Global Change Biology (2000) **6,** 247–254

Technical advance

A new global 1-km dataset of percentage tree cover derived from remote sensing

R. S. DEFRIES,* M. C. HANSEN,* J. R. G. TOWNSHEND,† A. C. JANETOS‡ and T. R. LOVELAND§

**Department of Geography, University of Maryland, College Park, MD 20742, USA, †Department of Geography, University of Maryland, College Park, MD 20742, USA, and University of Maryland Institute for Advanced Computer Studies, College Park, MD 20742, USA, ‡World Resources Institute, 10 G St. NE, Suite 800, Washington, DC 20002, USA, §United States Geological Survey, EROS Data Center, Sioux Falls, SD 57198, USA*

Abstract

Accurate assessment of the spatial extent of forest cover is a crucial requirement for quantifying the sources and sinks of carbon from the terrestrial biosphere. In the more immediate context of the United Nations Framework Convention on Climate Change, implementation of the Kyoto Protocol calls for estimates of carbon stocks for a base-line year as well as for subsequent years. Data sources from country level statistics and other ground-based information are based on varying definitions of 'forest' and are consequently problematic for obtaining spatially and temporally consistent carbon stock estimates. By combining two datasets previously derived from the Advanced Very High Resolution Radiometer (AVHRR) at 1 km spatial resolution, we have generated a prototype global map depicting percentage tree cover and associated proportions of trees with different leaf longevity (evergreen and deciduous) and leaf type (broadleaf and needleleaf). The product is intended for use in terrestrial carbon cycle models, in conjunction with other spatial datasets such as climate and soil type, to obtain more consistent and reliable estimates of carbon stocks. The percentage tree cover dataset is available through the Global Land Cover Facility at the University of Maryland at http://glcf.umiacs.umd.edu.

Keywords: carbon stocks, forest cover, global land cover, Kyoto Protocol, land use, remote sensing

Received 11 April 1999; resubmitted and accepted 10 June 1999

Introduction

Uncertainties in quantifying carbon fluxes due to changes in the Earth's vegetation are at the crux of the current inabilities to balance the carbon cycle (Schimel 1995). Vegetation can either be a carbon source in the case of decay or burning of plant material, or a sink in the case of carbon sequestration from regrowth, fertilization from enhanced atmospheric carbon dioxide concentrations, or nutrient deposition. The past, current and future roles of each of these factors in the carbon cycle remain a subject of ongoing research.

Correspondence: Fax: + 1–301–314–9299,
E-mail: rd63@umail.umd.edu

The role of land cover in the carbon cycle has recently been included in the implementation of the United Nations Framework Convention on Climate Change, which aims to stabilize greenhouse gas emissions at levels that prevent dangerous interference with the climate system. In addition to taking steps towards reducing fossil fuel emissions, Article 3.3 of the Kyoto Protocol agreed by signatory nations in 1997 states that 'The net changes in greenhouse gas emission by sources and removals by sinks resulting from direct human-induced land-use change and forestry activities, limited to afforestation, reforestation and deforestation since 1990, measured as verifiable changes in carbon stocks in each commitment period, shall be used to meet the

© 2000 Blackwell Science Ltd.

commitments under this Article of each Party included in Annex I'. Appropriate implementation of this article in principle requires an accounting system for the full carbon budget of forested ecosystems, including carbon contained in soils. The inclusion of other types of land use changes in addition to forestry activities is a matter of intense policy debate, but would also require, in principle, a complete accounting of the terrestrial carbon budget (IGBP 1998). Even in the more limited context of carbon stocks contained in forests within Annex I countries (primarily developed countries in temperate latitudes), there are currently a number of fundamental but unanswered questions that need to be addressed for successful implementation of this article of the Kyoto Protocol:

- What is the spatial extent of forest cover for any particular baseline year using a globally consistent definition of 'forest'?
- What carbon stocks are contained in these forests?
- How can changes in carbon stocks of these forests be monitored over time?

The dataset described in this note is intended to help address the first of these fundamental questions by using remotely sensed data to estimate the spatial extent of forest cover. Estimates of global forest cover have traditionally been obtained from country-level information based on ground surveys, combinations of maps, and in some cases remote sensing (FAO 1997; WCMC 1998). There are large discrepancies in ground-based estimates of land cover information arising from differences in forest definition, processing methodology, outdated sources of information, and confusion between potential and existing vegetation (DeFries & Townshend 1994a; Mayaux *et al*. 1998). Such problems make it difficult to obtain spatially and temporally consistent estimates of carbon stocks contained in the Earth's vegetation. Satellite data offer the prospect for internally consistent and repeatable assessment of land cover. Based on earlier work at coarser spatial resolutions (Tucker *et al*. 1985; Townshend *et al*. 1987; DeFries & Townshend 1994b; Stone *et al*. 1994), several efforts are underway to exploit satellite data for characterizing land cover (DeFries *et al*. 1998; Malingreau & *et al*. 1995; Loveland & Belward 1997). Regionally, high-resolution data from the Landsat sensors have been used to obtain rates of deforestation in the humid tropics (Skole & Tucker 1993; Townshend & *et al*. 1995)

Most efforts using satellite data to characterize land cover classify the vegetation according to a discrete number of classes. This approach labels each grid cell as a discrete cover type and does not allow for variability within broadly defined biome types (DeFries *et al*. 1995). Because there is important within-biome variation in vegetation attributes, representation of the heterogeneity in those attributes is an important requirement if carbon models are to estimate carbon stocks reliably. Furthermore, most human modifications of the landscape typically occur at spatial resolutions finer than 250 m (Townshend & Justice 1988). By indicating proportional tree coverage within each 1 km grid cell, the dataset we describe in this note captures information about the heterogeneity of vegetation at finer scales. For example, a grid cell containing trees with small patches of cleared areas would likely be categorized as 'forest' in a traditional classification. Estimates of proportional tree coverage, on the other hand, can represent the cleared areas as a reduced overall percentage tree cover for the grid cell. However, the dataset cannot provide information on the subpixel spatial arrangement of the vegetation required to characterize attributes such as fragmentation of forested ecosystems.

For the prototype dataset described in this note, we use remote sensing data to describe the proportional tree coverage within each grid cell as well as the proportional coverage of trees with different leaf longevity and leaf type. It is intended to be used within global carbon models, in conjunction with other spatial datasets providing information such as climate, soil type, and tree height, to improve estimates of carbon stocks and understanding of carbon dynamics at regional and global scales. The dataset can be obtained through the Global Land Cover Facility at http://glcf.umiacs.umd.edu.

Materials and methods

The prototype percentage tree cover product described here was obtained by combining two previously applied approaches for characterizing land cover from satellite data. Both of these datasets were derived from data from NOAA's Advanced Very High Resolution Radiometer (AVHRR) at 1 km spatial resolution acquired in 1992–93 and processed under the guidance of the International Geosphere Biosphere Programme (Eidenshink & Faudeen 1994). The two datasets are:

1 *Global maps of proportional cover for three vegetation characteristics:* leaf form (percentage woody vegetation, percentage herbaceous vegetation, and percentage bare ground), leaf type (percentage needleleaf and percentage broadleaf), and leaf longevity (percentage deciduous and percentage evergreen) as described in DeFries *et al*. (1999). These seven layers were developed as an alternative to classification schemes with discrete numbers of vegetation types. The method, described in detail in DeFries *et al*. (1999), employs a linear mixture model applied to the 1 km AVHRR data. Linear discriminants for input into the mixture model are derived from the same multitemporal metrics used for the classification product described

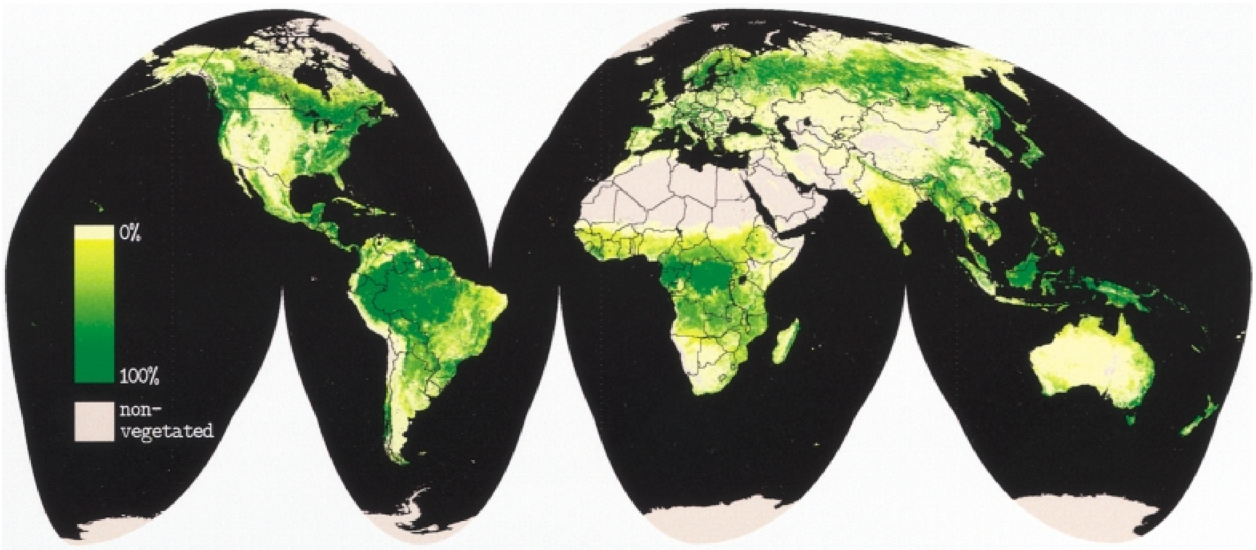


Fig. 1 A prototype dataset estimating percentage tree cover from 10 to 80% based on satellite data acquired by the Advanced Very High Resolution Radiometer in 1992–93. Percentage tree cover is likely to be underestimated in areas with significant cloud cover throughout the year. The spatial resolution of 1 km precludes the detection of finer forest fragments. Future satellites with higher spatial resolution will improve detection of forest patches as well as areas undergoing land cover change. Note that this is the Goode's equal area map projection and hence tree cover in high latitudes appears less extensive than in conventional maps.

below. Training data derived from a global network of Landsat scenes as described in DeFries *et al*. (1998) are used both to determine the weightings for the linear discriminants and to define 'endmember' or 'pure' pixels for calibration of the mixture model. A mixture model was developed to estimate proportional coverage of woody vegetation, herbaceous vegetation and bare ground. A separate mixture model was developed for each continent to determine mixtures of broadleaf evergreen, broadleaf deciduous, needleleaf evergreen, and needleleaf deciduous woody vegetation depending on which forest types are present in each continent. Linear discriminants were derived and the mixture model was calibrated using training data for these forest types in each continent to obtain estimates of the proportions of woody vegetation that are needleleaf, broadleaf, evergreen, and deciduous vegetation.

2 *Global land cover classification* (Hansen *et al*., in press). Details of the methodology for the classification are described in Hansen *et al*. (in press) and DeFries *et al*. (1998). Briefly, a global network of approximately 150 Landsat scenes were interpreted with ancillary data and consultation with experts on land cover in the particular area to obtain training data for the classifier. A decision tree algorithm was then employed using 41 metrics derived from the annual temporal profile of the Normalized Difference Vegetation Index (NDVI) and the five individual bands acquired by the AVHRR to obtain a classification of cover types. The classification contains 12 cover types based on requirements for land cover data identified by the International Geosphere Biosphere Program. Each of these cover types is defined in terms of dominant life form, percentage canopy cover of trees, height of woody vegetation, and seasonality (Belward & Loveland 1995; Hansen *et al*., in press).

The global maps of proportional cover from the mixture model results were observed to be overestimating tree coverage in locations known to have intermediate values of percentage canopy coverage characteristic of woodland (defined by the IGBP as 40–60% canopy cover) and wooded grasslands (10–40% canopy cover). Forest is defined as greater than 60% canopy cover. We adjusted the mixture model result so that the percentage tree coverage in each pixel is consistent with the classification result for that pixel. For each continent, we identified the maximum and minimum estimates of percentage tree cover from the mixture model for each cover type in the classification result. We then linearly scaled the mixture model results for all pixels identified as a given cover type by the maximum and minimum canopy cover values in the IGBP definition of that cover type. For example, if all pixels classified as woodland according to Hansen *et al*. (1999) in a particular continent ranged in percentage woody cover from 30 to 70% according to DeFries *et al*. (1999), the values were scaled to be within the 40–60% range as defined for woodland. Values obtained for leaf type and leaf longevity in DeFries *et al*. (in press) were adjusted so that the total percentage summed to the adjusted percentage woody value.

Pixels classified as agriculture in the tropics were adjusted to range from 10 to 25% tree cover because of the known presence of mixed farming with an overstorey. Islands not present in the 1 km dataset were taken from the AVHRR 8 km Pathfinder data (James & Kalluri 1994).

The mixture model results from DeFries *et al*. (1999) indicate that it is difficult to distinguish percentage

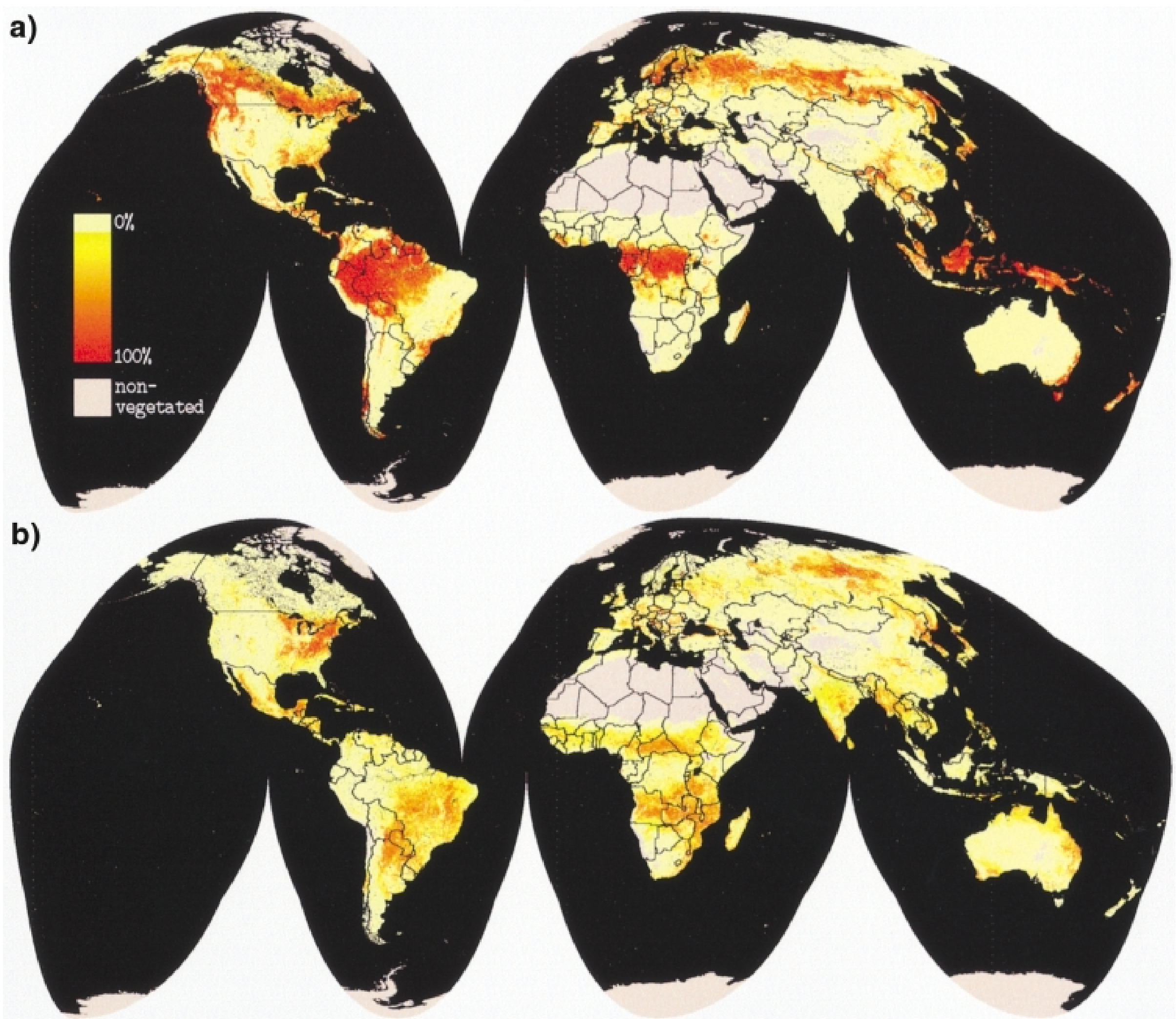


Fig. 2 A prototype dataset derived from satellite data estimating the percentage of tree cover with evergreen (Fig. 2a) and deciduous (Fig. 2b) woody vegetation.

tree cover at extreme high and low percentage coverage. At extreme high coverage, this difficulty is likely due to cloud contamination in humid forests and saturation of the spectral signature. At extreme low coverage, the difficulty is likely due to the overwhelming influence of soil and understorey background on the spectral signature. Consequently, we have depicted the percentage tree cover map as ranging from 10 to 80%, with a value of 80% representing tree coverage equal to or greater than 80% and a value of 10% indicating coverage equal to or less than 10%.

This prototype product was derived by combining two products previously generated from the 1 km AVHRR data. In future generations of such products, we envisage the generation of percentage tree cover in a single processing stream without the requirement for separate classification and mixture modelling results. This will involve the use of a decision tree classifier to stratify the AVHRR data and then

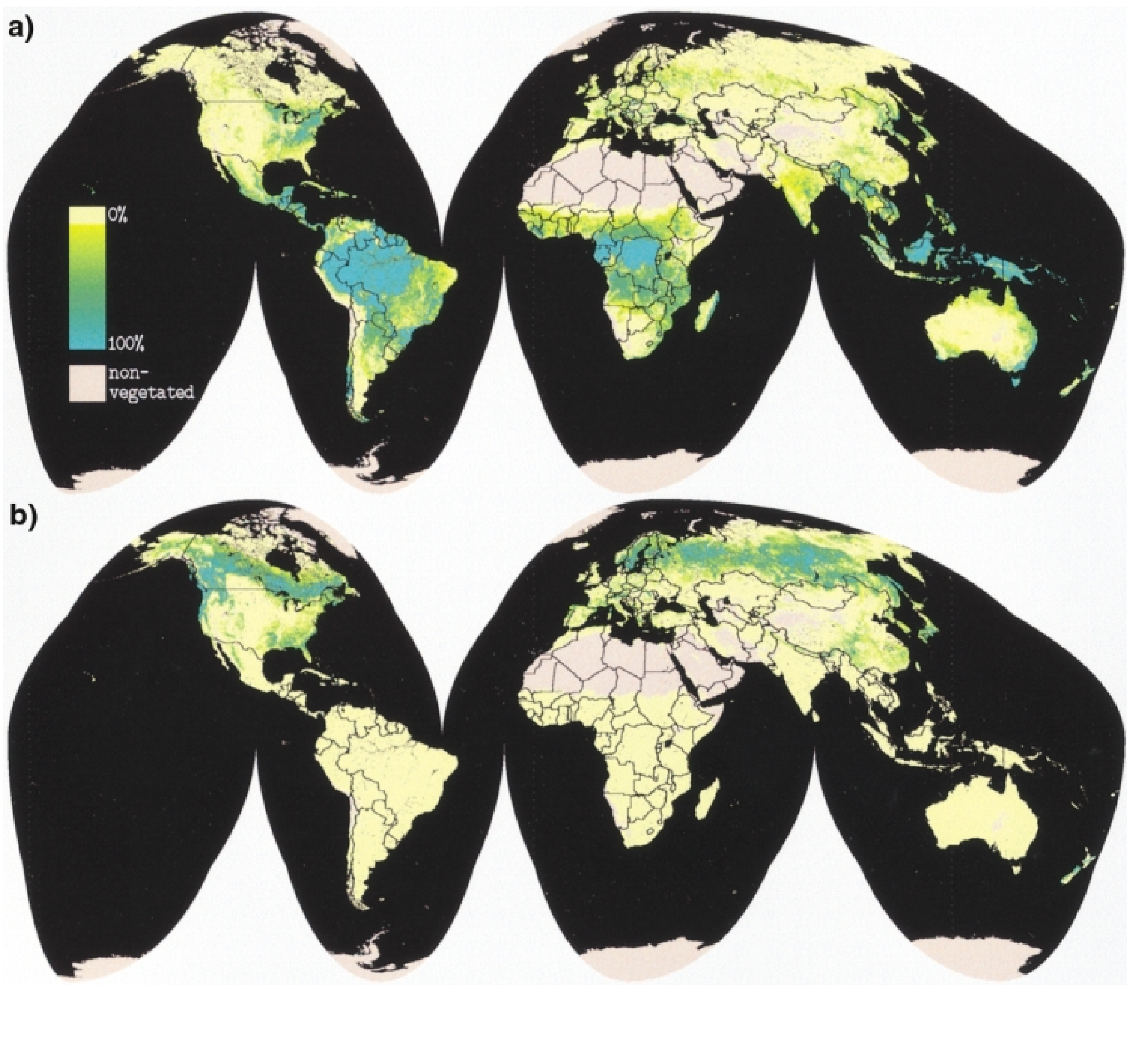
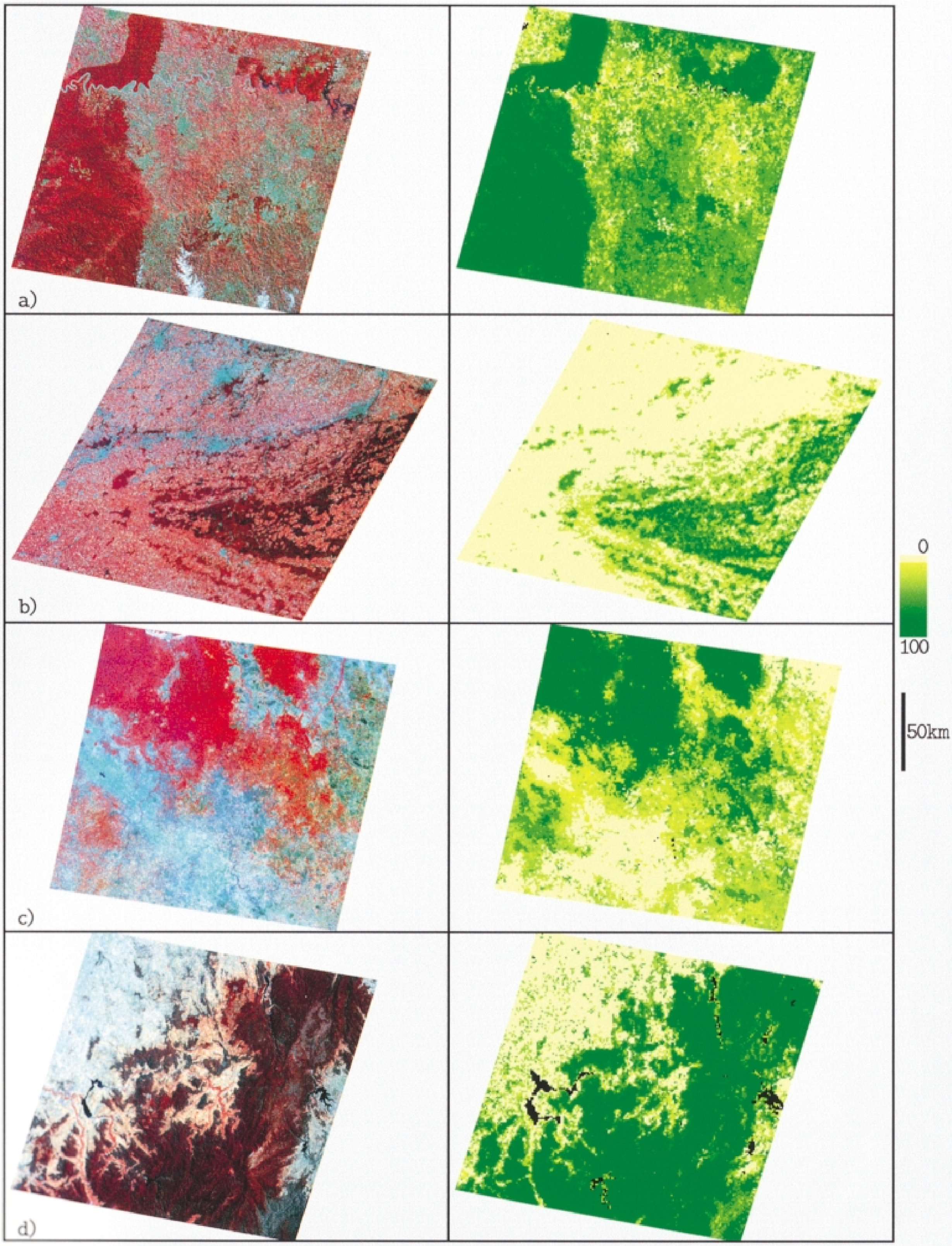


Fig. 3. A prototype dataset derived from satellite data estimating the percentage of tree cover with broadleaf (a) and needleleaf (b) woody vegetation.


a)
b)
c)
d)
0
100
50km

application of a mixture model within each stratified layer.

Results

The prototype product of percentage tree cover (Fig. 1) and accompanying percentages of different leaf longevity and leaf type (Figs 2 and 3, respectively) are at a spatial resolution of 1 km. The resolution can be aggregated to coarser resolutions for modelling applications if desirable. The product contains:

1 A layer with percentage tree cover estimated for each 1 km pixel. Each pixel has a value between 10 and 80%. Because it is not possible to resolve small differences in tree coverage at high and low density, a value of 80 is given for estimates between 80 and 100% tree cover and a value of 10 is given for estimates of percentage tree cover less than or equal to 10.

2 Two layers estimating leaf longevity (evergreen and deciduous). Each pixel has a value between 10 and 80% as does the percentage tree cover layer. The percentage evergreen and percentage deciduous values sum to the percentage tree cover value for each pixel.

3 Two layers estimating leaf type (broadleaf and needleleaf) with values between 10 and 80%. The percentage broadleaf and percentage needleleaf values sum to the percentage tree cover for each pixel.

These layers can be either used directly for parameterization in models or condensed into a more conventional land cover map. For the latter, the product offers the flexibility to derive land cover maps based on the user's requirements for a particular application. For example, forest can be defined as greater than 60% tree cover, as is the case in the classification developed by the IGBP (Loveland & Belward 1997), or as greater than 20% tree cover, as is the case for the Food and Agriculture Organization's forest assessment for developed countries (FAO 1997). The leaf type and leaf longevity layers allow the user to characterize forest type as, for example, needleleaf evergreen forest or broadleaf deciduous forest.

Rigorous quantitative validation of global land cover products is very challenging due to the absence of reliable reference datasets, particularly in the case of proportional coverage. Clearly this is essential and is the subject of major efforts in the future. Nevertheless, we can gain some confidence in this product through a comparison with the spatial distribution of forest cover observed from high resolution Landsat data. Figure 4 shows false colour composites of four Landsat Thematic Mapper scenes from the Argentina/Brazil border, Western Europe, Democratic Republic of Congo, and south-eastern Australia (Fig. 4). The red colour in the Landsat TM composite indicates high reflectance in the infrared wavelength characteristic of photosynthetically active vegetation. The spatial distribution of red patches in the Landsat TM scenes (left side) can be compared with the areas identified as high percentage tree cover (right side) to assess qualitatively the agreement between the high resolution images and the percentage tree cover estimates. While there are subtle differences in the Landsat images and the percentage tree cover product due to the differences in spatial resolution, overall the percentage tree cover estimates appear to depict the broad spatial patterns of forest distribution observed in the Landsat TM scenes.

Discussion and conclusions

The prototype product of percentage tree cover is a spatially explicit, internally consistent map independent of varying definitions of 'forest'. Such maps are a key requirement to be used in conjunction with other datasets within carbon models for estimating carbon stocks for a baseline or any subsequent year. The approach, coupled with analysis of higher resolution imagery and ground-based information, should prove useful in deriving a baseline for carbon stocks and assessing change in the context of the Kyoto Protocol. In particular, it is essential to develop approaches for extrapolating detailed field measurements from ground-based studies over larger areas using satellite data.

Future satellites with improved capabilities and higher spatial resolution, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Earth Observing System (EOS) AM1 platform (Justice & al. 1998) and Landsat 7 (Goward *et al.* 1996; Goward & Williams 1997), will enable more accurate determination of forest extent, particularly in

Fig. 4 (Previous page) Comparisons of four Landsat Thematic Mapper images (left) with 1 km tree cover product (right). The images are infrared, red, green (4–2-1) band composites for border of Argentina and Brazil centred on 53°25′W, 25°59′S (WRS2 path 223, row 078, 04/28/86) (Fig. 4a);. Belgium and France centred on 4°34′E, 50°15′N (WRS2 path 198, row 025, 07/05/87) (Fig. 4b); Southern Democratic Republic of the Congo centred on 23°44′E, 5°47′S (WRS2 path 176 row 064, 07/08/86) (Fig. 4c); and South-eastern Australia centred on 147°48′E, 36°2′S (WRS2 path 091, row 085, 11/22/82) (Fig. 4d). The colour bar represents 0–100% tree cover.

© 2000 Blackwell Science Ltd., *Global Change Biology*, **6**, 247–254

areas of fragmented forest. The dataset described in this note contributes to the activities being planned under the Committee on Earth Observations Systems pilot project on Global Observations of Forest Cover (Ahern *et al*. 1998) to apply remote sensing to the crucial need for accurate information on the spatial extent of forest.

Acknowledgements

We gratefully acknowledge NASA's support through the following grants: NASG56004, NAG56364, and NCC5300, and also the USGS for supporting this research. We are grateful for the assistance of Rob Sohlberg in producing the graphics.

References

Ahern F, Janetos A, Langham E (1998) Global Observations of Forest Cover: One component of CEOS' integrated global observing system strategy. *Proceedings of the 27th International Symposium on Remote Sensing of Environment, Tromso, Norway,.*

Belward AS, Loveland TR (1995) The IGBP DIS 1 km land cover project: Remote sensing in action. *Proceedings of the 21st Annual Conference of the Remote Sensing Society, Southampton, U.K.*, pp. 1099–1106.

DeFries RS, Townshend JRG, Hansen M (1999) Continuous fields of vegetation characteristics at the global scale at 1km resolution. *Journal of Geophysical Research*, **104,** 16,911–16,925.

DeFries RS, Field CB, Fung I *et al*. (1995) Mapping the land surface for global atmosphere-biosphere models: toward continuous distributions of vegetation's functional properties. *Journal of Geophysical Research*, **100**, 20, 867–20, 882.

DeFries R, Hansen M, Townshend JRG *et al*. (1998) Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat Imagery in decision tree classifiers. *International Journal of Remote Sensing*, **19,** 3141–3168.

DeFries RS, Townshend JRG (1994a) Global land cover: comparison of ground-based data sets to classifications with AVHRR data. In: *Environmental Remote Sensing from Regional to Global Scales* (Foody G., Curran P., eds). John Wiley, Chichester.

DeFries RS, Townshend JRG (1994b) NDVI-derived land cover classification at global scales. *International Journal of Remote Sensing*, **15,** 3567–3586.

Eidenshink JC, Faudeen JL (1994) The 1 km AVHRR global land data set: First stages in implementation. *International Journal of Remote Sensing*, **15,** 3443–3462.

FAO (1997) *State of the World's Forests: 1997*. Food and Agriculture Organization of the United Nations, Gland.

Goward S et al. (1996) Landsat-7 long term acquisition plan: Seasonal acquisitions. *Proceedings of Pecora-13, 20–22 August 1996*, Sioux Falls, South Dakota.

Goward S, Williams D (1997) Landsat and earth system science: Development of terrestrial monitoring. *Photogrammetric Engineering and Remote Sensing*, **63,** 887–900.

Hansen M, DeFries R, Townshend JRG *et al*. (in press) Global land cover classification at 1km resolution using a decision tree classifier. *International Journal of Remote Sensing*, in press.

IGBP Terrestrial Carbon Working Group (1998) The terrestrial carbon cycle: Implications for the Kyoto Protocol. *Science.*, **280,** 1393–1394.

James ME, Kalluri SNV (1994) The Pathfinder AVHRR land data set: an improved coarse resolution data set for terrestrial monitoring. *International Journal of Remote Sensing*, **15,** 3347–3364.

Justice CO, Vermote E, Townshend J *et al*. (1998) The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, **36,** 1228–1249.

Loveland TR, Belward AS (1997) The IGBP-DIS global 1 km land cover data set, DISCover: first results. *International Journal of Remote Sensing*, **18,** 3289–3295.

Malingreau JP, Achard F, D'Souza G *et al*. (1995) AVHRR for global tropical forest monitoring: The lessons of the TREES project. *Remote Sensing Reviews*, **12,** 29–40.

Mayaux P, Achard F, Malingreau JP *et al*. (1998) Global tropical forest area measurements derived from coarse resolution satellite imagery: a comparison with other approaches. *Environmental Conservation*, **25,** 37–52.

Schimel DS (1995) Terrestrial ecosystems and the carbon cycle. *Global Change Biology*, **1,** 77–91.

Skole D, Tucker CJ (1993) Tropical deforestation and habitat fragmentation in the Amazon: Satellite data from 1978 to 1988. *Science*, **260**, 1905–1910.

Stone TA, Schlensinger P, Houghton RA *et al*. (1994) A map of the vegetation of South America based on satellite imagery. *Photogrammetric Engineering and Remote Sensing*, **60,** 541–551.

Townshend JRG, Justice CO, Kalb VT *et al*. (1987) Characterization and classification of South American land cover types using satellite data. *International Journal of Remote Sensing*, **8,** 1189–1207.

Townshend JRG, Bell V, Desch A *et al*. (1995) The NASA Landsat Pathfinder Humid Tropical Deforestation Project. *Proceedings of the ASPRS Conference on Land Satellite Information in the Next Decade, 25–28 September 1995, Vienna, VA.*

Tucker CJ, Townshend JRG, Goff TE (1985) African land-cover classification using satellite data. *Science*, **227,** 369–375.

WCMC (1998) *Generalized World Forest Map*. World Conservation Monitoring Centre, Cambridge, U.K.