Remote Sensing of Environment 115 (2011) 3539–3553

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

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

Comparison and assessment of coarse resolution land cover maps for Northern Eurasia

Dirk Pflugmacher [a,*], Olga N. Krankina [a], Warren B. Cohen [b], Mark A. Friedl [c], Damien Sulla-Menashe [c], Robert E. Kennedy [a], Peder Nelson [a], Tatiana V. Loboda [d], Tobias Kuemmerle [e], Egor Dyukarev [f], Vladimir Elsakov [g], Viacheslav I. Kharuk [h]

[a] Department of Forest Ecosystems and Society, Oregon State University, 321 Richardson Hall, Corvallis, OR 97331, USA
[b] USDA Forest Service, Pacific Northwest Research Station, Forestry Sciences Laboratory, 3200 SW Jefferson Way, Corvallis, OR 97331, USA
[c] Department of Geography and Environment, Boston University, 675 Commonwealth Ave., 4th Floor, Boston, MA 02215, USA
[d] Department of Geography, University of Maryland, 2181 LeFrak Hall, College Park, MD 20742, USA
[e] Earth System Analysis, Potsdam Institute for Climate Impact Research (PIK), PO Box 60 12 03, Telegraphenberg A62, D-14412 Potsdam, Germany
[f] Institute of Monitoring of Climatic and Ecological Systems, Tomsk 634021, Russia
[g] Institute of Biology, Komi Science Center, Russian Academy of Sciences, Kommunisticheskaja st., 28, 167610 Syktyvkar, Russia
[h] V.N. Sukachev Institute of Forest, Krasnoyarsk, Russia

ARTICLE INFO

Article history:
Received 22 June 2010
Received in revised form 26 July 2011
Accepted 12 August 2011
Available online 28 September 2011

Keywords:
Eurasia
Land cover
Global
Validation
GLC-2000
GLOBCOVER
MODIS
LCCS

ABSTRACT

Information on land cover at global and continental scales is critical for addressing a range of ecological, socioeconomic and policy questions. Global land cover maps have evolved rapidly in the last decade, but efforts to evaluate map uncertainties have been limited, especially in remote areas like Northern Eurasia. Northern Eurasia comprises a particularly diverse region covering a wide range of climate zones and ecosystems: from arctic deserts, tundra, boreal forest, and wetlands, to semi-arid steppes and the deserts of Central Asia. In this study, we assessed four of the most recent global land cover datasets: GLC-2000, GLOBCOVER, and the MODIS Collection 4 and Collection 5 Land Cover Product using cross-comparison analyses and Landsat-based reference maps distributed throughout the region. A consistent comparison of these maps was challenging because of disparities in class definitions, thematic detail, and spatial resolution. We found that the choice of sampling unit significantly influenced accuracy estimates, which indicates that comparisons of reported global map accuracies might be misleading. To minimize classification ambiguities, we devised a generalized legend based on dominant life form types (LFT) (tree, shrub, and herbaceous vegetation, barren land and water). LFT served as a necessary common denominator in the analyzed map legends, but significantly decreased the thematic detail. We found significant differences in the spatial representation of LFT's between global maps with high spatial agreement (above 0.8) concentrated in the forest belt of Northern Eurasia and low agreement (below 0.5) concentrated in the northern taiga-tundra zone, and the southern dry lands. Total pixel-level agreement between global maps and six test sites was moderate to fair (overall agreement: 0.67–0.74, Kappa: 0.41–0.52) and increased by 0.09–0.45 when only homogenous land cover types were analyzed. Low map accuracies at our tundra test site confirmed regional disagreements and difficulties of current global maps in accurately mapping shrub and herbaceous vegetation types at the biome borders of Northern Eurasia. In comparison, tree dominated vegetation classes in the forest belt of the region were accurately mapped, but were slightly overestimated (10%–20%), in all maps. Low agreement of global maps in the northern and southern vegetation transition zones of Northern Eurasia is likely to have important implications for global change research, as those areas are vulnerable to both climate and socio-economic changes.

1. Introduction

Information on land cover at global and continental scales is critical for addressing a range of important science questions such as the effects of vegetation on the carbon cycle, surface energy, and water balance, and socioeconomic causes and consequences of land-use and land-cover change (Bonan et al., 2002; Running et al., 2004; Zhang et al., 2009; Foley et al., 2005). The need for accurate land-cover information is particularly acute in Northern Eurasia, which encompasses high diversity of ecosystems that range from arctic deserts to the steppes and deserts of Central Asia, and climates that encompass polar and boreal climates of Siberia, monsoon climate in the

* Corresponding author. Tel.: +1 541 750 7287; fax: +1 541 737 1393.
E-mail address: dirk.pflugmacher@oregonstate.edu (D. Pflugmacher).

0034-4257/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2011.08.016

east, and milder maritime climates in the Baltic Sea region. Because of this diversity, the area is a locus for climate change, and there is mounting evidence of significant recent changes in vegetation distribution, growing season duration, and patterns of snow cover and permafrost (e.g. Bulygina et al., 2010; Chapin et al., 2005; Randerson et al., 2006; Soja et al., 2007). The region also has unique and often poorly characterized land cover features, including vast expanses of larch (*Larix* spp.) forests, permafrost, wetlands, widespread disturbances (fire, harvest, pollution damage, insect damage), and drastic changes in land use following profound region-wide socioeconomic and institutional changes in the 1990s (e.g. Forbes et al., 2004; Groisman et al., 2009; Kuemmerle et al., 2008; Kuemmerle et al., 2009).

Since the mid 1990s substantial advances have been made toward the development of global vegetation and land cover datasets from moderate resolution satellite sensors. The first satellite-based global land cover maps were produced with data from the advanced high-resolution radiometer (AVHRR) (DeFries & Townshend, 1994; Hansen et al., 2000; Loveland et al., 2000). In 1998 and 1999, AVHRR was followed by VEGETATION-1 onboard the fourth Satellite Pour l'Observation de la Terre (SPOT) and the Moderate Resolution Imaging Spectroradiometer (MODIS), which allowed the concurrent development of two additional global land cover products: GLC-2000 (Bartholome & Belward, 2005) and the MODIS Global Land Cover Product (Friedl et al., 2002). More recently, two new global land cover datasets have been released: GLOBCOVER derived from Medium Resolution Imaging Spectrometer (MERIS) data (Arino et al., 2008) and the MODIS C5 Land Cover Product (Friedl et al., 2010). The most important change in the evolution of these datasets is the increase in spatial resolution from ~1 km (MODIS Collection 4 — hereafter MODIS C4 — and GLC-2000) to ~500 m (MODIS Collection 5, hereafter MODIS C5) and ~300 m (GLOBCOVER). Because many land cover features occur at a spatial resolution finer than 1 km (Gerlach et al., 2005; Krankina et al., 2008; Skinner & Luckman, 2004), the higher spatial resolution should improve the representation and accuracy of the GLOBCOVER and MODIS C5.

The availability of multiple, similarly structured land cover data sets provides the user community with choices, but for most users it is not clear which map suits their particular application best (Herold et al., 2008; Jung et al., 2006). Ultimately, the selection is often based on map legends rather than accuracy, in part because it is difficult to ascertain which map is the most accurate. Global accuracy estimates reported by map developers are very similar, but because of methodological differences used to perform the accuracy assessments, these estimates cannot be directly compared, e.g. GLC-2000: 68.6% (Mayaux et al., 2006), GLOBCOVER: 73.1% (Bicheron et al., 2008), MODIS Collection 3: 71.6% (MODIS land cover team, 2003), and MODIS Collection 5: 74.8% (Friedl et al., 2010).

While users often rely on overall measures of map accuracy to evaluate the quality of maps, map errors are rarely equally distributed (Strahler et al., 2006). Two maps can have the same overall accuracy, but a different spatial distribution of error. Studies that compared global land cover maps have found significant regional differences in spatial agreement (Fritz & See, 2008; Giri et al., 2005; Herold et al., 2008). Agreement tends to be lowest in regions with complex, heterogeneous land cover and for spectrally similar land cover classes (e.g. mixed versus pure broadleaf and conifer forests). Thus, the choice of map depends on how it will be used in a particular region of the map, which also needs to be considered in interpreting results of spatial modeling.

Differences among land cover maps have important implications for applications using these products, for example biogeochemical (e.g. Potter et al., 2008) or habitat models (e.g. Kuemmerle et al., 2011). A simple analysis that extrapolated results of biogeochemical modeling for the Arctic region of Northern Eurasia (North of 60°) showed that a very different picture of the regional carbon (C) balance emerged when different vegetation maps were used as model inputs: The estimate of C stock in live vegetation based on the GLC-2000 map (24 Pg C) was 40% higher than the estimate based on the MODIS plant functional type map (17 Pg C). Although the estimates of the total change in live vegetation C stocks were very similar for both maps (0.2 Pg yr^{-1} C sink), the attribution of the projected C sink was quite different depending on the map used: based on GLC-2000 map most of the C accumulation occurred in tree-dominated ecosystems while simulations using the MODIS map attributed most of the C sink to shrub vegetation (Krankina et al., 2011). The significant role of land cover map selection on forest biomass estimates in Russian forests was also reported by Houghton et al. (2007).

Despite the importance of Northern Eurasia for global change research, global maps have not been rigorously assessed in this region. For GLC-2000, Bartalev et al. (2003) compared estimates of percent forest cover with official forest cover statistics for administrative regions of the Russian Federation. The authors reported an R^2 of 0.93, indicating that forest cover was reliably mapped at the level of administrative units. The size of administrative divisions in Russia, however, varies considerably, i.e. from about 8×10^3 km^2 (Adygea) to 3×10^6 km^2 (Yakutia). Frey and Smith (2007) compared field observations in Western Siberia with two land cover and two wetland databases. Agreement between the field data and the two analyzed land cover data sets, the AVHRR Global Land Cover Characterization Database and the MODIS C3 Land Cover Product, was 22% and 11%, respectively. Other comparison studies have focused only on wetlands (Krankina et al., 2008; Pflugmacher et al., 2007).

Cross comparisons between global land cover maps help identify areas of potential high map uncertainty (Herold et al., 2008; See & Fritz, 2006), but independent validation studies are needed to reveal the sources of disagreement and provide local and regional scale estimates of classification accuracy. There are several challenges associated with this task:

- The collection of reference data is costly for large areas, particularly for remote regions such as Northern Eurasia. Consequently, it is crucial to design and implement validation methods that are not tailored to a single map product, i.e. a single classification system and spatial resolution, but that can accommodate a range of current and potentially future land cover maps (Olofsson et al., submitted for publication).
- Reference data is often collected at high spatial resolution and needs to be aggregated to the resolution of the coarse-scale map. The process of aggregation, however, can introduce biases towards dominant land cover types ("low resolution bias", Boschetti et al., 2004; Latifovic & Olthof, 2004; Moody & Woodcock, 1994).
- Maps differ with respect to spatial resolution, class definitions and thematic detail. This affects estimates of overall map accuracy (Latifovic & Olthof, 2004), and therefore makes cross-comparison of accuracy estimates difficult.

The objective of this study was to evaluate accuracy measures and procedures for validating global land cover datasets for Northern Eurasia using maps created from higher resolution satellite data (Landsat). This study is part of a broader effort to validate and improve land cover and land-cover change products for Northern Eurasia using a network of local test sites distributed throughout the region (http://www.fsl.orst.edu/nelda).

2. Methods

We compared four global land cover datasets with higher resolution (30-m) land cover maps developed from Landsat images at test sites distributed throughout the region of Northern Eurasia. Ideally, it is desirable to choose reference locations by random probability sampling (Strahler et al., 2006). However, the availability of reference data in this region represents a major constraint for validation studies. We therefore selected test sites across a range of climatic and

geographic conditions where high quality reference data and local expertise were available. This opportunistic selection of validation sites precludes a probability-based statistical assessment of map accuracy for the region, but permits different measures of map accuracy to be examined across a range of conditions within Northern Eurasia. A similar approach was used to validate MODIS Land Cover maps for selected biome types in North and South America (Cohen et al., 2003; Cohen et al., 2006; Latifovic & Olthof, 2004). For the purpose of this analysis we defined the region of Northern Eurasia to extend from 42°–74° N latitude (Fig. 1).

2.1. Global land cover data sources

The GLC-2000 effort was lead by the Joint Research Center of the European Commission in partnership with more than 30 institutions (Bartalev et al., 2003; Bartholome & Belward, 2005). The GLC-2000 map is based on Satellite Pour l'Observation de la Terre (SPOT) VEGETATION data acquired daily between November 1999 and December 2000. The map has a nominal spatial resolution of ~1 km and uses a 22-class legend based on the UN FAO's hierarchical Land Cover Classification System (LCCS, Di Gregorio, 2005). We downloaded the global data set (v1.1) from (http://www-gvm.jrc.it/glc2000/; last accessed 20 January 2008) in geographic coordinates, and projected it to an equal area projection (Lambert Azimuthal EA).

GLOBCOVER is the successor to GLC-2000, and as such also builds on an international network of partners. GLOBCOVER was developed using an annual mosaic of Envisat's Medium Resolution Imaging Spectrometer (MERIS) data from December 2004 to June 2006. GLOBCOVER relied in part on GLC-2000 as training data (Bicheron et al., 2008) and therefore the two datasets are not completely independent. With a nominal pixel size of 300 m, GLOBCOVER represents the highest spatial resolution global land cover dataset currently available. The map distinguishes between 22 classes that are compatible with the Land Cover Classification System (LCCS). We downloaded GLOBCOVER v2.2 from the European Space Agency GLOBCOVER Project (http://ionia1.esrin.esa.int, last accessed 4 May 2011).

The MODIS Land Cover Product is developed by scientists from Boston University for NASA's Earth Observing System (EOS) MODIS land science team (Friedl et al., 2002). The MODIS Land Cover Product is produced for each year since 2001 and provides global land cover in five different classification systems including the International Geosphere-Biosphere Programme (IGBP, Loveland & Belward, 1997) classification. The MODIS land science team periodically reprocesses the entire MODIS land product suite, including the yearly MODIS Land Cover Product. For this study we analyzed MODIS Land Cover Collection 4 data from 2001 and Collection 5 data from 2005 using the IGBP classification, hereafter referred to as MODIS C4 and MODIS C5, respectively. MODIS C5 has a spatial resolution of ~500 m and supersedes MODIS C4 with a spatial resolution of ~1 km. We included MODIS C4 here to explore differences between the two collections and because the 1-km resolution provided a suitable reference for comparisons with GLC-2000. The 17-class IGBP classification was selected because it is the most common and detailed land cover classification among the MODIS Land Cover layers. We obtained the data from the NASA Warehouse Inventory Search Tool (https://wist.echo.nasa.gov, last accessed 25 January 2010).

2.2. Differences among global land cover classifications

The global land cover classifications analyzed in this study are generally similar, but there are differences between maps that make a direct and uniform comparison challenging (Table 1). For example GLC-2000 differentiates between evergreen and deciduous shrub types, whereas MODIS IGBP does not distinguish between shrub leaf types, but does distinguish between open and closed shrub cover. Further, GLC-2000 and GLOBCOVER distinguish between non-vegetated, bare land cover and sparse vegetation with less than 10–20% and less than 15% vegetation cover, respectively. In comparison, MODIS IGBP combines these land cover types into a single class: barren and sparsely vegetated areas with less than 10% vegetation or snow cover. The maps also show different thematic information within tree life form classes. Generally, tree life form types are characterized by leaf type (needleleaf, broadleaf, and mixed) and leaf senescence (evergreen and deciduous). However, each map contains classes that cannot be unambiguously assigned to either of the two, e.g. the “Burnt Tree Cover”, “Regularly Flooded Tree Cover”, and “Mosaic Tree Cover” classes in GLC-2000, the “Open Needleleaved Deciduous or Evergreen Forest” class in GLOBCOVER, and the open tree vegetation cover classes in the IGBP layer of the MODIS land cover product (“Woody Savannas” and “Savannas”).

To minimize ambiguities in comparisons of accuracy statistics among global land cover datasets, map legends need to be converted to a common classification. For this paper, we generalized map legends to six classes on the basis of the dominant life form types (LFT) in each map legend (Table 1). Croplands were labeled as

Fig. 1. Locations of reference sites: 1 — Carpathians, 2 — St. Petersburg, 3 — Komi, 4 — Vasyugan, 5 — Priangarie, 6 — Chita.

Table 1
Aligning the legends of global maps: dominant life form type (LFT) and corresponding land cover classes from GLC-2000, GLOBCOVER & MODIS IGBP.

Dominant LFT	GLC-2000	GLOBCOVER	MODIS IGBP
Tree	[1] Tree cover; broadleaved; evergreen	[40] Broadleaved evergreen or semi-deciduous forest, closed to open	[1] Evergreen needleleaf forest
	[2] Tree cover; broadleaved; deciduous; closed	[50] Broadleaved deciduous forest, closed	[2] Evergreen broadleaf forest
	[3] Tree cover; broadleaved; deciduous; open	[60] Broadleaved deciduous forest/woodland, open	[3] Deciduous needleleaf forest
	[4] Tree cover; needle-leaved; evergreen	[70] Needleleaved evergreen forest, closed	[4] Deciduous broadleaf forest
	[5] Tree cover; needle-leaved; deciduous	[90] Needleleaved deciduous or evergreen forest, open	[5] Mixed forest
	[6] Tree cover; mixed leaf type	[100] Mixed broadleaved & needleleaved forest, closed-open	[8] Woody savannas
	[7] Tree cover; regularly flooded; fresh water	[160] Broadleaved forest regularly flooded, closed to open	[9] Savannas
	[8] Tree cover; regularly flooded; saline water		
	[9] Mosaic: Tree cover; other natural vegetation		
	[10] Tree cover; burnt		
Shrub	[11] Shrub cover; closed-open; evergreen	[130] Shrubland, closed to open	[6] Closed shrublands
	[12] Shrub cover; closed-open; deciduous		[7] Open shrublands
Herbaceous	[13] Herbaceous cover; closed-open	[11] Post-flooding or irrigated croplands (or aquatic)	[10] Grasslands
	[16] Cultivated and managed areas	[140] Herbaceous vegetation, closed to open	[12] Croplands
		[14] Rainfed croplands	
Barren	[14] Sparse herbaceous or sparse shrub cover vegetation cover between 1% and 10-20% vegetation cover	[150] Sparse vegetation (<15%)	[16] Barren or sparsely vegetated
	[22] Artificial surfaces and associated areas	[190] Artificial surfaces and associated areas	[13] Urban and built-up
	[19] Bare areas	[200] Bare areas	
	[21] Snow and ice	[220] Permanent snow and ice	[15] Snow and ice
Mosaic	[15] Regularly flooded shrub and-or herbaceous cover	[20] Mosaic cropland (50–70%)/vegetation (grassland/shrubland/forest) (20–50%)	[11] Permanent wetlands
	[17] Mosaic: Cropland; tree cover; other natural vegetation	[30] Mosaic vegetation (grassland/shrubland/forest) (50–70%)/cropland (20–50%)	[14] Cropland-natural vegetation mosaic
	[18] Mosaic: Cropland; shrub and-or herbaceous cover	[110] Mosaic forest or shrubland (50–70%)/grassland (20–50%)	
		[120] Mosaic grassland (50–70%)/forest or shrubland (20–50%)	
		[170] Broadleaved forest or shrubland permanently flooded, closed	
		[180] Grassland or woody vegetation on regularly flooded or waterlogged soil, closed-open	
Water	[20] Water bodies	[210] Water bodies	[0] Water bodies

herbaceous vegetation and classes that include mixtures of LFT types were labeled as “mosaic”. This level of generalization avoids most of the ambiguities described above, albeit at the cost of thematic detail. An alternative approach to bridging the differences among legends relies on “fuzzy logic” (See & Fritz, 2006) in contrast to “crisp logic” chosen for this study. We opted for crisp comparison to maintain objectivity by avoiding judging the severity of class disagreement.

2.3. Test sites

To assess the coarse resolution maps we selected six test sites that span wide latitudinal, longitudinal, and climatic gradients in Northern Eurasia (Table 2, Fig. 1). Mean annual temperature ranges from −5 °C at the Komi site to +6 °C at the Carpathian site and precipitation ranges from <400 mm per year in Chita to about 1000 mm in the Carpathians. While all sites have mild summers, the average winter temperature declines from west to east as the climate becomes increasingly continental. The terrain is flat at the St. Petersburg, Komi and Vasyugan site and mountainous with significant elevation gradients at the remaining three sites.

All sites are located within the forest belt that stretches across Northern Eurasia and vegetation cover in all sites includes a significant proportion of forest (tree-dominated land cover). Shrub cover is present in all sites as well, but its role tends to be small. The sites

Table 2
Test site information.

Site	Location	Latitude, longitude	Landsat path/row	Landsat image date	Mean temperature (°C)			Mean annual precipitation (mm)	Elevation range (m asl)
					January	July	Annual		
CARP	Carpathians (border region of Poland, Slovakia, Ukraine)	48.87, 22.40	186/20	6-Jun-2000 21-Aug-2000 30-Sep-2000	−7	19	6	900–1200	100–2000
STPB	St. Petersburg Region (Russia)	60.09, 31.27	184/18	2-Jun-2002	−10	17	4	600–800	0–250
KOMI	Komi Republic (Russia)	66.94, 57.10	171/13	1-Jun-2000 19-Jul-2000	−19	12	−5	360–500	20–240
VASY	Vasyugan Basin, Tomsk Region, West Siberia (Russia)	57.32, 82.09	150/20	16-Sep-1999	−18	18	0	450–550	110–130
PRIA	Priangarie, Krasnoyarsk Region, East Siberia (Russia)	57.30, 95.91	141/20	18-Aug-2000	−22	18	3	400–450	100–600
CHIT	Chita Region and Buriat Republic, East Siberia (Russia)	51.69, 111.64	129/24	11-Jun-2000	−26	15	−4	380	600–1500

represent different ecoregions including temperate forest (Carpathians), southern taiga or boreal forest (St. Petersburg and Vasyugan), northern taiga (Komi), and montane boreal (Priangarie and Chita) (Alexeyev & Birdsey, 1998). The sites also reflect the diversity of vegetation disturbance regimes in Northern Eurasia, including forest fires (Chita), insect outbreaks (Priangarie), timber harvest (Carpathians and St. Petersburg), and changes in agricultural land use (all sites except Komi). Wetlands occupy large areas at three test sites (St. Petersburg, Komi, and Vasyugan). Excessive moisture at these sites leads to accumulation of peat, reduced density of tree cover and greater shrubs and herbaceous vegetation cover.

Detailed land cover maps for each test site were developed based on Landsat imagery combined with local expert knowledge and ground data. Results of several completed and ongoing projects that mapped land cover and examined various types of land-cover, land-use change, and vegetation disturbance were also used (e.g., Heikkinen et al., 2004; Kharuk et al., 2003; Kharuk et al., 2004; Krankina et al., 2004a; Krankina et al., 2004b; Kuemmerle et al., 2007; Kuemmerle et al., 2008). The Landsat classification approaches varied by site and included supervised (e.g. decision trees) and unsupervised classifications. All site legends followed the Land Cover Classification System (LCCS, Di Gregorio, 2005) and included between 11 and 17 classes. These were recoded to match the LFT legend (Table 3). The LCCS classification defines tree-dominated cover as land with greater than 15% tree cover, where trees are any woody vegetation taller than 5 m or trees taller than 3 m. Similarly, shrub and herbaceous dominated vegetation types have a minimum of 15% plant cover, respectively. The dominant life form is defined hierarchically by the height of the canopy layer, which ranges from trees to shrubs to herbaceous plants. Lands with less than 15% vegetation cover are labeled as bare land and sparse vegetation.

Map accuracy at test sites was assessed using forest inventory data and ground observations from research projects. Where ground observations proved insufficient, additional reference data were collected through manual interpretation of aerial photos and high-resolution imagery in Google Earth. The overall accuracy of Landsat-derived maps for the LFT classification exceeded 92% for Carpathians, St. Petersburg, Vasyugan, and Priangarie sites; while it was somewhat lower for Chita and Komi sites (87% and 80%, respectively) (Table 3) (Krankina et al., 2011; Kuemmerle et al., 2006). For the purpose of this study the Landsat-based classifications are considered “ground truth”.

2.4. Comparison and assessment of global maps

We used multiple approaches to assess four global land cover datasets at the regional and local scale. First, we performed a cross comparison of global maps for the region of Northern Eurasia. Then, we compared land cover class frequencies and spatial agreement between global maps and reference data at local test sites. Spatial agreement was computed for all global map pixels, homogenous pixels, and for pixel blocks.

2.4.1. Cross comparison

Cross comparisons provide a means to reveal areas of uncertainty in global maps where reference data are sparse or not available (Herold et al., 2008). In this study, we first calculated spatial agreement between each map combination using the generalized LFT legend and then computed an overall agreement score as the average of the (six) pairwise agreement values. To bridge differences in spatial resolution and minimize misregistration errors we calculated fractional agreement for

Table 3
Error matrix of dominant life form types for Landsat-based reference maps at test sites.

Mapped		Reference class					
Class	Site	Tree	Shrub	Herb	Barren	Water	
Tree	STPB	*2176*		18			
	CARP	*466*	7	7		4	
	KOMI	*289*	22			1	
	PRIA	*851*	7	7			
	CHIT	*210*	3	3		2	
	VASY	*179*	3	2			
Shrub	STPB		*44*	6	1		
	CARP	2	*59*	19	2	2	
	KOMI	8	*39*	8	1		
	PRIA	2	*25*	13	1		
	CHIT	18	*26*	1			
	VASY	7	*45*	3			
Herbaceous	STPB	6	5	*58*	1		
	CARP	1	26	*346*	1		
	KOMI	12	1	*14*	15	3	
	PRIA	8	1	*658*			
	CHIT	2	7	*37*	8		
	VASY	1	10	*52*	1		
Barren	STPB		1	5	*380*		
	CARP		3	3	*29*	1	
	KOMI			15	*41*	1	
	PRIA		7	13	*199*		
	CHIT			3	*22*		
	VASY				*32*		
Water	STPB					*116*	
	CARP					*22*	
	KOMI	2				*34*	
	PRIA					*274*	
	CHIT					*32*	Overall accuracy
	VASY					*24*	
Percent correct	STPB	1.00	0.88	0.67	0.99	1.00	0.98
	CARP	0.99	0.62	0.92	0.91	0.76	0.92
	KOMI	0.93	0.63	0.38	0.72	0.87	0.8
	PRIA	0.99	0.63	0.95	1.00	1.00	0.97
	CHIT	0.91	0.72	0.84	0.73	0.94	0.87
	VASY	0.96	0.78	0.91	0.97	1.00	0.92
	All Sites	0.98	0.70	0.90	0.96	0.97	0.91

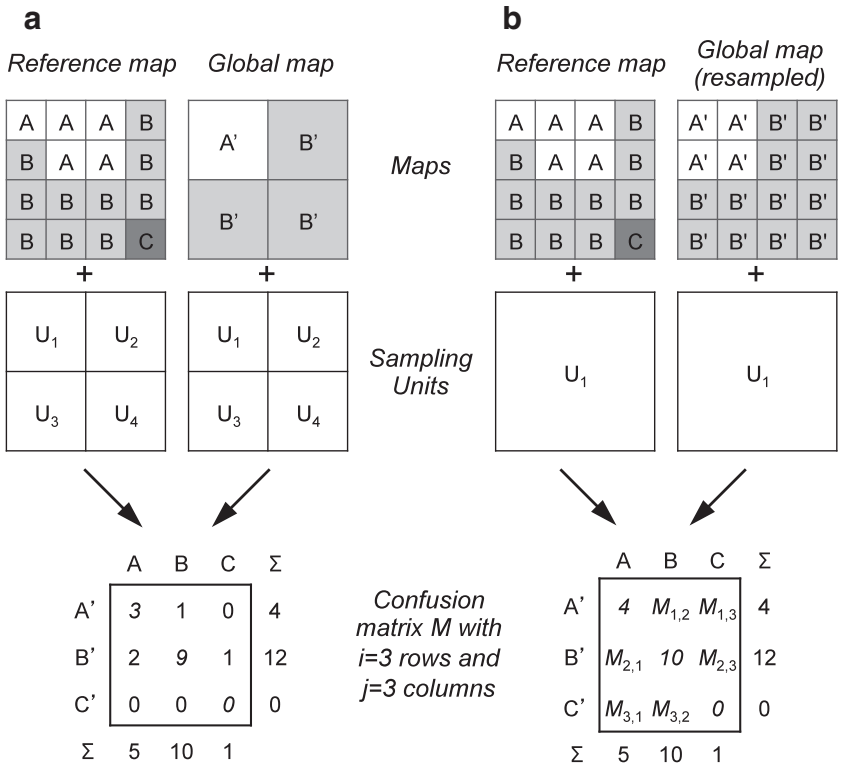

Fig. 2. Examples of constructing fractional error matrices for sampling units equal to a) the global map resolution and b) blocks of 2 × 2 coarse resolution pixels. For pixel blocks, row and column totals must be computed from the map statistics.

pixel-blocks of 3 × 3 km instead of pixel-by-pixel agreement as reported in previous global comparison studies (Giri et al., 2005; Herold et al., 2008). We computed overall fractional agreement from the relative frequency distributions of land cover within pixel-blocks as follows:

$$\textit{Overall fractional agreement} = \frac{1}{U}\sum_{u=1}^{U}\sum_{i=1}^{I}\min(X_{i;u}, Y_{i;u}) \quad (1)$$

where $X_{i;u}$ and $Y_{i;u}$ is the area fraction of pixel-block u occupied by class i in map X and Y, respectively.

2.4.2. Spatial agreement with Landsat-based maps

We calculated spatial agreement of global maps with Landsat-based maps at test sites using a fractional error method introduced by Latifovic and Olthof (2004). The method does not require spatial aggregation of the fine resolution reference maps to the coarse resolution of the global maps, and therefore takes full advantage of the observed land cover distribution within global map pixels. We populated fractional error matrices by counting the number of (fine resolution) reference pixels within global map pixels (Fig. 2a). Thus, instead of labeling a coarse resolution pixel either 100% correct or 100% incorrect, the "correctness" of that pixel was rated on a continuous scale based on the pixel fraction occupied by the "correct" reference class. We further computed standard accuracy measures such as the error of omission (exclusion) and commission (inclusion), and overall agreement and chance corrected agreement (kappa, Cohen, 1960) from these fractional error matrices. Omission error (OE) is the fraction of reference pixels that occupy coarse resolution pixels of other classes. Similarly, commission error (CE) is the total sub-pixel fraction of coarse resolution pixels occupied by reference pixels of other classes. Overall agreement and kappa range from 0 to 1, but the interpretation of kappa is less transparent. In general, kappa estimates between 0.6 and 0.8 are interpreted as good to very good, and between 0.4 and 0.6 as fair to moderate (Czaplewski, 1994). Because coarse resolution pixels usually contain a mixture of cover classes, total agreement scores obtained by this approach are generally less than 1 even if the classification were perfect.

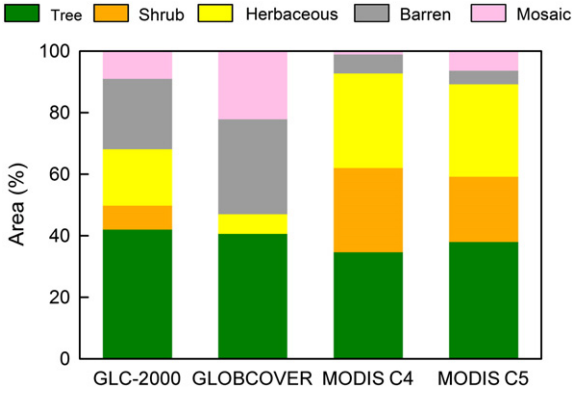

Fig. 3. Relative regional frequency of dominant life form types across Northern Eurasia based on global land cover maps.

2.4.3. Spatial agreement for homogenous land cover

Studies have shown an inverse relationship between map agreement and land cover heterogeneity in global maps (Herold et al., 2008). In heterogeneous landscapes, the proportion of a coarse resolution pixel occupied by a single, dominant land cover class can be significantly lower than in uniform landscapes. To measure map accuracy independent of this "coarse resolution bias" and landscape heterogeneity, we calculated accuracy statistics for "pure" pixels. We defined pure pixels as pixels that contained at least 95% of a single LFT class in the reference data.

2.4.4. Spatial agreement of pixel blocks

To explore how sampling affects estimates of map accuracy, we successively increased the size of sampling units from individual

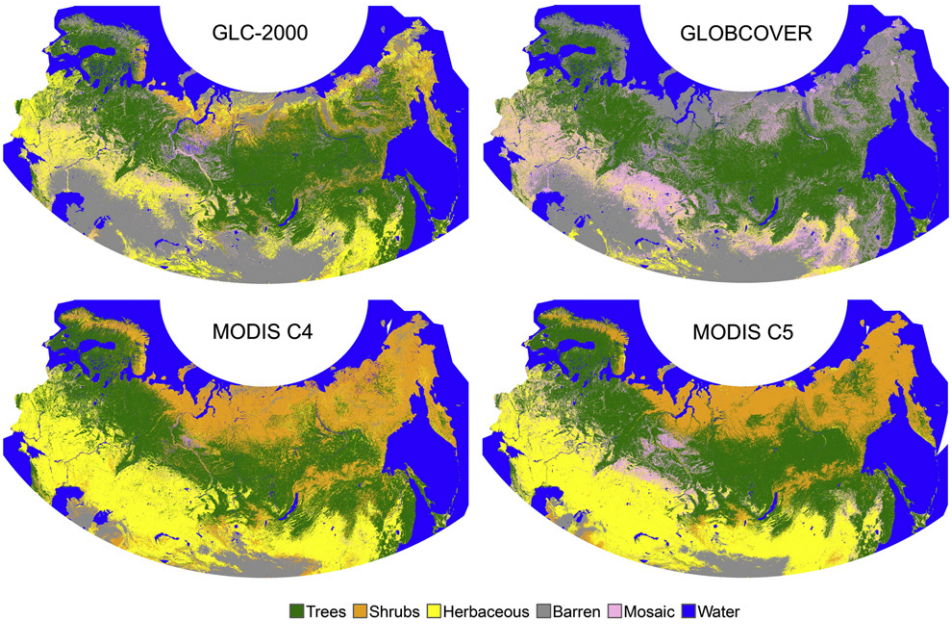

Fig. 4. Distribution of dominant life form types for Northern Eurasia based on GLC-2000, GLOBCOVER, MODIS C4 and MODIS C5.

pixels to blocks of multiple coarse resolution pixels (e.g. 2×2, 3×3 pixels, and so forth). First, we evaluated relative changes in agreement for each map as the size of analysis unit was increased. Then, we compared map accuracies among datasets using sampling units of equal area up to pixel blocks of ~5×5 km. The fractional error method has previously only been used with sampling units of individual pixels (Latifovic & Olthof, 2004); however, the method can also be applied to pixel blocks. To apply the method it was convenient to resample the coarse map to the spatial resolution of the reference data and compute the fractional confusion matrix in units of reference map pixels (Fig. 2b). The fractional confusion matrix was then populated by counting the minimum number of pixel units between map X and X' occupied by class i for each (pixel-block) sample u:

$$M_{i,i} = \sum_{u=1}^{U} \min\left(X_{i;u}, X'_{i;u}\right) \quad (2)$$

where i denotes the row and column (diagonal) of error matrix M, and $X_{i;u}$ and $X'_{i;u}$ denote the number of fine resolution pixels in pixel-block u occupied by class i in map X and X', respectively. Note that because the sampling unit can be greater than the map unit, non-diagonal elements do not have a unique solution. The calculation of agreement, omission and commission for pixel blocks is similar to that of individual global map pixels.

2.4.5. Treatment of mosaic classes

Mosaic classes of global land cover classifications contain mixtures of vegetation types, e.g. croplands, forests, shrub land, and grasslands that only occur at the coarse resolution scale (Table 1). This is problematic in accuracy assessment, because it is not clear when a pixel should be labeled as mosaic and when it should be labeled according to the dominant vegetation type. For example, an area with 30% tree cover could be classified either as mosaic or forest. To account for this ambiguity we assumed that mosaic pixels could contain any combination of tree, shrub and herbaceous vegetation types. Thus, we computed total agreement for mosaic pixels from the combined area of tree, shrub and herbaceous reference pixels. Similarly, when we calculated omission errors for these three classes, the area mapped as mosaic was added to tree, shrub, and herbaceous LFT, respectively, effectively decreasing the omission error. This approach reduced the uncertainty resulting from mosaics, but it potentially inflated the map accuracy measures in proportion to the area classified by global maps as mosaic or mixed vegetation.

3. Results

3.1. Comparison of global land cover for Northern Eurasia

There were significant differences among global maps with respect to the spatial distribution and relative frequency of vegetation types in Northern Eurasia (Fig. 3). Estimates of total tree dominated land area were relatively similar for the entire region, ranging from 35% (MODIS C4) to 42% (GLC-2000). However, differences in tree cover were not equally distributed but concentrated mainly in northern Siberia and the Far East (Fig. 4). For example, tree cover in GLC-2000 and GLOBCOVER was significantly higher in the transition

Table 4
Spatial agreement results from cross comparisons of global land cover maps for the region of Northern Eurasia (NE) and six validation sites.

	Spatial agreement for test sites and Northern Eurasia (NE)						
Maps compared	STPB	CARP	CHIT	KOMI	PRIA	VASY	NE
GLC2000–GLOBCOVER	0.75	0.62	0.76	0.38	0.77	0.80	0.65
GLC2000–MODIS C4	0.80	0.82	0.77	0.46	0.86	0.85	0.57
GLC2000–MODIS C5	0.86	0.65	0.69	0.66	0.91	0.86	0.60
GLOBCOVER–MODIS C4	0.71	0.67	0.67	0.16	0.71	0.76	0.42
GLOBCOVER–MODIS C5	0.77	0.74	0.57	0.37	0.76	0.84	0.48
MODIS C5–MODIS C4	0.82	0.71	0.75	0.69	0.88	0.81	0.79
Mean agreement	0.78	0.70	0.70	0.45	0.82	0.82	0.59

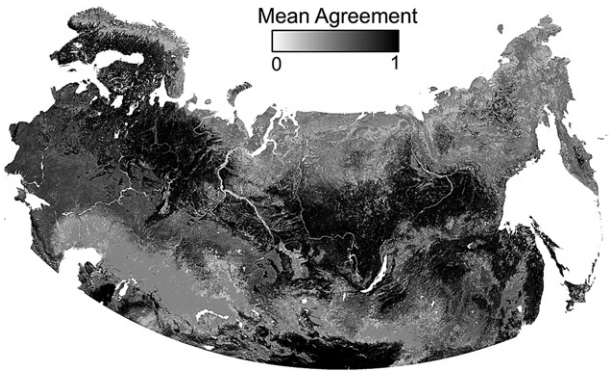

Fig. 5. Mean agreement of four land cover products (GLC-2000, GLOBCOVER, MODIS C4 and MODIS C5) after thematic aggregation to LFT legend (Table 1). Mean agreement is the average pairwise agreement between global maps and calculated as fractional agreement of land cover proportions within 3 × 3 km blocks, e.g. a value of 0.8 means that on average 80% of the land cover in a pixel block is in agreement.

zone between boreal forest and shrub tundra compared to MODIS C4 and MODIS C5.

Cross comparison of global land cover for Northern Eurasia revealed an average agreement in LFT classes of 0.59 (Table 4). Spatial agreement was highest between the related products MODIS C4 and C5 (0.79), and GLC-2000 and GLOBCOVER (0.65), but GLC-2000 and MODIS C5 also showed moderate agreement (0.6). Spatial agreement among all four maps (Fig. 5) was generally above 0.8 in the forest belt of Northern Eurasia and below 0.5 in the northern taiga-tundra zone and in the dry zones of the South. The European part of the region also showed fair (0.51–0.65) and moderate (0.66–0.8) agreement, in part because of the thematic ambiguity between the herbaceous and mosaic class in the LFT legend.

MODIS mapped a significant proportion of the northern taiga and the tundra as shrub dominated vegetation. According to MODIS C4 and C5, shrub land cover types make up about one quarter of the entire Northern Eurasia region (27% and 21%, respectively). This is about three times the total shrub dominated land area mapped by GLC-2000 (7%). Interestingly, GLOBCOVER classified less than 0.1% of Northern Eurasia as shrub lands, but mapped a significantly greater proportion of the region as sparse vegetation and barren land (31%) when compared to GLC-2000 (23%), and MODIS C4 (6%) and C5 (4%). The difference is most pronounced in the taiga-tundra ecotone and in the transition from steppe/grasslands to non-vegetated areas in the south, where MODIS mapped significantly more herbaceous vegetation than other maps. However, GLOBCOCVER classified 22% of the region as vegetation mosaics, which partly explains the low proportion of herbaceous and shrub dominated vegetation classes in this map.

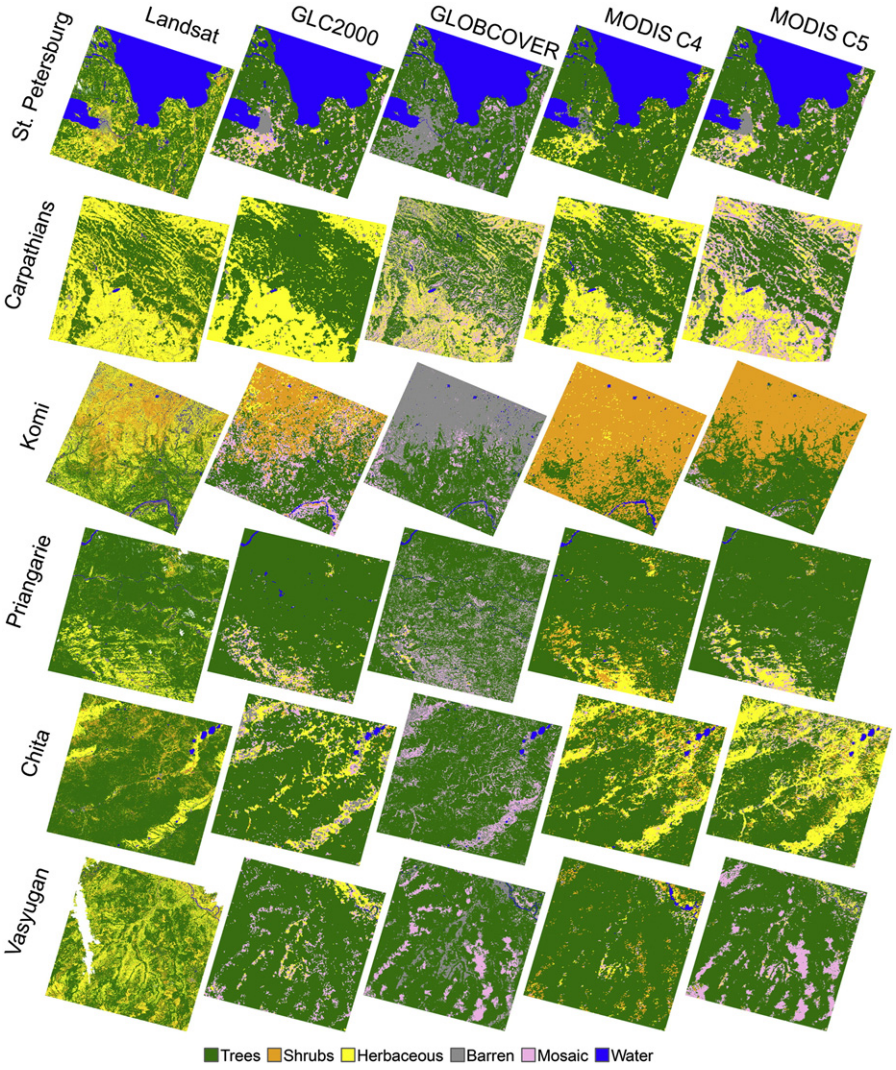

Fig. 6. Distribution of dominant life form types at reference sites.

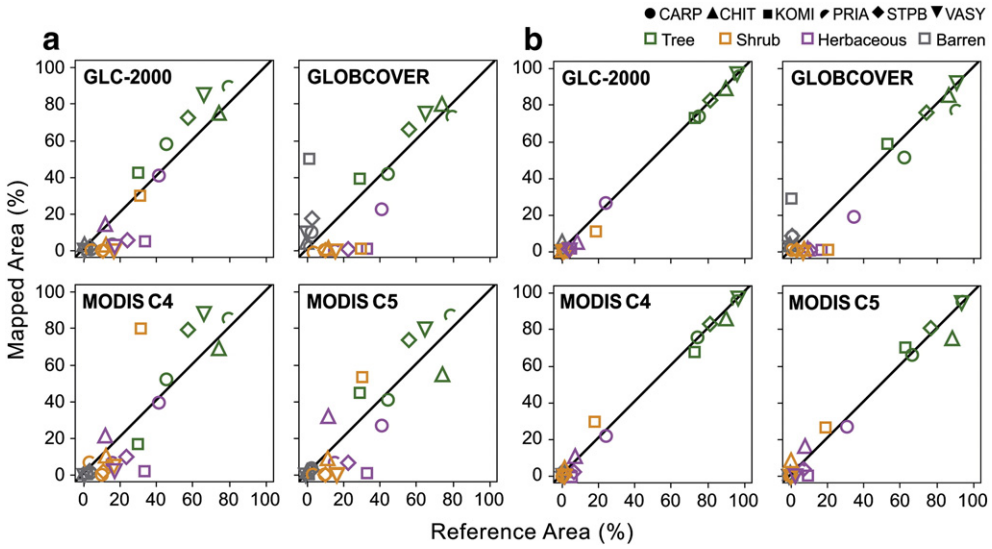

Fig. 7. Predicted (global map) versus reference land cover area, calculated from (a) all pixels and (b) pure pixels, in percent of total test site area for GLC-2000, GLOBCOVER, MODIS C4, and MODIS C5. Symbol shapes represent individual test sites. Symbol fill patterns and shades of gray represent individual LFT classes (tree: open light gray, shrub: filled light gray, herbaceous: open dark gray, barren: filled dark gray).

The proportion of mosaic pixels was significantly lower in GLC-2000 (9%), MODIS C4 (1%), and MODIS C5 (6%).

3.2. Comparison of global land cover at test sites

3.2.1. Distribution of global land cover at test sites

Comparisons of the four global land cover datasets with the Landsat-based reference maps (Fig. 6) revealed general patterns of agreement and disagreement at the test site. To quantify the representation of global land cover at this scale, we compared the relative proportions of LFT classes at each site without regard for their within-site spatial distribution (Fig. 7a). We found the percentage of coarse resolution pixels in the tree class matched most closely the percentage of reference pixels across all global maps. GLOBCOVER showed the highest overall agreement in this class (88%), followed by GLC-2000 (82%), MODIS C5 (80%), and MODIS C4 (78%). There was a general tendency in all four coarse resolution maps to overestimate the tree class by an average of 10-20% (Fig. 7a). Surprisingly, the pattern of under-reporting smaller classes did not improve on maps with finer resolution (GLOBCOVER and MODIS C5). The same analysis performed on coarse resolution pure pixels with uniform land cover (Fig. 7b) showed much higher agreement in total tree dominated area and other LFT classes for all global maps, although the improvement for non-tree classes was notably smaller for GLOBCOVER.

Differences in LFT class proportions between global maps and reference data were largest in areas with sparse tree cover. At Komi, GLC-2000 and GLOBCOVER overestimated tree cover by 30% and 26%, respectively, whereas MODIS C4 under-represented this class by 40%. The low percentage of tree cover in C4 was accompanied by a significantly higher proportion of shrub cover; 2.6 times higher than the reference data. In MODIS C5, classification of tree vegetation significantly improved and was more consistent with the reference data. Conversely, GLOBCOVER significantly overestimated barren/sparse vegetation at all test sites, but showed the best agreement for inland water bodies.

3.2.2. Spatial agreement at test sites

Spatial agreement scores derived from cross comparisons of global maps were generally above 0.7 at the test sites, and thus above the region-wide agreement of 0.59, except at Komi (0.45) (Table 4). However, agreement at Komi was moderate between MODIS C5 and

Table 5
Overall agreement and chance corrected agreement (Kappa) between reference maps and coarse resolution land cover products for per-pixel comparisons of dominant life form types at test sites.

	GLC2000		GLOBCOVER		MODIS C4		MODIS C5	
Site	Agreement	Kappa	Agreement	Kappa	Agreement	Kappa	Agreement	Kappa
All pixels								
Carpathians	0.69	0.45	0.73	0.57	0.75	0.56	0.87	0.79
Chita	0.76	0.40	0.83	0.51	0.72	0.37	0.67	0.35
Komi	0.58	0.41	0.34	0.22	0.42	0.16	0.49	0.28
Priangarie	0.84	0.39	0.85	0.49	0.80	0.35	0.86	0.51
St. Petersburg	0.74	0.51	0.69	0.48	0.67	0.36	0.75	0.54
Vasyugan	0.74	0.34	0.72	0.39	0.68	0.17	0.77	0.45
Total across sites	0.73	0.47	0.70	0.46	0.67	0.41	0.74	0.52
Pure pixels only								
Carpathians	0.95	0.87	0.89	0.78	0.98	0.94	0.99	0.97
Chita	0.94	0.69	0.94	0.71	0.92	0.65	0.85	0.52
Komi	0.91	0.77	0.66	0.48	0.86	0.70	0.84	0.68
Priangarie	0.98	0.68	0.94	0.62	0.97	0.54	0.98	0.82
St. Petersburg	0.97	0.91	0.91	0.78	0.96	0.88	0.96	0.89
Vasyugan	0.98	0.49	0.94	0.54	0.97	0.27	0.97	0.62
Total across sites	0.96	0.80	0.91	0.72	0.96	0.78	0.94	0.79

MODIS C4 (0.69) and between MODIS C5 and GLC-2000 (0.66). High global map agreement was observed at Priangare (0.82) and at Vasyugan (0.82).

Pixel-level agreement with Landsat-based references maps ranged from 0.34 at the Komi site with GLOBCOVER to 0.87 at the Carpathian site with MODIS C5 (Table 5). Mean agreement across global maps ranged from 0.46 ± 0.09 at Komi to 0.84 ± 0.02 at Priangarie. For all sites combined, total agreement for MODIS C4 and C5 was 0.67 and 0.74, respectively; and for GLOBCOVER and GLC-2000 total agreement was 0.70 and 0.73, respectively. MODIS C5 performed better than MODIS C4 on all test sites with the exception of Chita. The lower agreement at Chita may be attributable to widespread forest fires that burned about 17% of the total site area in 2002–2003, most of it tree-dominated. These disturbances occurred between the Landsat and MODIS C5 data acquisitions in 2000 and 2005, respectively. The time difference to the reference maps is similar to that for GLOBCOVER, but larger compared to GLC-2000 and MODIS C4. However, on average the disturbance rates did not exceed 0.1–0.2% per year over large areas. Thus, we expect the effect of differences in the data acquisition dates to be small on all sites with the exception of Chita. Chance corrected agreement at the pixel scale was fair to moderate for all maps (Kappa = 0.41-0.52).

The fractional error matrices for LFT classes reveal more detailed patterns of class confusions across test sites (Table 6). Map errors were low for tree vegetation with omission errors between 0.05 (GLC-2000 and GLOBCOVER), 0.08 (MODIS C5) and 0.13 (MODIS C4), and commission errors between 0.2 (GLOBCOVER and MODIS C5) and 0.24 (GLC-2000). Because fractional error estimates represent sub-pixel proportions of land cover, omission and commission errors correspond to the approximate boundaries of the mapped classes. For example, a 0.05 omission error indicates an average tree cover in non-tree pixels of 5%. Likewise, a commission error of 0.2 represents the mean fraction of non-tree cover included in coarse resolution tree pixels. Because global maps typically define a minimum tree cover threshold for tree dominated classes between 10 and 20%, it is reasonable to expect an average omission error between 0.1 and 0.2, depending on class definition.

Shrub and herbaceous classes were significantly underestimated by all global maps (OE > 0.48, Table 6). Omission of shrub vegetation was greatest in GLOBCOVER (OE = 0.82), and omission of herbaceous vegetation was greatest in MODIS C4 (OE = 0.61) and GLOBCOVER (OE = 0.60). A large proportion of the shrub and herbaceous vegetation in the reference maps was labeled as tree-dominated pixels in global maps (40–52% and 28–41%, respectively). The proportion was roughly 10% greater in GLC-2000 than in other maps. However, the confusion among shrub, herbaceous and barren classes was also considerable. About 35% and 30% of the shrub and herbaceous vegetation, respectively, was mapped as barren lands in GLOBCOVER; a significantly higher proportion than in the other maps where only 1–2% of shrub and herbaceous cover occurred in barren pixels. Conversely, MODIS C4 (27%) and MODIS C5 (19%) mapped a greater proportion of herbaceous vegetation as shrub land compared to GLC-2000 (9%) and GLOBCOVER (0%). Patterns of commission and omission varied among test sites (Fig. 8). For example, omission of herbaceous vegetation was relatively low at the Carpathian site and high at Komi. GLC-2000 estimated shrubs accurately at Komi, but missed shrub lands entirely at the Carpathian site.

3.2.3. Global map agreement for homogenous land cover

To quantify the effect of landscape homogeneity on map accuracy we performed a separate accuracy assessment for pure pixels. The

Table 6
Fractional error matrices based on dominant life form types of GLC-2000, GLOBCOVER, MODIS C4 and MODIS C5 (aggregated results with reference data from all test sites. Bold values represent class agreement.).

	Landsat land cover (km^2)						
	Trees	Shrubs	Herbaceous	Barren	Mosaics	Water	Commission
GLC-2000							
Trees	***88,307***	11,900	16,103	844	0	987	0.24
Shrubs	851	***4,180***	3,618	200	0	445	0.40
Herbaceous	3,570	3,277	***12,131***	1,054	0	223	0.30
Barren	160	162	867	***378***	0	60	0.77
Mosaic	***4,867***	***3,507***	***6,589***	389	0	468	–
Water	313	64	132	89	0	***1,226***	0.33
Omission	0.05	0.67	0.53	0.87	0.00	0.64	
GLOBCOVER							
Trees	***83,479***	10,010	11,752	589	0	796	0.20
Shrubs	1	***0***	6	0	0	0	0.00
Herbaceous	608	832	***4,831***	447	0	17	0.11
Barren	3,817	8,236	11,999	***1,198***	0	754	0.95
Mosaic	***11,311***	***4,230***	***11,248***	690	0	164	–
Water	125	46	94	72	0	***1,895***	0.15
Omission	0.05	0.82	0.60	0.60	0.00	0.48	
MODIS C4							
Trees	***83,933***	9,524	12,947	827	0	710	0.22
Shrubs	7,859	***9,275***	10,657	564	0	1,210	0.67
Herbaceous	4,044	3,409	***13,733***	1,038	0	180	0.36
Barren	133	152	254	***340***	0	54	0.64
Mosaic	***1,411***	***550***	***1,557***	74	0	51	–
Water	280	115	194	105	0	***1,182***	0.37
Omission	0.13	0.57	0.61	0.88	0.00	0.65	
MODIS C5							
Trees	***84,465***	9,207	11,100	775	0	1,209	0.20
Shrubs	2,943	***6,737***	7,354	428	0	747	0.52
Herbaceous	4,654	2,963	***12,020***	744	0	119	0.29
Barren	154	295	537	***592***	0	38	0.63
Mosaic	***6,553***	***4,043***	***8,737***	418	0	304	–
Water	20	8	14	26	0	***1,145***	0.06
Omission	0.08	0.54	0.48	0.80	0.00	0.68	

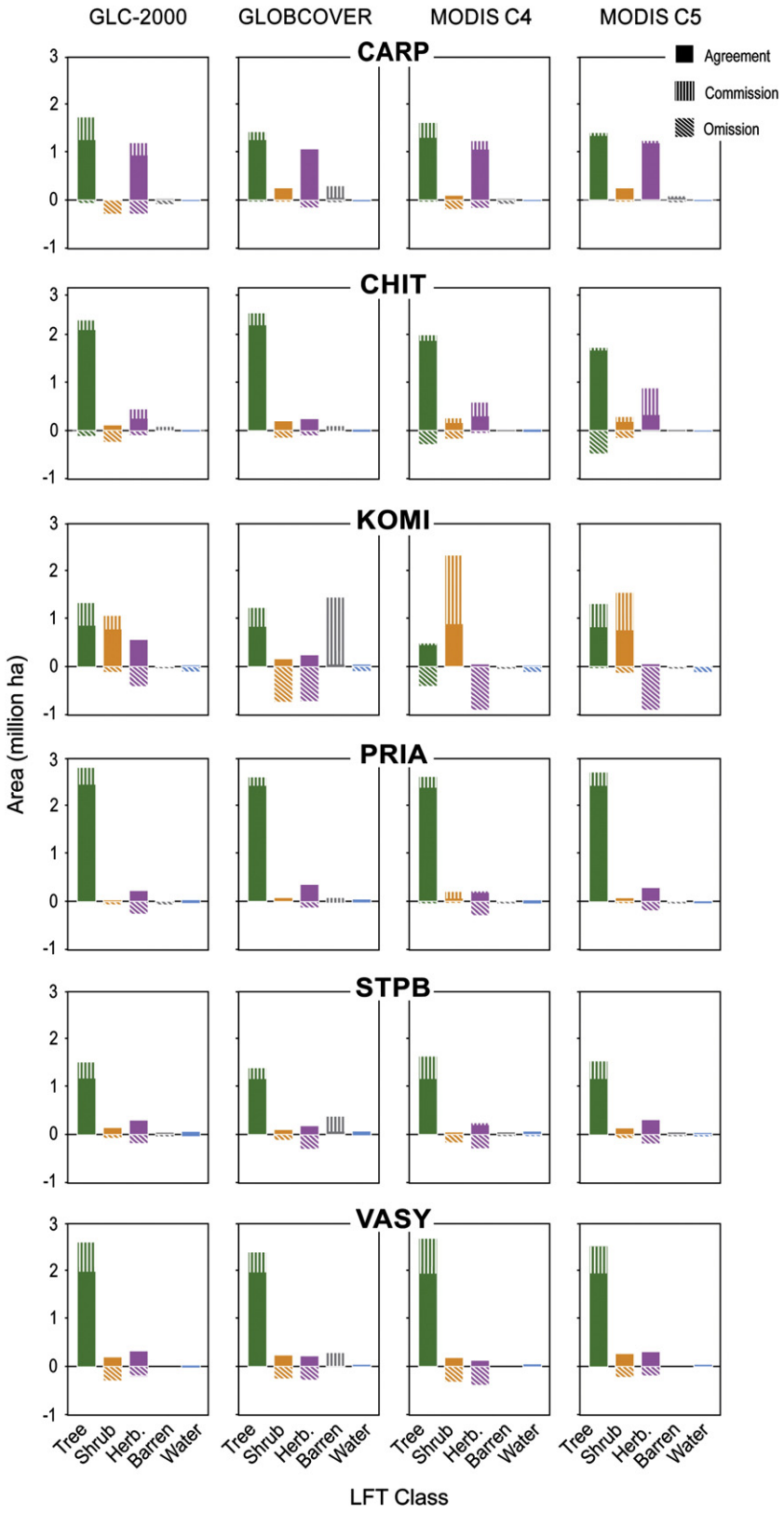

Fig. 8. Agreement, commission error, and omission error for LFT classes (tree: green, shrub: orange, herbaceous: magenta, barren: gray, water: blue) at six test sites (CARP, CHIT, KOMI, PRIA, STPB, and VASY) using 3 × 3 km pixel blocks as sampling units. For illustration, omission error is represented by negative values, whereas agreement and commission error are represented by positive values.

agreement of global maps was nearly perfect if only pure pixels were considered (Table 5). Except at Komi site, total agreement for LFT's exceeded 0.9. Kappa also indicated good and perfect agreement for pure pixels. Compared with results from all pixels total agreement increased between 0.09 and 0.45, with a higher increase at sites with heterogeneous land cover (e.g. 0.31–0.45 increase at Komi) and

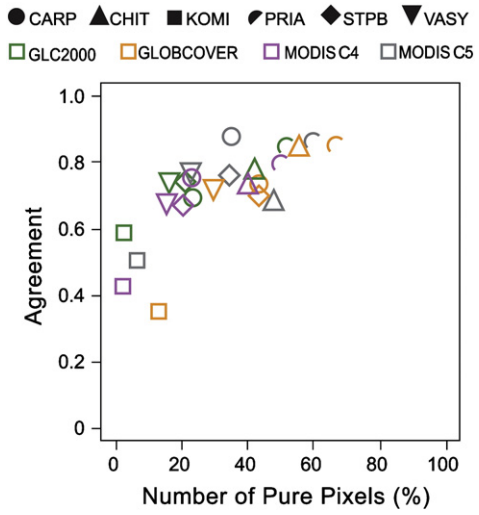

Fig. 9. Total agreement versus number of pure coarse resolution pixels at each site. Pure coarse pixels are defined to contain at least 95% of a single land cover class in the reference data. Symbol shapes represent individual test sites. Symbol fill patterns and shades of gray represent different global maps (GLC-2000: open light gray, GLOBCOVER: filled light gray, MODIS C4: open dark gray, MODIS C5: filled dark gray).

lower increase at more homogeneous sites (e.g. 0.09–0.17 at Priangarie). Further, the effect was greater for the 1-km resolution maps GLC-2000 (mean = 0.23 ± standard deviation = 0.06) and MODIS C4 (0.27 ± 0.10) than for GLOBCOVER (0.18 ± 0.08) and MODIS C5 (0.19 ± 0.08).

There was a clear relationship between map agreement and number of pure pixels at test sites (Fig. 9). The Priangarie site with the highest overall agreement also had the highest percentage of pure pixels ranging from 52% for 1-km pixels (GLC-2000 and MODIS C4) to 67% for 300-m pixels (GLOBCOVER). Conversely, the Komi site had the lowest map agreement and the lowest percentage of pure pixels (3%–13% for 1-km to 300-m pixels). Without Komi the relationship was linear with a more conservative slope estimate of 0.0026 (± 0.0017, 95% Confidence Interval (CI)) and an intercept of 0.659 (± 0.069 95% CI) (r = 0.6, p-value = 0.005). Thus, for a percentage of pure pixels greater than 15%, a 10% increase in numbers of pure pixels was associated with an increase in agreement of approximately 0.03 (± 0.02 95% CI). However, we did not find a clear relationship between number of pure pixels and total agreement within test sites and between maps indicating that factors other than spatial resolution affected map accuracy.

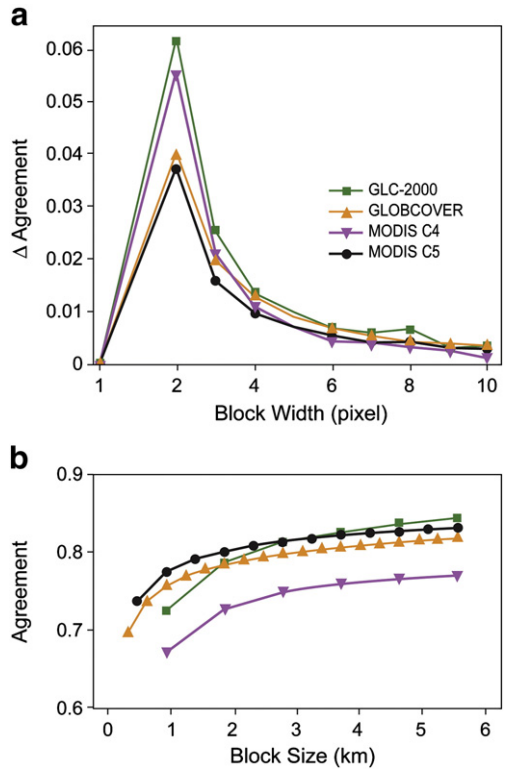

Fig. 10. Effect of increasing the size of the analysis unit on overall agreement across test sites: a) relative improvement in agreement associated with blocks in units of pixels, and b) increase in mean overall agreement associated with blocks sizes in units of area.

3.2.4. Relationship between map agreement and size of sampling units

Map accuracy increased significantly with the size of sampling units, but the effect decreased exponentially between 2 × 2 and 10 × 10 pixel blocks (Fig. 10a). The initial increase of overall accuracy scores at the lower range of block sizes was highest for the 1-km maps GLC-2000 and MODIS C4. Compared to pixel-scale results, blocks of 2 by 2 pixels increased the overall agreement by 0.062 ± 0.037 for GLC-2000, 0.056 ± 0.013 for MODIS C4, 0.040 ± 0.015 for GLOBCOVER and 0.037 ± 0.011 for MODIS C5. When a block size of 5 × 5 pixels was reached the subsequent improvement in agreement with increasing block size eventually dropped below a value of 0.01 (similar for all global datasets).

Comparison of overall agreement for sampling blocks of equal area revealed that the finer resolution datasets GLOBCOVER and MODIS C5 performed generally better than GLC-2000 for analysis units of ~1 km, but the advantage diminished for block sizes between 2 and 3 km, and even slightly decreased for larger blocks (Fig. 10b). MODIS C4 showed the lowest overall agreement across test sites and analysis units. At individual test sites, overall agreement generally differed among datasets across the observed range of block sizes such that agreement lines rarely crossed (Fig. 11).

4. Discussion

We presented an approach to assess recent global land cover maps consistently for the region of Northern Eurasia, and explored the effect of sampling unit, reference class assignment and landscape characteristics on accuracy statistics. The results showed that methodology and choice of sampling unit significantly influence accuracy estimates, which indicates that direct comparison of reported global map statistics may be misleading.

Results from this study revealed important similarities and differences among global land cover maps for Northern Eurasia. Disagreement among global maps was high in the taiga-tundra ecotone and in the southern dry zones of the region. At the test sites, all maps showed similar overall per-pixel agreement with Landsat-based reference data (0.67–0.74), although MODIS C5 and GLC-2000 produced slightly better results than MODIS C4 and GLOBCOVER. Map accuracy increased by 7% from MODIS C4 to MODIS C5, and decreased by 3% between GLC-2000 and GLOBCOVER. However, when sampling units of equal area were used, e.g. 3 × 3 km, then total agreement of GLOBCOVER (0.80), GLC-2000 (0.81), and MODIS C5 (0.81) was similar, and higher than MODIS C4 (0.75). Despite the comparable overall agreement scores, similarities at the test sites were mostly due to similar representations of tree dominated land cover. Though slightly overestimated, tree-dominated vegetation was the most accurate land cover class in all maps, whereas there were significant differences in the representation of non-tree land cover classes. Given that our test sites were predominantly located in the forest belt of the region and that the diversity of land cover types in Northern Eurasia is far greater than a classification of basic life form types, our assessment is probably a rather optimistic scenario and not representative for the entire region or full map legends.

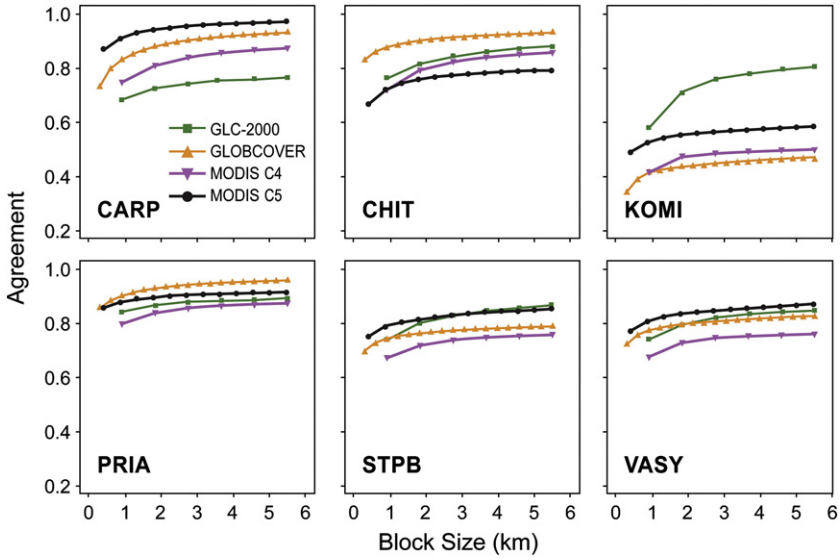

Fig. 11. Overall agreement for each test site associated with sequentially increasing the analysis unit from individual coarse resolution pixels to pixel blocks of ~6 km. The x-axis shows block size in units of km.

Subdominant and rare classes that exist largely at the sub-pixel scale were difficult to resolve with thematic coarse resolution maps. Despite their spectrally distinct signature, inland water bodies showed omission errors between 50 and 70%. Representation of subdominant and rare classes improved for pure pixels compared to all pixels (Fig. 7a,b), although the improvement was less significant for GLOBCOVER than for other maps. While the finer resolution of GLOBCOVER could be expected to improve the accuracy of small classes, water was the only small class more accurately mapped by GLOBCOVER at our test sites. It appears with finer resolution other sources of error including pixel registration, became more significant. In this context, it is important to note that GLC-2000 and GLOBCOVER were only available in a geographic coordinate system and had to be reprojected to an equal area projection prior to analysis. Thus, resampling and registration errors may have been more significant for these two maps. Further, GLC-2000 was included in the training database for GLOBCOVER and thus the two maps are inherently linked.

Our analysis of how the sampling unit affects accuracy suggests that blocks of 5×5 pixels are a good compromise between increased overall agreement and decreased spatial detail (Fig. 9a). Overall agreement increased between 0.07 and 0.11 for blocks of 5×5 pixels compared to per-pixel comparisons, but larger block sizes did not improve accuracies significantly. While map accuracies are usually reported for a single spatial unit, maps are often used at different spatial scales (e.g. pixels, polygons, and pixel blocks). More importantly, there is no standard sampling unit for global accuracy assessment. Reported statistics for global maps are often obtained from sampling units larger than the spatial resolution of the map, and samples are often selected to represent homogeneous or dominant land cover types (Bicheron et al., 2008; Friedl et al., 2002; Friedl et al., 2010; Mayaux et al., 2006). Because pixels are the basic mapping units of most global land cover data, users may not always be aware of the larger uncertainties associated with using these maps at the pixel scale. Map errors unrelated to the classification process such as spatial misregistration and convolution of the sensor signal with neighboring pixels can significantly influence the underlying remote sensing data (Townshend et al., 2000). This suggests that measurements of land cover properties are more robust at spatial resolutions coarser than the individual pixel.

To compare accuracy measures among maps, we generalized land cover classifications and re-labeled each pixel with its dominant LFT. Although other generalizations are possible, the LFT represents the most general and unambiguous common denominator of current global maps. However, the loss of thematic detail required to reconcile legend differences precludes evaluation of many environmentally important features of land cover such as the proportion of deciduous and evergreen tree species. As classifications of land cover products are more closely aligned with a common land cover language (LCCS, Di Gregorio, 2005) this type of analysis will become more feasible. For example, the MODIS Land Cover Product for Collection 6 plans to use an LCCS-based classification (Friedl et al., 2010), and there have been growing international efforts to promote LCCS and to provide translations between existing legends (Herold et al., 2006, 2008; Thomlinson et al., 1999).

In addition to differences in classification systems used in current land cover maps, several additional challenges present difficulties to assessing and comparing land cover maps at moderate spatial resolution. Mosaic classes representing mixtures of different land cover classes present substantial difficulties because they tend to be spectrally similar to multiple sub-pixel cover types. At the same time, because mosaic classes tend to be used as "catch-all" classes, they can artificially inflate estimates of map accuracy but not address underlying thematic uncertainties. As a result, hierarchical classification schemes such as the FAO LCCS system need to be designed to account for challenges associated with characterizing land cover in heterogeneous areas and ecotones. In this context, approaches that directly estimate sub-pixel proportions of land cover types provide a useful strategy (Hansen et al., 2002; Hansen et al., 2005).

Our approach was built on earlier global map validation efforts that relied on Landsat-based maps as reference data (Cohen et al., 2003; Cohen et al., 2006). The combination of higher resolution reference maps and a standardized LCCS-based classification provides a flexible validation framework for assessing coarse resolution maps of different spatial resolutions and thematic detail. A disadvantage is that our approach relies on limited number of test sites, which precludes comprehensive assessment of coarse resolution maps for the entire Northern Eurasian region. At the same time, our results provide useful information related to the differences, strengths and weaknesses of each map at specific locations, as well as the relative magnitude of different types of errors likely to be found in each map. Furthermore, comparison of different accuracy metrics at selected locations provides insights related to interpretation of commonly reported global map accuracies.

The fractional error approach provided a method to compare global maps of different spatial resolutions without sacrificing the fine spatial detail of the Landsat maps. Reference data for validating global

map products are often acquired at higher spatial resolution, but conventional error matrices (Foody, 2002) require the reference data to match the spatial resolution of the map being validated. As a consequence, when maps with differing spatial resolutions are compared, the map with the higher spatial resolution is often aggregated to the coarser resolution by assigning the dominant class to the aggregated pixel (Stehman & Czaplewski, 1998; Turner et al., 2000). The disadvantage of this approach is that it assumes a level of homogeneity in the landscape, which is rarely present, and as a result, subdominant and rare classes are often under-represented. The extent of this low-resolution bias depends on the patchiness of land cover. For example, in a study that assessed the representation of wetlands in global and continental land cover maps in the St. Petersburg region, we found that 50% of the wetland area was omitted when 30 m reference pixels were aggregated to 1-km pixels using the dominant class rule (Krankina et al., 2008).

The correlation between pure pixels and overall agreement demonstrated that landscape heterogeneity played a significant role in explaining differences in map accuracy among test sites. Because numbers of pure pixels are also a function of pixel resolution, this relationship could help explain the link between pixel size and classification accuracy. For example, GLOBCOVER (300 m) and MODIS C5 (500 m) have a higher spatial resolution than GLC-2000 (1 km) and MODIS C4 (1 km). Using the test site data we calculated that an increase in spatial resolution from 1 km to 300 m was associated with a 16% (10–20%) increase in number of pure pixels. In comparison, an increase in spatial resolution from 1 km to 500 m was associated with an increase of only 9% pure pixels (4-14%). According to the previously derived relationship this increase in the prevalence of pure pixels should yield a 0.04 and 0.02 improvement in overall agreement for the 300-m GLOBCOVER and 500-m MODIS C4 map, respectively, relative to the 1-km GLC-2000 and MODIS C4 land cover. While it is clearly an over-simplification, this result provides an approximate estimate of the order of magnitude of the effect.

5. Conclusions

Globally consistent vegetation maps with known accuracy are a fundamental requirement for global change research and sound policies to address global change. Growing recognition of the role of terrestrial ecosystems and land use in the global carbon cycle (IPCC 2007) places new demands on the accuracy of products derived from satellite observations. Methods for independent map validation have been actively studied in recent years but many issues remain unresolved, including design of map legends that lend themselves to uniform aggregation (i.e., LCCS), consensus on metrics of agreement between fine-resolution reference data and coarser global maps, and support for a globally distributed set of test sites against which all maps can be compared (Justice et al., 2000).

Global land cover maps differ significantly in their representation of Northern Eurasia. Improved understanding of global land cover map quality requires new approaches for accuracy assessment because accuracy statistics reported by global map developers are not directly comparable. Consistent comparisons of global maps are complicated by disparities in class definitions, thematic detail, and spatial resolution. We used LFT classes as a general and robust common denominator of current global maps but the loss of thematic detail precluded evaluation of map performance based on the thematic resolution required by many users.

Our analysis showed that agreement among the maps is highest in the forest belt (>0.8) and that disagreement among the maps is concentrated in transitions zones. Specifically, the maps provide very different representation of land cover in the northern transition zone between boreal forest and tundra and in the southern transition zone between forest and semi-arid grasslands, agriculture, and steppe. High uncertainty at biome borders and in areas of human activity is a major concern, because these zones are most vulnerable to climate and socio-economic changes.

A comparison of each global map with reference data from six test sites showed the effect of sampling unit area on accuracy estimates. Map agreement increased significantly from the scale of single coarse map pixels to block sizes of 5×5 pixels; for larger blocks, the improvement in agreement was marginal. Therefore, when the finer resolution datasets GLOBCOVER and MODIS C5 were scaled to the same unit area as GLC-2000 and MODIS C4 (~1 km pixel size) then GLOBCOVER and MODIS C5 performed better in terms of overall accuracy. For block sizes larger than 2 km overall accuracy of GLC-2000, GLOBCOVER and MODIS C5 was similar and better than MODIS C4. Global map errors were also positively associated with landscape heterogeneity: we found that the agreement was 0.09–0.45 higher when only homogenous pixels were considered, suggesting that global map accuracies based on samples located in homogenous areas may be too optimistic.

Tree dominated classes were the most accurately mapped land cover classes at all of the test sites, although all maps showed a tendency to overestimate tree dominated area by 10–20% as a result of the coarse spatial resolution bias. Conversely, lower map agreement with reference data was generally associated with shrub and herbaceous vegetation, and sparse tree cover. GLOBCOVER significantly underestimated shrub and herbaceous cover and overestimated barren and sparse vegetation. MODIS C4, and to a lesser degree MODIS C5, showed a tendency to map herbaceous tundra as open shrub land; and GLC-2000 mapped a greater proportion of shrub and herbaceous vegetation as tree vegetation compared to the other maps.

There are significant regional differences in the accuracy of global maps: no single map performed best across all test sites and land cover classes. Thus, the 'best' choice of map for Northern Eurasia likely depends on the geographic region and the land cover classes of interest. For example, the choice may be different if estimates of tree-dominated area are desired versus an accurate representation of shrub and/or herbaceous vegetation. More robust evaluation of map performance in different parts of Northern Eurasia and at finer thematic resolution requires further analysis and additional test sites.

Acknowledgments

The research was supported by the Land Cover/Land-Use Change Program of the National Aeronautics and Space Administration (grant numbers NNG06GF54G and NNX09AK88G) and in part by the Asia-Pacific Network for Global Change Research and the Alexander von Humboldt Foundation. We like to thank Dr. Curtis Woodcock for his advice in the early planning of this study, and Gretchen Bracher for preparing graphs. We are also thankful for the comments of two anonymous reviewers that helped to improve this manuscript.

References

Alexeyev, V. A., & Birdsey, R. A. (1998). Carbon storage in forests and peatlands of Russia. Radnor: U.S. Department of Agriculture, Forest Service, Northern Research Station.

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., & Weber, J. L. (2008). GLOBCOVER the most detailed portrait of Earth. *Esa Bulletin-European Space Agency*, 24–31.

Bartalev, S. A., Belward, A. S., Erchov, D. V., & Isaev, A. S. (2003). A new SPOT4-VEGETATION derived land cover map of Northern Eurasia. *International Journal of Remote Sensing*, *24*, 1977–1982.

Bartholome, E., & Belward, A. S. (2005). GLC2000: A new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, *26*, 1959–1977.

Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., et al. (2008). *GLOBCOVER: Products description and validation report* (pp. 47). Toulouse, France: POSTEL.

Bonan, G. B., Oleson, K. W., Vertenstein, M., Levis, S., Zeng, X. B., Dai, Y. J., et al. (2002). The land surface climatology of the community land model coupled to the NCAR community climate model. *Journal of Climate*, *15*, 3123–3149.

Boschetti, L., Flasse, S. P., & Brivio, P. A. (2004). Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sensing of Environment*, *91*, 280–292.

Bulygina, O. N., Groisman, P. Y., Razuvaev, V. N., & Radionov, V. F. (2010). Snow cover basal ice layer changes over Northern Eurasia since 1966. *Environmental Research Letters*, *5*.

Chapin, F. S., Sturm, M., Serreze, M. C., McFadden, J. P., Key, J. R., Lloyd, A. H., et al. (2005). Role of land-surface changes in Arctic summer warming. *Science*, *310*, 657–660.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Cohen, W. B., Maiersperger, T. K., Turner, D. P., Ritts, W. D., Pflugmacher, D., Kennedy, R. E., et al. (2006). MODIS land cover and LAI collection 4 product quality across nine sites in the western hemisphere. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 1843–1857.

Cohen, W. B., Maiersperger, T. K., Yang, Z. Q., Gower, S. T., Turner, D. P., Ritts, W. D., et al. (2003). Comparisons of land cover and LAI estimates derived from ETM plus and MODIS for four sites in North America: A quality assessment of 2000/2001 provisional MODIS products. *Remote Sensing of Environment*, *88*, 233–255.

Czaplewski, R. (1994). *Variance approximations for assessments of classification accuracy* (pp. 29). Fort Collins: US Department of Agriculture, Forest Service.

DeFries, R. S., & Townshend, J. R. G. (1994). NDVI derived land cover classifications at a global scale. *International Journal of Remote Sensing*, *15*, 3567–3586.

Di Gregorio, A. (2005). Land cover classification system: Classification concepts and user manual for software – version 2. Rome, Italy: Food & Agriculture Organization of the United Nations.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., et al. (2005). Global consequences of land use. *Science*, *309*, 570–574.

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

Forbes, B. C., Fresco, N., Shvidenko, A., Danell, K., & Chapin, F. S. (2004). Geographic variations in anthropogenic drivers that influence the vulnerability and resilience of social-ecological systems. *Ambio*, *33*, 377–382.

Frey, K. E., & Smith, L. C. (2007). How well do we know northern land cover? Comparison of four global vegetation and wetland products with a new ground-truth database for West Siberia. *Global Biogeochemical Cycles*, *21*.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, *83*, 287–302.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., et al. (2010). MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*, 168–182.

Fritz, S., & See, L. (2008). Identifying and quantifying uncertainty and spatial disagreement in the comparison of Global Land Cover for different applications. *Global Change Biology*, *14*, 1057–1075.

Gerlach, R., Skinner, L., Luckman, A., & Schmullius, C. (2005). *Deriving land cover information over Siberia: Comparing results from MERIS and MODIS.* MERIS and (A)ATSR user workshop, Frascati, Italy,26.-30. September 2005.

Giri, C., Zhu, Z. L., & Reed, B. (2005). A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, *94*, 123–132.

Groisman, P. Y., Clark, E. A., Kattsov, V. M., Lettenmaier, D. P., Sokolik, I. N., Aizen, V. B., et al. (2009). The Northern Eurasia Earth Science Partnership an example of science applied to societal needs. *Bulletin of the American Meteorological Society*, *90*, 671-+.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, *21*, 1331–1364.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Sohlberg, R., Dimiceli, C., & Carroll, M. L. (2002). Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data. *Remote Sensing of Environment*, *83*, 303–319.

Hansen, M. C., Townshend, J. R. G., Defries, R. S., & Carroll, M. (2005). Estimation of tree cover using MODIS data at global, continental and regional/local scales. *International Journal of Remote Sensing*, *26*, 4359–4380.

Heikkinen, J. E. P., Virtanen, T., Huttunen, J. T., Elsakov, V., & Martikainen, P. J. (2004). Carbon balance in East European tundra. *Global Biogeochemical Cycles*, *18*.

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

Herold, M., Woodcock, C., diGregorio, A., Mayaux, P., Belward, A. S., Latham, J., et al. (2006). A joint initiative for harmonization and validation of land cover datasets. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 1719–1727.

Houghton, R. A., Butman, D., Bunn, A., Krankina, O., Schlesinger, P., & Stone, T. A. (2007). Mapping Russian forest biomass with data from satellites and forest inventories. *Environmental Research Letters*, *2* (7pp.).

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

Justice, C. O., Belward, A., Morisette, J., Lewis, P., Privette, J., & Baret, F. (2000). Developments in the validation of satellite products for the study of the land surface. *International Journal of Remote Sensing*, *21*(17), 3383–3390.

Kharuk, V. I., Ranson, K. J., Kozuhovskaya, A. G., Kondakov, Y. P., & Pestunov, I. A. (2004). NOAA/AVHRR satellite detection of Siberian silkmoth outbreaks in eastern Siberia. *International Journal of Remote Sensing*, *25*, 5543–5555.

Kharuk, V. I., Ranson, K. J., Kuz'michev, V. V., & Im, S. (2003). Landsat-based analysis of insect outbreaks in southern Siberia. *Canadian Journal of Remote Sensing*, *29*, 286–297.

Krankina, O. N., Bergen, K. M., Sun, G., Masek, J. G., Shugart, H. H., Kharuk, V., et al. (2004a). Northern Eurasia: Remote sensing of boreal forests in selected regions. In G. Gutman, A. C. Janetos, C. O. Justice, E. F. Moran, J. F. Mustard, R. R. Rindfuss, D. Skole, B. L. Turner Ii, & M. A. Cochrane (Eds.), *Land change science: Observing, monitoring and understanding trajectories of change on the Earth's surface* (pp. 123–138). Berlin: Springer.

Krankina, O. N., Harmon, M. E., Cohen, W. B., Oetter, D. R., Zyrina, O., & Duane, M. V. (2004b). Carbon stores, sinks, and sources in forests of northwestern Russia: Can we reconcile forest inventories with remote sensing results? *Climatic Change*, *67*, 257–272.

Krankina, O. N., Pflugmacher, D., Friedl, M., Cohen, W. B., Nelson, P., & Baccini, A. (2008). Meeting the challenge of mapping peatlands with remotely sensed data. *Biogeosciences*, *5*, 1809–1820.

Krankina, O., Pflugmacher, D., Hayes, D. J., McGuire, A. D., Hansen, M., Haeme, T., et al. (2011). Vegetation cover in the Eurasian Arctic: Distribution, monitoring, and role in carbon cycling. In G. Gutman, & A. Reissell (Eds.), *Eurasian arctic land cover and land use in a changing climate* (pp. 79–108). Netherlands: Springer.

Kuemmerle, T., Chaskovskyy, O., Knorn, J., Radeloff, V. C., Kruhlov, I., Keeton, W. S., et al. (2009). Forest cover change and illegal logging in the Ukrainian Carpathians in the transition period from 1988 to 2007. *Remote Sensing of Environment*, *113*, 1194–1207.

Kuemmerle, T., Hostert, P., Radeloff, V. C., Perzanowski, K., & Kruhlov, I. (2007). Post-socialist forest disturbance in the Carpathian border region of Poland, Slovakia, and Ukraine. *Ecological Applications*, *17*, 1279–1295.

Kuemmerle, T., Hostert, P., Radeloff, V. C., van der Linden, S., Perzanowski, K., & Kruhlov, I. (2008). Cross-border comparison of post-socialist farmland abandonment in the Carpathians. *Ecosystems*, *11*, 614–628.

Kuemmerle, T., Radeloff, V. C., Perzanowski, K., & Hostert, P. (2006). Cross-border comparison of land cover and landscape pattern in Eastern Europe using a hybrid classification technique. *Remote Sensing of Environment*, *103*, 449–464.

Kuemmerle, T., Radeloff, V. C., Perzanowski, K., Kozlo, P., Sipko, T., Khoyetskyy, P., et al. (2011). Predicting potential European bison habitat across its former range. *Ecological Applications*, *21*, 830–843.

Latifovic, R., & Olthof, I. (2004). Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data. *Remote Sensing of Environment*, *90*, 153–165.

Loveland, T. R., & Belward, A. S. (1997). The IGBP-DIS global 1 km land cover data set, DISCover: First results. *International Journal of Remote Sensing*, *18*, 3291–3295.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., et al. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, *21*, 1303–1330.

Mayaux, P., Eva, H., Gallego, J., Strahler, A. H., Herold, M., Agrawal, S., et al. (2006). Validation of the Global Land Cover 2000 map. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 1728–1739.

MODIS land cover team (2003). Validation of the consistent year 2003. V003 MODIS land cover product. Available at:. http://geography.bu.edu/landcover/userguidelc/consistent.htm

Moody, A., & Woodcock, C. E. (1994). Scale-dependent errors in the estimation of land-cover proportions — Implications for Global Land-Cover datasets. *Photogrammetric Engineering and Remote Sensing*, *60*, 585–594.

Olofsson, P., Stehman, S.V., Woodcock, C.E., Friedl, M.A., Sibley, A.M., Newell, J.D., et al. (submitted for publication). A global land cover validation data set, I: Fundamental design principles. *International Journal of Remote Sensing*.

Pflugmacher, D., Krankina, O. N., & Cohen, W. B. (2007). Satellite-based peatland mapping: Potential of the MODIS sensor. *Global and Planetary Change*, *56*, 248–257.

Potter, C., Gross, P., Klooster, S., Fladeland, M., & Genovese, V. (2008). Storage of carbon in US forests predicted from satellite data, ecosystem modeling, and inventory summaries. *Climatic Change*, *90*, 269–282.

Randerson, J. T., Liu, H., Flanner, M. G., Chambers, S. D., Jin, Y., Hess, P. G., et al. (2006). The impact of boreal forest fire on climate warming. *Science*, *314*, 1130–1132.

Running, S. W., Nemani, R. R., Ann Heinsch, F., Zhao, M., Reeves, M., & Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *Bioscience*, *54*, 547–560.

See, L. M., & Fritz, S. (2006). A method to compare and improve land cover datasets: Application to the GLC-2000 and MODIS Land Cover products. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 1740–1746.

Skinner, L., & Luckman, A. (2004). Introducing a landcover map of Siberia derived from MERIS and MODIS data. *Geoscience and remote sensing symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International* (pp. 226).

Soja, A. J., Tchebakova, N. M., French, N. H. F., Flannigan, M. D., Shugart, H. H., Stocks, B. J., et al. (2007). Climate-induced boreal forest change: Predictions versus current observations. *Global and Planetary Change*, *56*, 274–296.

Stehman, S. V., & Czaplewski, R. L. (1998). Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sensing of Environment*, *64*, 331–344.

Strahler, A., Boschetti, L., Foody, G. M., Friedl, M. A., Hansen, M. C., Herold, M., et al. (2006). *Global land cover validation: Recommendations for evaluation and accuracy assessmlent of global land cover maps* (pp. 58). : European Communities.

Thomlinson, J. R., Bolstad, P. V., & Cohen, W. B. (1999). Coordinating methodologies for scaling landcover classifications from site-specific to global: Steps toward validating global map products. *Remote Sensing of Environment*, *70*, 16–28.

Townshend, J. R. G., Huang, C., Kalluri, S. N. V., Defries, R. S., Liang, S., & Yang, K. (2000). Beware of per-pixel characterization of land cover. *International Journal of Remote Sensing*, *21*, 839–843.

Turner, D. P., Cohen, W. B., & Kennedy, R. E. (2000). Alternative spatial resolutions and estimation of carbon flux over a managed forest landscape in Western Oregon. *Landscape Ecology*, *15*, 441–452.

Zhang, K., Kimball, J. S., Mu, Q. Z., Jones, L. A., Goetz, S. J., & Running, S. W. (2009). Satellite based analysis of northern ET trends and associated changes in the regional water balance from 1983 to 2005. *Journal of Hydrology*, *379*, 92–110.