Remote Sensing of Environment 114 (2010) 73–86

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices

Jean-Robert B. Bwangoy [a,*], Matthew C. Hansen [a], David P. Roy [a], Gianfranco De Grandi [b], Christopher O. Justice [c]

[a] Geographic Information Science Center of Excellence, South Dakota State University, Brookings, SD, 57006, USA
[b] European Commission, Joint Research Center, Institute for Environment and Sustainability, TP 440, 21020, ISPRA, Italy
[c] University of Maryland at College Park, Geography Department, College Park, MD, 20742, USA

ARTICLE INFO

Article history:
Received 18 September 2008
Received in revised form 8 August 2009
Accepted 10 August 2009

Keywords:
Wetlands
Land cover
Classification
Decision tree
Flooded forest
Multi-source data
Congo Basin
Central Africa

ABSTRACT

This study reports results of a classification tree approach to mapping the wetlands of the Congo Basin, focusing on the Cuvette Centrale of the Congo River watershed, an area of 1,176,000 km^2. Regional expert knowledge was used to train passive optical remotely sensed imagery of the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors, JERS-1 active radar L-band imagery, and topographical indices derived from 3 arc sec elevation data of the Shuttle Radar Topography Mission (SRTM). All data inputs were resampled to a common 57 m resolution grid. A classification tree bagging procedure was employed to produce a final map of per-grid cell wetland probability. Thirty bagged trees were ranked and the median result was selected to produce the final wetland probability map. Thresholding the probability map at <0.5 yielded a proportion of wetland cover for the study area of 32%, equivalent to 360,000 km^2. Wetlands predominate in the CARPE Lake Tele–Lake Tumba landscape located in the western part of the Democratic Republic of the Congo and the south-eastern Republic of Congo, where they constitute 56% of the landscape. Local topography depicting relative elevation for sub-catchments proved to be the most valuable discriminator of wetland cover. However, all sources of information (i.e. optical, radar and topography) featured prominently in contributing to the classification tree procedure, reinforcing the idea that multi-source data are useful in the characterization of wetland land cover. The method employed freely available data and a fully automated process, except for training data collection. Comparisons to existing maps and *in situ* field observations indicate improvements compared to previous efforts.

1. Introduction

Wetlands act as centers of biological productivity and perform many functions, including: regulation of the hydrological cycle and flood control, improvement and protection of water quality, erosion control and shoreline protection, conservation of biological diversity, habitat for wildlife, and fisheries and resources for human communities (Coughanowr, 1998). They are defined as "*lands transitional between terrestrial and aquatic systems where the water table is usually at or near the surface or the land is covered by shallow water*" (Cowardin et al., 1979; Middleton, 1999).

The role of tropical wetlands in the global climate system, ecosystem regulation, biodiversity conservation, water supply, water quality regulation, and food supply has long been recognized (Naiman et al., 1998; Segers, 1998). These ecosystem services have provided an important justification for promoting international wetland protection frameworks such as the Ramsar Convention (Davis, 1994) and other initiatives under the framework of the Convention for Biological Diversity (Balmford et al., 2005). Tropical and subtropical wetlands are also an important source of methane, contributing to 50–75% of annual methane emissions from natural wetlands (Matthews, 2000), and may contribute up to 22% of the present anthropogenically enhanced greenhouse gas effect (Gauci et al., 2004). However, despite the economic and ecological importance of wetlands, there are many uncertainties regarding their extent, distribution, and ecological and physical functions (Junk and Piedade, 2005; Matthews and Fung, 1987; Sahagian and Melack 1998). To better quantify the role of tropical wetlands in providing ecosystem services, an improved understanding of their distribution is required.

Remote sensing techniques are yielding promising results for mapping and quantifying tropical wetlands. However, limitations remain due to frequent cloud coverage and structural heterogeneity of the vegetation canopy (Saalovara et al., 2005). Although persistent cloud coverage often prevents data acquisition from passive optical wavelength remote sensing systems in equatorial areas (Asner, 2001; Ju and Roy, 2008; Roy et al., 2006), such data have widely been used for deforestation mapping of the humid tropics (Archard et al., 2001; Hansen and Defries, 2004; Hansen et al., 2008a). However, most of these

* Corresponding author.
E-mail address: Bwangoy.Bankanza@sdstate.edu (J.-R.B. Bwangoy).

0034-4257/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2009.08.004

studies have used photointepretation methods to identify wetlands. For example, the Brazilian PRODES (Amazon Deforestation Monitoring Project) project employs a manually interpreted hydrography layer to identify core wetland areas and remove them from their passive optical Landsat satellite-based forest monitoring analysis (INPE, 2002; Shimabukuro et al., 1998). Similarly, the United Nations Food and Agriculture Organizations Africover mapping program used manual methods with Landsat data to delineate wetland classes for the Democratic Republic of the Congo's (DRC) humid tropical forest zone (FAO, 2005). Algorithmic approaches to mapping tropical wetlands have been undertaken, most of them using active remote sensing data (De Grandi et al., 1998; Hoekman and Quinones, 2002; Hess et al., 2003; Hess et al., 1995; Melack, 2004; Simard et al., 2002; Townsend, 2002; Van der Sanden and Hoekman, 1999). To date, the only spatially explicit Congo Basin wetland map of this kind was generated using ERS1 and JERS1 SAR mosaics (De Grandi et al., 1998; Mayaux et al., 2004).

A number of studies have employed both passive optical and active remote sensing data for improved characterization of vegetation (Amarsaikhan and Douglas, 2004; Le Hegart-Mascle et al., 2003). Multi-source remote sensing data offer the potential for improved classification accuracy compared to the accuracy achieved by a single source classification (Amarsaikhan and Douglas, 2004; Briem et al., 2002; Ozesmi and Bauer, 2002), although the selection of appropriate analysis methods remains a challenge (Briem et al., 2002). The objective of this study is to employ multi-source data to classify the wetlands of the Congo Basin and to advance tropical wetland mapping methodologies. Past Congo basin mapping approaches have employed single-source data, including the Landsat-based Africover data set (FAO, 2005) and the radar-based map of De Grandi et al. (1998). Both of these works represent pathfinding research in the field of land cover mapping using remotely sensed data sets. The Africover map consisted of an international collaboration effort aimed at characterizing large areas of central and east Africa, and led to one of the first regional high spatial resolution land cover classifications. The product of DeGrandi et al. was created as part of the European Commission's Joint Research Centre's Global Rain Forest Mapping Project (JRC/GRFM), and was the first example of regional wetlands mapping in Central Africa. The 300 m spatial resolution JRC/GRFM map was later refined and incorporated into the 1 km Global Land Cover 2000 product (Mayaux et al., 2004).

In this paper, a multi-source statistical supervised classification approach is undertaken to classify wetlands in the Congo basin. Passive optical remotely sensed imagery provided by the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors, JERS-1 active radar L-band horizontal co-polarization imagery, and 3 arc sec elevation data derived from the Shuttle Radar Topography Mission (SRTM) are used as inputs. These data are related to wetland/non-wetland training data through a classification tree algorithm to yield a wetland probability map of the central Congo Basin for the 1986–2000 period, with the fundamental assumption that the limits between wetland and non-wetland areas were stable over this 15 year period.

This work is a contribution to the CARPE program (Central African Regional Program for the Environment), a long-term initiative funded by the United States Agency for International Development (USAID) to address the issues of forest management, human livelihoods, and biodiversity loss in the Congo Basin forest zone (CBFP, 2005). A number of landscapes were identified by the CARPE program as a focus for conservation activities beyond the protected area systems of Central Africa. Of these, a total of 12 landscapes were recognized as priority conservation areas based on their relative taxonomic importance, overall integrity, and resilience of existing ecological processes (CBFP, 2005). The study area in this paper encompasses six of these landscapes. It includes entirely four CARPE landscapes (Maringa–Lopori–Wamba, Lake Tumba–Lake Tele, Salonga–Lukenie–Sankuru and the Sangha Tri-National) and partially the Dja–Odjala–Minkebe Tri-National and Leconi–Bateke–Lefimi landscapes (Fig. 1). The methodology and resulting wetland map, hereafter referred to as the CARPE wetland map, serve as contributions to CARPE. The map is available for download through the CARPE website, either through the CARPE explorer (http://maps.geog.umd.edu/metadataexplorer/explorer.jsp) or directly at http://maps.geog.umd.edu/carpe/Products/swamp_congobasin.zip.

2. Materials and methods

2.1. Study area

The Congo Basin covers approximately three million square kilometers and includes the world's second largest tropical forest of approximately two million square kilometers (Laporte et al., 1998). The Basin's forest zone includes parts of six central African countries: Cameroon, Central African Republic, Congo, Equatorial Guinea, Gabon, and the Democratic Republic of the Congo. The study area includes the core wetlands of the Congo Basin, found within the forest zone from 5° North to 6° South and from 13° to 26° East (Fig. 1). These core wetlands are located in the Cuvette Centrale, or Central Basin, of the Congo River a Pliocene-epoch lake bottom that filled with sediment to form a shallow basin with little relief variation and traversed by slowly flowing rivers. The resulting lack of topographic variation complicates the discrimination of the extensive wetland areas from *terra firma* lowland forests.

The climate of the study area is warm and humid with two wet and two dry seasons (Bultot, 1974; Samba et al., 2008). The mean temperature is approximately 25 °C, and the difference between the temperature of the warmest month (March) and the coldest month (July) is only 2 °C (Bultot, 1974). The average rainfall is about 1800 mm per year in 115 days (Bultot, 1974). The basin is largely bisected by the Equator with little seasonal variation within 1° north or south of the Equator. Seasonality increases with distance from the equator and mirrored phenologies exist at similar latitudes north and south of the equator. The northern part of the basin experiences a minor rainy season from September to November and a major one from the first half of March to early May; in the south, the minor rainy season lasts from February to May and the major rainy season occurs between September and December (Samba et al., 2008).

The climax ecosystem is tropical evergreen forest with little seasonal variation (Lebrun and Gilbert, 1954; Mayaux et al., 2000). The forests typically have an irregular and very dense (70–100%) crown cover composed of heterogeneous evergreen and shade tolerant species that can reach 35–45 m height and that often preclude the development of understory shrubs and grasses (Mayaux et al., 2000). The lower layers typically have two individual strata composed of regenerating upper canopy layer species (Lebrun and Gilbert, 1954). At the border of the central basin, semi-deciduous forests become the dominant climax vegetation (Devred, 1958). Secondary forest consists of regrowth, fallow and complexes of regrowth, fallow and crops, with dense undergrowth and regular crown cover frequently dominated by *Musanga cecropioides* (Mayaux et al., 2000). Savannas are a significant part of the landscape in Central Africa and are thought to cover about 40% of the Congo Basin (Hansen et al., 2008b; Laporte et al., 1998). In this warm and humid climate, soil type and fire disturbance are important factors in regulating woody cover and savanna structure (Sankaran et al., 2005).

The wetland areas include flooded forests and inundated grasslands, both characterized by vegetation adapted to water saturated soils, anaerobic conditions and frequent flooding. Flooded forests cover large areas in the center of the study site, mainly in the western part of the Democratic Republic of Congo and the northeastern part of the Republic of Congo (Mayaux et al., 2000). They extend into various tributaries of the Congo River and feature fine-scale variation as they permeate the Cuvette Centrale. They have been poorly studied and

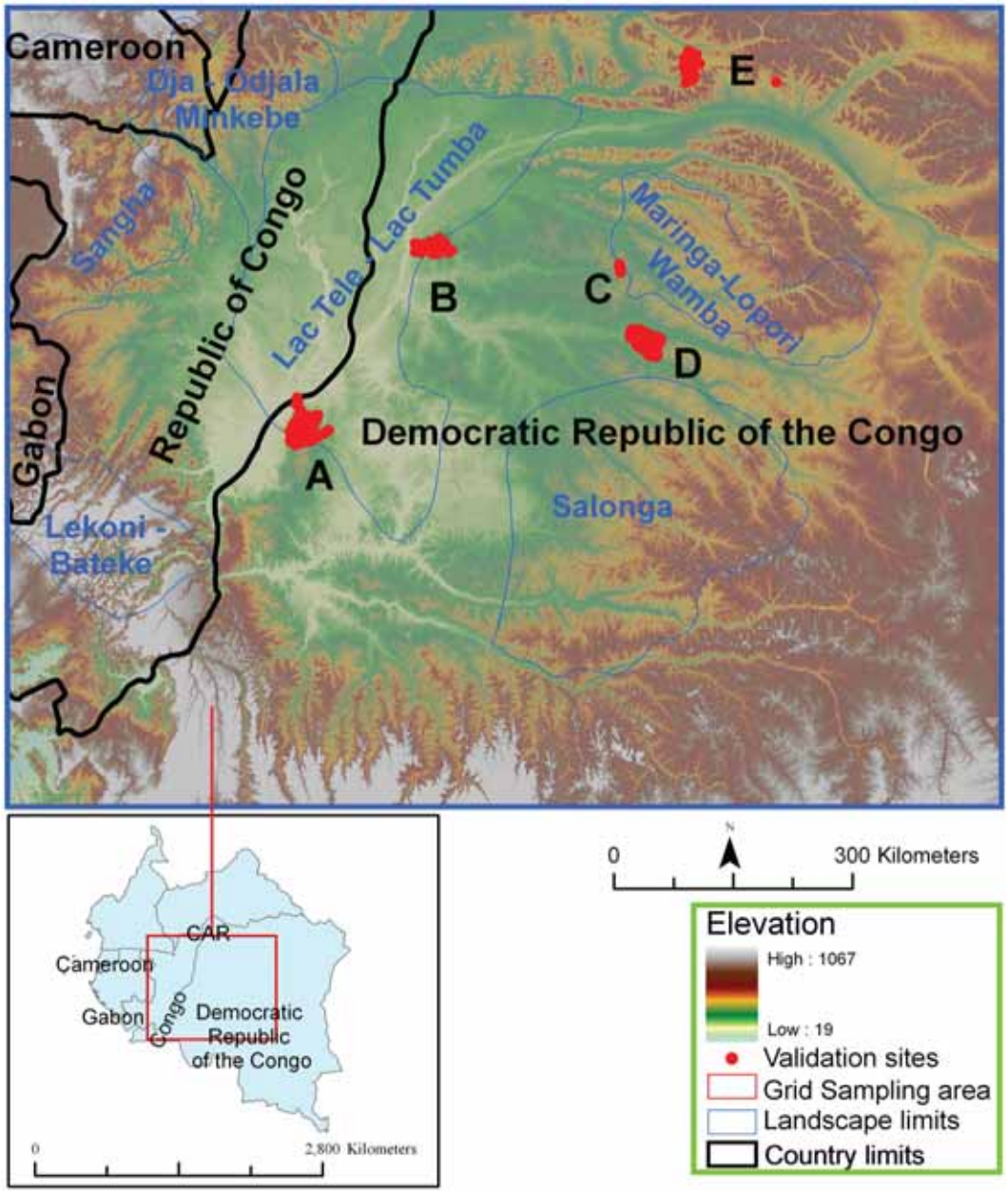

Fig. 1. Location of the study area and GPS validation sites. See text for descriptions of areas A–E.

sometimes not even mentioned in global vegetation maps (Campbell, 2005). They are temporally or permanently inundated, and in all cases characterized by soils with poor drainage (Mayaux et al., 2002). They are believed to have no marked period of drying out since there is no massive ENSO drought as in the Amazon (Saleska et al., 2007). Flooded forests have slightly heterogeneous vegetation, structure and composition (Campbell, 2005). They are multilayered with an upper canopy often dominated by Caesalpiniaceous species such as Bubinga (*Guibourtia demeusei*) (Campbell, 2005; Lebrun and Gilbert, 1954).

2.2. Data sets

2.2.1. Landsat data

Epochal Landsat mosaics consisting of radiometrically normalized Landsat TM and ETM+ imagery that conform to NASA's decadal Geocover orthorectified global Landsat dataset (Neigh et al., 2007) were used as inputs. Landsat mosaics for the 1990s (hereafter referred to as Landsat Epoch 1990) and the 2000s (Landsat Epoch 2000) were used, each covering a region of 100 Landsat World Reference System path/rows. Each Landsat acquisition (approximately 185×185 km) was normalized using a dark object subtraction method and cross-track anisotropy adjustments to reduce scene-to-scene and within scene radiometric variations associated with atmospheric, phenological, and sun-sensor-target geometric variations (Hansen et al., 2008b). In this process, a pre-existing 250-m MODIS Vegetation Continuous Field (VCF) forest/non-forest map (Hansen et al., 2003) was used to provide dark targets; in addition clouds and shadow were classified into low, medium and high confidence categories using a decision tree approach and photo interpreted training data (Hansen et al., 2008b). Up to 5 Landsat acquisitions per unique Landsat path/row were composited temporally to reduce the presence of cloud and shadow (Lindquist et al., 2008). In addition, for this study, the two epochal mosaics were composited together to further reduce cloud cover and shadow by discarding clouds. The mosaics had a 57-m pixel size and were defined in the Universal Traverse Mercator coordinate system (UTM zone 33). TM and ETM+ bands 4 (0.75–0.90 μm), 5 (1.55–1.75 μm), and 7 (2.09–2.35 μm) were

used as they are less sensitive to atmospheric effects than the shorter wavelength Landsat bands (Ouaidrari and Vermote, 1999). They are also absorptive over water targets and transpiring vegetation, making them useful for mapping wetlands. A band 6 (10.4–12.5 μm) 2000-epoch composite was generated by discarding cloudy pixels and taking the maximum brightness temperature (Roy, 1997).

2.2.2. Radar data

A mosaic of wide-area multi-resolution and multi-temporal JERS-1 Synthetic Aperture Radar (SAR) data of the Central African region compiled in the framework of the Global Rain Forest Mapping(GRFM) Project was used (De Grandi et al., 2000; Mayaux et al., 2002; Rosenqvist and Birkket, 2002). These data consisted of 100-m spacing bi-temporal georeferenced and calibrated wide area mosaics of both textural and radiometric information resampled from 12.5-m resolution JERS-1 SAR data. They were acquired in 1996 and timed to low water (January to March) and high water (October to November) stages of the Congo River, capturing seasonal variations in inundation. The seasonality captured in JERS-1 data may also help overcome issues related to the phenological variation between wetland and non-wetland vegetated areas (De Grandi et al., 2000; Mayaux et al., 2002; Rosenqvist and Birkket, 2002).

The JERS-1 satellite sensor scans the target at 23.5 cm (L-Band) with horizontal co-polarization between 37.4° and 41.5° incidence angles over Central Africa (De Grandi et al., 2000). The total backscattered volume can be modeled using the radiative transfer equation of Sun (1990), which treats the tree canopy, tree trunk and ground surface as layered media that scatter and attenuate the incoming microwave energy (Hess et al., 1995; Melack, 2004):

$$\sigma^{\circ}_{t} = \sigma^{\circ}_{c} + \sigma^{\circ}_{s} + \sigma^{\circ}_{m} + \sigma^{\circ}_{d} \quad (1)$$

where σ°_{t} is the total backscattered volume, σ°_{c} is the canopy volume scattering, σ°_{s} is the surface backscatter, σ°_{m} is the multiple path interaction of the canopy and the ground and, σ°_{d} is the double bounce trunk ground backscattering volume.

Using L-band observations with a 37° incidence angle and horizontal co-polarization, the canopy volume and the trunk-ground interaction contribute equally to the backscattering (Wang et al., 1995). The radiation hitting the surface in a typical wetland is reflected by the presence of the water layer and is backscattered by a double-bounce trunk-ground mechanism creating contrast between the flooded and non-flooded terrain and enabling wetland characterization (Mayaux et al., 2002).

For the present work, we used 8-bit binary image files representing backscattering amplitudes. These amplitude data were not converted to normalized radar cross section (σ°) to avoid compressing the high range of the data and boosting the low range, as wetlands have high returns.

Since the GRFM mosaic was provided in a Mercator projection (datum WGS84), it was reprojected to fit the Geocover data coordinate system, using the nearest neighbor resampling method. Although efforts were made to reduce range-dependent and processing dependent errors, some additional sources of calibration errors in the GRFM mosaic could not be corrected. These include along track calibration errors resulting from a lag time adjustment of the instrument's automatic gain control, and corrections for terrain-induced distortions (Hess et al., 2003).

2.2.3. Elevation data and derived topographic metrics

In this study, a 3 arc sec (approximately 92 m) spatial resolution Digital Elevation Model (DEM) derived from the NASA Shuttle Radar Topography Mission (SRTM), flown onboard the Space Shuttle mission Endeavour STS-99 from February 11 to February 22, 2000 (Hennig et al., 2001; Hamilton et al., 2007; Rabus et al., 2003) was used. The DEM was derived from interferometric processing of single-pass data collected by a C-band (5.6 cm) SAR (Brown et al., 2005; Hamilton et al., 2007; Hennig et al., 2001). It was reprojected from a spherical coordinate system (latitude and longitude) into the Geocover data coordinate system.

Tropical wetlands often occur as linear features corresponding to the alignment of valley bottoms (Saalovara et al., 2005). A number of metrics that quantify the effects of topography on hydrological processes have been considered for hydrological modeling and can be derived from digital elevation data, typically either as first order derivatives of elevation in the x and y axes (e.g. slope) or as second order derivatives (e.g. curvatures) (Wood, 1996; Wolock and McCabe, 1999). Non-systematic DEM errors will propagate adversely into the derivatives and so their accuracy depends upon the DEM's quality (Wolock and McCabe, 1999). The STRM data have an absolute estimated height error of 11.25 m (1.6 σ) and a relative height error of 1.6 m–3.3 m (Brown et al., 2005). The absolute height error is less important for this study as we do not threshold against height in an absolute sense but rather are concerned with the relative spatial change in elevation to quantify likely water presence. Metrics were derived from the first and second order derivatives of the DEM, and by estimating relative changes in height with respect to hydrological stream flow. These indices are discussed below. A total of 9 first and second order derivatives were derived at each DEM pixel: slope, aspect, shaded relief, profile convexity, plan convexity, longitudinal curvature, cross-sectional curvature, maximum and minimum curvatures (Wood, 1996). These were derived by fitting a quadratic surface to the digital elevation data using the ENVI software topographical modeling package. The quadratic surface is defined as:

$$Z(x,y) = ax^2 + by^2 + cxy + dx + ey + f \quad (2)$$

where $Z(x,y)$ is the height at pixel coordinate x,y and a, b, c, d, e and f are six quadratic coefficients defining the surface intersecting (x,y). The six quadratic coefficients were derived by ordinary least squares (OLS) fitting of 9 adjacent DEM pixel locations falling in a 3×3 kernel around each DEM pixel. The 9 first and second order derivatives were computed for each DEM pixel by solving established analytical partial derivatives of Eq. (2) (Wood, 1996; Zhou and Liu, 2002). The root mean square error of the OLS fitting was also derived to provide a metric indicative of the quality of the fit. A total of 6 metrics describing relative changes in height with respect to hydrological stream flow were derived at each DEM pixel. The hydrological topology, as defined by water flow paths and watershed regions, was considered using the D8 method implementing the steepest descent approach to compute the flow direction, the upslope contributing area (flow accumulation), and watersheds (MacMillan et al., 2000). In this method, the flow direction grid cell values are assigned by determining to which of eight possible neighboring cells (top, bottom, left, right, or one of the four neighboring corners) a hypothetical raindrop would flow (O'Callagan and Mark, 1984; Tarboton, 1997). The direction of steepest descent is determined and the cell to which that vector points defines the flow direction value. This method is referred to as D-8 because all flow will be assigned to one of the eight possible directional neighbors. Following O'Callagan and Mark (1984), drainage channels were extracted at all points with accumulated areas above a threshold defined by a support area (SA) parameter. In this method, all cells having at least as many cells flowing to them as the threshold value, are determined to be “streams”. In general, lower SA values yield high density drainage basins, and higher SA threshold values mean that slopes are averaged over longer distances, reducing the effects of DEM errors (Tarboton et al., 1989, Tarboton et al., 1992; Tarboton and Ames, 2001; Tarboton, 1997). For each DEM pixel, the absolute height difference between that pixel and the downstream closest pixel lying on a stream or river was derived (Deng, 2007; MacMillan et al., 2000; Williams et al., 2000). As the selection of an optimal SA threshold is poorly defined (Tarboton et al., 1992;

Tarboton et al., 1991), we repeated this six times using different SA thresholds of 1000, 1500, 5000, 10,000, 15,000 and 20,000 grid cells (each 3 arc sec grid cell is approximately 92 m) in order to encompass maximum stream density layers (with smaller SA values) and reduced DEM errors (higher SA values). Each of the derived variables represents relative elevations for the six different support area thresholds applied. Regarding SA thresholds, larger threshold values result in larger catchments and smaller thresholds in smaller catchments.

2.3. Training data

Training data were collected using manual photo-interpretation techniques from Landsat ETM+ mosaics along grid lines spaced 65 km apart in both *x* and *y* directions (systematic grid sampling). A total of 1,425,363 pixels were sampled, defining at each pixel the interpreted surface status (wetland or non-wetland), of which 607,551 pixels (or 42.6%) were interpreted as wetlands. The systematic grid sampling was performed in three CARPE landscapes: Lac Tele–Lac Tumba, Maringa–Lopori–Wamba and Salonga–Lukenie–Sankuru (Fig. 1) and provided an estimate of the actual class proportions, akin to prior probabilities. Systematic sampling does not generally ensure a high-fidelity map output due to the complexity of landscapes and input variable features. Consequently, additional, non-systematically collected training data were added to the initial training after a first algorithm run, to identify areas of obvious error by photo-interpretation of the initial product. Additional areas were labeled and used to augment the training data in iterating the algorithm.

2.4. Decision tree classification method

A classification tree algorithm (Breiman et al., 1984) was used to estimate the per-pixel probability of wetland occurrence, using the wetland and non-wetland training data, and the multisource data described in the previous section as independent variables. The independent variables used are summarized in Table 1. The classification tree uses a set of hierarchical rules that split data into two groups (child nodes) which are purer than the input group (parent nodes). The algorithm searches for the best univariate split (each decision rule is univariate) by computing the reduction of the deviance, a measure of node purity that reflects the misclassification error rate. The split that produces the greatest reduction of the deviance is used to divide the data and the process is repeated for the newly created subsets (Breiman et al., 1984). Decision trees have been used to characterize land cover using remotely sensed data sets (Hansen et al., 1996; Michaelson et al., 1994; Verbyla, 1987) and specifically for mapping tropical wetlands (Hess et al., 1995; Simard et al., 2000; Simard et al., 2002).

Thirty trees were grown independently using 10% of the training samples selected randomly with replacement (bagging) (Breiman et al., 1984) and a threshold deviance of 0.015% applied for terminating tree growth. Results from the 30 classification trees were ranked and the median result per pixel was used for estimating the probability of wetland class membership. In a two class case, the median represents the result that would typically occur under most voting methods used in classification tree bagging procedures (Hansen et al., 2008b). The final probability map was thresholded using a probability value that reproduced the proportions of the systematically acquired training data in order to preserve the prior probability proportions of wetland/non-wetland land cover, i.e., 42.6% wetland cover within the initial three landscapes.

2.5. Product evaluation

Field validation of regional-scale vegetation maps remains a challenge, particularly in Central Africa, because of the difficulty in conducting ground truth surveys in a representative way over large areas (Mayaux et al., 2002). The CARPE wetland map was first validated by comparison with in *situ* forest inventory data, and then compared to the Africover vegetation map of the Democratic Republic of the Congo (FAO, 2005) and the JRC/GRFM Regional Flooded Forest Map of Central Africa (De Grandi et al., 1998; Mayaux et al., 2004), both of which included wetlands as thematic classes. In this study, the producer's and user's accuracies and Kappa coefficient were used to intercompare map products and to compare map products with the *in situ* observations. These statistics were derived from a two-way (wetland and non-wetland) contingency table that was split into subclasses to further identify sources of disagreement per subclass. The Kappa coefficient measures the agreement between variables by adjusting the percentage of agreement to the amount of agreement due to chance alone (Cohen, 1968), although Kappa values may be biased in areas with uneven proportions of the different classes (Foody, 2004; Stehman, 2006). Kappa is bound from 0 to 1, where 0 equals no agreement beyond chance, and 1 perfect agreement (Titus et al., 1984). It was calculated as follows:

$$\text{Kappa} = (\text{observed agreement} - \text{chance agreement}) / (1 - \text{chance agreement}). \quad (3)$$

Within a two-way contingency table of wetland and non-wetland, the summed diagonal elements divided by the total sample size equals the observed agreement. The chance agreement equals the sum of the products of the per class marginal probabilities. The producer's accuracy

Table 1
Independent variables used to map Congo Basin wetlands, and the relative contribution of each variable in discriminating wetlands using the classification tree bagging procedure.

Metric name	Deviance	Relative contribution (%)
Relative elevation 20,000 grid support threshold area	492,483.73	46.25
Landsat band 5 epoch 2000	142,245.86	13.36
Radar high water	92,523.96	8.69
Relative elevation 1000 grid support threshold area	65,833.93	6.18
Landsat band 4 epoch 2000	60,909.21	5.72
Slope	42,339.41	3.98
Landsat band 5 epoch 1990	21,476.81	2.02
Relative elevation 1500 grid support threshold area	19,575.93	1.84
Landsat band 4 composite of epoch 1990 and 200	17,233.79	1.62
Relative elevation 5000 grid support threshold area	15,029.61	1.41
Landsat band 7 epoch 1990	13,386.83	1.26
Root mean square error	13,273.78	1.25
Lansat band 7 composite of epoch 1990 and 2000	10,835.47	1.02
Lansat band 6 composite of epoch 1990 and 2000	10,185.87	0.96
Landsat band 4 epoch 1990	8291.66	0.78
Landsat band 5 composite of epoch 1990 and 2000	7860.97	0.74
Relative elevation 10,000 grid support threshold area	7576.59	0.71
Radar low water	5921.56	0.56
Maximum curvature	4835.60	0.45
Relative elevation 15,000 grid support threshold area	4361.53	0.41
Landsat band 7 epoch 2000	4140.64	0.39
Minimum curvature	2482.36	0.23
Aspect	878.88	0.08
Longitudinal convexity	484.16	0.05
Profile convexity	362.81	0.03
Shaded relief	140.27	0.01
Cross sectional convexity	136.38	0.01
Plan convexity	53.53	0.01
Total	1,064,861.16	100.00

See Section 2.2.3 for explanation of the topographic metrics.

is a measure of the probability that a pixel observed in the field is correctly classified in the map. It is calculated by dividing the number of all correctly classified pixels for a class by the total number of pixels of that class observed in the field. The user's accuracy measures the probability that a pixel on the map correctly identifies the land cover category as it exists in the field. It is calculated by dividing the number of correctly classified pixels for a class by the sum of all pixels assigned to that class in the map (Congalton, 1991; Stehman, 1997).

2.5.1. Forest inventory data

Forest inventories were carried out in five areas within the Cuvette Centrale from February 2003 to December 2007. Tree diameters and quality were recorded for all trees with a diameter higher than 20 cm. In addition, geographic coordinates were recorded using a Global Positioning System every 100 m along each transect. For each point, vegetation characteristics and wetland/non-wetland status were recorded and in the case of wetlands, an approximate water depth and status was recorded, if present. A total of 6351 points were collected for five areas within the Cuvette Centrale. The areas (A through E shown in Fig. 1), cover 25 km^2, 11 km^2, 5 km^2, 20 km^2, and 17 km^2, respectively. A total of 3970 points were collected for area A during a 2 year forest inventory program from 2004 to 2005. For area B, 1143 points were collected from 2005 to 2007. Areas C and D were visited in November and December in 2005 and 389 points measured. The forest inventory data for area E consisted of 849 points collected in 2006. The easternmost and the northernmost validation sites (D, and E) are located in areas dominated by *terra firma* forests while the western sites (A and B) are located in the Lake Tele–Lake Tumba Landscape where wetlands constitute a majority of the land surface. All *in situ* sites were within the forest domain and did not include inundated grassland sites.

2.5.2. Map inter-comparison

In addition to the validation, a product inter-comparison with existing maps was undertaken, comparing the CARPE wetland map with the Africover vegetation map of the Democratic Republic of the Congo (FAO, 2005) and the JRC/GRFM Regional Flooded Forest Map of Central Africa (De Grandi et al., 1998), both of which included wetlands as thematic classes. The Africover vegetation map was derived by on screen drawing by local experts using Landsat TM and ETM+ data (FAO, 2005).

The JRC/GRFM map was derived from the JERS-1 GRFM project and the CAMP (Central Africa Mosaic Project) mosaics using statistical characterization of the radar signal (De Grandi et al., 1998). The characterization of the flooded forest was based on the postulate that since the flooded forest crown coverage appears more homogeneous, there would be a small (nearly constant) modulation of the local incidence angle, and therefore the radar cross section (RCS) or the shape (order) parameter (υ). These parameters will fluctuate over the *terra firma* primary forest suggesting an irregular crown canopy structure. Therefore, the shape parameter (υ-estimated using the method of moments) of a *K*-distribution function should be sensitive to the structural properties of the higher canopy layer and be able to separate the flooded forest from the *terra firma* lowland rain forest (De Grandi et al., 1998). The thematic goal of the JRC/GRFM map was to map the extent of the swamp forest and to monitor inundation conditions at two dates. Therefore, the map did not characterize wetlands outside the forest domain.

3. Results

3.1. Classification tree summary statistics

The classification tree algorithm was applied to the systematically derived training data from the three CARPE landscapes. The systematic sample was the initial training data set and revealed the percentage of wetlands within the area of the three landscapes. However, visual evaluation of the initial map product led to the addition of training pixels gathered in a non-systematic fashion in order to correct obvious errors. A second algorithm run led to the final wetland characterization. The overall proportion of wetlands within the three landscapes in the final map equaled the proportion of the systematic training sample, assuring that no bias was introduced in the regional depiction by the addition of new interpreter-driven training data selection.

A total of 9,533,392 training pixels were used for the final algorithm run, of which 5,274,826 were wetland training data (55%). These data were used in the bagging procedure. The thirty bagged classification trees explained on average 70% of the total root node deviance. The tree models contained an average of 670 terminal nodes per tree and had a mean classification accuracy of approximately 87%. In most cases, all the independent variables were used during the classification process, with each contributing different levels of deviance reduction. Table 1 lists the independent variables in order of decreasing contribution to deviance reduction in the suite of bagged trees. From Table 1, it is clear that reasonable classification results could be achieved with a greatly reduced list of input variables. For example, over 90% of the total deviance is explained by the first nine variables of Table 1. The first split in all 30 trees employed the relative elevation derived from a 20,000 grid support area threshold. This variable alone accounted for 46.2% of the total explained deviance. Landsat band 5 from the 2000 epoch mosaic was the second most important variable, accounting for 13.36% of the overall deviance reduction. The radar high water layer was third and explained 8.7% of the deviance.

The six relative elevation variables were important features for the classification, together accounting for 56.8% of the total explained deviance. This illustrates the relative importance of topographic position and landform for wetland occurrence. Of these relative elevation layers used in the model, the relative elevation derived using 15,000 and 10,000 grid SA contributed less to the classification tree showing perhaps a redundancy with the relative elevation derived from the 20,000 grid SA as the slope is averaged over long distances in all the three cases. The second order derivatives of the DEM together explained less than 1% (0.80%) of the deviance, and this is perhaps indicative of sensitivity to DEM errors that increases with higher order differentiation operators. The Landsat spectral bands (for the two epochs and the composite) accounted for 28.4% of the total explained deviance and the radar data accounted for 9.24%.

3.2. CARPE wetland map

The per pixel wetland probability classification derived from the classification tree bagging procedure is shown in Fig. 2. In general, and as expected, wetlands occur with a greater likelihood closer to the major rivers and the floodplains. The wetland probability classification was thresholded to produce the final CARPE wetland map. The final map was generated by defining wetlands as pixels with probabilities greater than or equal to 0.5; this threshold was derived as it gave the same proportion of wetlands (42.6%) as the systematically sampled training data from the three CARPE landscapes. This illustrates that even though non-systematically collected training samples were added, this process did not bias the proportion of wetlands characterized in the final product. In total, 359,556 km^2 (31.79%) of the study area was classified as wetlands. Within the Cuvette Centrale, wetlands predominate, particularly in the Lac Tele–Lac Tumba CARPE landscape where they cover 207,467 km^2 (56%) of the landscape, along the border between the Democratic Republic of the Congo and the Republic of Congo (Fig. 2). Another large wetland block is located east of Lake Mai-Ndombe in the Democratic Republic of the Congo, in the basin between the Lotoi and Lukuru rivers. Fig. 3 illustrates the CARPE wetland map in detail with polygons delineating wetland and non-wetland classes overlain on a 74-km by 56-km area of the Landsat 2000 mosaic for a portion of the Maringa–Lopori–Wamba landscape

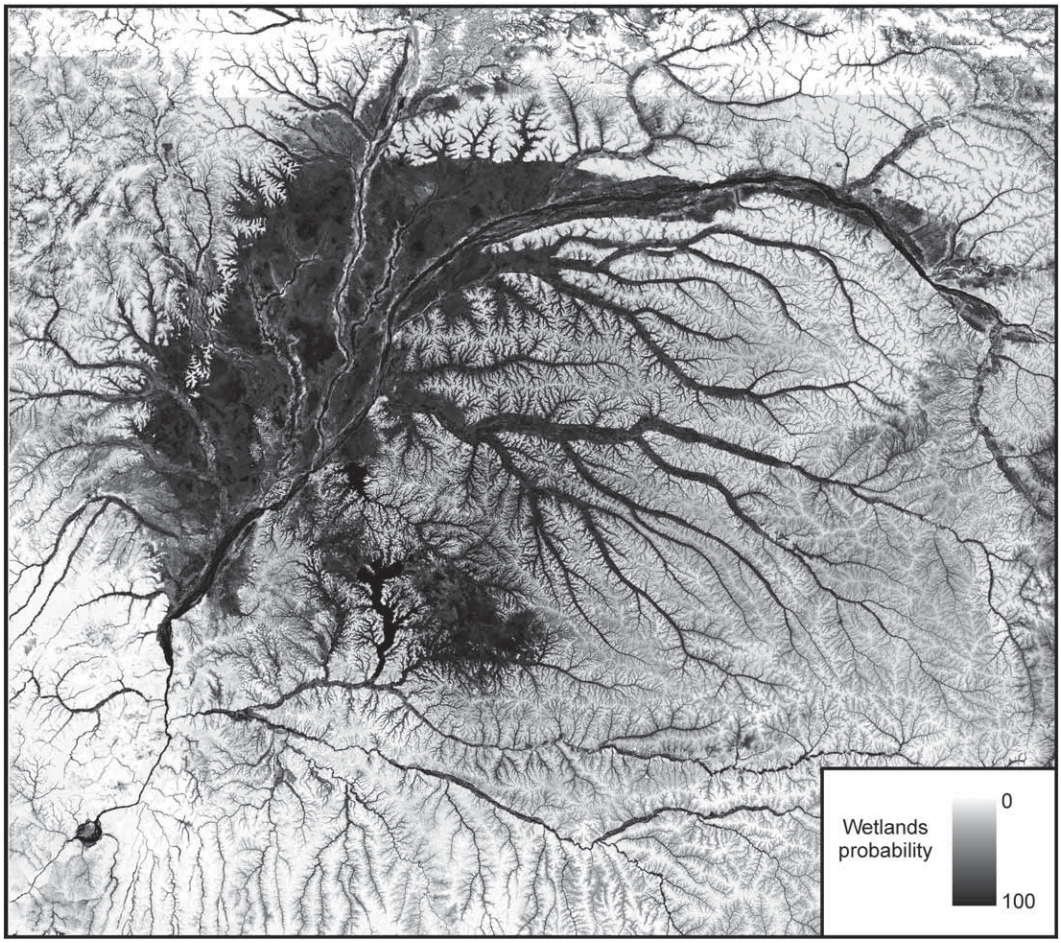

Fig. 2. Wetlands probability map (%) for the study area (1147 × 1025 km).

in the DRC. Landsat bands 4, 5 and 7 reveal the wetlands following drainage patterns and exhibiting a smoother spectral texture than the more structurally complex *terra firma* forests. Disturbance patterns associated with settlements and agricultural activities are brighter off-white in this color composite, occurring almost exclusively on the uplands, and in many cases delimiting the boundary between wetlands and *terra firma.*

3.3. Product evaluation

3.3.1. Validation

A contingency table was created showing agreement or disagreement between the field validation data and each of the maps described previously: the CARPE wetlands map, the Africover vegetation map of DRC and the JRC/GRFM Regional Flooded Forest map of Central Africa (JRC). Table 2 shows the results of these inter-comparisons.

3.3.1.1. CARPE wetland map validation. Out of 6361 in *situ reference* points, 722 were misclassified for an overall classification accuracy of 88.6% with user's and producer's accuracies of 63.3% and 92.0% in mapping wetlands and a Kappa coefficient of agreement of 0.680 (Table 2). Disagreement occurred in heterogeneous areas along wetland/*terra firma* interfaces or along small ribbons of wetlands that were incorrectly characterized by the mapping algorithm.

3.3.1.2. Africover validation. The Africover vegetation map of the Democratic Republic of the Congo included seven wetland vegetation classes (Table 3). These classes were merged and considered as wetlands and the other Africover classes were considered as non-wetlands. Classes 1 through 5 represent flooded forest classes and classes 6 through 8 inundated grassland classes. Africover non-wetland classes were grouped into dense humid forest, semi-deciduous forest, grass savannas, woody savanna, and secondary forest classes.

Comparison of the Africover vegetation map with *in situ* reference points (Table 2) revealed that a total of 4954 points were correctly classified both as wetlands and non-wetlands out of 6361points, yielding an overall classification accuracy of 77.8% and a Kappa coefficient of 0.2727. For the wetlands classes, 393 points were accurately classified out of a total 1713 points with user's and producer's accuracies of 81.9% and 22.9% respectively.

3.3.1.3. JRC/GRFM Regional Flooded Forest Map of Central Africa validation. The JRC/GRFM Regional Flooded Forest Map of Central Africa covered only a portion of the study site in north-western DRC and north-eastern Congo, including three of the five validation sites (B, C and D in Fig. 1) and only partially the fourth site (A; Fig. 1). Comparison of the JRC/GRFM map and *in* situ measurements indicated an overall agreement of 73.0% and a Kappa coefficient of 0.0817 (Table 2). For the wetland class, only 112 points were correctly classified out of 1563 *in situ* wetland points. The user's and producer's accuracies for mapping wetlands were 69.6% and 7.2% respectively.

3.3.2. Map inter-comparison

3.3.2.1. Comparison of the CARPE wetland and Africover maps. Fig. 4 shows the correspondence between the CARPE wetland and the Africover maps and Table 4 is the contingency matrix summarizing the map inter-comparisons.

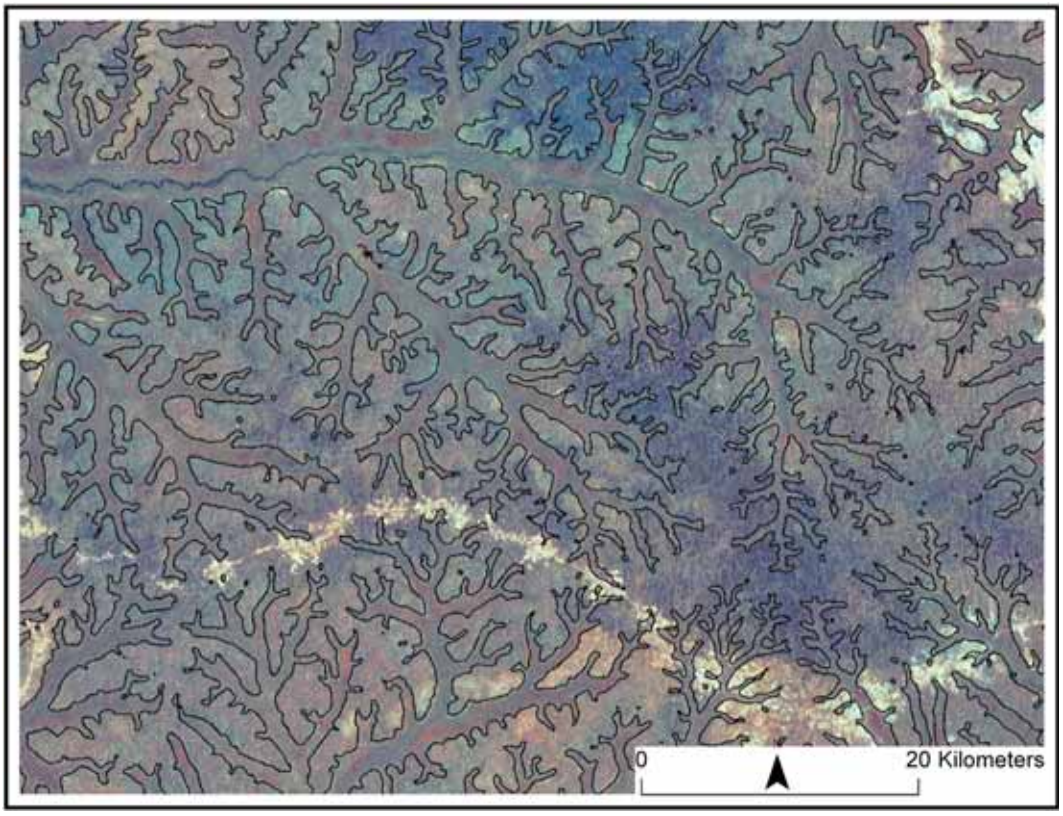

Fig. 3. A closer view (73 by 55 km) of the wetland mask, derived by thresholding the probability map at 50% (black lines), over a Landsat 4–5–7 image for a portion of the Maringa–Lopori–Wamba landscape in the Democratic Republic of the Congo, centered at 0.23N, 20.5E. Wetlands are located in the valley bottoms. Settlements and transportation corridors are found along the divides in the *terra firma* forest domain (brighter, nearly white areas).

The CARPE wetland map and the Africover vegetation map of DRC agreed in their wetland/non-wetland classification for 81.11% of the pixels with a Kappa coefficient of 0.509. The wetland user's and producer's accuracies were 51.13% and 80.97% respectively, reflecting a moderate agreement. The primary disagreement occurs in the wetlands forests of the Mankanza and Bomongo territories in the largely uninhabited core area of the Cuvette Centrale. This area was misclassified by Africover as dense semi-deciduous forest (Fig. 4A). Another source of disagreement occurs along ribbons of swamp forests adjacent to medium sized and small rivers in the Maringa–Lopori–Wamba CARPE landscape. These wetlands were omitted during the Africover visual interpretation process. In some areas, the CARPE map overestimated wetlands where gradual valley bottom topographies allow the wetlands to penetrate much farther into the upper reaches of watersheds within the Cuvette Centrale.

Table 5 summarizes the wetland and non-wetland areas mapped by CARPE with respect to the Africover vegetation classes and the degree of disagreement between the two maps. The CARPE map commission error (overestimation of wetlands) is 15.5% and the omission error (underestimation of wetlands) is 3.8%. The Africover semi-deciduous forest class was the source of greatest disagreement with the CARPE map, accounting for 58.1% of the total disagreement. Overestimation of

Table 2
Accuracy assessment results and comparison between the field validation data and the three maps.

	Map	Class	Truth (GPS points)	
			Non wetlands	Wetlands
Maps	JRC/GRFM	Non wetlands	3951	1451
		Wetlands	49	112
	Africover	Non wetlands	4561	1320
		Wetlands	87	393
	CARPE	Non wetlands	4554	628
		Wetlands	94	1085

Agreement						
Accuracy	JRC/GRFM		Africover		CARPE	
	Wetlands	Non-wetlands	Wetlands	Non-wetlands	Wetlands	Non-wetlands
Producer's	7.2	98.8	22.9	98.1	63.3	98.0
User's	69.6	73.1	81.9	77.5	92.0	87.9
Overall	73.0		77.9		88.6	
Kappa	0.0817		0.2727		0.680	

Table 3
Africover vegetation classes in the Cuvette Centrale.

Wetland classes
1 Broadleaved evergreen high forest with high shrubs on permanently flooded land
2 Broadleaved evergreen high forest on permanently flooded land
3 Broadleaved evergreen forest on permanently flooded land
4 Broadleaved evergreen high forest on temporarily flooded land
5 Broadleaved evergreen medium high woodland with high shrubs on permanently flooded land
6 Closed medium to high shrubs on permanently flooded land
7 Closed to open herbaceous vegetation on permanently flooded land
8 Open medium to high shrubs with medium to tall herbaceous vegetation on permanently flooded land
Non-wetland merged classes in the Cuvette Centrale
Cultivated and managed terrestrial area(s) (secondary forest/agricultural plantations)
Natural and semi-natural primarily terrestrial vegetation
- Open to closed herbaceous vegetation (grass savanna)
- Closed to very open herbaceous vegetation with trees and shrubs (woody savanna)
- Broadleaved semi-deciduous forest, continuous closed low forest, semi-evergreen high forest with shrubs (semi-deciduous forest)
- Broadleaved evergreen high forest with emergents and/or with shrubs (dense humid forest)

wetlands also occurred for the dense humid forest, the savanna and the secondary forest classes. Underestimation of wetlands was 18.9% for forested wetlands, while almost negligible for inundated grassland (Table 5).

3.3.2.2. Comparison of CARPE wetland and JRC/GRFM maps. Fig. 5 show the correspondence between the CARPE and JRC/GRFM map products. Both maps capture the core wetland areas, with more wetlands being mapped by CARPE adjacent to the core areas and farther into the system of Congo River tributaries within the Cuvette Centrale (Fig. 5A and B). An apparent disagreement is found in the inundated grasslands of the Sangha and the Ngiri Rivers (middle-left, Fig. 5) where the CARPE map depicts non-forested wetlands and the JRC/GRFM maps non-forest. This is not a true disagreement, but an artifact of the differing wetland definitions, as the JRC/GRFM map depicts only flooded conditions for forest, but not for grasslands. The results displayed in Fig. 5 indicate a relative underestimation of flooded forest extent by the JRC/GRFM map compared to the CARPE map. Comparing the two data sets quantitatively is problematic for this class as the JRC/GRFM map depicts only forested wetlands. Therefore, non-forested wetlands were not considered in the comparison.

The maps have an overall correspondence of 72.1% or 333,000 out of a total 462,000 km^2; with user's and producer's accuracies of 44.4% and 91.1% respectively in mapping forested wetlands, and a

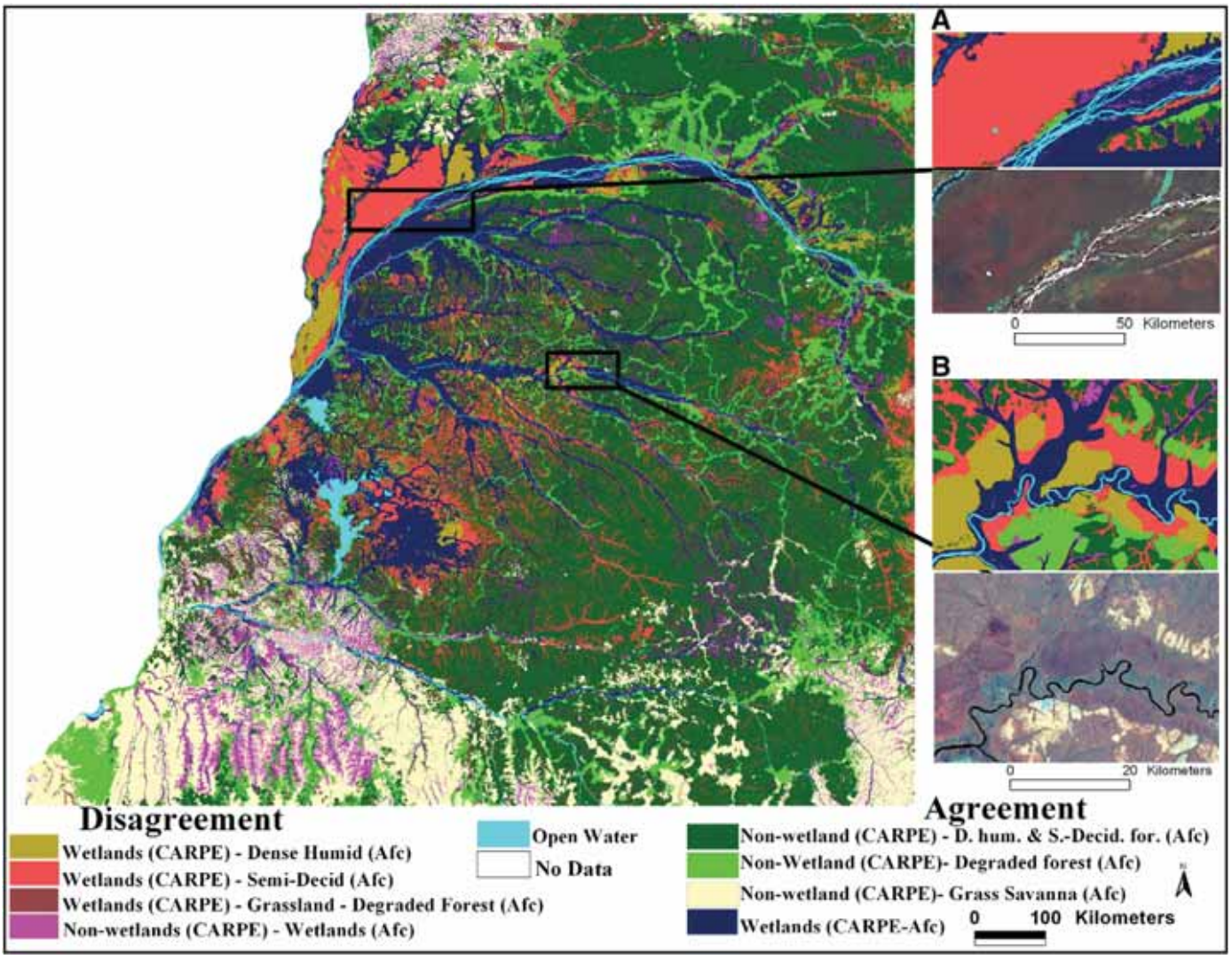

Fig. 4. Bivariate map for the CARPE wetland and Africover vegetation of the Democratic Republic of the Congo (Af) maps. Inset A shows a closer view of the Bomongo–Makanza swamps and B shows a portion of the Tshuapa River near Boende.

Table 4
Contingency table (top 4 rows) and summary statistics (bottom 5 rows) for map comparisons.

CARPE		Africover		JRC/GRFM		JRC/GRFM	Africover	
		Wetlands (km^2)	Non wetlands (km^2)	Wetlands (km^2)	Non wetlands (km^2)		Wetlands (km^2)	Non wetlands (km^2)
	Wetlands	143,404	137,086	96,000	120,000		27,900	31,700
	Non wetlands	33,701	589,875	9,000	237,000		39,900	113,600
	Accuracy	Wetlands (%)	Non-wetlands (%)	Wetlands (%)	Non-wetlands (%)		Wetlands (%)	Non wetlands (%)
	User's	51.13	94.6	44.4	96.2		41.1	78.2
	Producer's	80.97	81.14	91.1	66.3		46.8	74
	Overall		81.11		71.9			66.4
	Kappa coefficient		0.509		0.419			0.1995

Kappa coefficient of 0.419 (Table 4). The CARPE wetland map depicts nearly twice the wetland area as the JRC/GRFM map.

3.3.2.3. Comparison of Africover and the JRC/GRFM maps. Fig. 6 shows the correspondence between the JRC/GRFM and Africover maps. In general, they agree in floodplains along most major rivers, although there is disagreement along lesser tributaries of small and medium-sized rivers of the Maringa–Lopori–Wamba and Lake Tumba Landscapes, where the JRC/GRFM map does not depict wetlands. Significant disagreement occurs in the core wetland area of the Cuvette Centrale in the Lake Tumba–Lake Tele landscape, specifically in the Mankanza and Bomongo territories where the JRC/GRFM map depicts wetlands and Africover does not. As discussed previously, this area was misclassified by Africover as non-wetlands. The two datasets have an overall correspondence of 66%, a 0.2 Kappa and, wetland user's and producer's accuracies of 41% and 47% respectively. There is more agreement on mapping non-wetlands with user's and producer's accuracies of 78% and 74% respectively (Table 4).

4. Discussion

4.1. Multi-source data

Multi-source data have complementary value in mapping the wetlands of the Congo Basin. Table 1 illustrates the relative importance of 1) topographic indices derived from SRTM inputs, 2) optical wavelength Landsat imagery and 3) radar amplitude information, in characterizing wetlands. The complementary characteristics of both optical, radar and elevation data enabled the improved detection of both forested and non-forested wetlands. These cover types are difficult to map using single source data sets due to the high heterogeneity and structural complexity within tropical ecosystems (Hess et al., 2003; Mayaux et al., 2002). The comparative strength of the multi-source approach is evident in the comparisons of the three evaluated map products with *in situ* measurements. The CARPE wetland map had the highest percent agreement and Kappa coefficient of the three intercomparisons. Elevation data proved to be the most important information source within the classification process, illustrating the importance of landform and topographical position in wetland occurrence and identification. In particular, relative elevation for various catchment support thresholds proved to be most valuable. Even with the subtle nature of topographic variation within the Cuvette Centrale, and the uncertainties of the SRTM elevations, wetlands were reliably characterized using topographic indices. Anomalies in the DEM were more likely to affect the classification process in areas where the flooded forest canopy tended to have a higher elevation than adjacent upland vegetation. This occurred when degraded forests, settlements or savannas were directly connected to the floodplain and appeared as depressions in the DEM and topographic position derivatives. However, with the optical imagery, these land covers were often easily recognizable. Thus, the complementary nature of the inputs led to an improved characterization. Although it is reported in the literature that SAR data are highly suited for mapping wetlands, the backscattering response from upland forests sometimes overlaps with that of wetlands, making it difficult to discriminate between these two classes (Hess et al., 2003). Moreover, the decision tree analysis tool results showed that the DEM derivatives and optical data alone contributed more than 90% to the overall result. This suggests that the wetlands characterization may be undertaken using only the topographic and optical inputs.

Table 5
Contingency table, by Africover vegetation class, summarizing the correspondence between the Africover and the CARPE maps.

Africover vegetation classes	CARPE		% of total disagreement
	Wetlands (km^2)	Non-wetlands (km^2)	
Forested wetlands	120,007	$32{,}213^{(a)}$	$(a/t)\times 100 = 18.86$
Inundated grassland	4044	$1487^{(b)}$	$(b/t)\times 100 = 0.87$
Dense humid forest	$18{,}627^{(c)}$	43,884	$(c/t)\times 100 = 10.91$
Savannas	$8663^{(d)}$	102,692	$(d/t)\times 100 = 5.07$
Semi-deciduous	$99{,}199^{(e)}$	335,061	$(e/t)\times 100 = 58.08$
Secondary forest	$10{,}598^{(f)}$	107,632	$(f/t)\times 100 = 6.21$
Column totals:	261,138	622,969	100.00
Total disagreement (t)	$(a)+(b)+(c)+(d)+(e)+(f) = 170{,}787$ (km^2)		

The disagreement between the two maps is computed assuming that the Africover map is correct; the total disagreement is $(a)+(b)+(c)+(d)+(e)+(f)$. The commission error is 15.51% ($((c+d+e+f)\times 100$/total area) and the omission error is 3.81% ($((a+b)\times 100$/total area).

The classification exercise relied solely on existing data spanning 15 years apart. We therefore recognize that such a difference in image acquisition dates can be critical in wetland mapping, specifically in relation to seasonal and phenological changes in vegetation patterns. The JRC/GRFM radar mosaics that were termed to capture high water condition in either side of the basin were particularly old (1996's acquisition period) and may not reflect the actual condition on the ground. However, due to lack of data in the tropics, these mosaics have been widely used for wetland characterization (Hess et al., 2003; Mayaux et al., 2002; Simard et al., 2002). Moreover, we believe that the Congo Basin forested wetlands have not undergone major changes over the last 20 years (Archard et al., 2001; Hansen et al., 2008a; Hansen et al., 2008b; Mayaux et al., 2000; Defourny et al., 2006) due to political instability in the region, inaccessibility and isolation. Improved fine-scale time-series data, whether microwave or optical will help resolve this problem.

4.2. Training data and map outputs

The use of a systematic grid to initiate training labeling provided the prior probabilities of wetland and non-wetland classes within the region. However, relying solely on the systematically-acquired training data did not result in a geographically robust map. This illustrates the

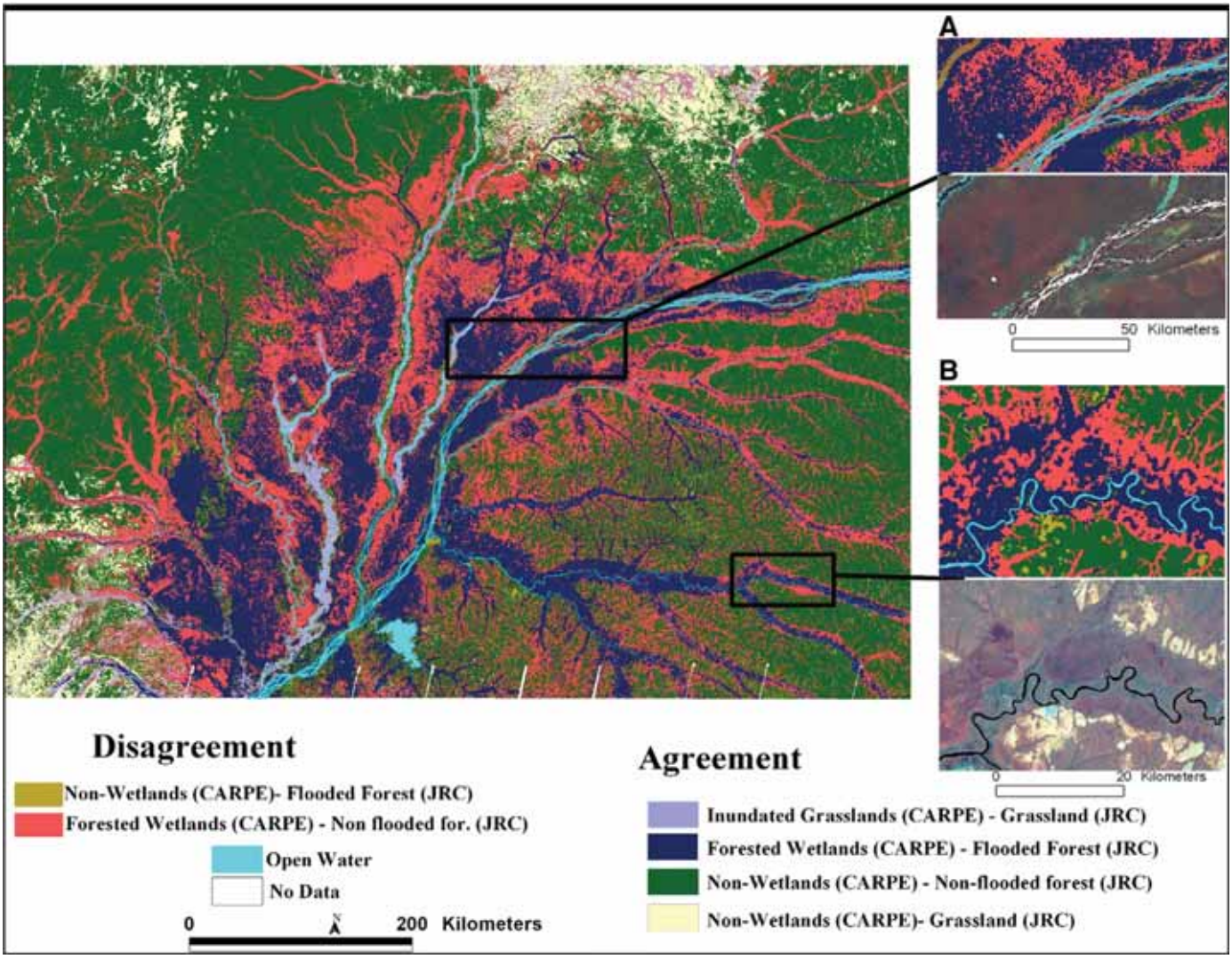

Fig. 5. Bivariate map for the CARPE Wetlands and the JRC/GRFM flooded forest of Central Africa maps. Inset A shows a closer view of the Bomongo–Makanza swamps and B shows a portion of the Tshuapa River near Boende.

tension between statistically valid estimates of cover and their geographic representation — a map may be 90% accurate, but if the 10% incorrectly labeled pixels are found within a geographically important region, the map's utility is quickly questioned by users. In order to avoid such problems, additional training data were obtained by evaluating the initial product derived solely from the systematically acquired training labels. Map interpretation of the initial product led to the addition of training labels to ensure a robust geographic characterization. The final training data set was considerably larger than the systematically sampled training data. In order to avoid biases to regional cover estimation introduced from the additional training, the output per pixel probabilities were used to match the final regional proportions of wetland/non wetland cover to the prior probabilities from the systematic grid. For this study, a threshold applied at 50% wetland probability yielded the same overall proportion of wetland land cover for the area where training data were systematically acquired.

4.3. Seasonality

Capturing seasonal variations in inundation is difficult as there is no high-temporal frequency imagery available at fine spatial scales over the Congo Basin. Arguably in this region the concept of seasonal high and low water is problematic as the Congo River crosses the equator twice, and the forest domain is nearly equally divided between the northern and southern hemisphere. Consequently, generalizing local inundation patterns with a regional mosaic, as attempted with the data used in this study, is difficult. Further, the seasonality of the vegetation leaf area is unknown and controversial at large scales for tropical forests (Myneni et al., 2007), and it remains unclear if any such non-wetland seasonality acted to increase or decrease the wetland discrimination.

The epochal CARPE Landsat composites are biased towards forest and against clouds and shadows without reference to within-year timing (Hansen et al., 2008b). However, from a photo-interpretation standpoint, both the radar and optical inputs have recognizable wetlands features. Combining these with the landform information derived from the SRTM data enabled the successful characterization of wetland land cover, as demonstrated by the *in situ* data comparisons.

5. Conclusion

Results from this study demonstrate the utility of multisource data in characterizing the wetlands of the Congo Basin. Passive optical, active microwave remotely sensed data and derived topographical indices were used as inputs, with topographical indices driving the overall characterization. This is of interest as the training data were

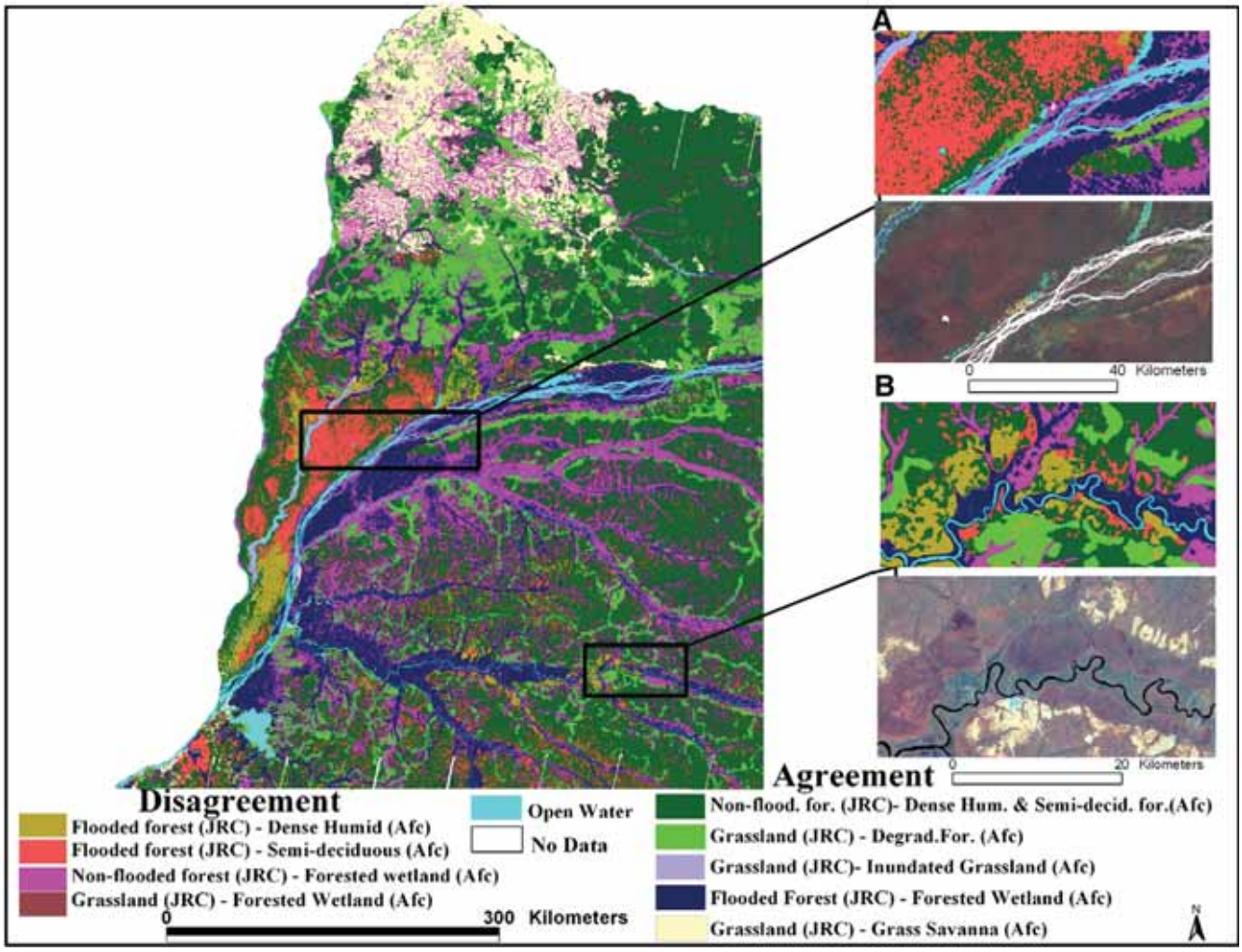

Fig. 6. Bivariate map of the JRC/GRFM flooded forest of Central Africa and Africover vegetation of the Democratic Republic of the Congo (Afc) maps. Inset A shows a closer view of the Bomongo–Makanza swamps and B shows a portion of the Tshuapa River near Boende.

derived using photointerpretation methods on the Landsat mosaic inputs. While photointerpretation is a common approach to mapping wetlands, it has proven difficult to digitally classify wetlands using multi-spectral data sets. The fact that the algorithm employed a relative elevation as the primary discriminator instead of the trained-upon optical data emphasizes that landscape-scale structure and not per pixel spectral information is most important in wetland characterization. However, each information source contributed to the overall characterization, with Landsat band 5 and radar high water amplitude ranked as the second and third most important inputs to mapping Congo Basin wetlands.

The result is the most spatially detailed Congo Basin wetland map to date. Comparison with *in situ* forest inventory measurements of wetland/non-wetland status indicate a more robust depiction than past pathfinding approaches that relied on single-source methods. All data sets used as inputs are freely available and the process is automated except for training data collection which is labor intensive and requires expert interpreter knowledge of the region's vegetation and landforms. Future work will include improving the thematic characterization of wetlands, including inundation duration and herbaceous versus forested wetlands. Improved depictions of the vegetation/hydrological/soil ecotones between terrestrial and aquatic systems will have more utility for value-added scientific applications. However, more intensive *in situ* measurements would be required to evaluate such characterizations.

References

Archard, F., Eva, H., & Mayaux, P. (2001). Tropical forest mapping from coarse resolution data: Production and accuracy assessment issues. *International Journal of Remote Sensing*, *22*(14), 2741−2762.

Amarsaikhan, D., & Douglas, T. (2004). Data fusion and image classification. *International Journal of Remote Sensing*, *25*(17), 3529−3539.

Asner, G. (2001). Cloud cover in Landsat observations of the Brazilian Amazon. *International Journal of Remote Sensing*, *18*, 3855−3862.

Balmford, A., Bennun, L., Brink, B., Cooper, D., Côté, C. M., Crane, P., et al. (2005). The convention on biological diversity's 2010 target. *Science*, *307*(5707), 212−213.

Breiman, L., Olshen, J. H., & Stone, C. J. (1984). *Classification and regression trees.* New York: Chapman and Hall.

Briem, G. J., Benediktsson, J. A., & Sveinsson, J. R. (2002). Multiple classifiers applied to multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(10), 2291−2299.

Brown, C. G., Sarabandi, K., & Pierce, L. E. (2005). Validation of the shuttle radar topography mission height data. *IEEE Transaction in Geoscience and Remote Sensing*, *43*(8), 1707−1715.

Bultot F. (1974). Atlas climatique du bassin zaïrois. Quatrième partie: pression atmosphérique, vent en surface et en altitude, température et humidité de l'air en altitude, nébulosité et visibilité, propriétés chimiques de l'air et des précipitations et classifications cl. Brussels: I.N.E.A.C: 193 maps.

Campbell, D. (2005). The Congo River basin. In L. Fraser (Ed.), *The world largest wetlands: Ecology and conservation* (pp. 149−165). Cambridge: Cambridge University Press.

CBFP. (2005). The forests of the Congo Basin: A preliminary assessment. Washington DC: The Congo Basin Forest Partnership. http://carpe.umd.edu/products/PDF_Files/FOC-B_AprelimAssess.pdf

Cohen, J (1968). Weighted Kappa: Nominal scale agreement with provision for scale disagreement or partial credit. *Psychological Bullentin*, *70*, 213−220.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely Sensed data. *Remote Sensing of Environment*, *37*, 35−46.

Coughanowr (1998). Wetlands of the humid tropics: Water related issues and problems of the Humid tropica and other warm humid regions. *IHP, Humid tropics program series 12* (pp. 47). Paris: UNESCO.

Cowardin, L.M., Carter, V., Golet, F.C., and LaRoe, E.T., (1979). Classification of wetlands and deep water habitat of the United States. Washington DC: US Department of the Interior, Fish and Wildlife Service, Biological Services, FWS/OBS-79/31.

Davis, T. J. (1994). *The Ramsar Convention manual.*A guide to the Convention of Wetlands of International Importance especially as waterfowl habitat Gland, Switzerland: Ramsar Convention Bureau.

Defourny, P., Vancutsem, C., Bicheron, P, Brockmann, C., Nino, F., Schouten, L., et al., 2006. GLOBCPVER: a 300m global land cover product for 2005 using ENVISAT MERIS Time Series. *Proceedings of ISPRS Commission VII Mid-Term Symposium: Remote Sensing: from Pixels to Processes*, Enschede (NL), 8–11 May, 2006.

De Grandi, G. F., Mayaux, P., Rauste, Y., Rosenqvist, A., Simard, M., & Saatchi, S. (2000). The Global Rain Forest Mapping Project JERS-1 radar mosaic of tropical Africa: Development and product characterization aspects. *IEEE Transaction on Geosciences and Remote Sensing*, *38*, 2218—2233.

De Grandi, F., Mayaux, P., Rosenqvist, A., Rauste, Y., Saatchi, S., Simard, M., et al., (1998). Flooded forest mapping at regional scale in the Central Africa Congo River Basin: First thematic results derived by ERS-1 and JERS-1 radar mosaics. Retrieval of Bio- and geo-physical parameters from SAR data for Land Applications Workshop: 21–23 October 1998. Noordwijk.

Deng, Y. (2007). New trend in digital terrain analysis: Landform definition, representation and classification. *Progress in physical geography*, *31*(4), 405—419.

Devred, R. (1958). La végétation forestière du Congo Belge et du Rwanda-Urundi. *Bulletin de la Société Royale Forestière de Belgique*, *65*, 409—468.

FAO. (2005). Land cover classification system (LCCS). Classification concepts and user manual. Software version 2.0. Rome: A. Di Gregorio, Food and Agriculture Organization of the United Nations (FAO).

Foody, G (2004). Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. *Photogrammetric Engineering and Remote Sensing*, *70*(5), 627—633.

Gauci, V., Matthews, E., Dise, N., Walter, B., Koch, D., Granberg, G., et al. (2004). Sulfate suppression of the wetland methane source in the 20th and 21st centuries. *Proceedings of the National Academy of Sciences*, *101*, 12583—12587.

Hamilton, S. K., Kellndorfer, J., Lehner, B., & Tobler, M. (2007). Remote sensing of floodplain as a surrogate for biodiversity in a tropical river system (Madre de Dios, Peru). *Geomorphology*, *89*, 23—38.

Hansen, M. C., & Defries, R. S. (2004). Detecting long term global change using continuous fields of trees-cover maps from 8-km advanced very high resolution radiometer (AVHRR) data for the year 1982–99. *Ecosystems*, *7*, 695—716.

Hansen, M. C., Defries, R. S., Townsend, R. R. G., Carroll, M., Dimicelli, C., & Sohlberg, R. A. (2003). Global percent of trees cover at a spatial resolution of 500 meters: First result of the MODIS vegetation continuous fields algorithm. *Earth Interactions*, *7*, 695—716.

Hansen, M., Dubayah, R., & DeFries, R. (1996). Classification trees: An alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, *17*, 1075—1081.

Hansen, M. C., Stehman, S. V., Potapov, P. V., Loveland, T. R., Townshend, J. R. G., & DeFries, R. S. (2008). Humid tropical forest clearing from 2000 to 2005 quantified using multi-temporal and multi-resolution remotely sensed data. *Proceedings of the National Academy of Sciences*, *105*, 9439—9444.

Hansen, M. C., Roy, D. P., Lindquist, E., Adusei, B., Justice, C. O., & Alstatt, 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.

Hennig, T. A., Kretsch, J. L., Pessagno, C. J., Salamonowicz, P. H., & Stein, W. L. (2001). The shuttle radar topography mission in. *Proceedings of First International Symposium, DEM, Lectures Notes in Computer Science, September 5–7, 2001* (pp. 65—77). Manno, Switzerland: Digital Earth Moving.

Hess, L. L., Melack, J. M., Filoso, S., & Wang, Y. (1995). Delineation of inundated area and vegetation along the Amazon floodplain with the SIR-C synthetic aperture radar. *IEEE Transaction on Geoscience and Remote Sensing*, *33*, 896—904.

Hess, L. L., Melack, J. M., Novo, E. M. L. M., Barbosa, C. C. F., & Gastil, M. (2003). Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. *Remote Sensing of Environment*, *87*, 404—428.

Hoekman, D. H., & Quinones, M. J. (2002). Biophysical forest type characterization in the Colombian Amazon by airborne polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(6), 1288—1300.

INPE (2002). *Deforestation estimates in the Brazilian Amazon.* São José dos Campo: Available at: http://www.obt.inpe.br/prodes/: INPE.

Ju, J., & Roy, D. P. (2008). The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. *Remote Sensing of Environment*, *112*, 1196—1211.

Junk, W. J., & Piedade, M. T. F. (2005). The Amazon River basin. In F. L. Keddy (Ed.), *The world largest wetlands: Ecology and conservation* (pp. 63—117). Cambridge: Cambridge University Press.

Laporte, N., Goetz, S. J., Justice, C. O., & Heinicke, M. (1998). A new land cover map of Central Africa derived from multiresolution, multitemporal AVHRR data. *International Journal of Remote Sensing*, *19*(18), 3537—3550.

Le Hegart-Mascle, S., Richard, D., & Ottle, C. (2003). Multi-scale data fusion using Dempster–Shafer evidence theory. *Integrated Computer Aided Engineering*, *10*, 9—22.

Lebrun, J., & Gilbert, G. (1954). Une classification écologique des forêts du Congo. *INEAC, Serie Scientifique N° 63* (pp. 44). Brussels: INEAC.

Lindquist, E., Hansen, H., Roy, D. P., & Justice, C. O. (2008). The suitability of decadal image data sets for mapping tropical forest cover change in the Democratic Republic of Congo: Implications for the global land survey. *International Journal of Remote Sensing*, *29*(24), 7269—7275.

MacMillan, R. A., Pettapiece, W. W., Nolan, S. C., & Goddard, T. W. (2000). A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy sets and Systems*, *113*, 81—109.

Matthews, E. (2000). Wetlands. In M. A. K. Khalil (Ed.), *Atmospheric methane: Its role in the global environment* Berlin: Spinger-Verlag.

Matthews, E., & Fung, I. (1987). Methane emission from natural wetlands: Global distribution, area and environmental characteristics of sources. *Global Biogeochemical Cycles*, *1*(1), 61—86.

Mayaux, P., Bartholome, E., Fritz, S., & Belward, A. (2004). A new land cover of Africa for the year 2000. *Journal of Biogeography*, *31*(6), 861—877.

Mayaux, P., De Grandi, G., & Malingreau, J. P. (2000). Central African forest cover revisited: A multisatellite analysis. *Remote Sensing of Environment*, *71*, 183—196.

Mayaux, P., De Grandi, G., RAUSTE, Y., Simard, M., & Saatchi, S. (2002). Large-scale vegetation maps derived fro the combined L-band GRFM and C-band CAMP wide area radar mosaics of Central Africa. *International Journal of Remote Sensing*, *23*(7), 1261—1282.

Melack, J. M. (2004). Remote sensing of tropical wetlands. In S. Ustin (Ed.), *Manual of remote sensing, 3rd edition Remote Sensing for Natural Resources Management and Environmental Monitoring*, *Vol. 4.* (pp. 319—343) New York: John Wiley & Sons.

Michaelson, J., Schimel, D. S., Friedl, M. A., Davis, F. W., & Dubayah, R. O. (1994). Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. *Journal of Vegetation Science*, *5*, 673—696.

Middleton, B. A. (1999). Flood pulsing in restoration: a feasible alternative for India? *Journal of the Ecological Society*, *12*, 10—14.

Myneni, R. B., Yang, W., Nemani, R. R., Huete, A. R., Dickinson, R. E., Knyazikhin, Y., et al. (2007). Large seasonal swings in leaf area of Amazon rainforests. *Proceedings of the National Academy of Sciences*, *104*(12), 4820—4823.

Naiman, R. J., Fetherston, K. L., McKay, S., & Chen, J. (1998). Riparian forests. In N. a. Bilby (Ed.), *River ecology and management: Lessons from the Pacific coastal ecoregion* New York: Spinger-Verlag.

Neigh, C. S. R., Tucker, C. J., & Townsend, J. R. G. (2007). North American vegetation dynamics observed with multi-resolution satellite data. *Remote Sensing of Environment*, *112*, 1749—1772.

O'Callagan, J. F., & Mark, D. M. (1984). The extraction of drainage networks from digital elevation data. *Computer Vision Graphic Image Processing*, *28*, 328—344.

Ouaidrari, H., & Vermote, E. F. (1999). Operational atmospheric correction of Landsat TM data. *Remote Sensing of Environment*, *70*, 4—15.

Ozesmi, S. L., & Bauer, M. E., 2002. Satellite remote sensing of wetlands. *Wetlands Ecology and Management, 10*, 381–402.

Rabus, B., Eineder, M., Roth, A., & Bamler, R. (2003). The shuttle radar topography mission— A new class of digital elevation models acquired by spaceborne radar. *Photogrammetric Engineering and Remote Sensing*, *57*, 241—262.

Rosenqvist, A., & Birkket, C. M. (2002). Evaluation of JERS-1 SAR mosaics for hydrological applications in the Congo River Basin. *International Journal of Remote Sensing*, *23*(7), 1283—1302.

Roy, D. P. (1997). Investigation of the maximum normalized difference vegetation index (NDVI) and the maximum surface temperature (Ts) AVHRR compositing procedures for the extraction of NDVI and Ts over forest. *International Journal of Remote Sensing*, *18* (11), 2383—2401.

Roy, D. P., Lewis, P., Schaaf, C., Devadiga, S., & Boschetti, L. (2006). The global impact of cloud on the production of MODIS bi-directional reflectance model based composites for terrestrial monitoring. *IEEE Geoscience and Remote Sensing Letters*, *3*, 452—454.

Saalovara, K., Thessler, S., Malik, R. N., & Tuomisto, H. (2005). Classification of Amazonian primary rain forest vegetation using Landsat ETM+ satellite imagery. *Remote Sensing of Environment*, *97*, 39—51.

Sahagian, D. and Melack, J.M., (1998). Global wetland distribution and functional characterization: Trace gases and the hydrologic cycle. Washington DC: IGBP Report 46, 92 p.

Saleska, S. R., Didan, K., Huete, A. R., & da Rocha, H. R. (2007). Amazon forests green-up during 2005 drought. *Science*, *318*(5850), 318—612.

Samba, G., Nganga, D., & Mpounza, M. (2008). Rainfall and temperature variations over Congo-Brazzaville between 1950 and 1998. *Theoretical and Applied Climatology*, *91*, 85—97.

Sankaran, M., Hanan, N. P., Scholes, R. J., Ratnam, J., Augustine, D. J., Cade, B. S., et al. (2005). Determinants of woody cover in African Savannas. *Nature*, *438*, 846—849.

Segers, R. (1998). Methane production and methane consumption: A review of processes underlying wetland methane fluxes. *Biogeochemistry*, *27*, 35—60.

Shimabukuro, Y. E., Batista, G. T., Mello, E. M. K., Moreira, J. C., & Duarte, V. (1998). Using shade fraction image segmentation to evaluate deforestation in Landsat thematic mapper images of the Amazon region. *International Journal of Remote Sensing*, *19*, 535—541.

Simard, M., De Grandi, J. F., Saatchi, S., & Mayaux, P. (2002). Mapping tropical coastal vegetation using JERS-1 radar data with a decision tree classifier. *International Journal of Remote Sensing*, *23*(7), 1461—1474.

Simard, M., Saatchi, S., & De Grandi, G. F. (2000). The use of decision tree and multiscale texture for classification of JERS-1 SAR data over tropical forest. *IEEE Transaction on Geoscience and Remote Sensing*, *38*(5), 2310—2321.

Stehman, S. V. (1997). Estimating standard errors of accuracy assessment statistics under cluster sampling. *Remote Sensing of Environment*, *60*, 258—269.

Stehman, S. V. (2006). Design, analysis, and inference for studies comparing thematic accuracy of classified remotely sensed data: A special case of map comparison. *Journal of Geographical Systems*, *8*, 209—226.

Sun, G., (1990). Radar backscattering modeling of coniferous forest canopies. Santa Barbara: Ph.D. dissertation. The University of California at Santa Barbara.

Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, *33*(2), 300—319.

Tarboton, D. G., & Ames, D. P. (2001). Advances in mapping of flow networks from digital elevation data. *World Water and Environmental Resources Congress, May24. Orlando, Florida, USA.*

Tarboton, D. G., Bras, R. L., & Rodriguez-Iturbe, I. (1992). A physical basis for drainage density. *Geomorphology*, *5*, 59−76.

Tarboton, D. G., Bras, R. L., & Rodriguez-Iturbe, I. (1989). Scaling and elevation in river networks. *Water Resources Research*, *25*(9), 2037−2051.

Tarboton, D. G., Bras, R. L., & Rodriguez-Iturbe, I. (1991). On the extraction of channel networks from digital elevation data. *Hydrologic Processes*, *5*(1), 81−100.

Titus, K., Mosher, J. A., & Williams, B. K. (1984). Chance-corrected classification for use in discriminant analysis: Ecological applications. *The American Midland Naturalist*, *111*(1), 1−7.

Townsend, P. (2002). Estimating forest structure in wetlands using multitemporal SAR. *Remote Sensing of Environment*, *79*, 288−304.

Van der Sanden, J. J., & Hoekman, D. H. (1999). Potential airborne radar to support the assessment of land cover in tropical rain forest environment. *Remote Sensing of Environment*, *68*, 26−40.

Verbyla, D. L. (1987). Classification trees: A new discrimination tool.*Canadian Journal of Forest Research*, *17* 1150,1152.

Wang, Y., Hess, L. L., Filoso, S., & Melack, J. M. (1995). Understanding the radar backscattering from flooded and non flooded forest: Results from canopy backscatter modeling. *Remote Sensing of Environment*, *54*, 324−332.

Williams, W. A., Jensen, M. E., Winne, C. J., & Redmond, R. L. (2000). An automated technique for delineating and characterizing valley-bottom settings. *Environmental Monitoring and ssessment*, *64*, 105−114.

Wolock, D. M., & McCabe, G. J. (1999). Estimates of runoff using water-balance and atmospheric general circulation. *Journal of the American Water Resources Association*, *35*, 1341−1350.

Wood, J., (1996). The geomorphological characterization of digital elevation models. Leicester, UK.: Ph. D. Thesis, University of Leicester, Department of Geography.

Zhou, Q., & Liu, X. (2002). Error assessment of grid-based flow routing algorithm used in hydrologic models. *International Journal of Geographic Information Science*, *16*(8), 819−842.