) Estimating global land use change over the past 300 years: the HYDE database. *Global Biogeochemical Cycles*, **15**, 417–434.

Krinner G, Viovy N, de Noblet-Ducoudre N *et al*. (2005) A dynamic global vegetation model for studies of the coupled atmosphere–biosphere system. *Global Biogeochemical Cycles*, **19**, GB1015, doi: 10.1029/2003GB002199.

Kuemmerle T, Hostert P, St-Louis V, Radeloff VC (2009) Using image texture to map farmland field size: a case study in Eastern Europe. *Journal of Land Use Science*, **4**, 85–107.

Kuhn W (2001) Ontologies in support of activities in geographical space. *International Journal of Geographical Information Science*, **15**, 613–631.

Lambin EF (1999) Monitoring forest degradation in tropical regions by remote sensing: some methodological issues. *Global Ecology and Biogeography*, **8**, 191–198.

Lotze-campen H, Mueller C, Bondeau A, Rost S, Popp A, Lucht W (2008) Global food demand, productivity growth, and the scarcity of land and water resources: a spatially explicit mathematical programming approach. *Agricultural Economics*, **39**, 325–338.

Lucas R, Rowlands A, Brown A, Keyworth S, Bunting P (2007) Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, **62**, 165–185.

Lund HG (2004) *Considerations for developing U.S. standard definitions of forest and rangeland*. Report prepared for Meridian Institute, Contract No. 0045-A. Project No. 9147.8, Forest Information Services, Rev. Gainesville, VA,108 pp. Available at http://fhm-server.lv-hrc.nevada.edu//fia/ab/issues/Final_report_F-R_7May05.doc (accessed 3 April 2009)

Lund HG (2007) Accounting for the world's rangelands. *Rangelands*, **29**, 3–10.

MacDonald D, Crabtree JR, Wiesinger G *et al*. (2000) Agricultural abandonment in mountain areas of Europe: environmental consequences and policy response. *Journal of Environmental Management*, **59**, 47–69.

Macleod R, Congalton R (1998) A quantitative comparison of change detection algorithms for monitoring eelgrass from remote sensing data. *Photogrammetric Engineering and Remote Sensing*, **64**, 207–216.

Marceau DJ, Hay GJ (1999) Remote sensing contributions to the scale issue. *Canadian Journal of Remote Sensing*, **25**, 357–366.

McConnell W, Moran EF (2001) *Meeting in the Middle: The Challenge of Meso-Level Integration*. LUCC Focus 1 Office, Anthropological Center for Training and Research on Global Environmental Change, Indiana University.

Meijl HV, van Rheenen T, Tabeau A, Eickhout B (2006) The impact of different policy environments on agricultural land use in Europe. *Agriculture, Ecosystems and Environment*, **114**, 21–38.

Monfreda C, Ramankutty N, Foley JA (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, **22**, GB1022, doi: 10.1029/2007GB002947.

Moody A, Woodcock CE (1994) Scale-dependent errors in the estimation of land-cover proportions: implications for global land-cover data sets. *Photogrammetric Engineering and Remote Sensing*, **60**, 585–596.

Mottet A, Ladet S, Coque N, Gibon A (2006) Agricultural land-use change and its drivers in mountain landscapes: a case study in the Pyrenees. *Agriculture, Ecosystems and Environment*, **114**, 296–310.

Nepstad D, Carvalho G, Barros AC *et al*. (2001) Road paving, fire regime feedbacks, and the future of Amazon forests. *Forest Ecology and Management*, **154**, 395–407.

Neumann K, Herold M, Hartley A, Schmullius C (2007) Comparative assessment of CORINE2000 and GLC2000: spatial analysis of land cover data for Europe. *International Journal of Applied Earth Observation and Geoinformation*, **9**, 425–437.

Neumann K, Verburg PH, Stehfest E, Mueller C (2010) The yield gap of global grain production: a spatial analysis. *Agricultural Systems*, **103**, 316–326.

Nol L, Verburg PH, Heuvelink GBM, Molenaar K (2008) Effect of land cover data on nitrous oxide inventory in fen meadows. *Journal of Environmental Quality*, **37**, 1209–1219.

Ozdogan M, Woodcock CE (2006) Resolution dependent errors in remote sensing of cultivated areas. *Remote Sensing of Environment*, **103**, 203–217.

Park MH, Stenstrom MK (2008) Classifying environmentally significant urban land uses with satellite imagery. *Journal of Environmental Management*, **86**, 181–192.

Pelorosso R, Leone A, Boccia L (2009) Land cover and land use change in the Italian central Apennines: a comparison of assessment methods. *Applied Geography*, **29**, 35–48.

Petit C, Scudder T, Lambin E (2001) Quantifying processes of land-cover change from remote sensing observations: resettlement and rapid land-cover changes in south-eastern Zambia. *International Journal of Remote Sensing*, **22**, 3435–3456.

Petit CC, Lambin EF (2001) Integration of multi-source remote sensing data for land cover change detection. *International Journal of Geographical Information Science*, **15**, 785–803.

Pielke RA, Avissar R, Raupach M, Dolman AJ, Zeng X, Denning AS (1998) Interactions between the atmosphere and terrestrial ecosystems: influence on weather and climate. *Global Change Biology*, **4**, 461–475.

Polsky C, Easterling WE III (2001) Ricardian climate sensitivities: accounting for adaptation across scales. *Agriculture, Ecosystems and Environment*, **85**, 133–144.

Pongratz J, Reick C, Raddatz T, Claussen M (2008) A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, **22**, GB3018, doi: 10.1029/2007GB003153.

Portmann FT, Siebert S, Doll P (2010) MIRCA2000 – global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, **24**, GB1011, doi: 10.1029/2008GB003435.

Potter P, Ramankutty N, Bennett EM, Donner SD (2010) Characterizing the spatial patterns of global fertilizer application and manure production. *Earth Interactions*, **14**, 1–22.

Quaife T, Quegan S, Disney M, Lewis P, Lomas M, Woodward FI (2008) Impact of land cover uncertainties on estimates of biospheric carbon FLUXES. *Global Biogeochemical Cycles*, **22**, GB4016, doi: 10.1029/2007GB003097.

Ramankutty N, Evan AT, Monfreda C, Foley JA (2008) Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, **22**, GB1003, doi: 10.1029/2007GB002952.

Ramankutty N, Foley JA (1998) Characterizing patterns of global land use: an analysis of global croplands data. *Global Biogeochemical Cycles*, **12**, 667–685.

Ramankutty N, Foley JA (1999) Estimating historical changes in global land cover: croplands from 1700 to 1992. *Global Biogeochemical Cycles*, **13**, 997–1027.

Rambaldi G, Muchemi J, Crawhall N, Monaci L (2007) Through the eyes of hunter-gatherers: participatory 3D modelling among Ogiek indigenous peoples in Kenya. *Information Development*, **23**, 113–128.

Richards JA, Jia XP (2006) *Remote Sensing Digital Image Analysis; An Introduction*. Springer, Berlin.

Ridd MK, Liu J (1998) A comparison of four algorithms for change detection in an urban environment. *Remote Sensing of Environment*, **63**, 95–100.

Robinson VB (2003) A perspective on the fundamentals of fuzzy sets and their use in geographic information systems. *Transactions in GIS*, **7**, 3–30.

Rounsevell MDA, Ewert F, Reginster I, Leemans R, Carter TR (2005) Future scenarios of European agricultural land use: II. Projecting changes in cropland and grassland. *Agriculture, Ecosystems and Environment*, **107**, 117–135.

Ruddiman WF, Ellis EC (2009) Effect of per-capita land use changes on Holocene forest clearance and CO_2 emissions. *Quaternary Science Reviews*, **28**, 3011–3015.

Sabor AA, Radeloff VC, McRoberts RE, Clayton M, Stewart SI (2007) Adding uncertainty to forest inventory plot locations: effects on analyses using geospatial data. *Canadian Journal of Forest Research-Revue Canadienne de Recherche Forestiere*, **37**, 2313–2325.

Sacks WJ, Deryng D, Foley JA, Ramankutty N (2010) Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, **19**, 607–620, Available at http://blackwell-synergy.com/doi/abs/10.1111/j.1466-8238.2010.00551.x

Schmit C, Rounsevell MDA, La Jeunesse I (2006) The limitations of spatial land use data in environmental analysis. *Environmental Science and Policy*, **9**, 174q–188.

Schneider A, Friedl MA, Potere D (2009) A new map of global urban extent from MODIS satellite data. *Environmental Research Letters*, **4**, 044003, doi: 10.1088/1748-9326/4/4/044003.

Schulp CJE, Nabuurs G-J, Verburg PH (2008a) Future carbon sequestration in Europe – effects of land use change. *Agriculture, Ecosystems and Environment*, **127**, 251–264.

Schulp CJE, Nabuurs G-J, Verburg PH, de Waal RW (2008b) Effect of tree species on carbon stocks in forest floor and mineral soil and implications for soil carbon inventories. *Forest Ecology and Management*, **256**, 482–490.

See LM, Fritz S (2006) A method to compare and improve land cover datasets: application to the GLC-2000 and MODIS land cover products. *IEEE Transaction on Geoscience and Remote Sensing*, **44**, 1740–1746.

Siebert S, Doll P, Hoogeveen J, Faures JM, Frenken K, Feick S (2005) Development and validation of the global map of irrigation areas. *Hydrology and Earth System Sciences*, **9**, 535–547.

Sirén AH, Brondizio ES (2009) Detecting subtle land use change in tropical forests. *Applied Geography*, **29**, 201–211.

Sohl TL, Sayler KL, Drummond MA, Loveland TR (2007) The FORE-SCE model: a practical approach for projecting land cover change using scenario-based modeling. *Journal of Land Use Science*, **2**, 103–126.

Sterling S, Ducharne A (2008) Comprehensive data set of global land cover change for land surface model applications. *Global Biogeochemical Cycles*, **22**, GB3017, doi: 10.1029/2007GB002959.

Tasser E, Walde J, Tappeiner U, Teutsch A, Noggler W (2007) Land-use changes and natural reforestation in the eastern central Alps. *Agriculture, Ecosystems and Environment*, **118**, 115–129.

van de Steeg J, Verburg PH, Baltenweck I, Staal S (2010) Characterization of the spatial distribution of farming systems in the Kenyan highlands. *Applied Geography*, **30**, 239–253.

Verburg P (2006) Simulating feedbacks in land use and land cover change models. *Landscape Ecology*, **21**, 1171–1183.

Verburg P, Eickhout B, van Meijl H (2008) A multi-scale, multi-model approach for analyzing the future dynamics of European land use. *The Annals of Regional Science*, **42**, 57–77.

Verburg PH, Chen YQ, Veldkamp A (2000) Spatial explorations of land-use change and grain production in China. *Agriculture, Ecosystems and Environment*, **82**, 333–354.

Verburg PH, Denier van der Gon HAC (2001) Spatial and temporal dynamics of methane emissions from agricultural sources in China. *Global Change Biology*, **7**, 31–47.

Verburg PH, Van Bodegom PM, Denier van der Gon HAC, Bergsma AR, Van Breemen N (2006) Upscaling regional emissions of greenhouse gases from rice cultivation: methods and sources of uncertainty. *Plant Ecology*, **182**, 90–106.

Verburg PH, van de Steeg J, Veldkamp A, Willemen L (2009) From land cover change to land function dynamics: a major challenge to improve land characterization. *Journal of Environmental Management*, **90**, 1327–1335.

Wadsworth R, Balzter H, Gerard F, George C, Comber A, Fisher P (2008) An environmental assessment of land cover and land use change in Central Siberia using quantified conceptual overlaps to reconcile inconsistent data sets. *Journal of Land Use Science*, **3**, 251–264.

Wang A, Price DT, Arora V (2006) Estimating changes in global vegetation cover (1850–2100) for use in climate models. *Global Biogeochemical Cycles*, **20**, GB3028, doi: 10.1029/2005GB002514.

Woodcock CE, Gopal S (2000) Fuzzy set theory and thematic maps: accuracy assessment and area estimation. *International Journal of Geographical Information Science*, **14**, 153–172.

Wriedt G, van der Velde M, Aloe A, Bouraoui F (2009) A European irrigation map for spatially distributed agricultural modeling. *Agricultural Water Management*, **96**, 771–789.

Zomer RJ, Trabucco A, Coe R, Place F (2009) *Trees on farm: analysis of global extent and geographical patterns of agroforestry.* ICRAF Working Paper No. 89. World Agroforestry Centre, Nairobi, Kenya.