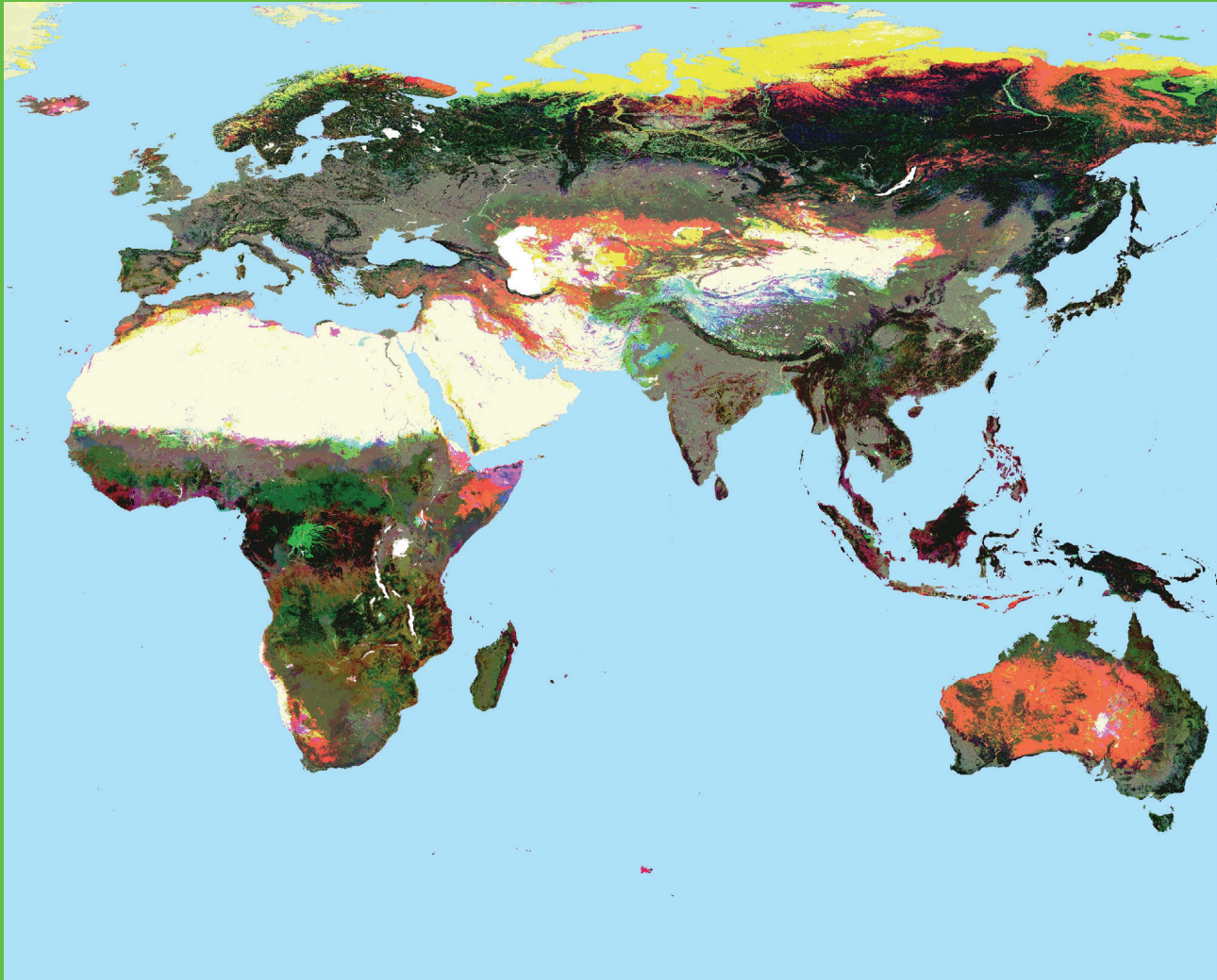


Global land cover map validation, comparison and integration for different user communities



Nandin-Erdene Tsendbazar

Global land cover map validation, comparison and integration for different user communities

Nandin-Erdene Tsendbazar

Thesis committee

Promotor

Prof. Dr M. Herold
Professor of Geo-information Science and Remote Sensing
Wageningen University

Co-promotor

Dr S. de Bruin
Associate professor, Laboratory of Geo-information Science and Remote Sensing
Wageningen University

Other members

Prof. Dr H.B.J. Leemans, Wageningen University

Prof. Dr P. Verburg, VU University Amsterdam, The Netherlands

Dr L. See, International Institute for Applied System Analysis, Laxenburg, Austria

Dr Z. Szantoi, Joint Research Centre, European Commission, Ispra, Italy

This research was conducted under the auspices of the C.T. de Wit Graduate School of Production Ecology & Resource Conservation (PE&RC)

Global land cover map validation, comparison and integration for different user communities

Nandin-Erdene Tsendbazar

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 20 May 2016
at 4 p.m. in the Aula.

Nandin-Erdene Tsendbazar

Global land cover map validation, comparison and integration for different user communities

164 pages

PhD thesis, Wageningen University, Wageningen, NL (2016)

With references, with summary in English

ISBN 978-94-6257-781-7

Table of Contents

1	Introduction.....	1
2	Assessing global land cover reference datasets for different user communities	13
3	Comparative assessment of thematic accuracy of GLC maps for specific applications using existing reference data	45
4	Spatial accuracy assessment and integration of global land cover datasets	71
5	Integrating global land cover datasets for deriving user specific maps.....	95
6	Synthesis	119
	References	135
	Summary	155
	Acknowledgements	158
	List of publications	160
	Short biography	162
	PE&RC Training and Education Statement	163

Chapter

1

Introduction

1.1 Background

Land cover (LC) is an essential environmental variable in understanding the interactions between mankind and nature. Change in LC is one of the main drivers of changes in earth systems such as climate (Verburg et al. 2011). For instance, the expansion of cropland and pasture at the cost of forests result in an increase of atmospheric carbon dioxide, and LC change can cause changes in land surface parameters and thus air temperature (Feddema et al. 2005; Verburg et al. 2011). Because of its importance, LC is identified as one of the Essential Climate Variables by the Global Climate Observing Systems (FAO 2008).

The observation of global-scale LC is of importance to international initiatives, governments, and scientific communities in their understanding and monitoring of the changes affecting environment and the coordination of actions to mitigate and adapt to global change. As such, reliable and consistent global LC (GLC) datasets are sought to be used for various applications such as climate models, ecosystem modelling, hydrological models and so on (Verburg et al. 2011).

In response to this need, a number of GLC maps have been produced from various initiatives and they differ in used input data, methodologies and validation (Herold et al. 2008). Due to the discrepancies among the GLC maps coupled with diverse user requirements, it is challenging for the users to select the GLC map that is the most suitable for their applications.

1.2 Global land cover maps

The advancement of remote sensing technologies in the last two decades has enabled the production of several GLC datasets. The first attempts to map GLC using remote sensing created 1 degree and 8 km coarse spatial resolution maps from AVHRR data (De Fries et al. 1998; DeFries and Townshend 1994).

Table 1.1. A summary of existing GLC maps and integrated maps

Dataset name	Spatial resolution	Temporal frame	Sensor	Classification scheme and classes	Accuracy %	References
<i>GLC maps</i>						
UMD				AVHRR		
1 degree	1 degree	1987			IGBP 12	(DeFries and Townshend 1994)
8 km	8 km	1984			IGBP 14	85 (DeFries et al. 1998)
1km	1 km	1992-1993			IGBP 14	(Hansen et al. 2000)
IGBP-DISCover	1 km	1992-1993		AVHRR	IGBP 17	67 (Loveland et al. 2000)
MODIS				MODIS	IGBP 17	
Collection 4	1 km	2001				71.6 (Friedl et al. 2002)
Collection 5	500 m	Yearly since 2001				74.8 (Friedl et al. 2010)
GLC2000	1 km	2000		SPOT4 VGT	LCCS 22	68.6 (Bartholomé and Belward 2005)
Globcover				MERIS	LCCS 22	
2005	300 m	2004-2006				73.1 (Bicheron et al. 2008)
2009	300 m	2009				67.5 (Bontemps et al. 2011a)
GLCNMO				MODIS	LCCS 20	
Version 1	1 km	2003				81.2 (Tateishi et al. 2011)
Version 2	500 m	2008				82.6 (Tateishi et al. 2014)
Land cover CCI	300 m			MERIS /SPOT VGT	LCCS 22	(CCI-LC 2014)
2000		1998-2002				
2005		2003-2007				
2010		2008-2012				70.8
FROM-GLC	30 m	circa 2010		Landsat TM, ETM+	9	63-65.5 (Gong et al. 2013)
GLC250m_CN	250m			MODIS	11	(Wang et al.)
2001		2000-2001				74.9
2010		2009-2011				75.1
Globeland30	30 m			LANDSAT TM, ETM7, HJ-1	10	80.3 (Chen et al. 2015)
2000		2000				78.6
2010		2010				80.3
<i>Integrated GLC maps</i>						
SYNMAP	1 km	2000		IGBP, GLC2000, MODIS	44	(Jung et al. 2006)
GLC-share	1 km	-		Various products	LCCS 11	82 (Latham et al. 2014)
Hybrid GLC	300 m	2005		Globcover 2005, MODIS, GLC2000	10	87.9 (See et al. 2015)
FROM-GLC Hierarchy	0.03-100 km	Circa 2010		FROM-GLC-agg, MODIS, Globcover	11	69.5-73.5 (Yu et al. 2014)
Integrated map	1 km	2000		MODIS, GLC2000, UMD	14	61.3 (Iwao et al. 2011)
Consensus LC	1 km	-		IGBP, GLC2000, MODIS2005, GlobCover	LCCS 12	(Tuanmu and Jetz 2014)
Integrated map	1 km	2000		MODIS, GLC2000, UMD, GLCNMO	6	74.6 (Kinoshita et al. 2014)

Following these efforts, several 1 km resolution GLC maps such as the IGBP-DISCover, UMD, MODIS and GLC2000 have been developed (Bartholomé and Belward 2005; Friedl et al. 2002; Hansen et al. 2000; Loveland et al. 2000). Medium resolution time-series satellite data enabled the development of 300m and 500m resolution GLC maps such as Globcover, MODIS and LC-CCI (Arino et al. 2007; CCI-LC 2014; Friedl et al. 2010). Recently, the release of Landsat archive further led to the production of high resolution (30m) GLC maps namely FROM-GLC and Globeland30 (Chen et al. 2015; Gong et al. 2013). A summary of existing GLC maps is provided in Table 1.1.

Efforts to improve GLC mapping have also focused on integrating existing GLC maps. Several integrated GLC maps were created by assigning LC classes based on properties such as map or per-class accuracy and agreement with other GLC maps (Iwao et al. 2011; Jung et al. 2006; Kinoshita et al. 2014; Tuanmu and Jetz 2014). Regional and global LC datasets were harmonized and combined to create an integrated GLC map: GLC-Share (Latham et al. 2014). See et al. (2015) used a reference dataset collected by volunteers to integrate GLC maps based on their local accuracies.

The available GLC maps differ from one another in terms of used data, methods and as well as spatial and thematic details (Table 1.1). Therefore, understanding the differences of available GLC map is important for the users. For this purpose, GLC maps are often harmonized in terms of thematic and spatial properties and they are compared against each other to assess spatial agreement (Fritz and See 2005; Herold et al. 2008; McCallum et al. 2006).

1.3 Global land cover map validation and reference datasets

Accuracy assessment of a thematic map is an important part of the map-making process (Olofsson et al. 2012). This may be done for different purposes such as to provide an overall measure of the quality of map and to evaluate the utility of the maps for specific applications (Congalton and Green 2009; Stehman 2000; Stehman et al. 2011).

In general, creating reference datasets and validating GLC maps follow three different steps: the sampling design, response design and analysis and estimation protocols (Stehman and Czaplewski 1998). Firstly, in sampling design protocol, aspects such as sampling units, sample size and sample selection scheme are defined to determine both the cost and statistical rigor of the accuracy assessment (Congalton and Green 2009). Secondly, aspects determining the agreement

between the map and reference classifications are reflected in the response design protocol. These include information sources for the reference classification, labelling and the definition of agreement between the reference and map classification. Since it is impractical to visit all the samples on the ground, satellite images are used to determine reference LC classes over large geographic areas (Wulder et al. 2006). Lastly, map accuracies are estimated in the analysis and estimation protocol. The standard way of reporting thematic accuracy of categorical maps is in the form of an error/confusion matrix to derive statistics such as overall and class specific accuracies and their confidence intervals (Congalton 1991; Foody 2002).

The quality of existing GLC maps has been assessed in different ways such as using independent reference datasets (Bontemps et al. 2011a; Mayaux et al. 2006; Scepan et al. 1999; Tateishi et al. 2011), cross validation with calibration datasets (Friedl et al. 2010; Strahler et al. 2003) or comparing with existing regional maps (Hansen et al. 2000). The thematic accuracy of current GLC maps has been reported to vary between 61 and 87% (Table 1.1).

Several independent reference datasets have been created to validate GLC maps and these include the IGBP-DIS, GLC2000, GlobCover and FROM-GLC and so on (Bontemps et al. 2011a; Gong et al. 2013; Mayaux et al. 2006; Scepan et al. 1999). In addition, other datasets that provide information on reference LC have also been generated. Examples are the training datasets for MODIS and GLCNMO maps, crowd-sourcing Geo-Wiki and View-IT datasets as well as other datasets such as FAO-FRA remote sensing survey datasets (Clark and Aide 2011b; Friedl et al. 2010; Fritz et al. 2009; Potapov et al. 2011; Tateishi et al. 2011). These datasets were created with different strategies in terms of sampling and response design protocols. The characteristics of existing and evolving reference datasets are summarized in Table 1.2.

Table 1.2. Characteristics of existing and evolving GLC reference datasets

Dataset name	Temporal frame	Classification scheme and class	Number of sample	Sampling scheme	Sample unit size	Labelled by	Reference
IGBP-DISCover	1990-1993	IGBP 16	379	Stratified random	1x1km	Regional experts	(Scepan et al. 1999)
GLC2000	1999-2002	LCCS 22	1265	2 stage stratified cluster	3x3km	Regional experts	(Mayaux et al. 2006)
Globcover		LCCS 22		Stratified random		International experts	
	2005 circa 2005		4258		1.5x1.5km		(Bicheron et al. 2008)
	2009 circa 2009		4164		0.9x0.9 km		(Bontemps et al. 2011a)
GLCNMO						Regional experts	
Version 1 validation	circa 2000	LCCS 20	600	Stratified random	1x1km		(Tateishi et al. 2011)
Version 1 training	2000-2003	LCCS 14	1607	No sampling	> 3x3km		
Version 2 validation	circa 2008	LCCS 20	904	Stratified random	500 x500 m		(Tateishi et al. 2014)
Version 2 training	circa 2008	LCCS 14	2080	No sampling	> 500x500m		
MODIS/STEP training	till 2014	IGBP 17	2762	No sampling	1 to 376 pixels	Regional experts	(Friedl et al. 2010)
LC-CCI	2000, 2005, 2010	LCCS 22	13000	2 stage stratified cluster	SSU ~1x1km	Automated pre-labelling and experts	(Achard et al. 2011)
Boston U./GOFC-GOLD	circa 2010	LCCS 12	500	Stratified random	5x5 km	Automated classification	(Olofsson et al. 2012)
VIIRS	Circa 2010	IGBP 17	circa 5000	Stratified and cluster	1 x1km	International experts	(GOFC-GOLD 2014)
FROM-GLC Training	Circa 2010	10	91433	No sampling	250x250m	International experts	(Gong et al. 2013)
FROM-GLC Validation			38664	Cluster			
Globeland30	2010	9	150000	2 stage stratified cluster	30x30m	International experts	(Chen et al. 2015)
GLC-Share	-	11	1000	Stratified random		experts	(Latham et al. 2014)
DCP		6	4200	Systematic	1 x1km	Volunteers and experts	(Iwao et al. 2006)
GEO-WIKI	Circa 2010	10	circa 18600	similar to FROM-GLC	1 x1km	Volunteers and experts	(Fritz et al. 2009; Schepaschenko et al. 2015)
VIEW-IT	2000-2010	7	46207	Stratified random and random	250x 250m	Volunteers	(Clark and Aide 2011b)
FAO-FRA	1990, 2000, 2005, 2010	9	13689	Systematic	10 x 10 km	Automated pre-labelling	(Potapov et al. 2011)

1.4 GLC map users and their requirements

Information on the state and dynamics of LC is crucial for a wide range of scientific purposes (Bontemps et al. 2011b). Appropriate GLC maps are increasingly sought for many applications, e.g. climate modelling (Herold et al. 2011; Verburg et al. 2011), forest resources assessments (FAO 2001), crop distribution modelling (You et al. 2009), agricultural land availability assessments (Cai et al. 2011) and biodiversity assessments (Joseph et al. 2009). As the purposes of applications vary, they have diverse requirements on GLC maps (e.g. spatial, temporal and thematic detail) and their quality assessments.

In terms of spatial detail, climate modellers typically use GLC maps at 1km spatial resolution or coarser (Kooistra et al. 2010) whereas this resolution is too coarse for LC change studies to detect small-scale changes, e.g. by forest logging (GOFCC-GOLD 2011). In terms of thematic detail, some applications require a single LC class such as forest, cropland and built-up area distribution (Feldpausch et al. 2012; Fritz et al. 2015; Klein Goldewijk et al. 2010). While multiple LC classes are used for applications such as carbon density mapping to create a global scale map based on per-class biomass values (Ruesch and Gibbs 2008). A more detailed information on specific LC classes can also be required by the users. For instance, LC classes representing the degree of human interactions such as cropland, urban and mosaics of these classes with other LC classes are useful for land system modelling (Asselen and Verburg 2012; Letourneau et al. 2012). In terms of temporal domain, studies dealing with land change and then climate change require multi-date GLC maps, while this is not a strict requirement for other studies (Herold et al. 2011; Hibbard et al. 2010).

The selection of GLC maps and their quality have a significant influence on the outcomes of these models (Hibbard et al. 2010; Nakaegawa 2011). For instance, Benítez et al. (2004) noted that the choice of GLC map influenced the model results by as much as 45% on the global cumulative carbon sequestration. Lower quality LC datasets can have strong effects on user applications (Ge et al. 2007; Sertel et al. 2010). The diverse user requirements should also be reflected in accuracy assessment of GLC maps and the reference datasets used for validation. To determine the fitness of GLC maps for certain applications, map accuracy assessments need to consider the perspectives of the users. Certain errors of GLC maps may be significant for specific users but not for others (DeFries and Los 1999). For example, confusion error between water and snow/ice may not be as important for biomass estimation, as it is for albedo estimation.

1.5 Research needs

Given the importance of GLC maps and better understanding of their quality for different applications, the issues of accuracy and inter-comparability of the different GLC maps have been a concern to many studies (Defourny et al. 2010; Herold et al. 2008; Herold et al. 2009b; Mayaux et al. 2006; McCallum et al. 2006). However, gaps still exist in consistency, comparability and validation of GLC maps, particularly from user's perspective (Foody 2002; Fritz et al. 2011b; Herold et al. 2008; Olofsson et al. 2012).

Considerable efforts have been made to validate GLC maps using independent validation datasets to provide statistically rigorous assessments (Bontemps et al. 2011a; Mayaux et al. 2006; Scepan et al. 1999; Tateishi et al. 2011). Despite efforts put into generating these reference datasets, their use is mainly limited to the original intended use and only a few studies reported re-using these datasets for other uses (FAO 2001; Göhmann et al. 2009). Currently, there is no assessment providing information on how these datasets can be used beyond their original scope and what the implications would be for specific user applications having different requirements on GLC maps and their validations. Such analyses is particularly important as international initiatives such as GOF-C-GOLD released several GLC reference datasets and proper use of these datasets is recommended (GOF-C-GOLD 2014).

Since GLC maps are used for a large number of applications, user-oriented accuracy reporting can help in understanding the uncertainty and limitations of GLC maps on specific applications (DeFries and Los 1999). However, studies on accuracy reporting from users' perspective of GLC maps are limited (DeFries and Los 1999; Mayaux et al. 2006). Furthermore, comparability of different GLC maps should be analysed considering the perspectives of the users. Most comparison studies focused on spatial (dis)agreement between different maps (Fritz et al. 2011b; Hansen and Reed 2000; McCallum et al. 2006). However, information on which map provides better quality for specific application is still lacking as a comparative accuracy assessment of GLC maps for specific applications does not exist till date.

The quality of GLC maps are commonly expressed by descriptive measures of overall accuracy and class specific accuracies from a confusion matrix (Mayaux et al. 2006; Scepan et al. 1999). These accuracies vary from region to region (Congalton 1988), and little information on the spatial variability of the GLC map accuracy is available so far. Spatial variation in map accuracies is important to

understand the regional GLC map quality and this information can be used to further improve the GLC maps. In particular, improved accuracy is required by the users as errors in GLC maps add to the uncertainties in the output of the applications (Bontemps et al. 2011b; Nakaegawa 2011). To improve GLC maps, existing GLC maps have been integrated based on their relative strengths (Iwao et al. 2011; Jung et al. 2006; Kinoshita et al. 2014). More research is needed to explore the benefit of using available reference datasets and spatial variation in map accuracy in improving GLC maps.

Another aspect of GLC mapping needing more research concerns map thematic flexibility. Current GLC maps provide a single fixed legend. These fixed legends do not always meet the requirements of users. Moreover, translating fixed legends, particularly mosaic classes, to required legends can be ambiguous or even impossible (Friedl et al. 2002; Herold et al. 2008). Therefore, the efforts to improve GLC maps should also consider the thematic requirements of different users.

1.6 Research objectives

This research aimed to address the gaps and problems identified in the previous section. The main objective of this thesis is to assess, compare and improve GLC maps while accounting for user requirements. Accordingly, the following research objectives and related research questions were considered:

1. To assess the re-usability of GLC reference datasets for different user communities
 - *What is the suitability of GLC reference datasets from the perspectives of different user communities?*
2. To compare the recent GLC maps for specific applications using existing reference data
 - *Which GLC map has the highest accuracy considering the perspectives of specific applications?*
3. To compare integration methods based on spatial variation in map accuracy and class presence using available GLC maps and reference datasets
 - *What are the local accuracies of the recent GLC maps?*
 - *Which integration method performs best for improving the recent GLC maps?*
4. To integrate GLC datasets to derive improved user-specific GLC maps

- *How can improved GLC maps be produced to address the thematic requirements of different users?*

1.7 Thesis outline

This thesis consists of six content chapters, including this Introduction Chapter. Chapters 2 to 5 address the research questions presented in Section 1.6. Figure 1.1 is a graphical outline of this thesis.

Chapter 2 addresses research question 1 by reviewing the characteristics of existing reference datasets and assess the suitability of re-using these datasets considering the requirements of different user applications.

Chapter 3 explores to answer research question 2 by reanalysing an existing reference dataset to make it suitable for comparing multiple maps and comparatively assessing the thematic accuracy of several GLC maps. In this, the importance of misclassification errors were also considered from the perspective of specific user applications.

Chapter 4 deals with research question 3 and 4 by assessing the spatial variation of accuracy of recent GLC maps and comparing different integration methods to improve GLC maps. Several existing reference datasets were used in this analysis.

Chapter 5 addresses research question 5 by creating an improved GLC map based on available GLC maps and reference dataset and demonstrating a concept of creating GLC maps with user-specific legends.

Chapter 6 summarizes the main results of this thesis in relation to the research questions, discusses its added value and limitations and presents recommendations for further research.

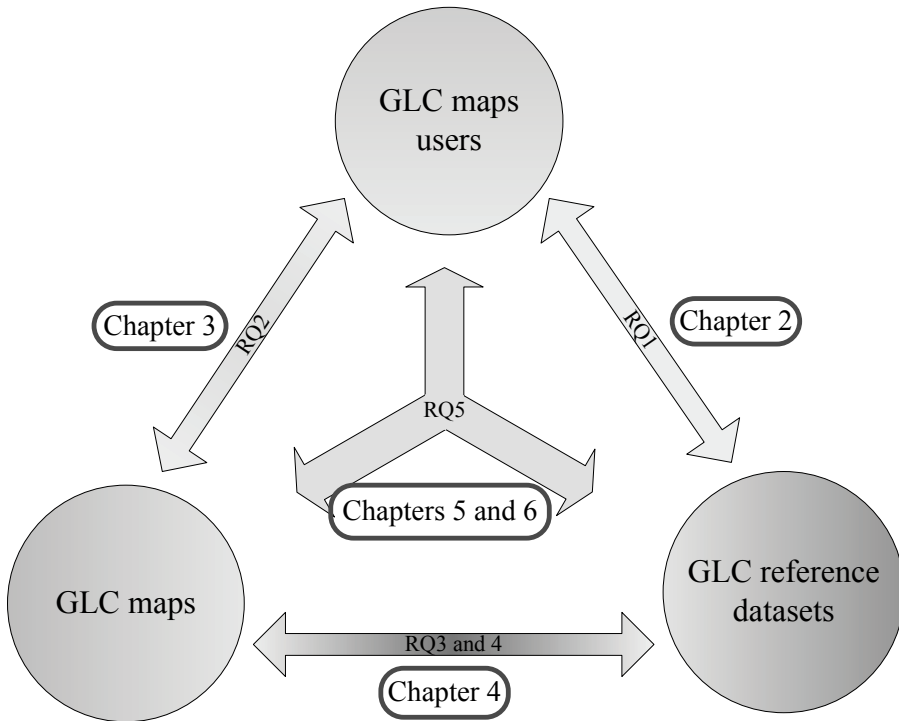


Figure 1.1. Overview of and links between the chapters of this thesis. RQ denotes research question

Chapter

2

Assessing global land cover reference datasets for different user communities

Nandin-Erdene Tsendbazar, Sytze de Bruin and Martin Herold

*Published in ISPRS Journal of Photogrammetry and Remote Sensing, Volume 10,
May 2015, Pages 93-114*

Appendices to this chapter can be found in the online publication.

Abstract

Global land cover (GLC) maps and assessments of their accuracy provide important information for different user communities. To date, there are several GLC reference datasets which are used for assessing the accuracy of specific maps. Despite significant efforts put into generating them, their availability and role in applications outside their intended use have been very limited. This study analyses metadata information from 12 existing and forthcoming GLC reference datasets and assesses their characteristics and potential uses in the context of 4 GLC user groups, i.e., climate modellers requiring data on Essential Climate Variables (ECV), global forest change analysts, the GEO Community of Practice for Global Agricultural Monitoring and GLC map producers. We assessed user requirements with respect to the sampling scheme, thematic coverage, spatial and temporal detail and quality control of the GLC reference datasets. Suitability of the datasets is highly dependent upon specific applications by the user communities considered. The LC-CCI, GOF-C-GOLD, FAO-FRA and Geo-Wiki datasets had the broadest applicability for multiple uses. The re-usability of the GLC reference datasets would be greatly enhanced by making them publicly available in an expert framework that guides users on how to use them for specific applications.

Keywords: Global land cover; validation; reference datasets; user groups; dataset suitability for users

2.1 Introduction

Observation of land cover at the global scale is essential for understanding and monitoring global change and for coordinating actions to mitigate and adapt to climate change (Herold et al. 2008). Information on these assessments is used by governments, scientific communities and international initiatives (Bontemps et al. 2011b). Many climate models, dynamic vegetation models, hydrological models, and carbon (stock) models use land cover datasets as one of the model inputs (Hibbard et al. 2010; Verburg et al. 2011). These users have many different requirements on global land cover (GLC) maps and their accuracy assessments. Climate modellers, for example, typically use GLC maps at 1-km spatial resolution or coarser (Kooistra et al. 2010) whereas this resolution is too coarse for GLC change studies to detect small-scale changes, e.g. by forest logging (GOFC-GOLD 2011). Accuracy assessment of GLC maps should account for these different user requirements and use suitable reference datasets in the assessments.

Over the last two decades, several global land cover (GLC) maps have been produced using remote sensing data, and GLC mapping is progressing towards higher spatial resolution datasets (Mora et al. 2014). The summary of current GLC products and their future trends are discussed in detail by Mora et al. (2014). GLC maps are commonly validated using higher-quality reference data, such as independent validation datasets and regional maps, or they are cross validated against training datasets (Friedl et al. 2002; Hansen et al. 2000; Mayaux et al. 2006). Currently, there are several independently-validated GLC datasets, namely IGBP-DIS, GLC2000, GlobCover5, GlobCover9, GLCNMO, and FROM-GLC (Bontemps et al. 2011a; Defourny et al. 2006; Gong et al. 2013; Mayaux et al. 2006; Scean et al. 1999; Tateishi et al. 2011). The accuracy of existing GLC maps typically varies between 67 and 81%, and while it is lower (10-50%) in some regions of the world (Frey and Smith 2007; Mora et al. 2014). The users of the GlobCover map and the Land Cover-Climate Change Initiative (LC-CCI) maps have stressed that the current quality of GLC maps should be improved (Bontemps et al. 2011b; Herold et al. 2011). The reason is that errors in GLC datasets add to modelling uncertainties; thus lower quality GLC datasets can have a strong impact on the final model outcomes (Nakaegawa 2011; Sertel et al. 2010).

The generation of reference datasets for accuracy assessment of GLC maps is a difficult task. The current reference datasets were generated by visual interpretation of satellite images, regional maps and geo-tagged photos. Enormous effort is required to analyse a large number of satellite images and interpret the land cover type. For example, 39 international experts interpreted 379 confidence sites of the IGBP-DIS dataset during a two-week workshop (Scean et al. 1999);

253 Landsat images were pre-processed, and international experts interpreted 1265 sample sites for the GLC2000 dataset (Mayaux et al. 2006); and 16 international experts completed on-screen collection of ground data for 4258 sample sites for the GlobCov5 validation dataset (Defourny et al. 2011b). Despite the efforts put into generating these reference datasets and the scarcity of validation data, their use is mainly limited to the original intended use and only a few studies reported re-using these datasets for other uses (FAO 2001; Göhmann et al. 2009).

In addition to the existing GLC validation datasets, there are a number of existing and evolving datasets that provide reference information for GLC maps and other global level assessments such as forest and agricultural studies. There are training datasets for MODIS and GLCNMO product generations (Friedl et al. 2010; Tateishi et al. 2011). Volunteer-based Geo-Wiki and View-IT datasets also provide an inexpensive way of collecting potentially useful land cover reference data (Clark and Aide 2011b; Fritz et al. 2009). Other datasets such as the FAO-FRA remote sensing survey focusing on a statistical sample-based assessment of global forest change have also been generated (Potapov et al. 2011).

The above mentioned reference datasets on GLC have potential of being used for applications outside their original scope. A thorough analysis of the efficient use of all available data for GLC map validation and calibration has not been previously investigated. To move towards the efficient use of reference datasets for GLC map validation, Olofsson et al (2012) proposed a new reference dataset that is created independent of any GLC map. Boston University and the Global Observation for Forest Cover and Land Dynamics (GOF-C-GOLD) are jointly generating a database for GLC map validation that can be augmented and used for different map validations (Olofsson et al. 2012; Stehman et al. 2012). However, the question of efficient use still remains for available and other upcoming GLC reference (GLCR) data. Currently, there is no assessment providing information on how these datasets can be used beyond their original scope and what the implications would be for specific user applications having different requirements on GLC maps and their validations. This situation can limit the value of the existing GLCR datasets for various applications and can hinder informed decision making concerning the usefulness of GLCR datasets for particular studies.

The objectives of this chapter are to (1) analyse the published literature to provide information on GLCR datasets and their user requirements, and (2) assess the potential uses and limitations of different GLCR datasets for four targeted GLC map user groups, namely the climate modelling community, global forest change analysts, the Group on Earth Observation (GEO) community of practice (CoP) for global agricultural monitoring and producers of improved GLC maps. Analysing

the characteristics of the GLCR datasets and their re-usability is important for understanding and reducing inconsistencies of reference datasets for the GLC mapping community. This will also help users of GLC datasets to make better-informed selections of reference datasets and reduce the uncertainty in their applications.

Section 2.2 reviews the characteristics of the GLCR datasets as well as the main requirements of GLCR for different users groups and describes the methods for assessing the suitability of the GLCR datasets for different user groups. The results of the assessments are provided in Section 2.3. Section 2.4 discusses the results and the main findings of the study are concluded in Section 2.5.

2.2 Methods

The general procedure followed for the analysis is shown in Figure 2.1. We reviewed the main dataset characteristics which are related to GLCR datasets (Section 2.2.1). Next, we compared the GLCR metadata (Section 2.2.2). The main user requirements were identified through a literature review (Section 2.2.3). Finally, the suitability of GLCR datasets for user groups was assessed by using the

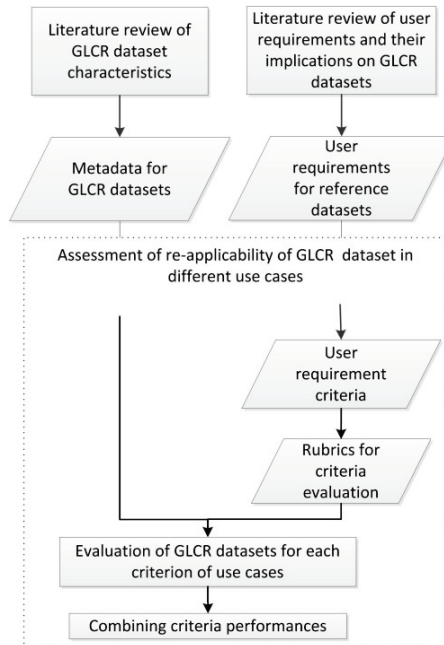


Figure 2.1. General analysis procedure

metadata of the GLCR datasets and the identified user requirements (Section 2.2.4).

2.2.1 Characteristics of reference datasets

Different characteristics related to the sampling and response design protocols need to be considered when generating and assessing reference datasets. Information from these protocols is used in the analysis and estimation protocols (Stehman and Czaplewski 1998). In addition, the thematic classification scheme determining the class legend, hierarchical classifiers and thematic detail of the land cover information need to be considered (Herold et al. 2006).

Sampling design is an important aspect determining both the cost and statistical rigor of the thematic map accuracy assessment (Congalton and Green 2009). An appropriate sampling unit, sample size and sample selection scheme should be considered carefully depending on the purpose, budget, and extent of the study. Area sampling units of pixels, blocks of pixels, and polygons can be used for accuracy assessment, and their choice may have an impact on sampling schemes and accuracy estimates (Stehman and Czaplewski 1998). The minimum mapping unit (MMU) is also specified since it can influence the size of the area sampling units (Stehman and Wickham 2011). For statistically sound accuracy assessments using limited resources, different approaches are adopted to determine the required sample size, e.g. using confidence intervals, allowable error, and significance levels (Foody 2009; Janssen and Van der Wel 1994). Design-based probability sampling schemes, e.g. simple random, stratified and systematic samples, are frequently used to select the samples for a credible accuracy assessment (Stehman 2000). Samples should have a known sample inclusion probability that is greater than zero (Stehman and Czaplewski 1998). The issue of choosing the right sampling scheme is extensively discussed in the literature (Stehman 2000; Stehman and Czaplewski 1998; Strahler et al. 2006).

All aspects of determining the agreement between the map and reference classifications are reflected in the response design protocol (Stehman and Czaplewski 1998). This includes information sources for the reference classification, the labelling protocol and the definition of agreement between the reference and map classification. Since it is impractical to visit all the samples on the ground, satellite images with temporal coverage closer to the target map are used to determine land cover classes over large geographic areas (Wulder et al. 2006). Generally, each sampling unit is assigned to a class using visual image interpretation. However, a visual interpretation can be subject to interpreter variability (different interpretation) and bias (incorrect interpretation) (Strahler et al. 2006). To reduce such errors, the confidence in interpretation is recorded, and it

is verified by other interpreters (Strahler et al. 2006). The issues of positional accuracy and heterogeneous areas also make reference data problematic and lead to uncertainty in the accuracy estimates of final map (Foody 2010; Powell et al. 2004). Therefore, transparency in reference data generation is useful for further analysis.

Final map accuracies are estimated in the analysis and estimation protocol (Stehman and Czaplewski 1998). The standard way of reporting thematic accuracy is in the form of an error/confusion matrix for categorical maps to derive statistics such as overall and class specific accuracies and their confidence intervals (Congalton 1991; Foody 2002). Detailed information on accuracy estimates can be found in the literature (Card 1982; Foody 2008; Stehman 1997).

2.2.2 Metadata comparison

This study examined the characteristics of the following existing GLCR dataset: IGBP-DIS, GLC2000, GlobCover 2005 and 2009, and GLCNMO. In addition, we considered the MODIS and GLCNMO training datasets as well as the FAO-FRA remote sensing survey datasets, which could also serve as a reference dataset for GLC map validation. The datasets under development for LandCover CCI map validations and by Boston University/GOFC-GOLD as well as volunteer-based “crowd-sourcing” datasets, such as Geo-Wiki and View-IT, were also included. For detailed information about these datasets, references are provided in Table 2.1.

Table 2.1. Reference datasets used for assessment

Dataset	Abbreviation	Current state	Source of information
IGBP-DISCover	IGBP-DIS	Existing	(Scepan et al. 1999)
GLC 2000	GLC2000		(Mayaux et al. 2006)
GlobCover 2005	GlobCov5		(Bicheron et al. 2008; Defourny et al. 2011b)
GlobCover 2009	GlobCov9		(Bontemps et al. 2011a)
GLCNMO validation and training dataset	GLCNMO-tr/val		(Tateishi et al. 2011)
MODIS training (STEP dataset)	MODIS-tr		(Friedl et al. 2010; Strahler et al. 2003)
FAO-FRA	FAO-FRA		(FAO and JRC 2012; Potapov et al. 2011)
GEO-WIKI	Geo-Wiki	On-going	(Fritz et al. 2009; Fritz et al. 2011a)
VIEW-IT	View-IT		(Clark and Aide 2011b)
LC-CCI	LC-CCI		(Achard et al. 2011)
Boston U./GOFC-GOLD	GOFC-GOLD		(Olofsson et al. 2012)

Based on the characteristics discussed in Section 2.2.1, we reviewed and compared the characteristics of the GLCR datasets. We also reviewed the current use of the

existing GLCR datasets and derived accuracy estimates. We considered the following characteristics of the GLCR datasets for comparison (Table 2.2).

Table 2.2. The characteristics of reference datasets used for assessment

Legend	Sampling design	Response design	Current use
<ul style="list-style-type: none"> • Classification scheme • Number of class • Classifier information provided 	<ul style="list-style-type: none"> • Sample unit type and size • Sample size • Sample selection scheme • Sample stratification • Inclusion probability • MMU 	<ul style="list-style-type: none"> • Source of information • Temporal coverage • Location accuracy • Labelling procedure • Sample verification • Confidence in interpretation • Majority classes and their fraction 	<ul style="list-style-type: none"> • Intended application • Other applications • Applied pre-processing • Derived accuracy estimates

2.2.3 Users of GLC validation and their requirements

Information on the state and dynamics of land cover is crucial for a wide range of scientific purposes (Bontemps et al. 2011b). Appropriate GLC maps are increasingly sought for many applications, i.e. climate modelling (Herold et al. 2011; Verburg et al. 2011), forest resources assessments (FAO 2001), crop distribution modelling (You et al. 2009), agricultural land availability assessments (Cai et al. 2011) and biodiversity assessments (Joseph et al. 2009). These applications have different requirements for GLC datasets and their validation (e.g., in terms of thematic and spatial details). In addition to varying user requirements, the existing GLC maps and reference datasets also have various characteristics resulting from different interests, mapping standards, methodologies and so on (Herold et al. 2008). It is important for users to make an informed decision when selecting appropriate GLC datasets to meet their requirements. This section provides a review of the requirements of the main users of GLC datasets and the implications for the map validation procedure, particularly for GLCR datasets. Such information can be useful for increasing the re-usability of GLCR datasets for applications by different user groups.

2.2.3.1 *Climate modelling community (ECV)*

Systematic land cover observation is essential for the climate modelling community (CMC) as land cover is identified as 1 of the 13 terrestrial Essential Climate Variables (ECV) by the Global Climate Observing Systems (GCOS) (FAO 2008). To support the work of UNFCCC and other international bodies on mitigation and adaptation to climate change, continuous monitoring of ECV to reduce uncertainties of the global climate system is required. In general, CMCs use land cover observations in the following ways (Herold et al. 2011):

- As a proxy for plant functional types (PFTs) related to land surface parameters
- As a proxy for tracking human activities
- As datasets for validating model outcomes or studying feedback effects.

Herold et al. (2011) systematically analysed the user requirements of CMCs. The main CMCs such as GCOS, Global Terrestrial Observing System (GTOS) and Climate Modelling User Group (CMUG) stressed several requirements for GLC maps and their validation to improve climate modelling and reduce uncertainties. Firstly, statistically rigorous and independent GLC map validation is a prerequisite to determine the map fitness for CMCs (Bontemps et al. 2011b). Thus, a map validation procedure should follow internationally agreed and repeatable methods following the best practices suggested by the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration and Validation (WGCV) (Strahler et al. 2006). The uncertainties of GLC data should be quantified and reduced as much as possible for the best and continuous observation of land cover ECV. A maximum of 15% error for the GLC maps is allowed by the CMCs (Herold et al. 2011). Furthermore, CMCs require continuous monitoring of GLC, and the need for stability in accuracy should be reflected by implementing multi-date accuracy assessments (Bontemps et al. 2011b). The stability of multi-date data needs to be at least 85-90% (Herold et al. 2011). This highlights that reference datasets with stable multi-date records (no un-realistic land cover changes in time) that can provide accurate error estimates of the GLC data and uncertainty assessments are essential for CMCs (Herold et al. 2011; Hibbard et al. 2010).

As GLC datasets are used as a proxy for PFTs, the relative importance of different class accuracies strongly varies depending on which parameter and PFT are estimated. However, the CMCs stressed that class-specific accuracies should be improved, which is very critical in grassland, shrub-land and wetland areas (Frey and Smith 2007; Poulter et al. 2011; Quaipe et al. 2008). Moreover, CMCs also require consistency and comparability of land cover characterisation among the GLC datasets, and the use of a Land Cover Characterisation System (LCCS) is supported since it is currently the most comprehensive, internationally applied and flexible framework for land cover characterization (Herold et al. 2011). To summarize, CMC applications require GLC datasets and their accuracy assessments that are statistically rigorous, stable in multi-date assessment and flexible thematic characterization. Reference datasets used for the accuracy assessments should also meet these requirements.

2.2.3.2 Global forest change analysts

Several globally consistent assessments of forest area and change have been carried out using wall-to-wall (FAO 2001; Hansen et al. 2010; Townshend et al. 2012) and sample-based approaches (Achard et al. 2002; FAO 2007). As forest change is an important factor for the global carbon cycle, the information obtained from such assessments may be used in efforts to reduce emissions from deforestation and forest degradation (GOFC-GOLD 2011) as well as for different applications, e.g., global carbon sequestration modelling and carbon budget accounting and emission estimations (Benítez et al. 2004; Harris et al. 2012). Therefore, accuracy estimates and uncertainty information on forest change area are crucial for dataset comparability and reducing the uncertainty for further applications (Olofsson et al. 2013).

Forest area and change estimation can be conducted using different approaches, e.g. (1) map pixel counting that is adjusted and improved by accounting for misclassification error using reference samples (de Bruin 2000a; GOFC-GOLD 2011), (2) post-classification comparison of multiple maps (Olofsson et al. 2013; Van Oort 2005), (3) a dedicated sampling survey with area sample units (e.g. 5x5 km, 10x10 km) (Eva et al. 2010; Gallego and Stibig 2012), (4) change detection analyses using multi-date satellite images (Townshend et al. 2012) and (5) calibration of a wall-to-wall assessment using sample-based estimates (Hansen et al. 2010). The GLCR datasets can be used for assessing global forest change products, as well as improving their area estimation (FAO 2001; GOFC-GOLD 2011; Olofsson et al. 2013). Accuracies of such forest area and change estimates should be assessed and incorporated into the final estimates using reference datasets, and the uncertainty estimates should also be provided in the form of confidence intervals and sampling standard errors, etc. (McRoberts 2010; Olofsson et al. 2013).

The preferred sample selection schemes are clustered, stratified sampling or systematic sampling for forest area and change analyses (Achard et al. 2002; Gallego and Stibig 2012; GOFC-GOLD 2011). Stratification could be based on knowledge driven or remote sensing based characterisation of possible forest change areas (Achard et al. 2002; Hansen et al. 2010; Strahler et al. 2006). A suitable forest definition should be considered since various forest definitions exist (GOFC-GOLD 2011). Reference datasets based on medium to high-resolution satellite images would be suitable as 1 to 5 ha MMU is commonly accepted for forest change monitoring (GOFC-GOLD 2011). The errors in the reference data have a great influence on the change detection accuracy (Foody 2010), thus statistically rigorous and high-quality reference data should be promoted for forest monitoring (GOFC-GOLD 2011).

In summary, global forest change analysts require reference datasets that have statistical rigorousness, stability in multi-date records, a suitable sample selection scheme and forest definition, and spatial detail suitable for change detection.

2.2.3.3 GEO Global Agricultural Monitoring CoP

Agriculture is one of the main concerns and areas of societal benefits of the GEO (GEO 2012). It is essential for meeting the food demands of a rapidly growing population (Cassman et al. 2005) and is a factor in many environmental issues, e.g. tropical deforestation, biodiversity loss (Foley et al. 2005; Sachs et al. 2010). Aiming to support sustainable agriculture, the GEO implementation plan targets global mapping, monitoring and modelling of changes in agricultural land use and type (Béquignon et al. 2010). As such, international organizations, e.g. Joint Experiment for Crop Assessment and Monitoring and the Global Agriculture Monitoring Project, are working to facilitate the comparison and assessments of monitoring and modelling methods and to provide a continuous global agricultural monitoring system (Becker-Reshef et al. 2010; JECAM 2012).

GLC maps are often used in combination with other statistical surveys to create global cropland products, e.g. the M3-Cropland layer for 2000 (Ramankutty et al. 2008), the global map of yield and type of 175 distinct crops for 2000 (Monfreda et al. 2008) and a consolidated community global cropland map by synergizing GLC maps with various regional maps (Fritz et al. 2013; See et al. 2012). GLC datasets can provide the distribution of agricultural land cover; however, the capability to obtain information about agricultural land-use practices, such as crop type, yield and fertilizer use, is very limited from remote sensing based datasets (Monfreda et al. 2008; Pittman et al. 2010). Furthermore, the accuracy of GLC maps in capturing cropland classes varies from 56-76% for some GLC products, and inconsistencies in their estimates of cropland distribution and area are also evident (Fritz et al. 2011b).

Accuracy assessments of GLC products providing information on the confidence of cropland area estimates and problematic areas are crucial for the uncertainty estimation of global cropland products. GLCR datasets can be used for improving and validating global cropland products. However, attention has to be paid to the definition of cropland and pasture (Goldewijk et al. 2007; Ramankutty et al. 2008). In addition, reference datasets having information on the main crop and management types are desirable since information on crop type is often lacking in GLC maps. Mosaic cropland classes cause uncertainty in cropland area estimation (Goldewijk et al. 2007), thus reference datasets with improved representation of mixed-unit classes or information on majority classes and their fraction could be useful (Fritz et al. 2011c; Herold et al. 2009a). Attention also needs to be given to

quality assurance as well as the rigorous validation and comparison of global agricultural datasets (Becker-Reshef et al. 2010; Fritz et al. 2013).

Thus, global agricultural monitoring needs a sample that affords area estimation and have thematic detail in agricultural land cover, suitable definitions of cropland and pasture, and information on mosaic classes (majority classes).

2.2.3.4 Producers of improved GLC maps

The producers of GLC maps can also be regarded as users of GLC validation and reference data when the data are used for improving land cover maps. The GLC map producers underlined that the GLC products only reached an overall accuracy around 70% despite the significant developments in technology and methodology as well as the user requirement for improved accuracies (Defourny et al. 2012; Fritz et al. 2011b; Herold et al. 2011). General inability and inconsistency of GLC mapping approaches to clearly discriminate mixed trees, shrubs, and herbaceous vegetation have been noted (Herold et al. 2008; Huttich et al. 2011; Jung et al. 2006).

To overcome this issue, in addition to analysing and improving the uncertainty components of GLC maps summarized by Defourny et al. (2012), the available reference datasets could also be used in different ways. Reference datasets can be utilized to map the spatial accuracy and classification uncertainty of thematic maps for understanding error sources and describing map limitations (Kyriakidis and Dungan 2001; Powell et al. 2004; Strahler et al. 2006). The GLCR datasets could also be used to improve land cover maps using various approaches, i.e., map training using a large number of samples, especially at the global scale where one class can have different spectral signatures (Li and Zhang 2011; Loveland et al. 2000), map calibration by analysing the posterior probability and reference data (Magnussen and de Bruin 2003) and creating synthetic maps based on several GLC maps (Fritz et al. 2009; Göhmann et al. 2009). A large number of samples having uniform spatial coverage and representation are promising for these purposes (Brus and De Gruijter 1997; Brus 2011).

The key areas of improvement in GLC mapping also include a proper definition of mixed-unit classes using LCCS and an appropriate treatment of rare classes, e.g. urban and wetland classes (Defourny et al. 2011b; Herold et al. 2009a; Herold et al. 2008). Moreover, inconsistency among different GLC datasets is often attributed to landscape heterogeneity (Jung et al. 2006; McCallum et al. 2006; Pérez-Hoyos et al. 2012). Thus, reference datasets derived from fine-resolution satellite data and information on class proportions in heterogeneous samples could

be helpful to deal with this issue. The quality of reference datasets is also the main issue to examine when the aim is to improve GLC maps and their validation.

Based on the above, a high sampling density and spatial coverage, detailed representation and flexible characterisation of land cover classes including rare classes and the characterisation of heterogeneous areas were deemed essential requirements from the map producer's perspective.

2.2.4 Re-usability assessment of GLCR datasets for different use cases

This section focuses on the assessment of the GLCR datasets for their re-usability for the applications of four targeted GLC map user groups, e.g., CMC, global forest change analysts, GEO Global Agricultural Monitoring CoP and improved GLC map producers. Applications included calibration, validation and parameter (e.g. area) estimation of user-oriented global-scale maps that are available or being produced. The specific cases of four user-groups are summarized in Table 2.3.

We used the upcoming Land cover-ECV GLC products as an example application for CMC since these map are being generated specifically for this group. Three consecutive maps are being created under the European Space Agency programme, Climate Change Initiatives, for global monitoring of ECVs to fulfil the requirements of GCOS (Defourny et al. 2011a). The currently available Global Forest Cover Loss (GFLC) map was adopted as an example case for global forest change estimation. Annual MODIS images for 2000 to 2005 with 500 m pixels were used to provide forest change hotspot areas on this map (Hansen et al. 2003). The beta version of the Global Hybrid Cropland Map (GHCM) was selected for the user group of GEO Global Agricultural Monitoring CoP. This map was generated through a mapping subtask of the GEO-Global Agricultural Monitoring System task led by IIASA (Fritz et al. 2013). A number of national and regional land cover products and crop statistics were combined synergized with GLC products to provide a global cropland map (Fritz et al. 2011c). The specifications of the upcoming GLC map, using recently launched Visible Infrared Imaging Radiometer Suite (VIIRS) satellite data, were used as an example of producing an improved GLC map. We analysed the suitability of GLCR datasets for calibrating and validating the VIIRS GLC map with the focus of increasing the quality of the map.

Table 2.3. Example cases of different user groups

	Climate modelling community	Global forest change analysts	GEO Global Agricultural Monitoring CoP	Producers of improved GLC maps
Focus/aim of evaluation	Map validation	Map validation and calibration of area estimates	Map validation and calibration of area estimates	Map calibration and validation
Map used for evaluation	Land Cover-ECV	Global Forest Cover Loss map (GFCL)	Beta version of Community based Global Hybrid Cropland map	VIIRS
Corresponding year	2000, 2005, and 2010	2000-2005	2000, 2005, and 1995-2000, 2010 (various temporal extents)	2012
Spatial resolution	300 m for 2005 and 2010; 1 km for 2000	500 m	1 km	1 km
Thematic detail and classification scheme	LCCS 22 classes	Tree percentage	Cropland percentage	IGBP 17 classes
Reference	(Defourny et al. 2011a)	(Hansen et al. 2003; Hansen et al. 2010)	(Fritz et al. 2013; See et al. 2012)	(NASA 2013)

We evaluated the characteristics of the datasets against the user requirements. We selected user specific requirements which are reviewed in Section 2.2.3. We also included a general requirement namely the quality flag information of the GLCR datasets that is relevant to all the user groups considered. Moreover, map related requirements such as suitability in temporal coverage and spatial resolution also apply to all the user groups and are included in the analysis. Criteria for evaluating according to these user requirements were derived and the GLCR datasets were assessed for each criterion. User requirements and criteria of four user groups are summarised in Table 2.4.

Table 2.4. Requirements and specific criteria on reference datasets of the four GLC dataset user groups

		User groups						
Theme	Climate modelling community		Global forest change analysts		GEO Global Agricultural Monitoring CoP		Producers of improved GLC map	
	Requirements	Criteria	Requirements	Criteria	Requirements	Criteria	Requirements	Criteria
Sampling scheme	Statistically rigorous (accurate and precise) overall and class specific accuracy estimates	Probability sampling	Statistically rigorous and high sampling density	Probability sampling	Statistically rigorous	Probability sampling	High sampling density and spatial coverage	Sample size and spatial coverage
		Class representation		Class representation				
Sampling scheme	Compatibility in combining and augmenting the samples	Easily combined and augmented sampling scheme	Sample selection scheme	SRS or SYS ² scheme used. Stratification based on potential forest change areas	Affords area estimation	Affords area estimation	Detailed representation of thematic classes especially rare classes	Class representation
			Affords area estimation	Affords area estimation				
Thematic detail	Flexible thematic cover characterization	Classification scheme and thematic detail of the legend (W = 0.167)	Thematic detail (forest and non-forest classes)	Forest / non forest classes available	Thematic detail classification scheme for agricultural land covers	Classification scheme for agricultural land cover classes (W = 0.25)	Flexible thematic cover characterization	Classification scheme and thematic detail of the legend (W = 0.167)
		Hierarchical classifiers	Forest definition	Forest definition	Definition of cropland and	Classifier with information on crop-type and management-type (W = 0.75)	Cropland definition	Hierarchical classifiers

		provided (W = 0.833)			pasture			provided (W = 0.833)
					Majority classes and its fraction cover	Majority classes and its fraction cover		
Spatial and temporal detail	Temporal coverage	Temporal coverage	Multi-date sample	Multi-date sample	Spatial resolution	Spatial resolution	Spatial resolution	Temporal coverage
	Stability over multi-date datasets	Stable multi-date sample	Temporal coverage	Temporal coverage	Temporal coverage	Temporal coverage	Temporal coverage	Temporal coverage
	Spatial resolution	Spatial resolution	Suitable spatial detail for change detection	Suitable spatial detail for change detection	High (6.25 ha) spatial detail for change	High (6.25 ha) spatial detail for change	Spatial resolution	Spatial resolution
							Ability to characterise heterogeneous areas	Spatial detail and homogeneity (W = 0.833) Major classes and their fraction recorded for heterogeneous areas (W = 0.167)
Quality checking	Quality flag information available	Verified (W = 0.833) Interpretation confidence recorded (W = 0.167)	Quality flag information available	Quality flag information available	Verified (W = 0.833) Interpretation confidence recorded (W = 0.167)	Quality flag information available	Quality flag information available	Verified (W = 0.833) Interpretation confidence recorded (W = 0.167)

¹ W – Criterion weights for trade off criteria

² SRS – Stratified random sampling; SYS – Systematic sampling

We assigned scores ranging from 1 (unsuitable) to 5 (highly suitable) to the GLCR datasets for each criterion of the user groups according to detailed rubrics that are provided in Table A.1 (see Appendix in the online publication).

The criteria scores were aggregated to derive one indicator value representing the dataset re-usability for given use cases. Firstly, all criteria were divided into trade-off and non-trade-off criteria since some criteria can be so essential that the strong performance of one criterion cannot compensate for the weak performance of another criterion. For instance, good performance of the class representation criterion cannot compensate for the weak performance of the stable multi-date sample criterion. We aggregated trade-off criteria scores using a weighted linear combination method (Malczewski 2000), and weights indicating the criteria importance were derived using Analytical Hierarchy Process (pairwise comparisons) developed by Saaty (1977). For example, the requirement of a flexible thematic land cover characterisation is represented by criteria on the classification scheme and thematic detail and presence of information on the classifier (Table 4). Since presence of information on classifiers is very important to make the legend adaptable to other legends, this criterion received five times more weight than the classification scheme and thematic detail. For example, if a dataset is assigned a score 3 (moderately suitable) to criterion of classification scheme and thematic detail and a score 2 (marginally suitable) to presence of classifiers information, the composite score for flexible thematic land cover characterisation equals $3 \times 0.167 + 2 \times 0.833 = 2.167$ (rounded of to 2: marginally suitable). The weights of the trade-off criteria are calculated similarly; details are provided in Table 2.4.

Finally, composite codes of trade-off criteria and all non-trade-off criteria were aggregated using the Ordered Weighted Averaging procedure of Yager (1988). With the trade-off and ORness measures of this approach, analysts can control the criterion substitutability and involved risk of the analysis results respectively. For non-trade-off criteria, the Boolean operators AND (minimum) and OR (maximum) can be used (Malczewski 2006). The AND-type combination can be associated with a pessimistic or risk-averse approach while the OR-type combination represents an optimistic or risk-taking strategy (Drobne and Liseč 2009). In this analysis, the AND operator that yields a low-risk result with no trade-off was used. The criterion with a minimum score was given a weight of 1, and all others had zero weights in the aggregation process. In other words, if a dataset is evaluated as unsuitable on one criterion, the final composite code will also be unsuitable.

2.3 Results

2.3.1 Metadata comparison

The thematic classification scheme is important for legend harmonization for integrating the datasets. Table 2.5 shows different thematic classification schemes adopted for the datasets. Among them, IGBP and LCCS were the main classification schemes employed.

Table 2.5. Classification schemes of the datasets

Datasets	IGBP	LCCS	Other	Number of classes
IGBP-DIS	✓			16
GLC 2000		✓		22
GlobCov5		✓		22
GlobCov9		✓		22
GLCNMO-val		✓		20
GLCNMO-tr		✓		14
MODIS-tr	✓			17
FAO-FRA		✓		9
LC-CCI		✓		22
GOFC-GOLD		✓		12
GEO-Wiki	✓	✓		17-22
VIEW-IT			✓	7

Some datasets such as GLC2000, GlobCover and IGBP were interpreted according to the hierarchical classifiers and translated into the map legends to enable the use of the datasets for other purposes or maps. In terms of the thematic detail, most reference datasets using LCCS have 20-22 classes (except GLCNMO-tr and FAO-FRA), but the IGBP-DIS and MODIS-tr datasets have 16-17 classes based on the IGBP scheme (Table 2.5). FAO-FRA, GOFC-GOLD and View-IT datasets have less thematic detail. Metadata comparisons of the GLCR datasets are described in detail, i.e., in terms of sampling and response design protocols as well as the current uses.

2.3.1.1 Sampling design

In this section, we present the results from the comparison of the characteristics related to the sampling design protocol. Figure 2.2 demonstrates the sample unit type and sizes of the datasets. Pixels, blocks of pixels, polygons and area units were used as sampling units for the reference datasets. IGBP-DIS, GLCNMO-val, VIEW-IT and Geo-Wiki datasets have a pixel sampling unit, whereas GLC2000 and GlobCover datasets used blocks of pixels. GLCNMO-tr and MODIS-tr datasets have polygon units with areas ranging from 0.09 km² to over 100 km² and varying shapes. The FAO-FRA and GOFC-GOLD datasets adopted

area units. Because these area units are not dependent on a pixel or block of pixels of any GLC maps, they can be easily applied for validating different GLC maps. These area sampling units have large spatial support areas, but they can be flexibly decreased since they have small MMUs (4 m and 5 ha respectively).

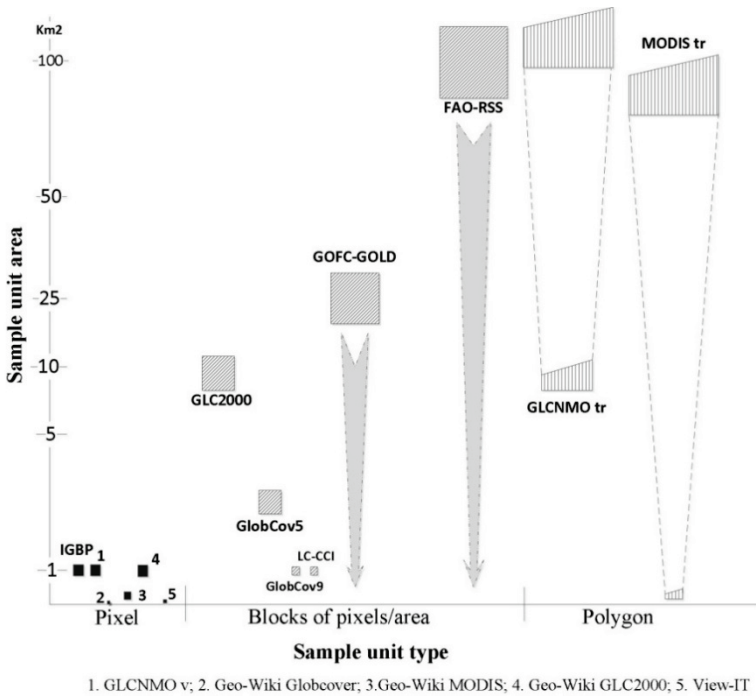


Figure 2.2. Sample unit type and spatial support area of GLCR datasets (dashed arrows - flexible spatial support area, dashed lines - varying spatial support area)

The sample size and sample selection scheme of the datasets can be found in Table A.2 (see Appendix in the online publication). Most GLCR datasets used stratified random sampling emphasizing class specific accuracies. IGBP-DIS, GOFC-GOLD and GLCNMO-val have sizes of 379 to 600 samples. The sample size of the GOFC-GOLD dataset can be increased if small secondary sampling units are used. The GlobCover datasets have around 4200 samples each, and View-IT has the highest sample size number (~46000). The FAO-FRA dataset employed systematic sampling of 13000 samples at each latitude and longitude intersection.

Two-stage stratified clustered sampling, highlighting the priority classes and landscape complexity in a cost-effective way, was implemented for the GLC2000

dataset. A similar scheme was also adopted for the LC-CCI dataset. 2600 primary sampling units of this dataset were selected by sub-sampling of the FAO-FRA data and 5 secondary sampling units were selected systematically from each primary samples accounting total sample size of around 13000. Some datasets did not use a probability sampling scheme and thus statistical sample-based assessments cannot be drawn. These include the training datasets for GLCNMO and MODIS which have selective samples for training of land cover classification. However, they have substantial sample sizes ranging from 1600 to 1860. The Geo-Wiki dataset has a large sample size (10 500 by October 2012) that is updated and increased as volunteers interpret more samples.

Stratum allocation of the samples varies depending on the class abundance and priority, except for the IGBP-DIS and GLCNMO datasets that allocate 25 to 30 samples in each stratum. In all probability sampling schemes, sample inclusion probabilities are different per stratum (Table A.2). All stratified sampling based datasets (except GOFD-GOLD) used stratifications that are targeted to particular land cover maps (Table A.2).

2.3.1.2 Response design

The comparison of information sources for the GLCR datasets shows that satellite imagery, open source maps, geo-tagged photos and other regional maps are the main sources of information (Table 2.6). Mid-resolution Landsat and Spot images together with time-series coarse-resolution NDVI profiles from Spot-VGT, MERIS and MODIS sensors are used to address spatial and temporal variability of land cover classes. The possibility of very-high resolution satellite images in conjunction with Virtual or Google Earth/maps and other open source maps complements the reference data sources as well. Table 2.6 also shows the importance of geo-tagged photos as source data for the GLCR datasets.

Other than the spatial resolution of source data, the quality of these data, particularly positional accuracy, is important. Positional accuracies of source data for the GLCR datasets were not documented except for the IGBP-DIS, GLC2000 and FAO-FRA datasets (Table A.2). Spatial coverage of reference datasets is still limited in some classes or parts of the world, for example, for the IGBP-DIS (Snow/Ice and Water body classes) and GlobCover datasets (Pakistan, Afghanistan, Iran, Japan, Colombia, Central America and Eastern Brazil).

Table 2.6. Reference data sources of the validation datasets.

	NDVI profile from Spot-VGT, MERIS, MODIS	Landsat, Spot, Aster imagery	Classification of Quickbird, Geo-Eye imagery	Open source Google Earth, Open street, Bin and Yahoo maps	Geo-tagged photos and Confluence photos	Other regional maps and aerial photographs
IGBP-DIS		✓				
GLC 2000		✓				✓
GlobCov5	✓			✓	✓	
GlobCov9	✓	✓		✓	✓	✓
GLCNMO-val	✓			✓	✓	✓
GLCNM- tr	✓	✓		✓	✓	✓
MODIS-tr		✓		✓		
FAO-FRA	✓	✓				
LC-CCI	✓	✓		✓		
GOFC-GOLD			✓			
GEO-WIKI	✓			✓	✓	
VIEW-IT	✓			✓	✓	

Figure 2.3 compares the temporal coverage of source data for the GLCR datasets. The GLCR datasets used source data from 1987 onwards. Source data for many datasets are centred around the year 2000. Geo-Wiki and View-IT datasets provide greater temporal coverage; however, land cover changes may occur within this time frame, and this should be taken care. Time stamp on each sample record would be useful for this, however, information on the availability of such time stamps in GLCR datasets are not provided in the literature.

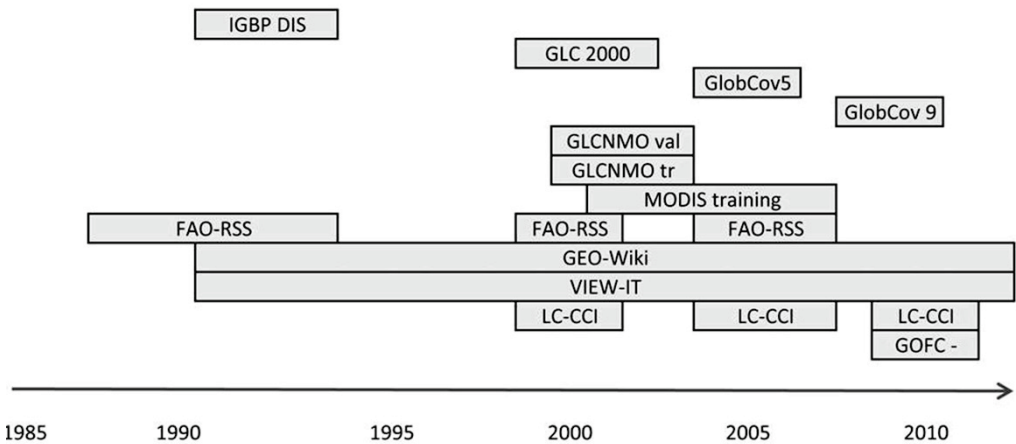


Figure 2.3. Temporal coverage of reference data sources for the validation datasets

A procedure for sample labelling is commonly carried out by visual image interpretation for the GLCR datasets except the upcoming datasets of LC-CCI and GOF-C-GOLD, which are using automated classification (Table A.2). International and national/regional experts interpreted most GLCR datasets. In contrast, Geo-Wiki and View-IT datasets are interpreted by volunteers.

To decrease the interpretation variability and bias, samples were also verified by different interpreters, and the confidence level of class assignments was recorded for GLCR datasets. Verification was done for many datasets internally during the creation of the datasets and externally consolidated by independent researchers (Table A.2). Among the datasets, IGBP-DIS, GlobCov5, and MODIS-tr datasets were verified and the interpretation confidence was recorded. LC-CCI and GOF-C-GOLD datasets are planned to contain such information. GLC2000, GLCNMO-tr and FAO-FRA datasets were verified, but the interpretation confidence was not recorded. GlobCov9 and Geo-Wiki datasets are not verified yet, whereas the interpretation confidence information is available. There is no information provided for the GLCNMO-val dataset.

Determination of the agreement between the map and reference classification is different for homogeneous and heterogeneous sampling units. In the homogeneous case, class coverage from 75% (GlobCover) to 80% (GLC2000) of a sampling unit is considered homogeneous. For heterogeneous samples consisting of two or more land cover types, often 2-3 main land cover types are recorded. Their fraction in the sample unit area was visually estimated for GLC2000 as well as GlobCov 5 and 9 datasets (Table A.2). A more objective way to quantify the fraction could be to use an automated approach. LC-CCI and GOF-C-GOLD plan to provide such information calculated from automatic segmentation and classification.

2.3.1.3 Current use of the existing GLCR datasets

The GLCR datasets were used to calibrate and validate the target GLC maps. The literature review revealed that the reference datasets were also used for other purposes. Table 2.7 summarizes the current uses of the GLCR datasets, required pre-processing steps and calculated accuracy estimates.

The IGBP-DIS dataset classes were translated into four generic classes and used to validate the global forest cover map of the FAO (Table 2.7). The GLC2000 dataset was processed again to have 11 general classes for comparing the accuracies of available GLC maps. The GlobCov5 dataset was re-interpreted once more for the GlobCov9 datasets, and the MODIS-tr dataset is updated and

reused. The use of the reference datasets for other applications is limited to these few applications.

Table 2.7. Current use of GLCR datasets

Dataset	Intended application (map accuracy assessment)	Other application	Pre-processing	Estimates	Source
IGBP-DIS	IGBP map	FAO Global forest cover	Translated into 4 general classes	Overall, class specific accuracy, standard error, also for continental level	(FAO 2001)
GLC2000	GLC2000 MAP	Validating IGBP, GLC2000, MODIS maps and their synergy	Translated into 5 and 11 generic classes; quality and consistency were checked	Overall accuracy, class specific accuracy	(Göhmman et al. 2009)
GlobCov5	GlobCover map 2005	Some samples fed to the validation datasets for GlobCover 2009	Re-interpretation	Overall accuracy, class specific accuracy	(Bontemps et al. 2011a)
GlobCov9	GlobCover map 2009				
GLCNMO-val	GLCNMO map				
GLCNMO-tr	GLCNMO training				
MODIS-tr	MODIS GLC map	MODIS: Global Urban Area mapping	Revised for training		(Schneider et al. 2009)
FAO-FRA	Forest resources assessment				(Potapov, et al., 2011)
LC-CCI					
GOFC-GOLD		Validating MODIS IGBP map in Europe	Re-stratification	Overall and class specific accuracies	(Stehman et al. 2012)
GEO-WIKI		African hybrid cropland map; Biofuel land availability map	The percentage of cropland within a 1 km pixel; extent of human impact and abandoned land as well as land cover type were recorded with confidence levels	Overall accuracy, error of omission and commission for cropland delineation of 5 GLC maps	(Fritz et al. 2011c; Perger et al. 2012)
VIEW-IT	Land change of Latin America and the Caribbean	Land change of Bolivia; deforestation and reforestation of Latin America and the Caribbean; forest change of Guatemala; land use, land cover map of Uruguay	Generalized into 5 classes	Overall and class specific accuracies	(Aide et al. 2012; Clark and Aide 2011a; López-Carr et al. 2011; Redo et al. 2012)

The situation is slightly better for Geo-Wiki and View-IT as they were used for several other applications.

2.3.2 Use cases

The final combined result from the re-usability assessment is shown in Figure 2.4. The suitability level provided in Figure 2.4 can be a pessimistic evaluation since the lowest criterion score was defined to be the final suitability level. The detailed performance of the datasets for each criterion is provided in Table A.3 (see Appendix in the online publication).

2.3.2.1 Climate modelling community (ECV)

Among the GLCR datasets, only the LC-CCI dataset seems most suitable for validating upcoming Land cover-ECV products (Figure 2.4). This is expected since the LC-CCI reference dataset is developed for this purpose. These maps and their validation procedure were designed to meet the requirements of CMC.

Most datasets would be suitable from a statistical sampling point of view, except GLCNMO-tr, MODIS-tr and Geo-Wiki due to the absence of probability sampling. We assessed the datasets providing hierarchical classifiers as highly suitable for the flexible land cover characterisation requirement since different map legends can be drawn from this information. As the Land cover-ECV dataset will have consecutive maps for 2000, 2005 and 2010, most datasets were identified to be temporally suitable for one of the three maps except the IGBP-DIS dataset for 1993.

The stability over multi-date datasets is an important user requirement, and many datasets did not satisfy this requirement. The reason is that datasets were mostly designed for one time period. A few datasets provide records in multiple dates. For example, the Globcov5 and Globcov9 datasets have some samples that have two records in time. However, around 30% of such samples showed differences in land cover interpretation. This differences could be related to inconsistency of the interpretation procedure of the two datasets. The MODIS-tr dataset may have two records in time since it is regularly updated. It may be possible that reference samples could be analysed in different time frames in the case of the Geo-Wiki and View-IT datasets. However, stability remains of concern. In contrast, FAO-FRA and LC-CCI samples from different epochs meet this requirement. The low performance of most datasets regarding the stability over multi-date datasets resulted in low overall performance. The criterion of an easily combined and augmented sampling scheme was not included for the final evaluation. If users are interested in combining the datasets, they are referred to the criterion performances listed in Table A.3.

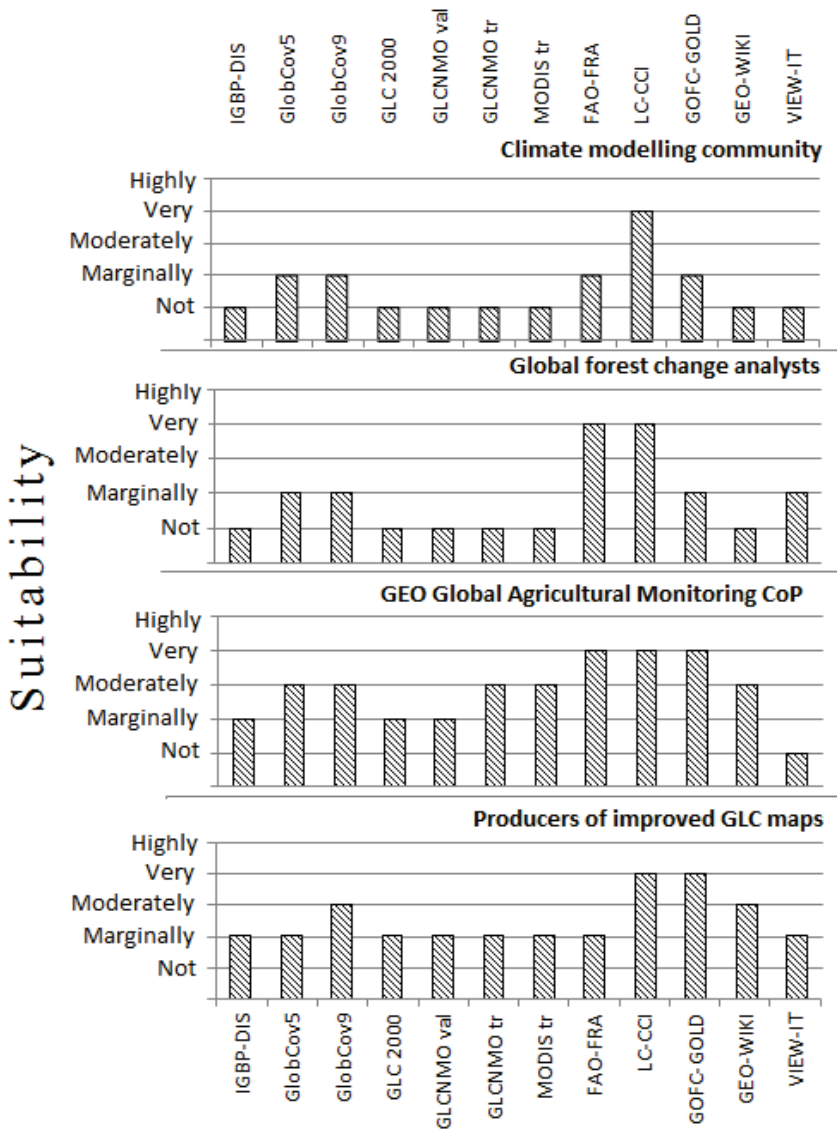


Figure 2.4. Usability of GLC validation datasets for different user groups

2.3.2.2 Global forest change analysts

The FAO-FRA and LC-CCI datasets were identified as very suitable for assessing the GFCL map according to the requirements of global forest change analysts (Figure 2.4). This is mainly because the sole purpose of the FAO-FRA is to estimate forest change at the global scale, and the LC-CCI will be generated based mostly on the FAO-FRA datasets. One of the main requirements of this user group is the availability of a multi-date sample, and only FAO-FRA and LC-CCI fully meet this requirement. In contrast, other datasets do not satisfy this requirement due to the absence of multi-date samples. For area estimation, all the datasets were evaluated as moderate to highly suitable depending on whether they used a block of pixels or area sampling unit. Most stratified sampling based datasets did not use stratification of potential forest change areas, which is important for this user group.

In terms of thematic detail and forest definition, all the datasets were identified as moderately suitable to highly suitable as information on forest/non-forest classes, forest type, and forest density are provided. Datasets assessed as moderately suitable do not use the same tree height threshold in the forest definition. In terms of spatial detail, GLC2000 and GLCNMO-tr datasets were evaluated as unsuitable due to their large sample unit size, and the others were evaluated as moderately to highly suitable.

2.3.2.3 Geo Global Agricultural Monitoring CoP

The assessment shows that FAO-FRA, LC-CCI, and GOF-C-GOLD datasets are very suitable to validate or calibrate the GHCM map. GlobCov 5 and 9 were identified as moderately suitable. GLCNMO-tr, MODIS-tr and Geo-Wiki datasets showed moderate suitability. A model-based approach can be adopted to utilize these datasets since statistical sampling-based assessment is not possible. A few datasets, such as IGBP-DIS, GLC2000 and GLCNMO-val, were evaluated as marginally suitable due to insufficient cropland sample size (less than 100). The large sample unit size of GLC2000 may be difficult to apply in heterogeneous areas. The View-IT dataset was identified as unsuitable owing to its small sample unit size compared to the target map.

For the class representation criterion, we compared the number of cropland samples and most datasets were found to be moderately to highly suitable. We evaluated the datasets that were suitable for the target map in terms of the thematic detail requirement because the target map has only cropland and non-cropland classes. Another important criterion is the information on majority classes and its fraction since map classes are represented as a percentage (cropland percentage) for each pixel. For this criterion, we assessed IGBP-DIS,

GlobCov5, and Geo-Wiki datasets as moderately suitable and the rest were very suitable to highly suitable. These datasets do not contain majority class information. Nevertheless, some information can be drawn from the legend. The GlobCov5 has information on majority classes, but their fraction was not quantified.

2.3.2.4 Producers of improved GLC maps

From the perspective of GLC map producers, the LC-CCI and GOF-C-GOLD datasets were evaluated as very suitable for validating and calibrating the VIIRS map. Moreover, GlobCov9 and Geo-Wiki datasets were also evaluated as moderately suitable, and the rest were marginally suitable.

Geo-Wiki and View-IT datasets were shown to be very suitable for the spatial coverage and class representation criterion due to a large number of samples. The only drawback of these datasets is that some regions and classes might have been represented poorly. Other datasets, except IGBP-DIS, were identified as moderately suitable for this criterion. Many datasets were found to be suitable for the flexible thematic characterisation criterion except the datasets which were identified as less suitable due to the absence of classifier information and a LCCS-based legend. This is because the VIIRS map will use the IGBP characterisation system. Most datasets generated before around 2010 were also evaluated as being difficult to re-use since the VIIRS map will be for 2012, and land cover change may have occurred within this time frame. The datasets were assessed as moderate to highly suitable for the characterisation of heterogeneous area requirements except for the GLC2000 dataset, which has a 3x3 km sample unit size. In terms of quality flag information, the datasets were evaluated as moderately to very suitable. The GlobCov5 dataset was evaluated as highly suitable since independent verification has been done for this dataset. In contrast, the GLCNMO-val dataset was assessed as marginally suitable due to the unavailability of information on verification and confidence recording.

2.4 Discussion

2.4.1 The characteristics of the GLCR datasets

The comparison of characteristics of the GLCR datasets shows large differences in aspects such as sample unit size and type, sample selection scheme and temporal coverage (Table A.2). In contrast, thanks to the joint efforts made earlier on the harmonization of different GLC map legends (Herold et al. 2006), most datasets used LCCS in the interpretation (Table 2.5). LCCS is adopted more commonly due to its flexibility and ability to present various land cover types with a more standardized description (Di Gregorio 2005). Moreover, LCCS is

recognized as a main land cover language, and new GLC maps such as new GLC maps of MODIS collection 6 products will be migrated into the LCCS (Friedl et al. 2010; Herold et al. 2006).

Most datasets used a stratified random or systematic sampling scheme because of the robustness in class-specific accuracies and cost (Table A.2) (Defourny et al. 2011b; Mayaux et al. 2006; Scepan et al. 1999). In contrast, datasets such as GLCNMO-tr, MODIS-tr and Geo-Wiki were not based on probability sampling. Nevertheless, they provide a large number of samples. All datasets, except GOF-C-GOLD, using stratified sampling were based on a stratification that is map-dependent. This restricts the use of datasets for other maps unless they are re-stratified. The GOF-C-GOLD dataset used a stratification that is independent of any GLC map, thus aiming to be applicable to many maps (Olofsson et al. 2012).

The data source comparison showed that most datasets made use of available satellite images (coarse to medium resolution), open source maps, geo-tagged photos and regional maps (Table 2.6) with temporal coverage concentrated around the year 2000 (Figure 2.3). Since most datasets were visually interpreted, the quality flag information, such as interpretation confidence and verification, is useful to have an estimate of the interpreters' variability and bias (Strahler et al. 2006). Such information is available for some of the datasets (Table A.2). Some datasets were only verified or they only report an interpretation confidence, or neither was done. When re-using GLCR datasets, quality flag information is important as reference datasets are often the result of classifications (interpretations) which are not error-free (Congalton and Green 2009; Herold et al. 2009b). Information on the confidence in interpretation could be used for analysing spatial variation in the accuracy of reference data. In addition, a certain amount of human variability is inevitable, especially in fragmented and seasonal landscapes (Gardin et al. 2011; Mayaux et al. 2006; Powell et al. 2004; Scepan et al. 1999). The users should also keep in mind that reporting an interpretation confidence may not be sufficient as some areas such as savannah and grasslands are commonly interpreted with a low confidence level. The confidence level in labelling could account for as much as a 10% difference in the classification accuracy of the IGBP-DISCOVER map (Scepan et al. 1999).

This study showed that the GLCR datasets have limitations to applications outside the intended uses (Table 2.7). This could be due to the inaccessibility of the GLCR datasets to the public and inconsistencies among them. Currently, few datasets such as the IGBP confidence sites, GLCNMO-tr and Geo-Wiki are openly available to the public. Recent efforts by the GOF-C-GOLD to make existing GLCR datasets available to the public can increase the re-use of GLCR

data substantially (Mora et al. 2014). Generating a dataset that can be used efficiently for validating multiple maps is also proposed for the improved utilization of GLCR datasets (Olofsson et al. 2012). Nevertheless, considering the enormous efforts made to generate existing GLCR datasets, their efficient use should also be emphasized and supported. Providing systematic information on the existing and forthcoming reference datasets is essential for such purposes.

Due to inconsistencies in the accuracy assessment methods, a comparison of the recorded accuracies of GLC maps is problematic (Olofsson et al. 2012). Even though several studies focused on GLC map comparisons using per-pixel disagreement analysis (Giri et al. 2005; Hansen and Reed 2000; McCallum et al. 2006), the inconsistencies of the accuracy assessment methods, in particular the characteristics of the reference datasets, have not been analysed previously. Our study focused on this and revealed considerable inconsistencies in the GLCR dataset characteristics, i.e. sampling units, sampling selection scheme as well as spatial, thematic and temporal details. Researchers and organizations are working towards reducing the inconsistencies of reference datasets and assessment methods, for example, CEOS-WGCV introduced guidelines on conducting statistically rigorous repeatable accuracy assessments (Strahler et al. 2006). The importance of such internationally accepted methods should be further emphasized.

2.4.2 The suitability of the GLCR datasets for different user communities

The suitability assessment of the GLCR datasets for the applications of four targeted user groups showed that the LC-CCI, GOF-C-GOLD, FAO-FRA and Geo-Wiki datasets were generally more suitable for re-use than the other datasets (Figure 2.4). For the CMC and global forest change analysts use cases, the requirement of stable multi-date samples could not be satisfied by most datasets. Most GLCR datasets were not designed this way; however, high density sampling in transition zones could be helpful since the change mostly occurs in transition zones (Herold et al. 2008). The requirement on the thematic detail also varied across the user groups. For example, datasets such as the FAO-FRA were less suitable for the CMC and improved GLC map production use cases (Figure 2.4). In contrast, they were considered very suitable in the use cases of the Global forest change analysts and GEO Global Agricultural Monitoring CoP where there was no strict requirement on the thematic detail. Since forest change areas and cropland areas are not uniformly distributed in space, potential spatial distribution of these areas should be considered when selecting samples for reference data. We assessed whether this was considered for the GLCR datasets from the perspective of Global forest change analysts. Only a few datasets used

stratifications based on potential forest change areas when selecting samples (Table A.3). In general, most datasets were identified as moderately suitable or higher for the agricultural use case, whereas many were assessed as marginally suitable for the improved GLC map production case.

Other than the user-group specific criteria, we also considered criteria (e.g. spatial resolution and temporal coverage) that are particularly related to target maps. The dataset suitability could vary depending on which map is being targeted within the same user group. For the production of an improved GLC map, for example, most datasets were identified as marginally suitable for validating and calibrating the VIIRS map due to temporal coverage. The dataset suitability can be different when another target map is considered, e.g. a GLC map for 2000.

The re-usability of crowd-sourcing GLCR datasets for the improved GLC map producers and GEO Global Agricultural Monitoring CoP was demonstrated in our analysis (Figure 2.4). These datasets have been re-used more frequently than other datasets (Table 2.7). The quality of such datasets may be challenged, but preliminary results from the Geo-Wiki datasets show that most volunteers have experience in interpreting satellite data (Fritz et al. 2011a) and volunteers were trained in the case of the View-IT dataset (Clark and Aide 2011b). Nevertheless, a successful arrangement of the Geo-Wiki, providing a growing database of samples resulting from the collective efforts of volunteers in a cost-effective way, is appealing for GLCR data generation. Such efforts should be supported in the new generation of GLCR datasets.

2.5 Conclusion

A variety of applications with different purposes and requirements use GLC maps as an input, thus no single map can be optimal for all applications. This is also true for the GLCR datasets aiming to calibrate and validate GLC maps. This study therefore considered the varying requirements of user groups of GLC maps and assessed the re-usability of available and forthcoming GLCR datasets. Varying GLCR dataset suitability levels were obtained depending on the reference data characteristics, user requirements and target maps. The following datasets were identified as being most suitable for multiple user groups: LC-CCI, GOFC-GOLD, FAO-FRA and Geo-Wiki. We emphasise the potentiality of existing and forthcoming GLCR datasets for multiple use cases in this study. In particular, most datasets were found to be at least moderately suitable for the GEO Global Agricultural Monitoring CoP case and at least marginally suitable for the improved GLC map production case.

The analysed GLCR datasets revealed some limitations in meeting specific requirements of the user groups. Stable multi-date sample requirement of CMC and Global forest change analysts groups could not be satisfied by the GLCR datasets since most dataset were developed for a single time frame. The requirement on sample selection scheme and stratification based on potential forest change areas also resulted in limited suitability of GLCR datasets for Global forest change analysts. This is mainly because sampling stratifications did not consider potential forest change areas. The requirement on thematic detail of GLCR datasets were less strict for Global forest change analysts and GEO Global Agricultural Monitoring CoP use cases than for the CMC and improved GLC map production use cases. The scores on this criterion differed accordingly for the two groups. This exemplifies the importance of considering GLC validation from different user perspective.

Coordinated international efforts are working to increase the integrity of GLC maps and reference datasets and to support the re-usability of available GLCR datasets by making them accessible to the public (Mora et al. 2014). By providing systematic information about GLCR datasets and their re-usability in different use cases, this study can contribute to better-informed decision making for the efficient use of GLCR datasets. Apart from assessing the re-usability of the GLCR datasets, this chapter presented a method that can be adopted for dataset evaluations in similar applications. The scope of the study was also to demonstrate the capability of the GLCR datasets and to emphasize an efficient recycling of available reference datasets.

The authors conclude that the following items should be prioritized in reference dataset generation to increase the general usability of GLCR datasets for multiple user applications:

- Use of a probability sampling scheme with representations from each class and region is advantageous.
- A LCCS-based legend with classifier information making the datasets adaptable to different map legends.
- Use of a small MMU and a large sample unit area and updating the dataset regularly
- Quality assurance information
- Publicly available in an expert framework guiding proper usage.

Acknowledgements

This study was supported by the EU FP7 GEOCARBON, ESA-ECVLCover and ESA GOFC-GOLD projects. We would like to thank S. Keshkamat, R. Chavez, R. Fuchs and C. Brown for reviewing the manuscript and providing useful comments.

Chapter

3

Comparative assessment of thematic accuracy of GLC maps for specific applications using existing reference data

*Nandin-Erdene Tsendbazar, Sytze de Bruin, Brice Mora, Léon Schouten and
Martin Herold*

*Published in International Journal of Applied Earth Observation and
Geoinformation, Volume 44, February 2016, Pages 124-135*

Supplementary data to this chapter can be found in the online publication

Abstract

Inputs to various applications and models, current global land cover (GLC) maps are based on different data sources and methods. Therefore, comparing GLC maps is challenging. Statistical comparison of GLC maps is further complicated by the lack of a reference dataset that is suitable for validating multiple maps. This study utilizes the existing Globcover-2005 reference dataset to compare thematic accuracies of three GLC maps for the year 2005 (Globcover, LC-CCI and MODIS). We translated and reinterpreted the LCCS (Land Cover Classification System) classifier information of the reference dataset into the different map legends. The three maps were evaluated for a variety of applications, i.e., general circulation models, dynamic global vegetation models, agriculture assessments, carbon estimation and biodiversity assessments, using weighted accuracy assessment. Based on the impact of land cover confusions on the overall weighted accuracy of the GLC maps, we identified map improvement priorities. Overall accuracies were $70.8\pm 1.4\%$, $71.4\pm 1.3\%$ and $61.3\pm 1.5\%$ for LC-CCI, MODIS and Globcover respectively. Weighted accuracy assessments produced increased overall accuracies (80-93%) since not all class confusion errors are important for specific applications. As a common denominator for all applications, the classes mixed trees, shrubs, grasses and cropland were identified as improvement priorities. The results demonstrate the necessity of accounting for dissimilarities in the importance of map classification errors for different user application. To determine the fitness of use of GLC maps, accuracy of GLC maps should be assessed per application; there is no single-figure accuracy estimate expressing map fitness for all purposes.

Keywords: Global land cover; accuracy assessment; comparison; reference dataset; user applications; weighted accuracy assessment

3.1 Introduction

Global land cover maps (GLC) provide essential information for policies and scientific applications such as climate modelling, food security, biodiversity and environmental modelling (Verburg et al. 2011). Currently, several GLC maps exist, including the IGBP-DISCover (Loveland et al. 2000), GlobCover maps (Defourny et al. 2012), MODIS GLC maps (Friedl et al. 2010), LC-CCI maps (CCI-LC 2014) and FROM-GLC maps (Gong et al. 2013). These maps originate from various initiatives and they differ in used input data, methodologies and validations (Herold et al. 2008). Consequently, for users it is difficult to compare and select the GLC map that fits their application best.

Several studies have compared GLC maps to highlight their strengths (Fritz et al. 2011b; Giri et al. 2005; McCallum et al. 2006). These studies analysed spatial agreement among GLC maps, however, they do not provide information on comparative accuracies of the maps. Fritz et al. (2011b) and Herold et al. (2008) compared the accuracy estimates of GLC maps after harmonizing reported confusion matrices with different legends into confusion matrices with a common legend. As this harmonization only concerns the confusion matrices, it remains unclear how the GLC maps are compared relative to the same reference dataset. Except for a few regional comparisons (Frey and Smith 2007; Wu et al. 2008), there are no reports on comparative accuracy assessments of GLC maps that use the same reference dataset.

Comparative accuracy assessment of GLC maps is challenging because of lack of reference datasets that are suitable for multiple maps. Several GLC reference datasets were produced and used to validate specific GLC maps (Bontemps et al. 2011a; Mayaux et al. 2006; Scepan et al. 1999). However, Olofsson et al. (2012) questioned the use of these datasets to validate multiple maps due to inconsistencies in their legends, sampling and response design protocols. Olofsson et al. (2012) proposed a new GLC reference dataset with flexible thematic and spatial representations and an adjustable sampling scheme to validate multiple maps. This dataset is under development. Existing datasets could be useful for validating multiple GLC maps if they satisfy the following criteria: (1) based on a probability sampling design, (2) employ UN Land Cover Classification System (LCCS) classifiers, (3) are quality flagged and (4) are publicly available (Tsendbazar et al. 2015a).

Conventional accuracy reporting from confusion matrices assumes that all confusion errors are equally important. However, confusion between certain classes may have more impact on applications of land cover (LC) maps than other

confusions (DeFries and Los 1999). For example, confusion between water and snow/ice may not be as important for biomass estimation, as it is for albedo estimation. Several studies accounted for such differences and calculated GLC map accuracy for specific applications using weights derived from class similarities in terms of certain parameters (DeFries and Los 1999; Mayaux et al. 2006). To date, however, a comparative accuracy assessment of GLC maps for specific applications does not exist.

The objectives of this study are (1) to use the existing Globcover-2005 reference dataset for comparative assessment of thematic accuracy; (2) to compare the recent GLC maps and assess their strengths and weaknesses for specific applications; and (3) to identify priorities for improving the GLC maps for those applications. We quantitatively compared GLC maps for the year 2005, namely the Globcover, LC-CCI and MODIS using the Globcover-2005 reference dataset that has flexible thematic information (i.e. LCCS-classifiers). Furthermore, we compared the thematic accuracies of the GLC maps to analyse which maps are the most accurate for applications of climate models namely General Circulation Models (GCM) and Dynamic Global Vegetation Models (DGVM), agriculture assessments, carbon estimation and biodiversity assessments. We then identified map improvement priorities (specific LC classes and LC confusions) by analysing LC classes and confusions with high impact on the weighted accuracy. Comparative assessment of GLC map accuracies for different applications can help users to better select GLC maps and to understand their uncertainty as model inputs. Moreover, the analysis of improvement priorities for the GLC maps is also beneficial to the GLC map producers.

3.2 Material and methods

3.2.1 Global land cover maps

The GLC maps considered in this study are: the Globcover, LC-CCI and MODIS maps for the year 2005. These maps were selected because they are temporally similar to the reference dataset.

The Globcover project of the European Space Agency (ESA) provides GLC maps of 2005 and 2009 based on 300m resolution MERIS satellite data (Defourny et al. 2011b). The Globcover-2005 map has an LCCS based legend with 22 classes. The thematic overall accuracy of the Globcover-2005 was reported as 73.1% (Bicheron et al. 2008).

Building on the experience of Globcover maps, the ESA's Land Cover-CCI (LC-CCI) project produced three consecutive GLC maps for the epochs of 2000, 2005

and 2010 at 300m resolution using the MERIS data archive (CCI-LC 2014). These maps were specifically targeted to meet the requirements of climate modellers. The maps have an LCCS based legend with 22 classes. The LC-CCI project will conduct an accuracy assessment for the 2005 epoch map.

Using data from the Moderate Resolution Imaging Spectroradiometer (MODIS), Boston University provides annual MODIS Collection-5 GLC maps at 500m resolution (Friedl et al. 2010). MODIS GLC maps provide five legends including a legend based on the International Geosphere-Biosphere Programme (IGBP) classification scheme. The overall accuracy of the 2005 map is 74.8% based on the cross-validation using the training dataset (Friedl et al. 2010). We used the GLC map with the IGBP legend (17 classes) in this study. Table 3.1 provides a summary of the GLC maps.

Table 3.1. A summary of GLC maps used for comparison

GLC map	Globcover	LC-CCI	MODIS
Temporal frame	2005		
Spatial resolution at the Equator	300 m	300 m	500 m
Input data	MERIS: Bi-monthly from 10-day composites	MERIS global SR composite, SPOT-VGT time series (for updating)	MODIS: Monthly EVI, LST and 7 bands from 8-day composites
Time of data collection	2005-2006	2003-2007	2005
Classification method	(Un)supervised spatio-temporal clustering; expert-based labelling	Unsupervised spatio-temporal clustering; machine learning classification	Supervised decision tree boosting
Classification scheme	LCCS based:22 classes	LCCS based: 22 classes	5 different legends including the IGBP (17 classes)
Reference	(Bicheron et al. 2008; Defourny et al. 2011b)	(CCI-LC 2014)	(Friedl et al. 2010)

We harmonized the map legends into 13 general classes for comparative assessment (Table 3.2). The Globcover legend does not have a separate class for DNL trees. Therefore, its harmonized legend has 12 classes. The harmonization followed the approach of Herold et al. (2008) that provides a table for harmonizing into 13 general LC classes based on the LCCS-based legend translation protocols.

Table 3.2. The lookup table for 13 general class legend and corresponding classes of the GLC maps

Class	Generalized land cover class	Abbreviation	Globcover*	LC-CCI	MODIS-IGBP
1	Evergreen Needleleaf Trees	ENL trees	70	70	1
2	Evergreen Broadleaf Trees	EBL trees	40, 160, 170	50, 160, 170	2
3	Deciduous Needleleaf Trees	DNL trees		80	3
4	Deciduous Broadleaf Trees	DBL trees	50, 60	60	4
5	Mixed/Other Trees (incl. mixtures)	Mixed trees	90, 100, 110	90, 100	5, 8, 9
6	Shrubs	Shrubs	130	120	6,7
7	Herbaceous vegetation	Grasses	120, 140	110, 130, 140	10
8	Cultivated and managed vegetation /agriculture (incl. mixtures)	Cropland	11, 14, 20, 30	10, 20, 30, 40	12, 14
9	Wetland vegetation	Wetland	180	180	11
10	Urban/built up	Urban	190	190	13
11	Snow and Ice	Snow and Ice	220	220	15
12	Bare/sparse vegetation	Bare/sparse vegetation	150, 200	150, 200	16
13	Open water	Water	210	210	17

*The legend descriptions corresponding to the Globcover, LC-CCI and MODIS maps can be found in Table S2 in the supplementary data of the online publication.

3.2.2 Global land cover reference dataset

Comparative accuracy assessment was based on the reference dataset that was developed to validate the Globcover-2005 GLC map. This reference dataset was selected to compare the GLC maps since it has a probability sampling design and LCCS-classifier information while it is also readily accessible through the GOF-GOLD Land-Cover office.

The design procedures to derive the Globcover-2005 reference dataset are described in Bicheron et al. (2008). The reference sample units were selected using a stratified random sampling scheme (Bicheron et al. 2008). The sample units correspond to 5x5 MERIS pixels (1.5km x1.5km area at the equator). This area matches to 5x5 pixels on the Globcover and LC-CCI maps and 3x3 pixels on the MODIS map. The reference LC information of more than 4000 sample sites was interpreted by international experts using web mapping capabilities including Google-Earth, Virtual-Earth and 8-years of NDVI profiles (Defourny et al. 2011b).

To enhance the potential future use of the reference dataset, the LC at sample locations was characterised by LCCS-classifiers that categorize sets of diagnostic attributes of LC (e.g. vertical and horizontal structure, floristic composition)(Di Gregorio 2005). LC classes are defined by applying combinations of the LCCS-classifiers. Depending on diagnostic attributes of LC, up to 6 levels of LCCS-

classifiers were used for characterizing LC at the sample locations. The corresponding lookup table for the LCCS-classifiers is provided in Table S1 in the supplementary data of the online publication. In heterogeneous areas, a maximum of three dominant LC classes occupying the largest part of the sample unit area were reported.

3.2.3 Translation and re-interpretation of the Globcover-2005 reference dataset

Translation and reinterpretation steps were followed to ensure the reference dataset's quality and consistency (Figure 3.1). The LCCS-classifiers of the reference dataset were translated to the legends of the Globcover, LC-CCI and MODIS maps following a fixed protocol based on the translation guideline provided in Bicheron et al. (2008).

For each legend, main LC classes were translated from the LCCS-classifier information per sample unit. For heterogeneous sample units (2 or 3 LC classes), two possible mosaic classes were created (e.g. mosaic of natural vegetation and cropland). The lookup table for translating main and mosaic LC classes can be found in Table S3-5 in the supplementary data of the online publication.

The Globcover-2005 reference dataset does not provide quantitative class proportion information indicating relative dominance of LC classes within composite classes. Therefore, it is difficult to identify the final LC classes of the reference sample, particularly in heterogeneous areas. To overcome this, we detected the reference sample sites with multiple general classes after harmonizing the translated LC classes (main and mosaic) of the reference datasets into 13 general classes. For example, a site with rain-fed cropland, irrigated cropland and mixed tree classes could have two possible general LC classes after the harmonization (cropland and mixed trees). The detected sample sites were analysed to identify the source of uncertainty in the translation and they were eventually re-interpreted.

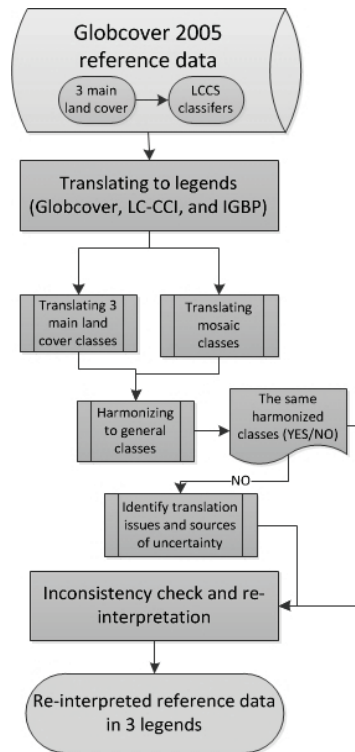


Figure 3.1. Translation and re-interpretation steps

We re-interpreted the identified sample sites as well as other directly translatable sites with the availability of more and better source materials (i.e., very high resolution images). In the re-interpretation, we used images from Google-Earth and Bing maps, geotagged photos, the Global Land Survey Landsat-2005 reflectance and NDVI data, tree cover information from Vegetation continuous field (MODIS-VCF) data and the global map of irrigated areas (Siebert et al. 2013). Sample sites were labelled according to the legends of the Globcover, LC-CCI and MODIS maps to obtain a GLC reference dataset in three legends. Figure 3.2 shows the spatial distribution of the sample processed in this step.

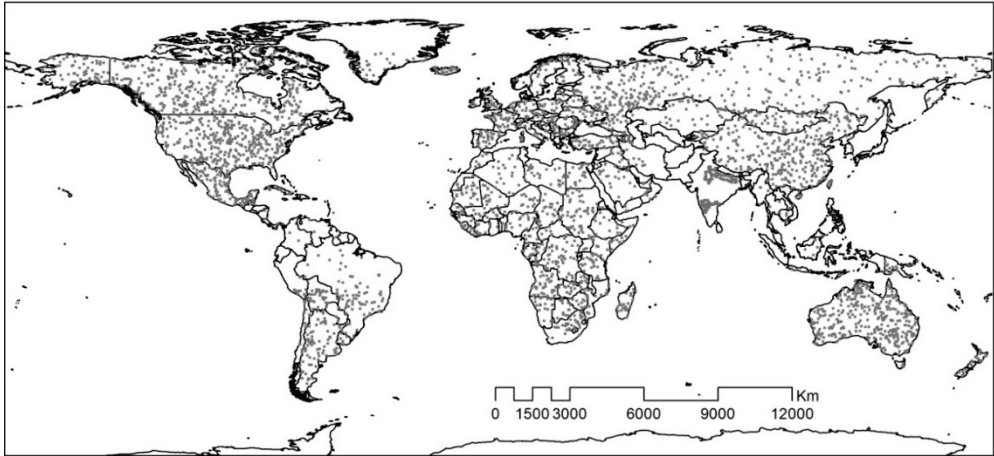


Figure 3.2. Spatial distribution of the Globcover-2005 reference dataset (grey dots)

3.2.4 Thematic accuracy assessments

The thematic accuracies of the GLC maps were calculated using confusion matrices. To build the confusion matrices, map categories corresponding to the reference sample locations were extracted; any heterogeneous sites with two or more map categories were assigned to the class occupying the majority of the site area.

In the accuracy assessments, we adjusted confusion matrices from sample counts by accounting for the class proportions of the map (Card 1982). The overall accuracy and class-specific accuracies were derived following the method described by Card (1982). The variance and confidence level of the accuracies were calculated using the equations for stratified random sampling (Card 1982; Olofsson et al. 2013). Figure 3.3a illustrates the inputs to adjust the confusion matrix using class proportions.

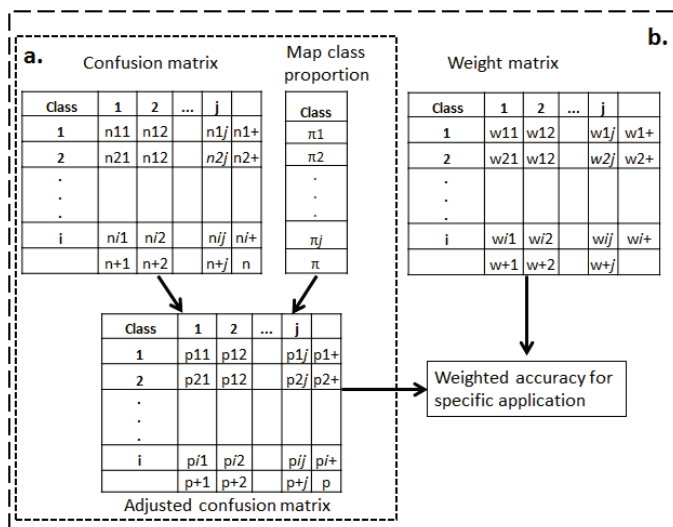
3.2.5 Weighted accuracy assessments and weight calculation for specific applications

Weighted accuracies considering class similarities for specific applications were assessed for the GLC maps using the method described by Cohen (1968) and Rossiter (2004). Weighted overall accuracies (A_{ow}), accuracy standard deviation (Sa_{ow}) and confidence intervals (at 95% confidence level) were calculated based on the Equations 3.1-3.3. The matrices used for this calculation are shown in Figure 3.3b.

$$A_{ow} = \sum_{i=1}^r \sum_{j=1}^r w_{ij} \cdot \rho_{ij} \quad (3.1)$$

$$s_{aow} = \sqrt{\frac{A_{ow} \cdot (1 - A_{ow})}{n}} \quad (3.2)$$

$$A_{ow} \pm [s_{aow} \cdot Z_{1-\alpha}] \quad (3.3)$$



n_{ij} - number of observation in row i , column j ; n_{i+} - marginal sum of row (mapped class) i ; n_{+j} - marginal sum of column (reference class) j ; n - total number of observations; p_{ij} -proportion of map area in row i , column j ; p_{i+} - proportion of map area in row i (classified); p_{+i} -proportion of map area in column i (adjusted by the reference data); W_{ij} - weight correspond to map area P in row i and column j .

Figure 3.3. Procedures to derive adjusted confusion matrix (a) and weighted accuracies for specific application (b)

In this study, we assessed the accuracy of GLC maps in terms of thematic similarity, climate models such as GCM and DGVM, agriculture assessments, carbon estimation and biodiversity assessments. Once the weights were defined, weighted accuracies were assessed for these applications using the confusion matrix and the adjusted confusion matrix (Figure 3.3). The following subsections describe the weight calculations for specific applications.

3.2.5.1 Thematic similarity

Thematic similarity between LC classes was defined based on thematic proximity from the legend. We assessed the thematic similarity between 13 LC classes following the method described in Mayaux et al. (2006). Similarity indices were assigned to each pair of LC classes for each of 7 LCCS-classifiers namely

(non)vegetated, terrestrial/aquatic, managed/natural, life form, cover, height, leaf type and leaf cycle. Thematic similarities were then calculated as the average of the 7 scores. The scale of the similarity indices ranges from zero (meaning that the two LC classes are completely dissimilar) to one (meaning that the two LC classes are assumed to be identical for given applications). As the index increases, the seriousness of the confusion between the two classes is considered to diminish. Hence, the diagonals of the similarity matrix are composed of ones, while the off-diagonals are between zero and one. This similarity matrix was used to provide the weights for assessing the GLC map accuracy in terms of thematic similarity.

3.2.5.2 Agriculture assessments

GLC maps are often used in combination with other statistical surveys to create global cropland maps (See et al. 2012). We included this application to illustrate an application that uses only one or few classes of GLC maps. As GLC maps do not have a separate class for pastureland, the similarity between the LC classes was defined based on the class “cropland”. In this case, the confusion of a cropland class with non-cropland classes or vice versa is considered a serious error. In contrast, the confusion within non-cropland classes does not have any impact. To reflect such differences, the similarity between cropland and non-cropland classes was set at zero (completely dissimilar) and similarity between non-cropland classes was set at one, meaning that they are identical for agriculture assessments.

3.2.5.3 Global Circulation Models

GLC maps are commonly used for climate models such as GCM and DGVM (Kooistra et al. 2010). We assessed the similarity between the LC classes for 9 land surface parameters that are used for GCM. The global datasets of 9 land surface parameters of ecosystem types provided by Hagemann (2002) were used as reference. These are:

- background surface albedo (1)
- surface roughness length due to vegetation (2)
- fractional vegetation cover for the growing (3) and dormancy seasons (4)
- leaf area index for the growing (5) and dormancy seasons (6)
- forest ratio (7)
- plant-available total soil water holding capacity (8)
- volumetric wilting point (9)

The mean values per LC class were estimated by reclassifying the ecosystem types into 13 LC classes. After calculating LC class similarity for the 9 land surface parameters, the average similarity score from the similarity matrices was used as

the weights. The LC class similarity was calculated by modifying the method of Herold et al. (2011)(Equation 3.4).

$$W_{aij} = \text{Max} \left(0, 1 - \frac{|\text{Par}_{ai} - \text{Par}_{aj}|}{SD(\text{Par}_a)} \right) \quad (3.4)$$

where, W_{aij} is the similarity index for pair of classes (i, j), Par_{ai} and Par_{aj} are values for the variable of interest for classes i and j , and $SD(\text{Par}_a)$ is the standard deviation of variable of interest. Herold et al. (2011) related the absolute difference of values of variable of interest to the range of variable of interest based on Gower's similarity coefficient (Gower 1971), we used standard deviation for standardization to avoid excessive sensitivity to a single LC class having an extreme variable value. Note that such approach is commonly used for standardizing dissimilarity (i.e. the inverse of similarity) measures (Milligan and Cooper 1988).

3.2.5.4 *Dynamic Global Vegetation Models*

Plant functional types datasets, commonly derived from GLC maps, are used for DGVMs to simulate the effects of future climate change on natural vegetation and its carbon and water cycles (Kooistra et al. 2010; Lapola et al. 2008). To assess the similarity of LC classes in terms of DGVM, we used a table provided by Poulter et al. (2015) that lists fractions of 12 different plant functional types in LC classes. These plant functional types include: trees and shrubs (broadleaf evergreen, broadleaf deciduous, needle leaf evergreen and needle leaf deciduous (for trees only)), grass (natural and managed), bare soil, water and snow/ice. For each of these plant functional types, the similarity of LC classes was calculated using the fraction of plant functional types in each LC class and Equation 3.4. These similarity matrices were then averaged to derive a general similarity matrix for DGVM.

3.2.5.5 *Carbon or biomass estimation*

One way of obtaining consistent forest carbon information at global scale is using biome-average approach that estimates average forest carbon stocks for broad forest categories and applies these estimates to GLC maps to derive total carbon stocks at global level (Ruesch and Gibbs 2008). We calculated the similarity of LC classes for carbon estimation using the global carbon density map developed by Ruesch and Gibbs (2008). This map is based on IPCC Tier-1 carbon density values and the GLC2000 map. As this map is based on a GLC map, we estimated median carbon density per LC class based on the reference sample. Median values rather than means were used because the former are less sensitive to extreme values caused by possible misclassification in the GLC2000 map. For each legend, a

similarity matrix of LC classes in terms of carbon estimation was calculated based on Equation 3.4.

3.2.5.6 Biodiversity assessments

In biodiversity assessments, LC datasets provide proxies for expected biodiversity and ecological processes and they are used for different species distribution models using known species-land-cover relationships (Tuanmu and Jetz 2014). To assess the accuracy of the GLC maps for biodiversity assessments, we used the publicly-accessible species richness maps provided by Jenkins et al. (2013). The species richness map was created by summing up the species richness of birds, mammals and amphibians maps at 10x10km resolution. Numbers of species were extracted per reference sample location and median species richness per LC class was calculated for three legends. Using Equation 3.4, for each legend, a similarity matrix LC classes was calculated in terms of biodiversity applications.

3.2.6 Identifying priorities for improving GLC maps

The impact of LC classes on the overall weighted accuracy for each application was assessed to identify the priorities (specific LC classes and LC confusions that needs to be improved) for improving the GLC map accuracies. For this, we assumed that the confusion error of each class would be reduced by half by efforts to update the GLC maps. This 50% error reduction is used only to illustrate the impact of the LC class on the overall weighted accuracy. Any reduction percentage will produce similar priorities since the overall weighted accuracy calculation is a linear function (Equation 3.1).

We created confusion matrices of the GLC maps with 50% less confusion (omission) error and more agreement (the 50% omissions were assumed to be corrected) for one LC class at a time, and transformed these into adjusted confusion matrices. Subsequently, the latter were used to assess the overall weighted accuracies based on the methods explained in Sections 3.2.4 and 3.2.5. The improvement on the overall weighted accuracy was analysed by comparing the overall weighted accuracy with reduced errors and the results from Section 3.2.5.

For a detailed assessment of improvement priorities, we repeated the above analysis on pairwise 50% reductions of confusion error between each LC class pairs.

3.3 Results and discussion

3.3.1 Reference dataset translation and re-interpretation

The LCCS-classifier information of the Globcover-2005 reference dataset was translated into LCCS-based Globcover and LC-CCI, IGBP-based MODIS legends. About half of the 3857 sites could be directly translated into the different legends. For the remaining sample sites, translation was problematic owing to several issues. Figure 3.4 shows the uncertainty sources of sample sites in the legend translation process. For translation to the Globcover and LC-CCI legends, one of the main issues was lack of classifier information related to trees and cropland types. For the MODIS-IGBP legend, differences in class definition were an important issue. Forest coverage classifiers of the LCCS such as open forest (10-20%-60-70%) and very open to closed forest (15-100%) did not match with the definition of woody vegetation classes of the Globcover and the MODIS-IGBP legends.

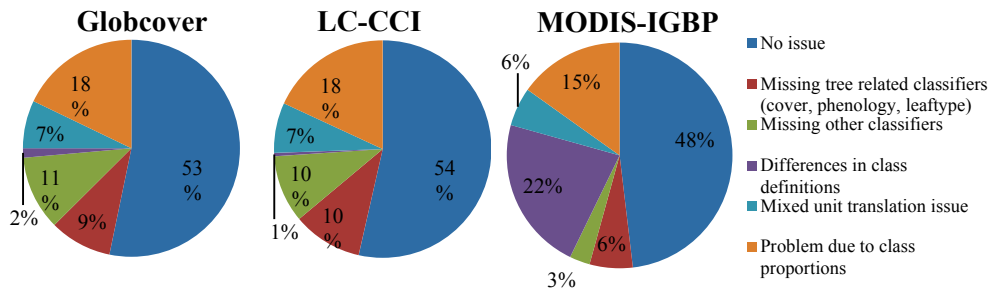


Figure 3.4: Sources of uncertainty of sample sites with translation issues

In total, 21-25% of the sample sites had translation issues due to heterogeneous landscapes such as translating mixed units (mosaic classes) (6-7%) and problems due to lacking quantitative class proportion information (15-18%) (Figure 3.4). As mentioned in Section 3.2.3, information on class proportion in the sample unit area was not provided in the reference data. This caused problems in characterising heterogeneous landscapes. These issues were resolved by re-interpreting the dataset and by presenting the land cover in three legends that reflect the differences in class definitions. Although interpreter variability and bias are inevitable in human interpretation (Foody 2010), our re-interpretation created a consistent reference dataset for the purpose of this study.

Although adapting an existing dataset for different applications requires significant effort, the result is a valuable contribution to the scarce reference datasets available

for comparative accuracy assessments of GLC maps. The use of existing reference dataset was previously discussed by (Tsendbazar et al. 2015a) who assessed the suitability of existing reference datasets for different applications. The efficient use of existing reference dataset is promoted by GOFC-GOLD (2014) which provides open access to several GLC reference datasets. These experiences of re-using reference datasets and analysing their suitability for re-use can be helpful in generating a new reference datasets (Fritz et al. 2011a; Olofsson et al. 2012) that can be directly applicable to multiple map validation. Our reference dataset will be made available through the GOFC-GOLD reference dataset portal (GOFC-GOLD 2014).

3.3.2 Thematic accuracy assessments of the GLC maps

The confusion matrices of the GLC maps are provided in Table 3.3. The result shows that the overall accuracies of the LC-CCI and MODIS maps were similar ($70.8\pm 1.4\%$ and $71.4\pm 1.3\%$ respectively). The overall accuracy of the Globcover map ($61.3\pm 1.5\%$) was less than the accuracy reported by Bicheron et al. (2008). The latter was based on a subset of 3167 sample sites where land cover could be determined with certainty while confusion errors concerning mosaic classes were disregarded in the assessment (Bicheron et al. 2008). In contrast, our accuracy estimate was based on the entire set of 3857 sample sites. Our results are consistent with those of previous studies that also reported around 70% overall accuracy for different GLC maps (Bontemps et al. 2011a; Mayaux et al. 2006).

The class-specific accuracies were generally the lowest for the Globcover and the highest for the MODIS (Table 3.3). The mixed trees class in the Globcover map includes open DNL and ENL trees. Although, this difference was reflected in the reference datasets, mixed trees deemed to be over-represented at the cost of ENL trees in this map (producer's accuracy is higher than user's accuracy) (Table 3.3). The LC-CCI map had higher class-specific accuracies for tree classes other than mixed trees which was significantly under represented (low producer's accuracy and high user's accuracy) (Table 3.3). The LC-CCI map had higher class-specific accuracies for tree classes other than mixed trees which was significantly under represented (low producer's accuracy and high user's accuracy) (Table 3.3). The MODIS map had higher class-specific accuracies for tree classes except for DBL trees, which was considerably under-represented and confused with mixed trees class.

All maps had high confusion errors for shrubs, grass and cropland classes. The Globcover map under-represented shrubs, whereas the MODIS over-represented this class (Table 3.3). The Globcover and LC-CCI maps had a high confusion

between grass and bare/sparse vegetation classes. In tundra regions, grass was under-represented, whereas it was over-represented in some areas of the Tibetan plateau and in Central Australia. Rare classes of urban and wetland areas were under-represented in the Globcover and MODIS maps (Table 3.3). Bare/sparse vegetation area was over-represented in the Globcover and LC-CCI maps. In contrast, the MODIS characterizes less bare/sparse vegetation area which was committed mostly to shrubland areas. These findings are consistent with previous studies that show inaccuracies in characterizing shrubs, grass and cropland classes due to low spectral separability and mixed vegetation components (Fritz et al. 2011b; Herold et al. 2008).

The stratification of the Globcover 2005 reference dataset was the Globcover map (Bicheron et al. 2008), and this makes the dataset map dependent. Since, the purpose of a stratification is to increase the precision of the accuracy assessment (Stehman 2009), the precision of the accuracy estimates for the GLC maps other than the Globcover may not be optimal, and it is likely to be over-estimated. Nonetheless, the calculated accuracy estimates are statistically sound as they are based on a probability sampling which has a large number of sample units.

The weighted overall accuracies in terms of thematic similarity were $83.0 \pm 1.2\%$, $86.5 \pm 1.1\%$ and $89.7 \pm 1\%$ for the Globcover, LC-CCI and MODIS maps, respectively. These figures are higher than overall accuracies in Table 3.3 since confusion between thematically similar LC classes such as mixed trees, shrubs and grass were deemed of less consequence (Figure 3.5a).

Comparative assessment of thematic accuracy of GLC maps for specific applications using existing reference data

Table 3.3. Confusion matrices of the GLC maps

Globeover	Reference classes*													Correct	Total	UA** %	± c.i. UA**	
	1	2	4	5	6	7	8	9	10	11	12	13						
Map classes	1	53	7	6	19	3	1	7	1	1	0	0	1	53	99	53.5	9.8	
	2	0	165	6	6	0	6	22	2	0	0	0	1	165	208	79.3	5.5	
	4	2	13	151	41	10	12	44	3	0	0	0	0	151	276	54.7	5.9	
	5	114	5	27	207	44	66	38	13	3	0	7	5	207	529	39.1	4.2	
	6	1	13	20	22	121	40	36	0	2	0	11	0	121	266	45.5	6.0	
	7	6	2	8	24	84	130	82	3	5	1	46	0	130	391	33.2	4.7	
	8	0	22	23	26	43	116	869	4	14	0	13	1	869	1131	76.8	2.5	
	9	1	3	0	0	0	7	0	24	0	0	0	0	24	35	68.6	15.4	
	10	0	0	0	0	0	0	5	0	20	0	0	0	20	27	74.1	16.5	
	11	0	0	0	0	0	4	0	0	38	6	3	38	51	74.5	12.0	3.3	
	12	2	0	5	13	27	134	41	10	4	3	509	4	509	752	67.7	7.7	
	13	2	1	1	5	1	1	2	1	1	1	76	0	76	92	82.6	7.7	
	Correct	53	165	151	207	121	130	869	24	20	38	509	76	2363				
Total	181	231	247	363	333	517	1146	61	50	42	593	93	3857					
PA %	26.7	84.3	59.8	57.8	39.1	23.1	66.2	29.8	21.3	93.1	89.3	80.0						
c.i. PA±	7.8	4.7	6.8	4.2	5.5	3.3	1.7	27.8	30.8	23.6	1.5	16.4						
Overall accuracy: 61.3 ± 1.5																		
LC-CCI	Reference classes*													Correct	Total	UA** %	± c.i. UA**	
	1	2	3	4	5	6	7	8	9	10	11	12	13					
Map classes	1	186	5	7	5	83	3	13	12	2	0	0	0	186	316	58.9	5.4	
	2	1	172	0	5	6	1	0	3	0	0	0	0	172	189	91.0	4.1	
	3	1	0	52	0	12	0	8	0	0	0	0	1	52	74	70.3	10.4	
	4	2	5	1	182	40	10	3	21	1	0	0	0	182	265	68.7	5.6	
	5	8	1	4	11	39	2	4	2	1	0	0	0	39	72	54.2	11.5	
	6	1	10	0	3	21	185	40	20	1	0	0	19	0	185	300	61.7	5.5
	7	2	4	0	12	14	50	237	117	1	5	0	61	1	237	504	47.0	4.4
	8	0	34	0	26	19	56	83	942	11	9	0	7	2	942	1189	79.2	2.3
	9	2	0	2	3	11	2	8	3	30	0	0	2	1	30	64	46.9	12.2
	10	0	0	0	0	0	2	0	1	15	0	0	0	0	35	53	66.0	12.7
	11	0	0	0	0	0	0	0	0	0	0	41	4	0	41	45	91.1	8.3
	12	4	0	4	0	12	25	118	9	10	1	1	496	1	496	681	72.8	3.3
	13	0	0	0	1	6	0	2	2	4	0	0	3	87	87	105	82.9	7.2
Correct	186	172	52	182	39	185	237	942	30	35	41	496	87	2684				
Total	207	231	70	248	265	334	517	1146	61	50	42	593	93	3857				
PA %	85.7	88.0	83.5	74.5	18.1	56.1	39.2	78.5	48.1	62.7	98.1	88.9	92.4					
c.i. PA±	5.0	5.1	14.2	6.0	5.1	6.4	3.7	1.4	29.4	48.5	14.5	1.5	12.4					
Overall accuracy: 70.8 ± 1.4																		

MODIS	Reference classes*													Correct	Total	UA** %	± c.i. UA**	
	1	2	3	4	5	6	7	8	9	10	11	12	13					
1	70	1	1	2	19	1	0	0	1	0	0	0	0	0	70	95	73.7	8.9
2	2	165	0	0	7	0	1	5	1	0	0	0	0	0	165	181	91.2	4.1
3	0	0	18	0	2	0	0	0	1	0	0	0	0	0	18	21	85.7	15.0
4	0	0	0	44	4	0	1	0	0	0	0	0	0	0	44	49	89.8	8.5
5	39	16	1	51	523	27	21	62	12	3	0	1	4	4	523	760	68.8	3.3
6	0	0	2	0	106	277	131	18	15	1	0	46	6	6	277	602	46.0	4.0
7	2	1	0	2	44	71	237	77	8	2	0	106	6	6	237	556	42.6	4.1
8	2	6	0	17	54	36	30	923	7	13	0	3	0	0	923	1091	84.6	2.1
9	1	2	0	0	8	0	3	0	22	1	0	1	4	4	22	42	52.4	15.1
10	0	0	0	0	1	1	0	2	0	30	0	0	0	0	30	35	85.7	11.6
11	0	0	0	0	0	0	0	0	0	0	0	42	14	42	56	75.0	11.3	
12	0	0	0	0	0	1	1	1	1	0	0	285	2	2	285	290	98.3	1.5
13	2	0	0	1	0	2	0	1	0	0	0	3	70	70	79	88.6	7.0	
Correct	70	165	18	44	523	277	237	923	22	30	42	285	70	70	2706			
Total	118	191	23	116	770	414	426	1088	67	50	42	459	93	93				
PA %	51.9	93.3	88.6	37.2	71.6	68.8	56.9	74.8	33.9	52.5	100.0	75.1	71.9	71.9				
c.i. PA±	18.6	4.8	42.3	21.4	2.2	4.1	4.2	2.1	25.7	53.4	0.0	3.3	20.3	20.3				
															Overall accuracy: 71.4 ±1.3			

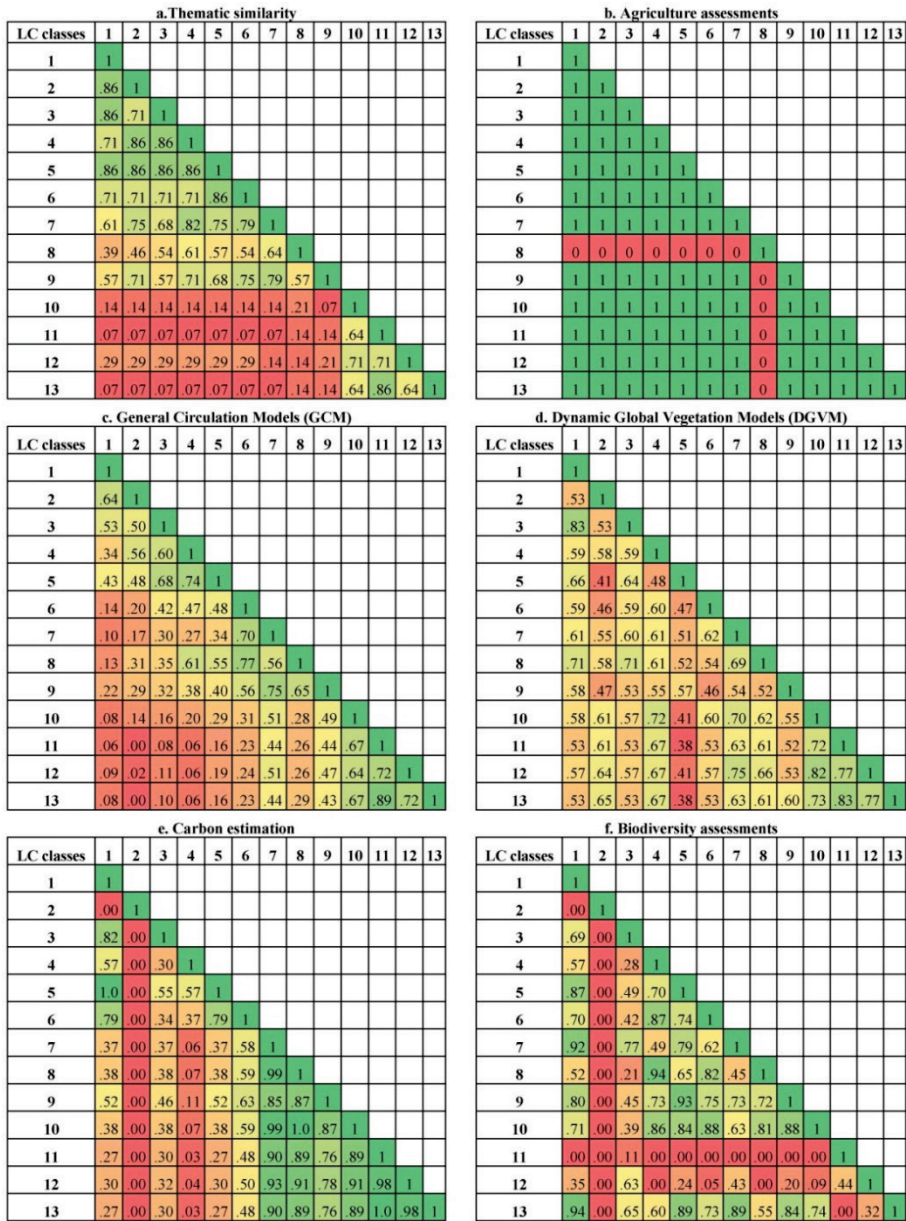
* 1. ENL trees, 2. EBL trees, 3. DNL trees, 4. DBL trees, 5. Mixed trees, 6. Shrubs, 7. Grasses, 8. Cropland, 9. Wetland, 10. Urban, 11. Snow and ice, 12. Bare/sparse vegetation, 13. Water

** UA. User's accuracy, PA. Producer's accuracy, c.i. Confidence intervals

3.3.3 Thematic accuracy assessments of the GLC maps in specific applications

Figure 3.5 shows the used similarity matrices for the applications considered in this work. GLC-map specific similarity matrices for carbon estimation and biodiversity assessments are provided in Table S6 in the supplementary data of the online publication; the averages given in Figure 3.5 are provided to illustrate the main findings.

Figure 3.5 shows that similarity between LC classes varies according to the considered applications. The idea is easiest illustrated for agriculture assessments, where all non-cropland classes were considered mutually equivalent but jointly completely dissimilar from the cropland class (Figure 3.5b). For GCM, forest classes were moderately (0.4-0.6) similar to one another (Figure 3.5c). For DGVM, mixed tree class was less similar (0.4-0.5) from other classes as it represents several plant functional types (Figure 3.5d). For carbon estimation, classes of EBL and DBL trees with high carbon density were dissimilar (<0.1) from other classes (Table S6). For biodiversity assessments, the classes with high and low species richness (EBL tree and bare/sparse vegetation or snow/ice respectively) stand out as being most dissimilar (typically <0.1) from the other classes (Table S6). In weighted accuracy analysis, confusions between similar classes were largely waived (regarded less important), whereas confusions between dissimilar LC classes were identified as severe errors (regarded more important). These differences in the importance of confusion errors per application highlight the need of assessing GLC map accuracy for specific applications (DeFries and Los 1999).



* 1. ENL trees, 2. EBL trees, 3. DNL trees, 4. DBL trees, 5. Mixed trees, 6. Shrubs, 7. Grass, 8. Cropland, 9. Wetland, 10. Urban, 11. Snow and ice, 12. Bare/sparse vegetation, 13. Water.

Figure 3.5. The similarity matrices (weights) used for weighted accuracy assessments

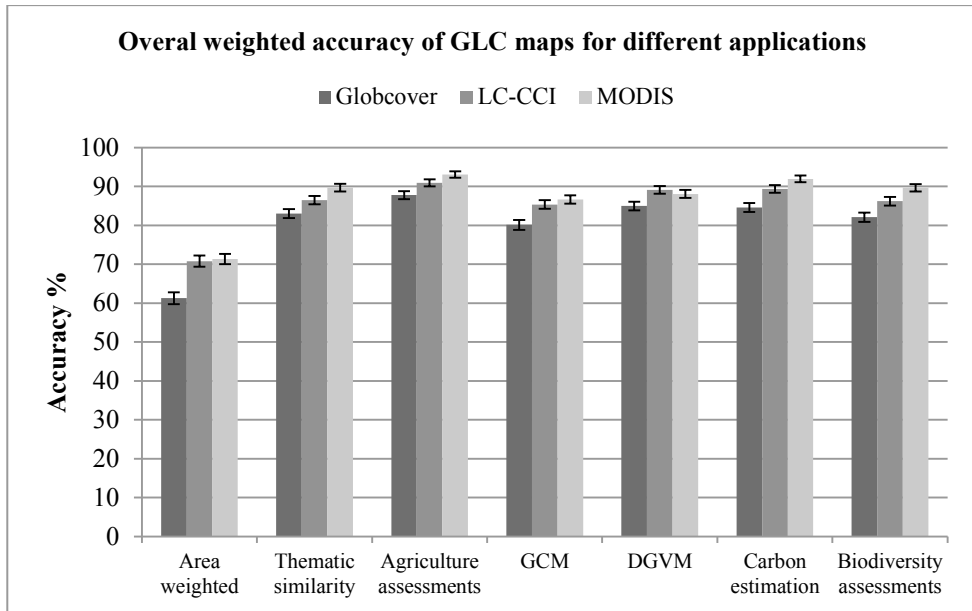


Figure 3.6. Weighted accuracies of the GLC maps for different applications

Figure 3.6 shows the weighted overall accuracies (and accuracy confidence intervals) of the GLC map for specific applications. The overall weighted accuracy of the GLC maps varied between 80-92% for the applications considered. An accuracy increase was observed because confusions between easily confused classes such as mixed trees, shrubs and grass tended to be inconsequential for the applications (Figure 3.5). The MODIS map had the highest accuracy for all applications except the DGVM, where less credits were given to confusion errors between mixed trees, shrubs and grass. The LC-CCI map had slightly higher accuracy for DGVM.

DeFries and Los (1999) and Mayaux et al. (2006) also reported higher overall weighted accuracies in their assessments for simple biosphere modelling and mapping perspective. Such user oriented map accuracy assessment will lead to more user confidence and thus underpins the credibility and usefulness of the land cover mapping program.

The approach of deriving weights used in this study is similar to the work of Herold et al. (2011) and Knight (2002) in which the weights were derived by analysing the similarity between LC classes for specific applications. Other studies have used weighting based on expert opinion on the LC class importance for

certain applications (Kooistra et al. 2010) or LC class separability in case of producer's perspective (Fritz and See 2005). To make weighted accuracy assessments reproducible for different studies, objectively derived weights are recommended (Knight 2002). In this study, the weights were derived quantitatively based on LC class similarities for specific applications.

This study assessed the weighed thematic accuracies at the original resolution of the GLC maps. As GLC maps are often resampled to coarser resolution and used for different applications, differences in resolution should be taken into consideration when using accuracy estimates for uncertainty assessments of the applications.

3.3.4 Improvement priorities of the GLC maps for specific applications

Potential accuracy increase was analysed to identify improvement priorities of the GLC maps. Figure 3.7 shows the improvements of overall weighted accuracy when confusion errors of LC classes are reduced by 50%. The higher the accuracy improvement, the more the impact a LC class has on the overall weighted accuracy. In general, improvement was largest for the Globcover map and smallest for the MODIS map (Figure 3.7). With 50% reduced confusion errors of a LC class, the weighted accuracy improved up to 2% for all applications except agriculture assessments, which had up to 4% improvement (Figure 3.7).

Figure 3.7 also indicates the classes that would contribute most to improving the GLC maps in terms of specific applications if confusion errors were reduced. According to our assessments, these priority classes of the GLC maps are: cropland vegetation for agriculture assessments, grass for GCM, mixed trees, grass and shrubs for DGVM, shrubs and mixed trees for carbon estimation, and grass and bare/sparse vegetation for biodiversity assessments. Besides the difference in improvement priority classes per application, the main priority classes for specific applications were in general the same as the highly confused classes in the confusion matrices (i.e. mixed trees, shrubs and grass) (Table 3.3). These classes occupy large areas in the world.

The most pronounced effects of 50% reduction of confusion between LC pairs on the weighted accuracies are listed in Table S7 in the supplementary data of the online publication. Across all considered applications, reduced confusion between grass, cropland and bare/sparse vegetation classes showed largest improvements (up to 2%) on the weighted accuracy of the Globcover and LC-CCI maps. For the MODIS map, confusion between shrubs, grass and mixed trees was deemed to have a high impact on the overall weighted accuracy.

The results on the map improvement priorities can directly be used to enhance the GLC maps, particularly the LC-CCI and MODIS maps. Since the LC-CCI maps were developed specifically for climate modellers (CCI-LC 2014), priorities for improvement are set by the results obtained for GCM and DGVM. The MODIS map is regularly updated with new collections of MODIS data (Friedl et al. 2010). Although this map was developed without targeting specific applications, developers can target the classes that have high impact on the overall accuracy for multiple applications. For example, the efforts made on mixed tree and cropland classes would significantly benefit agriculture assessments, carbon estimation and DGVM applications.

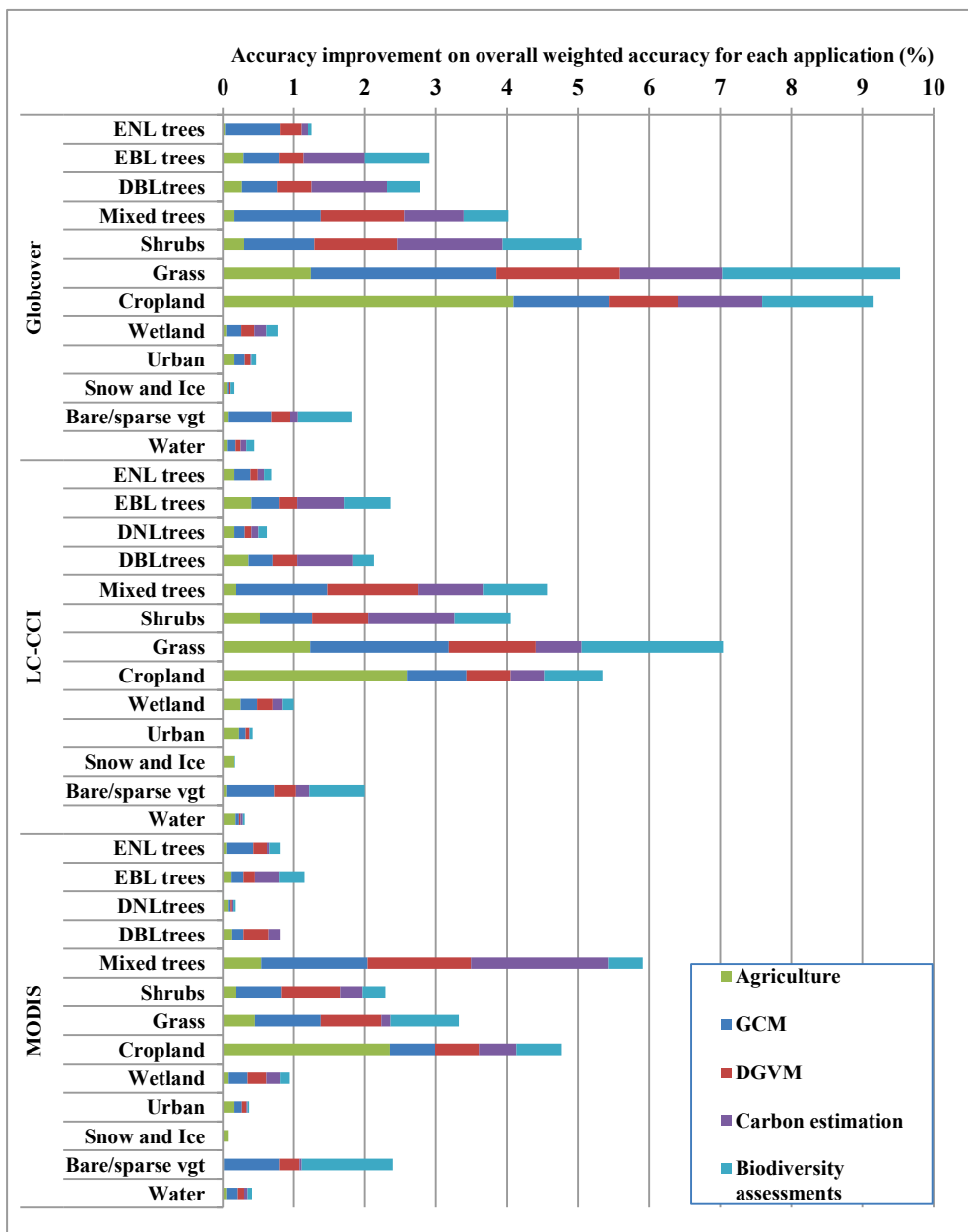


Figure 3.7. Accuracy improvement on overall weighted accuracy when the class confusion errors are reduced by 50%.

3.4 Conclusions

This study compared the thematic accuracy of the GLC maps using the Globcover-2005 GLC reference dataset. Our analysis showed the possibility of using an existing GLC reference dataset for multiple map validations with additional considerations such as legend translation and re-interpretation. The LCCS-classifier information of the reference dataset was useful for addressing the thematic differences of the three legends. The study highlights the use of existing GLC reference dataset for multiple map validation if the dataset has flexible thematic classifiers and quantitative class proportion information that are consolidated and verified.

The overall accuracies of the LC-CCI and MODIS maps were around 70% which are similar to the previous GLC maps, despite several significant improvements in technology and methodology of GLC mapping. Moreover, the inability of the GLC mapping approaches to clearly discriminate mixed trees, shrubs and grass continue to be a problem even in recent GLC maps. Emphasis should be put on these classes to improve GLC maps as suggested by Herold et al. (2008). Future studies will focus on combining these GLC maps and existing reference datasets to create an improved GLC map.

We demonstrated the weighted accuracy assessment of the GLC maps for five different applications by analysing the importance of LC confusion errors for the applications. Our results suggests the accuracy of GLC maps should be assessed per applications when the aim is to determine the fitness of GLC maps for an application; there is no single-figure accuracy estimate expressing map fitness for all purposes. For this reason, deriving the weights that denote the importance of confusion errors for an application is an important process, and this process should be well documented and reproducible.

This study presented an approach to identify priorities for improving GLC maps by assessing the impact of LC class and confusions on the overall weighted accuracies. This approach can be applicable to other similar studies. The results from this analysis can be used to further improve the GLC maps such as the LC-CCI and MODIS, which are progressively being updated.

Acknowledgements

This study was supported by the EU FP7 project GEOCARBON, the ESA Land Cover CCI project and the ESA GOF-C-GOLD project office. We are grateful to the experts involved in the Globcover-2005 reference data collection for their input in developing the dataset. We would like to thank Sukhad Keshkamat, Sarah Carter and Peter Griffith for reviewing the manuscript and providing useful comments.

Chapter

4

Spatial accuracy assessment and integration of global land cover datasets

Nandin-Erdene Tsendbazar, Sytze de Bruin, Steffen Fritz and Martin Herold

Published in Remote Sensing, Volume 7, Issue 12, December 2015, Pages 15804-15821

Abstract

Along with the creation of new maps, current efforts for improving global land cover (GLC) maps focus on integrating maps by accounting for their relative merits,

e.g., agreement amongst maps or map accuracy. Such integration efforts may benefit from the use of multiple GLC reference datasets. Using available reference datasets, this study assesses spatial accuracy of recent GLC maps and compares methods for creating an improved land cover (LC) map. Spatial correspondence with reference dataset was modeled for Globcover-2009, Land Cover-CCI-2010, MODIS-2010 and Globeland30 maps for Africa. Using different scenarios concerning the used input data, five integration methods for an improved LC map were tested and cross-validated. Comparison of the spatial correspondences showed that the preferences for GLC maps varied spatially. Integration methods using both the GLC maps and reference data at their locations resulted in 4.5%–13% higher correspondence with the reference LC than any of the input GLC maps. An integrated LC map and LC class probability maps were computed using regression kriging, which produced the highest correspondence (76%). Our results demonstrate the added value of using reference datasets and geostatistics for improving GLC maps. This approach is useful as more GLC reference datasets are becoming publicly available and their reuse is being encouraged.

Keywords: global land cover; data integration; spatial accuracy; geostatistics; reference dataset.

4.1 Introduction

Multiple global land cover (GLC) maps have been produced over the past decades. These maps are used for various applications such as climate modeling, food security, biodiversity, ecosystem services and environmental monitoring (Verburg et al. 2011). Currently, GLC map production is progressing towards higher resolution maps, namely the Land Cover-CCI (LC-CCI) maps at 300 m resolution and the Fine Resolution Observation and Monitoring (FROM-GLC) and Globeland30 maps at 30 m resolution (Chen et al. 2015; Mora et al. 2014). However, these maps were developed using different input data and methods (Herold et al. 2008), and as a consequence, considerable disagreements amongst GLC maps have been found (Fritz et al. 2011b; Herold et al. 2008). Despite efforts in advancing GLC mapping approaches, the accuracy of GLC maps has not improved significantly and continues to be around 70% (Tsendbazar et al. 2016). Such accuracies mostly do not meet the requirements of GLC map users (Kooistra et al. 2010) and thus, there is a need to improve GLC maps.

A common approach to improving GLC maps has been the integration of existing GLC maps using a variety of methods (Jung et al. 2006; See et al. 2015). In map integration, pixels are assigned to land cover classes based on class labels from multiple GLC maps, sometimes in combination with other data sources. For example, Jung et al. (2006) created the SYNMAP by assigning the land cover (LC) class that multiple GLC maps agreed upon. Iwao et al. (2011) adopted a LC class favoured by the majority of GLC maps and a LC class with highest accuracy in case of no majority. Tuanmu and Jetz (2014) created a GLC map specifically for biodiversity and ecosystem modeling applications by integrating the reported LC class accuracies and the map resolution. Other researchers focused on map integration for cropland and forest biomass datasets (Fritz et al. 2015; Ge et al. 2014). For example, Fritz et al. (2015) created a 1 km global cropland percentage map by integrating several cropland maps at global to national scales along with national crop statistics. Ge et al. (2014) generated a biomass map for Eastern Africa by fusing existing biomass maps using weights associated with the accuracy of source maps. This approach was improved and applied to a larger area to create an integrated pan-tropical biomass map using multiple reference datasets (Avitabile et al. 2015).

Existing reference datasets that were built for calibrating and validating GLC maps can be re-used in the integration of GLC maps. However, only a few studies have considered these datasets for integration. For example, Kinoshita et al. (2014) assessed the presence probability of LC classes using logistic regression and the Degree Confluence Project (DCP) dataset and used this for integration.

See et al. (2015) created hybrid GLC maps using Geo-Wiki reference data within a geographically weighted kernel approach (Comber et al. 2013). Similarly, Schepaschenko et al. (2015) created a global hybrid forest cover map based on different forest and land cover maps and a dataset collected through the Geo-Wiki platform. The above studies made limited use of existing GLC reference datasets for integration and reported improvements on the integrated maps. Currently, several GLC reference datasets are being made accessible via the Global Observation of Forest and Land Cover Dynamics (GOFC-GOLD) reference data portal and Geo-Wiki platform (Fritz et al. 2011a; GOFC-GOLD 2014) and this enables assessment of their utility for improving existing GLC maps.

The accuracy of GLC maps is often expressed in terms of global accuracies assessed from statistical sampling. Global accuracies do not inform about spatial variability in map accuracy, yet classification errors are not distributed evenly across the map (Carneiro and Pereira 2012). Spatial variation of map accuracy can be modeled using indicator kriging (de Bruin 2000a; Kyriakidis and Dungan 2001). Carneiro and Pereira (2012) and Kyriakidis and Dungan (2001) used indicator kriging to assess spatial accuracy of regional scale land cover maps. These types of assessments require a large number of reference sample sites with a good geographical spread, which explains why spatial variability of GLC map accuracies has hardly been studied. However, See et al. (2015) assessed spatial correspondence of GLC maps with the Geo-Wiki volunteer based reference data using geographically weighted kernel approach. With the available GLC reference datasets from the GOFC-GOLD and Geo-Wiki platform, the number of reference sample sites increases substantially and a combined reference dataset could be used to model the spatial variability of accuracy of large-scale LC maps.

The objective of this chapter is to analyze and compare the spatial correspondence of recent GLC maps and to integrate available GLC maps and reference datasets for generating a LC map with improved correspondence to reference LC. Firstly, we assess the spatial correspondence of the recent GLC maps for the year 2010 ± 1 with available GLC reference data. Our analysis involved the Globcover 2009, LC-CCI 2010, MODIS 2010 and Globeland30 maps. The assessment focuses on Africa—a continent with complex heterogeneous landscapes that are known to causes inconsistencies among GLC maps (Huttich et al. 2011; Kaptué Tchuenté et al. 2011). Secondly, we test five different integration methods to create an improved LC map. Three of these methods are based on integration of GLC maps and reference datasets, one method is based on the GLC maps only and the other method is based on the reference datasets only. We assess the performance of the integration methods by cross-validation. Finally, we create an improved LC map

using the method selected by cross-validation and discuss the use of available reference datasets for integration.

4.2 Data

4.2.1 Global land cover maps

The GLC maps included in this study are: Globcover, LC-CCI, MODIS and Globeland30 maps for the year 2010 (2009 for Globcover). The Globcover project of the European Space Agency (ESA) provided a GLC map for 2009 based on 300 m resolution MERIS satellite data (Defourny et al. 2011b). This map has an LCCS (United Nations Land Cover Classification System) based legend with 22 classes and the thematic overall accuracy of the Globcover-2009 was reported as 70.7% based on 1484 homogenous sample sites (Bontemps et al. 2011a).

Recently, ESA's Land Cover-CCI (LC-CCI) project delivered three consecutive GLC maps for the epochs of 2000, 2005 and 2010 at 300 m resolution using the MERIS data archive (CCI-LC 2014). These maps were specifically targeted to meet the requirements of climate modelers. The maps also have an LCCS based legend with 22 classes. The overall thematic accuracy of the LC-CCI 2010 map was 74.4% based on the same reference sample as the Globcover-2009 validation.

Using data from the Moderate Resolution Imaging Spectroradiometer (MODIS), Boston University provided annual MODIS Collection-5 GLC maps at 500 m resolution (Friedl et al. 2010). The MODIS GLC maps have five legends including a legend based on the International Geosphere-Biosphere Programme (IGBP) classification scheme with 17 classes. The accuracy of the MODIS GLC map of 2010 has not yet been assessed. However, based on cross-validation using the training dataset, Friedl et al. (2010) reported an overall accuracy of 74.8% for the 2005 map.

The GlobeLand30 project of the Ministry of Science and Technology of China generated a GLC map for the year 2010. The GlobeLand30 map was derived from 30 m resolution multispectral images of Landsat TM and ETM+ as well as the Chinese Environmental Disaster Alleviation Satellite (HJ-1). We used the 250 m resolution version of Globeland30, which contains LC class fraction information. This map has 10 LC classes. The overall accuracy of the map has been reported to be 83.5% (Chen et al. 2015). All above-mentioned map accuracies concern the global extent, and for specific regions such as Africa different accuracies are expected. Table 4.1 provides a summary of the used GLC maps.

All the GLC maps were cropped to the extent of Africa. The MODIS and Globeland30 maps were resampled to 0.00278 degrees resolution using nearest neighbor assignment to match the resolution of the Globcover and LC-CCI maps. The legends of the GLC maps were harmonized into eight general LC classes following the approach of Herold et al. (2008), which provides a table for harmonizing input classes into 13 general LC classes using LCCS-based legend translation protocols. Since the Globeland30 map does not have detailed forest classes, we used a single general forest class only (Table 4.2). Figure 4.1 presents the four GLC maps with the harmonized legend.

Table 4.1. A summary of GLC maps used for comparison.

GLC Map	Globcover	LC-CCI	MODIS	Globeland30
Spatial resolution at the Equator	300 m	300 m	500 m	250 m
Input data	MERIS: Bi-monthly from 10-day composites	MERIS global SR composite, SPOT-VGT time series (for updating)	MODIS: Monthly EVI, LST and 7 bands from 8-day composites	Landsat TM, ETM+ and HJ-1 multispectral images
Time of data collection	2009	2008–2012	2010	2010 ± 1 year
Classification method	(Un)supervised spatio-temporal clustering; expert-based labeling	Unsupervised spatio-temporal clustering; machine learning classification	Supervised decision tree boosting	Integration of pixel and object based classification and Knowledge based interactive verification
Classification scheme	LCCS based: 22 classes	LCCS based: 22 classes	5 different legends including the IGBP (17 classes)	10 classes
Reference	(Bontemps et al. 2011a)	(CCI-LC 2014)	(Friedl et al. 2010)	(Chen et al. 2015)

Table 4.2. General land cover classes and corresponding classes of the GLC datasets.

Code	Land Cover Class	Globcover	LC-CCI	IGBP (MODIS, STEP and VIIRS)	GLC2000	Geo-Wiki	GLCNMO
1	Forest	40–110, 160, 170	50–100, 160, 170	1–5, 8, 9	1–10	1	1–5
2	Shrubland	130	120	6, 7	11, 12	2	7
3	Grassland	120, 140	110, 130, 140	10	13	3	8, 9
4	Cropland (incl. mixtures)	11–30	10–40	12, 14	16–18	4	11, 12, 13
5	Wetland vegetation	180	180	11	15	6	15
6	Urban/built up	190	190	13	22	7	-
7	Bare/sparse vegetation	150, 200	150, 200	16	14, 19	9	10, 16, 17
8	Water and Snow/Ice	210, 220	210, 220	15, 17	20, 21	8, 10	-

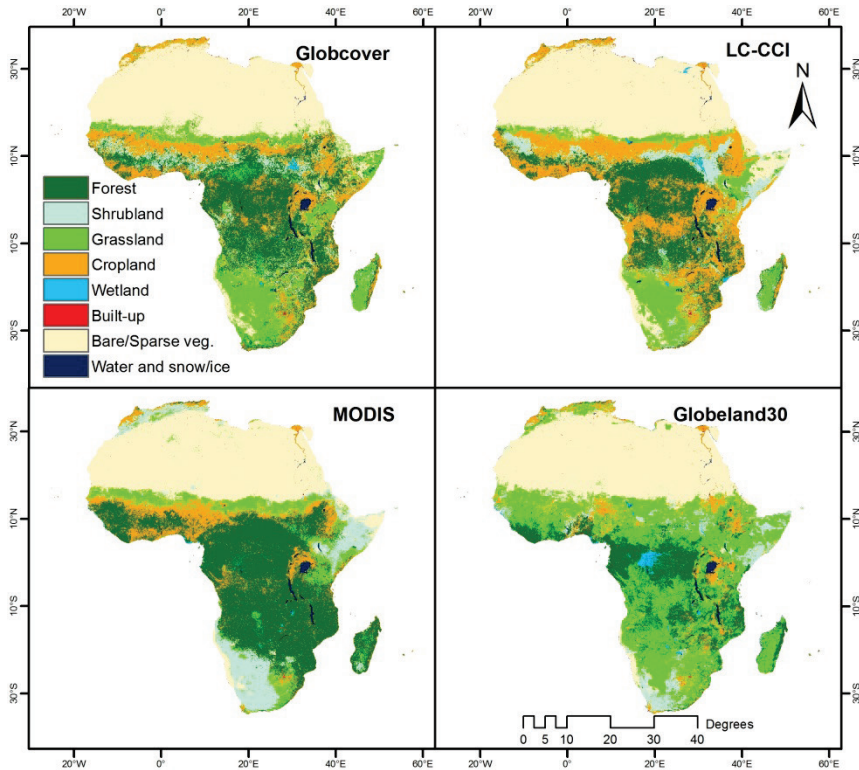


Figure 4.1. Global land cover maps used in the analyses.

4.2.2 Reference datasets

The reference datasets used in this work are denoted as: GLC2000rd, GLCNMOrd, Geo-Wikird, MODIS/STEPrd, VIIRSrd and the Globcover-2005rd. The subscript “rd” is added here to avoid potential confusion between LC maps and reference datasets. GLC2000rd concerns the consolidated version (11 LC classes) of the reference dataset generated for validating the Global Land Cover 2000 map (Mayaux et al. 2006; Schultz et al. 2015). GLCNMOrd refers to the calibration dataset of Global Land Cover by National Mapping Organizations, which was used to generate a GLC map for 2003 (Tateishi et al. 2011). This dataset employs 14 LC classes, which were assigned to sample sites by international experts. MODIS/STEPrd has been used to calibrate the MODIS collection 4 and 5 GLC maps (Friedl et al. 2010). This dataset was developed and updated by Boston University and it has 17 LC classes according to the IGBP legend. Boston university also created VIIRSrd (Visible Infrared Imaging Radiometer Suite), which was used to validate the VIIRS surface type products (GOFC-GOLD 2014;

Olofsson et al. 2012). The reference LC of this dataset was assigned by visual interpretation of very high-resolution images using the same classes as MODIS/STEPrd. Geo-Wikird was developed through a volunteer based online platform and volunteers' interpretation of the reference LC was validated by a group of experts (Fritz et al. 2009). Globcover 2005rd is a re-interpreted version of the reference dataset that was built for validating the Globcover 2005 GLC map (Bicheron et al. 2008; Tsendbazar et al. 2016). Detailed information on the characteristics of the available reference datasets are provided in (Tsendbazar et al. 2015a). Although there are temporal differences between the used datasets, we deemed these to be of minor importance, since errors owing to LC changes over the time frame are negligible compared to misclassification errors of the GLC maps.

These reference datasets are publicly accessible through the GOFC-GOLD Reference data portal, Geo-Wiki portal and International Steering Committee for Global Mapping (Fritz et al. 2009; GOFC-GOLD 2014; Tateishi et al. 2011).

To cope with differences in sample site areas across the reference datasets, we assumed that the LC of the sample site corresponds to the LC of the centroid of that sample sites. Reference data were then compared to the LC classes of the GLC maps at the centroids of the reference sites. For the combined reference dataset, the legends of all reference datasets were harmonized into the eight general classes listed in Table 4.2 to correspond with GLC map harmonization as given in Section 4.2.1.

In total, 3887 sample sites within Africa were used in this study. Based on this reference dataset, model-based geostatistical analysis was used since in contrast to design-based inference it does not require a probability sampling design. Figure 4.2 shows the sample distribution of each reference dataset (left) and the reference LC classes (right).

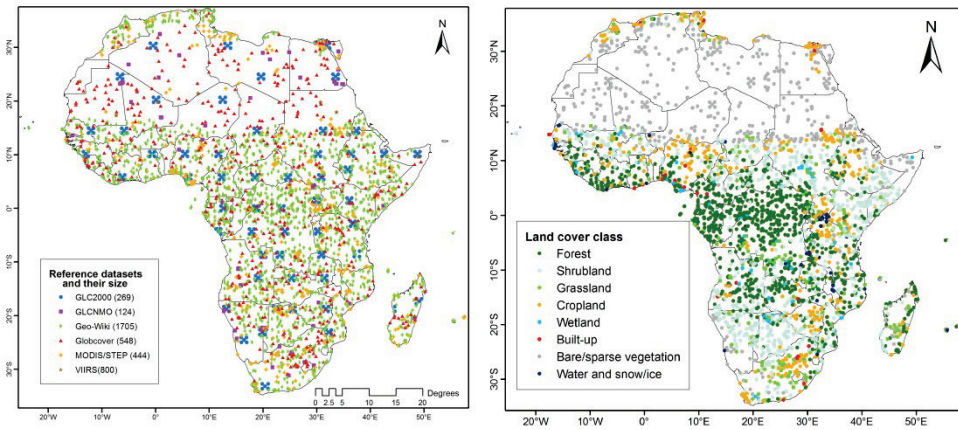


Figure 4.2. Spatial distribution of the reference datasets (**left**) and reference LC classes (**right**).

4.3 Method

4.3.1 Spatial correspondence assessment

To assess spatial accuracy (spatial variation in map accuracy), we analyzed the spatial correspondence of the GLC map with the reference dataset. Correspondences between GLC maps and reference data were indicator coded. If the LC class of the reference sample site matched with that of a map, an indicator code 1 was assigned to that sample site. Conversely, an indicator code 0 was given to sites where the mapped LC differed from the reference class. Next, we analyzed spatial autocorrelation of the indicator-coded data (correspondence with reference LC) using indicator semivariograms. Nested variogram models were fitted to experimental semivariogram data obtained by the method of moment approach with binning of 3–5, 10–15 and intervals of 25 km (Pebesma and Wesseling 1998). Variograms were fitted by weighted least squares using N_j/h_j^2 as weights, where N_j denotes the number of point pairs in the j -th lag and h_j is the corresponding lag distance.

Spatial correspondence maps were created for each GLC map for Africa at 0.00278 degrees resolution (300 m at the Equator) by indicator kriging (Bierkens and Burrough 1993) using the *gstat* package in R (Pebesma 2004). The spatial correspondence maps depict the local correspondence values ranging between 0 and 1, which denotes the local probability that a particular map is correct. Figure 4.3 demonstrates the semivariograms of spatial correspondence for the GLC maps

and used fitted models for the indicator kriging. To restrict the size of the kriging system, kriging was done with the nearest 50 observations.

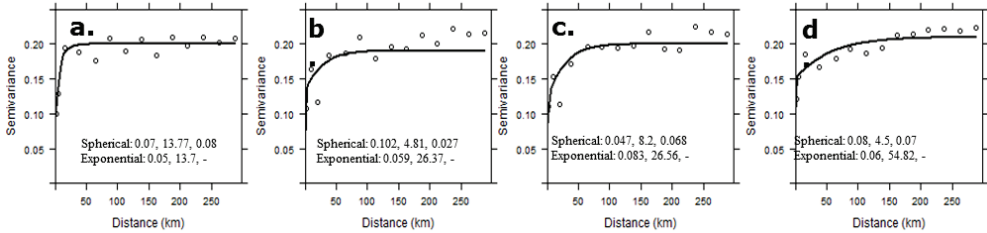


Figure 4.3. Semivariograms and fitted models for spatial correspondence of the Globcover (a), LC-CCI (b), MODIS (c), and Globeland30 (d) maps (model parameters: partial sills, range and nugget).

4.3.2 GLC dataset integration

Analyzing the local variation in map accuracy is useful for obtaining information on where a map is accurate and where not, and this information can be valuable in creating an improved GLC map. Previous integration efforts of GLC maps did not focus on the local variation in map accuracy except the work of See et al. (2015), who analyzed GLC maps with highest correspondence at a coarser grids of 0.25 degrees using geographically weighted kernel approach. However, the resulting integrated maps have artifacts in the pattern of LC classes that are caused by the coarse grid kernels (See et al. 2015).

Our study extends the principle of considering local variation of map accuracies and LC class probabilities for creating an improved LC map. We used a geostatistical approach to assess and model the spatial dependence of map accuracy and class probabilities. We compared different integration methods, as depicted in Figure 4.4, which represent a variety of choices concerning the use of input datasets. These include methods based on spatial correspondence of the GLC maps, agreement amongst input maps and the LC class presence probabilities, *i.e.*, using both the GLC maps and the reference datasets. In addition, methods based on a conventional voting approach (Iwao et al. 2011), *i.e.*, without using reference data, and a geostatistical method that relies only on the reference data, *i.e.*, without using the GLC maps, were also compared. We first applied all methods to the sample locations. After selecting the integration method with highest correspondence by cross-validation (see Section 4.3.2.6), the latter was applied to the full extent of Africa.

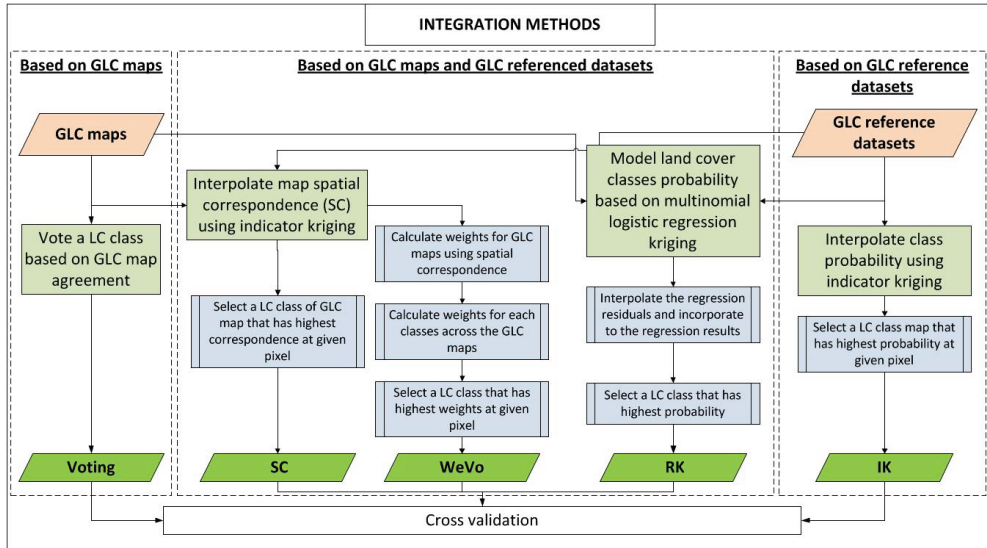


Figure 4.4. Conceptual diagram of different integration methods.

The following subsections describe the integration methods used in this study.

4.3.2.1 Voting

This integration method only uses the GLC maps as input. At each pixel location, the LC class corresponding to the majority of the mapped LC classes of the four input maps was assigned. In case of a tie, the LC class of a map that has the highest overall reported accuracy was assigned. Since there is no information on the accuracy of these maps in Africa, the reported global confusion matrices of the maps (see Table 4.1 for reference) were converted into confusion matrices for the eight generalized classes and the corresponding overall accuracy was calculated. The global accuracies at generalized class level were computed as 66%, 75.3%, 85.4% and 83.5% for the Globcover, LC-CCI, MODIS and Globeland30, respectively. Accordingly, the MODIS LC class was assigned in case of ties.

4.3.2.2 Spatial Correspondence (SC)

This method (SC) uses both the GLC maps and reference datasets as inputs. Based on the spatial correspondence map for each GLC map resulting from the method described in Section 4.3.1, we selected the LC class of the map that has the highest spatial correspondence value at a pixel location.

4.3.2.3 Weighted Voting (WeVo)

Weighted voting (WeVo) also uses both the GLC maps and the reference datasets. We created normalized weight maps using the spatial correspondence maps of the GLC maps. Let $sc_i(x)$ denote the spatial correspondence of the i -th GLC map ($i = 1, \dots, 4$) at location x . $W_i(x)$, the weight assigned to map i at location x , is then:

$$W_i(x) = \frac{sc_i(x)}{\sum_{i=1}^4 sc_i(x)} \quad (4.1)$$

LC classes were dummy coded into multiple 1 or 0 indicators, where 1 indicates that a LC class k ($k = 1, \dots, 8$) is present and 0 means k is absent. Using these indicator values, we assigned the weights to the classes mapped on each of the GLC maps. For each LC class k , a total weight of the LC class at x location was created by summing the class weights of the four GLC maps (Equations 4.2 and 4.3).

$$w_{i,k}(x) = W_i(x) * k_i(x) \quad (4.2)$$

$$W_k(x) = \sum_{i=1}^4 w_{i,k}(x) \quad (4.3)$$

where k is the LC class, $W_k(x)$ is the total weight of the LC class at location x , and $w_{i,k}(x)$ is class weight of the GLC map. A LC class with highest total weight at a location ($W_k(x)$) was then selected for this method.

4.3.2.4 Regression Kriging (RK)

Regression kriging (RK) similarly uses both the GLC maps and the reference datasets. The general trend of probabilities of presence of LC classes were predicted using a multinomial logistic (MNL) regression model. These were locally adjusted by interpolating indicator residuals by simple kriging (Equation 4.4).

$$p_k(x) = \pi_k(x) + \varepsilon_k(x) \quad (4.4)$$

where $p_k(x)$ denotes the presence probability of a LC class at location x , $\pi_k(x)$ is a predicted probability trend of an LC class that was obtained by MNL regression (Kempen et al. 2009) and $\varepsilon_k(x)$ is the indicator residuals for that class. The latter was obtained by simple kriging. MNL regression also uses indicator values of the LC classes. There is an indicator variable for all but one class (Kempen et al. 2009). The MNL regression estimated a separate binary logistic regression model

for each of these indicator variables. For each indicator variable ($k = 2, \dots, 8$), the log odds function for predicted probability is:

$$\eta_k(x) = \log\left(\frac{\pi_k(x)}{1 - \pi_k(x)}\right) = \beta_{0k} + \beta_{1k}h_1(x) + \beta_{2k}h_2(x) + \dots + \beta_{jk}h_j(x) \quad (4.5)$$

where h_j (with $j = 1, \dots, 4$) are the explanatory variables (LC class of the four GLC maps at sample locations), $\beta_{1k} \dots \beta_{jk}$ are the regression coefficients and β_{0k} is the intercept. To ensure that all probabilities are in the interval $[0,1]$ and that the probabilities sum to 1, Equations 4.6 and 4.7 were used (Kempen et al. 2009).

$$\pi_k(x) = \frac{\exp(\eta_k(x))}{\sum_{k=2}^8 \exp(\eta_k(x))} \quad (4.6)$$

$$\pi_1(x) = 1 / \sum_{k=2}^8 \exp(\eta_k(x)) \quad (4.7)$$

where $\exp(\eta_k(x))$ denotes the odds of class k at location x . This was implemented using the `nnet` package in R (Ripley et al. 2014).

Next, regression residuals at sample locations were calculated and simple kriging was used to interpolate the regression residuals ($\varepsilon_k(x)$) at un-sampled locations for all classes except water. For the water class, no spatial correlation was observed on the regression residuals based on the experimental semivariogram. Semivariograms were fitted using the same method as described in Section 4.3.1. Figure 4.5 demonstrates the semivariograms of regression residual for the LC classes and fitted variogram models used for kriging.

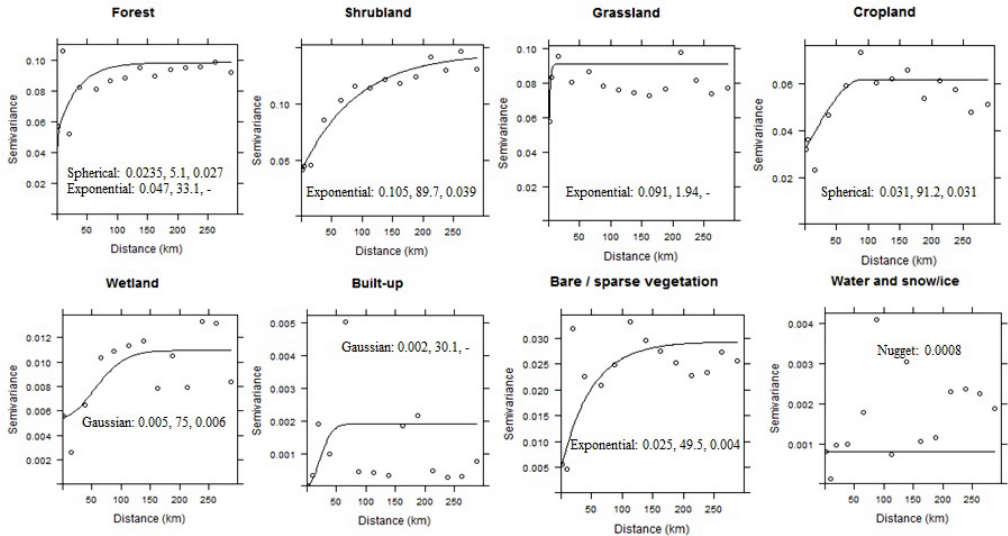


Figure 4.5. Semivariograms and fitted models for residual kriging.

After adjusting the predicted probabilities with residual kriging, any probability outside the interval $[0, 1]$ was set to the closest bound, zero or one. Subsequently, the estimates $pk(x)$ $k = 1, \dots, K$ were normalized by their sum to meet the condition $\sum_{k=1}^K p_k(x) = 1$ (de Bruin 2000a). A pixel was assigned to the LC class having the highest probability.

4.3.2.5 Indicator Kriging (IK)

For comparison, the last integration method was based on indicator kriging that uses only the reference datasets. Based on these indicator variables for LC classes, the presence probability of LC classes was modeled at the test locations of the cross validation (see next section).

Figure 4.6 shows the semivariograms and the fitted models used for modeling LC classes presence probability based on indicator kriging (Section 4.3.1). A LC class with highest modeled probability was selected for this method.

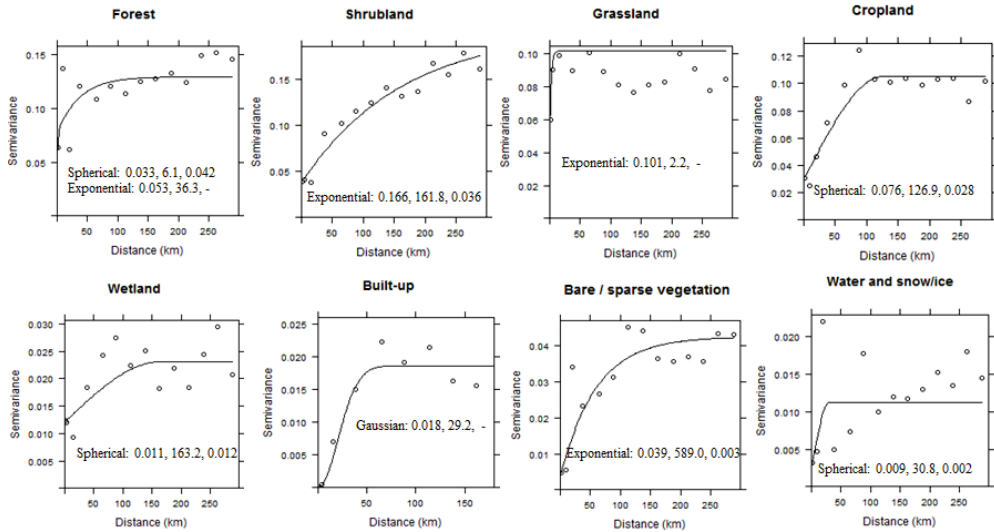


Figure 4.6. The semivariograms and fitted models used for Indicator Kriging.

4.3.2.6 Cross-validation

The performance of these methods was analyzed using 10 fold cross-validation. The reference sample sites were partitioned into 10 random subsamples. Nine subsamples (3498 ± 1 sample sites) were used to train the integration methods and one subsample (389 ± 1 sample sites) was used to validate the method performance by assessing the overall correspondence between the reference LC and LC from method outputs. This step was repeated 10 times so that each subsample was used for method training as well as validation and each sample site was used for validation exactly once. The median percentage of integrated LC classes locally corresponding with reference subsamples was then calculated. Note that these values should not be confused with the overall accuracy of LC maps since they are based on cross-validation using a heterogeneous sample rather than comparison against an independent reference dataset obtained by probability sampling. Based on the cross-validation results, the integration method having the highest correspondence with the reference LC was selected for creating an improved LC map.

4.4 Results and discussions

4.4.1 Spatial correspondence of GLC maps in Africa

The spatial correspondences of the GLC maps based on indicator kriging are provided in Figure 4.7. In terms of the spatial correspondence with reference LC, all four maps show similar trends. The Sahara desert and tropical rainforest regions were mapped with high correspondence, whereas the Sahel, and dry and moist savannah regions were generally mapped with low correspondence. In the latter regions, some differences in terms of spatial correspondence of the maps could be observed. For instance, the LC-CCI showed higher spatial correspondence related to cropland areas in Morocco and northern Algeria, Ethiopia, Eritrea, Sudan, Zambia, Zimbabwe and Angola. In other regions, the LC-CCI map tends to over-represent the cropland class. The MODIS map had higher correspondence in Somalia, Kenya, Mozambique, Namibia, Botswana and western part of South Africa as it has more shrubland areas. The Globeland30 map had higher correspondence in the tropical forest regions of western Africa, Chad, Uganda, Tanzania, Madagascar, and eastern part of South Africa related to grassland areas. A general tendency of over-representing the grassland class was also observed for the Globeland30 in other regions. These differences are also highlighted in Figure 4.7f, which illustrates the maps with highest correspondence at a given location. The strengths of the GLC maps over one another in different regions show the potential of creating an improved GLC map by integrating them.

Figure 4.7e shows the maximum spatial correspondence of the four maps and this demonstrates that the Sahel and dry savannah regions of Africa were mapped with the lowest spatial correspondence in all four maps. This could be due to the presence of multiple LC classes (*i.e.*, heterogeneous landscapes) in transition zones of major ecosystem, which are difficult to classify correctly, owing to spectral and thematic similarity. GLC maps often do not agree in these regions (Herold et al. 2008; Kaptué Tchuenté et al. 2011). These regions should be the main focus of map improvement efforts including the development of new GLC maps.

Information on the spatial variation in map correspondence is useful in uncertainty assessments of applications that use GLC maps and in map improvement efforts. It also provides confidence in using the GLC maps for regions with high map spatial correspondence and limited regional data availability.

In terms of correspondence with reference LC classes aggregated over the entire African continent, MODIS had highest correspondence (63%) followed by Globeland30 (57.2%), LC-CCI (55.5%) and finally Globcover (50.8%).

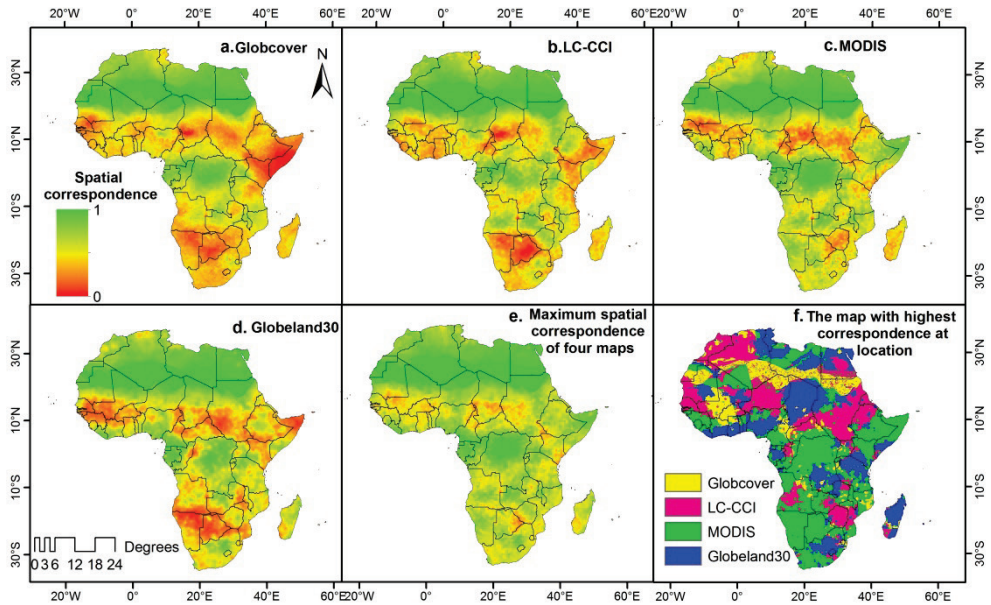


Figure 4.7. Spatial correspondence of the GLC maps (a–d), maximum correspondence (e) and the map with highest correspondence at location (f).

4.4.2 GLC dataset integration methods

The result of the 10 fold cross-validation assessing the performances of the integration methods for an improved GLC map is presented in Figure 4.8. The medians of correspondence of integrated LC with reference data varied from 62.3%–76% across different integration methods (Figure 4.8). The integration method based on only the GLC maps (Voting) resulted in the lowest correspondence of 62.3%, which is less than the 63% of the MODIS map. A possible explanation is that the voting rule will assign a pixel to a wrong class if the majority of input maps agrees to that class.

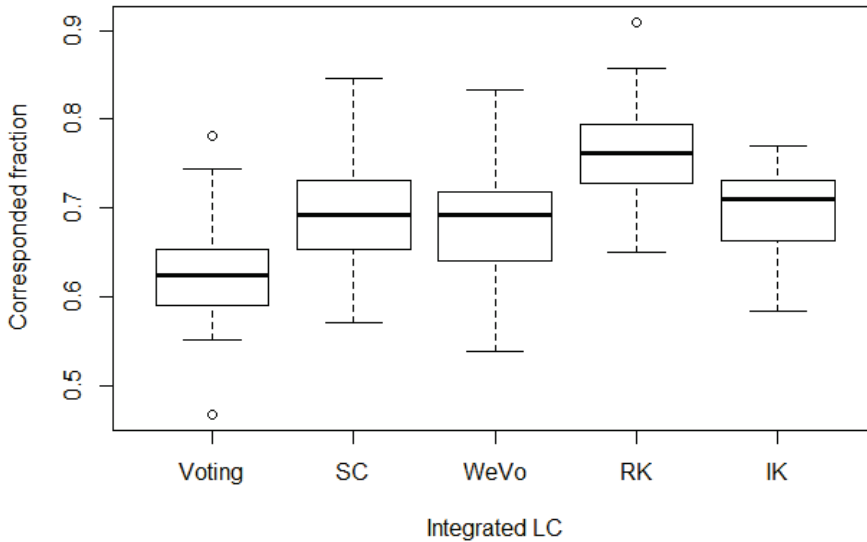


Figure 4.8. Correspondence of integrated LC with reference sample LC (10-fold cross validation).

The integration methods based on both the GLC maps and reference datasets resulted in 67.5%–76% correspondence with the reference LC, which is at least 4.5%–13% higher than the correspondence of the input maps. The RK method produced the highest correspondence (76%) compared with the other integration methods. The RK method ensures to reduce the classification errors as much as possible by exploiting the “best” of the available data *i.e.*, modeling global trends of the LC class probabilities using the GLC maps as explanatory variables and calculating the local deviations from the global trends near reference points using spatial correlation of the residuals between the trends and reference data (Hengl et al. 2004). The smaller sill values of the fitted models for residual kriging compared to that of the indicator kriging (Figure 4.5 and Figure 4.6) are indicative of the contribution of MNL regression in explaining the LC class probabilities (Hengl et al. 2004). This also justifies the use of residual kriging to model the remaining unexplained spatial variation of LC class probabilities.

Figure 4.8 shows that all methods using reference data produced higher correspondence than the Voting method. This could have been expected, since more data are being used. However, even IK that uses only reference data produced better correspondence than Voting. This underlines the importance of reference data in map improvement efforts. The spread in the cross-validation results obtained by IK is expected, since cross-validation repeatedly removes difference subsets of the reference data while IK is based on the reference data only. The

intermediate positions of SC and WeVo can be explained by the fact that they employ map spatial correspondence and agreements amongst input maps, rather than class specific probabilities as considered in the RK method. Using different methods,

See et al. (2015) also observed limitations in using map spatial correspondence and agreement amongst map for integration. Our results demonstrate the advantage of using both the GLC maps and the reference data for integration where data abound while relying on the GLC maps only in places where the reference data is sparse.

4.4.3 Integrated LC and LC probability maps of Africa

Since the RK integration method had the highest correspondence with reference LC (see Section 4.4.2), we used this method to create an integrated LC map of Africa using the input GLC maps and reference datasets (Figure 4.9). The integrated map had similar pattern to the input maps in terms of forest and bare/sparse vegetation classes. The main difference between the integrated map and the input maps is the fact that more area of shrubland and relatively less area of cropland and grassland are present. On the other hand, the general patterns of the LC classes were similar to those of the reference data (Figure 4.2 right).

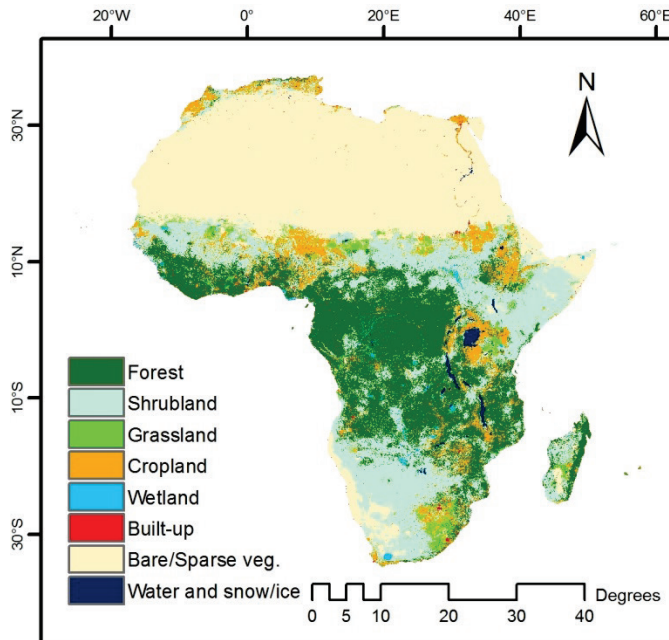


Figure 4.9. Integrated LC map based on RK method.

Table 4.3 compares class-specific correspondences of RK integration and the input maps with the reference dataset. All LC class correspondences were derived by cross-validation (see Section 4.4.2). The RK method improved class correspondences for LC classes excluding forest, cropland, grassland and built-up. The forest, grassland and cropland classes were over-represented in the MODIS, Globeland30 and LC-CCI maps, respectively.

Table 4.3. Class-specific correspondences of RK integration and the input GLC maps with reference data.

	Globcover	LC-CCI	MODIS	Globeland30	RK
Forest	71.1	67.3	90.2	63.7	84.9
Shrubland	11.9	21.3	26.9	17.3	70.8
Grassland	18.4	18.9	27.1	70.4	41.1
Cropland	57.7	79.2	66.7	76.0	75.0
Wetland	25.0	31.5	59.8	52.2	67.0
Built-up	74.5	91.5	78.7	91.5	89.4
Bare/sparse vegetation	76.0	78.5	75.0	72.0	87.6
Water and snow/ice	80.0	80.0	70.0	78.0	86.7
Total	50.7	55.4	62.8	57.1	76.3

Probability maps for each LC class produced by means of RK are shown in Figure 4.10. While distinct high probability areas of forest, bare/sparse vegetation and water and snow/ice classes can be observed in Figure 4.10, the Sahel and savannah areas are represented by multiple classes such as shrubland, grassland and cropland, which had similar probabilities.

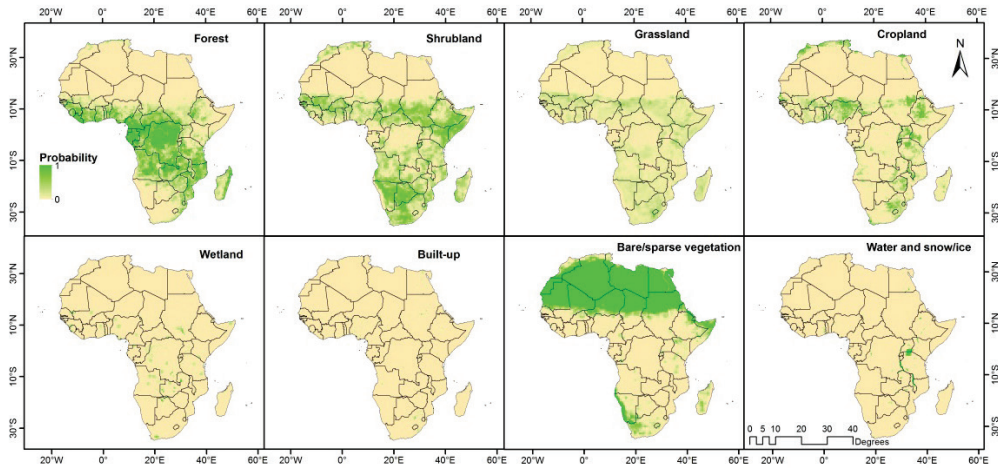


Figure 4.10. Probability maps of LC classes.

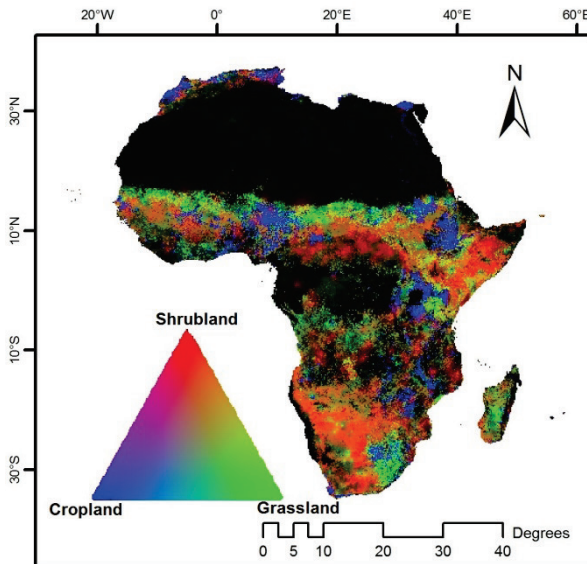


Figure 4.11. RGB image of class probabilities of shrubland, grassland and cropland. Dark shades represent areas where none of these three classes has a presence probability.

As the LC classes of the integrated map were selected based on the maximum presence probability, shrubland class superseded the grassland class by having a higher probability value in these regions and therefore more area of shrubland is

observed in Figure 4.9. This can be observed in Figure 4.11, which shows the combination of class probabilities of shrubland (red), grassland (green) and cropland (blue). Substantial areas in orange color highlight the combination of shrubland and grassland as probable classes with the presence probability of shrubland is higher than that of grassland (Figure 4.11). In contrast, the extent of areas with only shrubland as probable class (red) is considerably less. The combination of grassland and cropland as probable classes is shown in cyan color that can mostly be observed in the northern part of Sahel and eastern part of South Africa. Figure 4.11 illustrates the complexity of landscape with multiple probable class in the Sahel and savannah areas. For studies regarding these areas, consulting with the presence probability maps of the LC classes are recommended.

The presence probability maps of the LC classes are helpful in understanding the uncertainties in class assignment in the integrated maps as well as the complexity of heterogeneous landscapes.

4.4.4 On the use of available reference datasets for integration

This study made use of GLC reference datasets that were developed from different initiatives. The combined reference dataset has rather dense spatial distribution over a large portion of the African continent, which is beneficial for geostatistical interpolation. In the Sahara desert, sample density was lower. Nevertheless, correspondence with the reference LC was high in this region (Figure 4.7) as bare areas are usually mapped with high accuracy (Herold et al. 2008). One should be cautious when integrating different reference datasets as they may have discrepancies in their legends, sampling design and response design (*i.e.*, sample site area) (Tsendbazar et al. 2015a). To reduce the legend discrepancies of the reference datasets, we harmonized their legends into a common system with 8 general classes. However, there may be some inconsistencies in the reference datasets due to the discrepancies in the definition of LC classes.

Another issue is that reference datasets use different spatial supports. Our approach of using the centroids of reference sites provided a practical solution. However, differences in spatial support among reference data sets (and maps) are often a source of uncertainty about the true land cover. Block kriging and area-to-point (Goovaerts 2010; Kyriakidis and Yoo 2005) have been proposed for dealing with different spatial supports. Note that area-to-point kriging requires semivariograms at the fine spatial resolution, which may be difficult to acquire. Last, but not least, the integration approach of the reference datasets demonstrated in this study can be used for other studies that use geostatistical approaches. Since some reference datasets are not based on probability sampling, design-based statistical inference cannot be used. Moreover, design-based statistical inference using multiple

reference datasets with different statistical sampling designs requires known inclusion probabilities (Stehman 2014).

4.5 Conclusion

This study utilized the available GLC reference datasets from the GOF-C-GOLD, Geo-Wiki and the International Steering Community for Global Mapping. These datasets were originated from various institutions and the diversity of the reference datasets characters (e.g., legend and sample site area) makes them challenging to be integrated and reused for other studies. Our study provides an example of dealing with such diversities by harmonizing the thematic and spatial support differences of the reference datasets and using them for model-based geostatistical estimations. Further initiatives on generating better and more consolidated GLC maps can be useful to reduce discrepancies and uncertainty caused by legend harmonization. The advantages of including different reference datasets for integration were demonstrated in this study. Such information is useful as more reference datasets are becoming available to the public thanks to GLC mapping and validation communities (Fritz et al. 2011a; GOF-C-GOLD 2014).

Our study analyzed and compared the spatial variation in thematic correspondence of GLC maps, namely the Globcover 2009, LC-CCI 2010, MODIS 2010 and Globeland30, with the reference datasets. Based on the spatial autocorrelation structure of map correspondence, we modeled the spatial correspondence of the GLC maps as a measure of spatial accuracy. The comparison of the spatial correspondence maps demonstrated generally uncertain areas in LC mapping in Africa that need attention for improvement efforts while the preferences for GLC maps varied spatially. This finding demonstrates a motivation of integrating GLC maps based on their spatial variation in map correspondence in order to create an improved GLC map.

Aiming to create an improved LC map, we tested five different methods which are based on multiple GLC maps and reference datasets. The integration methods that employed both the GLC maps and the reference datasets resulted in 4.5%–13% higher correspondence with the reference LC classes than that of the input GLC maps. These methods exceeded the other two methods by making best use of the available data by calibrating the GLC maps with the help of reference datasets and relying on the GLC maps in places where the reference dataset is sparse. This result illustrates the benefit of using existing reference datasets and geostatistical approaches for map integration. In contrast, integration based on the agreement amongst the input maps without questioning their spatial correspondence did not

result in improved correspondence with reference LC. Nevertheless such approaches are commonly adopted for map integration efforts.

The RK method, which ensures to reduce the classification errors as much as possible through MNL regression and kriging of the regression residuals, showed the highest correspondence with reference LC. This method was selected to create an integrated LC map and the LC class probability maps of Africa. Uncertainty in class assignment was higher in heterogeneous areas with mixtures of different LC classes than in homogenous areas. In heterogeneous areas such as the Sahel and dry and moist savannahs, the LC probability maps can be useful. This study was done for the extent of Africa. With increasing computational power and more data coming available, the approach can be extended to global coverage and other datasets can also be included as covariates.

Acknowledgement

This study was supported by the ESA Land Cover CCI project. The authors would like to thank the ESA GOF-C-GOLD project office, Geo-Wiki platform and International Steering Committee for Global Mapping for providing the access to reference datasets on global land cover. We are grateful to the experts involved in the generation of the reference datasets for their valuable input. We are also thankful to the anonymous reviewers who contributed to enhancing this manuscript.

Chapter

5

Integrating global land cover datasets for deriving user specific maps

Nandin-Erdene Tsendbazar, Sytze de Bruin and Martin Herold

Submitted to Remote Sensing of Environment

Abstract

Global scale land cover (LC) mapping has interested many researchers for last two decades as it is an input data source for various applications. However, current global land cover (GLC) maps often do not meet the accuracy and thematic requirements of specific users. This study aimed to create an improved GLC map by integrating available GLC maps and reference datasets. We also address the thematic requirements of multiple users by demonstrating a concept of producing GLC maps with user-specific legends. We used a regression kriging method to integrate Globcover-2009, Land Cover-CCI-2010, MODIS-2010 and Globeland30 GLC maps along with several publicly available GLC reference datasets. Overall correspondence of the integrated GLC map with reference LC was 80% based on 10-fold cross validation using 24681 sample sites. This is globally 10% and regionally 6-13% higher than the correspondence of the input GLC maps. Furthermore, based on LC class probability maps expected area fraction maps for LC classes at coarser resolution were created and used for characterizing mosaic classes for land system modelling and biodiversity assessments. As more reference datasets are becoming available to the public, GLC mapping can be further improved by using the pool of all available reference datasets. Class probability and area fraction maps allow tuning LC products to specific user needs. Future GLC mapping efforts should take this into consideration.

Keywords: global land cover; data integration; user-specific legend, land cover fraction

5.1 Introduction

Land cover (LC) is an essential environmental variable for understanding interactions between mankind and nature. Global scale LC mapping has therefore interested many researchers for last two decades. After the first attempts to produce a global land cover (GLC) map at 1 degree resolution using remote sensing (De Fries et al. 1998), multiple medium resolution (300-1000m) GLC maps such as the GLC2000, MODIS, Globcover and LC-CCI have been produced (Arino et al. 2007; Bartholomé and Belward 2005; CCI-LC 2014; Friedl et al. 2010). Recently, the release of Landsat archive further led to the production of high resolution (30m) GLC maps namely the FROM-GLC and Globeland30 (Chen et al. 2015; Gong et al. 2013). Despite all advances, the overall accuracy of existing GLC maps typically varies between 67% and 81% (Mora et al. 2014).

GLC maps are used as an input to various applications such as climate modelling, hydrological modelling, agricultural monitoring and biodiversity monitoring (Costa et al. 2003; Fritz et al. 2015; McCarthy et al. 2012; Verburg et al. 2011). It is challenging for the users to select the GLC map that is the most suitable for their applications because of diverse requirements and discrepancies among GLC maps in terms of used input data, methodologies and validation (Herold et al. 2008). Errors in GLC maps add to the uncertainties in the output of user applications (Nakaegawa 2011); therefore, improved accuracy of GLC map is required by the users (Bontemps et al. 2011b). This is particularly crucial in some regions of the world where GLC map accuracies range between a mere 10-50% (Frey and Smith 2007) and in areas with mixed LC classes (e.g. mosaics of trees, shrubs, and herbaceous vegetation) that are typically mapped with low accuracy (Herold et al. 2008; Tsendbazar et al. 2016).

To create improved GLC maps, existing GLC maps are often integrated based on properties such as map or class specific accuracy and agreement with other GLC maps (Iwao et al. 2011; Jung et al. 2006; Kinoshita et al. 2014). Reference datasets that were produced for calibrating and validating GLC maps can also be used for integration since some reference datasets are publicly accessible through the GOF-C-GOLD Reference Data portal, Geo-Wiki platform and International Steering Committee for Global Mapping (Fritz et al. 2009; GOF-C-GOLD 2014; Tateishi et al. 2011). See et al. (2015) used the Geo-Wiki crowd-sourcing reference data to model the spatial correspondence of GLC maps and used this for map integration. Tsendbazar et al. (2015b) compared multiple integration methods that consider the spatial variation in map correspondence and LC class probabilities based on recent GLC maps and several publicly accessible reference datasets for the African continent.

Depending on the purpose of application, users of GLC maps may require different thematic classes. For example, some applications concern only a single LC class such as forest, cropland or built-up area (Feldpausch et al. 2012; Fritz et al. 2015; Klein Goldewijk et al. 2010), while multiple LC classes are needed for carbon density mapping based on per-class biomass values (Ruesch and Gibbs 2008). Other applications may specifically focus on a subset of LC classes. For instance, land system modelling needs information on LC classes indicating the degree of human interactions such as cropland, urban and mosaics of these classes with other LC classes such as natural vegetation (Asselen and Verburg 2012; Letourneau et al. 2012).

Except for the MODIS GLC map, which has 5 different legends, current GLC maps employ a single fixed legend. These fixed legends do not always meet the requirements of users. Mosaics of urban with other LC which characterise many rural landscapes, for example, are not included in current GLC map legends, while such classes can be useful for understanding human interactions with the environment (Asselen and Verburg 2012). Moreover, translating fixed legends — particularly mosaic classes— to required legends can be ambiguous or even impossible (Friedl et al. 2002; Herold et al. 2008). For example, in dynamic global vegetation models (DGVM), LC classes are translated to fractional plant function types (PFT). This translation assumes certain fractions of PFTs in different LC areas based on the definition of LC classes. Poulter et al. (2015) assumed that a mosaic class composed of cropland and natural vegetation contained four different PFTs (60% managed grass, 15% natural grass, 15% shrubs and 10% trees). These conversions were derived from expert knowledge and the assumed fractions can be subject to error.

GLC maps could address thematic requirements of several user applications if the area fraction of LC classes is provided. With area fraction maps of LC classes, users can characterize LC classes (e.g. mosaic classes) according to their requirement. This has been demonstrated by Asselen and Verburg (2012) who created a global scale land system map characterizing different levels of human interaction with the environment. In their study, specific mosaic LC classes representing different land systems were created based on fraction information of main LC types such as tree cover, cropland, urban and bare areas. In addition to improving map accuracy, GLC map integration should provide area fraction maps of LC classes in order to meet thematic requirements of different user applications.

In this study, we created an integrated GLC map and LC class probability maps and demonstrated the proof of concept of producing GLC maps with user-specific legends based on expected area fraction maps of LC classes. We used a regression

kriging method as described in Tsendbazar et al. (2015b) which was applied at global scale to create an integrated GLC map. Based on LC class probability maps produced from this integration, we created expected area fraction maps for LC classes at coarser resolution as many applications use GLC maps at coarse resolution (Herold et al. 2011). The LC expected area fraction maps were then used to create user-specific GLC maps from the perspectives of land system models and biodiversity assessment.

5.2 Data and methods

A flowchart of methodological steps adopted in this study to generate integrated GLC maps user-specific legends is provided in Figure 5.1. The following subsections describe the used GLC maps and reference datasets as well as methods for integrating GLC datasets and creating user-specific legends.

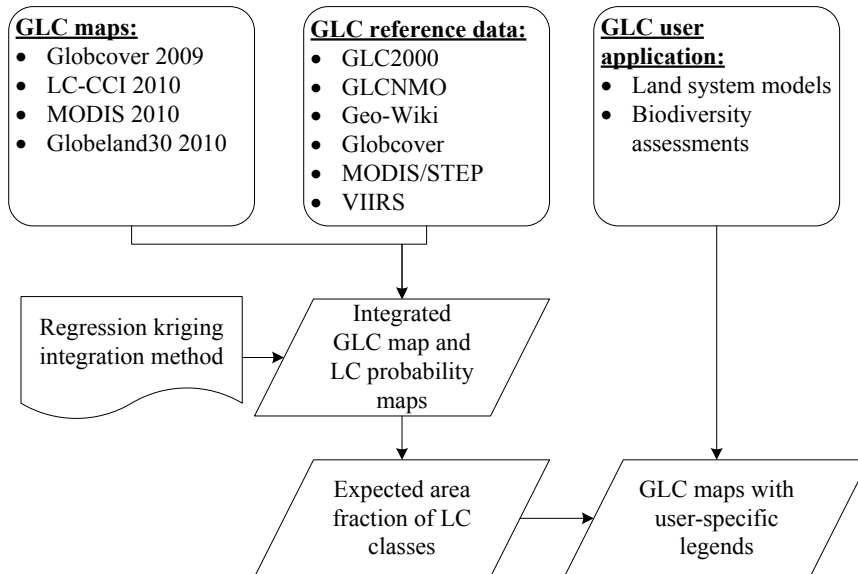


Figure 5.1. Flowchart of creating GLC maps with user-specific legends

5.2.1 Global land cover maps

The following recent GLC maps were used as inputs for our analysis: Globcover (2009), LC-CCI (2010), MODIS (2010) and Globeland30 (2010). The Globcover project of the European Space Agency (ESA) produced a GLC map for 2009 using MERIS satellite data at 300m resolution (Defourny et al. 2011b). This map characterizes 22 LC classes based on an LCCS (United Nations Land Cover Classification System). This map was validated using an independent reference

dataset and the thematic overall accuracy was 70.7% based on 1484 homogenous sample sites (Bontemps et al. 2011a). The Land Cover-CCI (LC-CCI) project, funded by the ESA, produced GLC maps for the epochs of 2000, 2005 and 2010 at 300m resolution using the MERIS data archive (CCI-LC 2014). These maps were specifically produced to meet the requirements of climate modellers. The maps have 22 LC classes according to the LCCS. The LC-CCI 2010 map was validated using the same reference sample as the Globcover-2009 validation and the overall thematic accuracy of the LC-CCI 2010 map was 74.4%. Boston University provided annual GLC maps (MCD12Q1) at 500m resolution based on Moderate Resolution Imaging Spectroradiometer (MODIS) data (Friedl et al. 2010). The MODIS GLC maps have five different legends including a 17 class legend based on the International Geosphere-Biosphere Programme (IGBP) classification scheme. We used the GLC map of 2010 with IGBP legend. The accuracy of this map has not yet been assessed. Based on cross-validation using the training dataset, an overall accuracy of the 2005 map was 74.8% (Friedl et al. 2010). The GlobeLand30 project of the Ministry of Science and Technology of China generated GLC maps for the year 2000 and 2010. These maps were developed using 30m resolution multispectral images of Landsat TM and ETM+ as well as the Chinese Environmental Disaster Alleviation Satellite (HJ-1). The maps are accessible at 30 m and 250 m resolution. We used the 250 m resolution version of Globeland30, which contains LC class fraction information. This map has 10 LC classes. The overall accuracy of the map was reported to be 83.5% (Chen et al. 2015).

The MODIS and Globeland30 maps were resampled to 0.00278 degrees (300m at the Equator) resolution using nearest neighbor assignment to match the resolution of the other two maps. The map legends were harmonized into nine general LC classes based on the approach of Herold et al. (2008), which provides a table harmonizing the input classes into 13 general LC classes using LCCS-based legend translation protocols. Since the Globeland30 map does not have detailed forest classes, the five forest classes of Herold et al. (2008) were merged to one (Table 5.1). The Globeland30 map has a “tundra” class that is a combination of shrubland and grassland. Since this class does not match with the legends of the other input maps, it was excluded from the analysis. Figure 5.2 presents the four GLC maps with the harmonized legend.

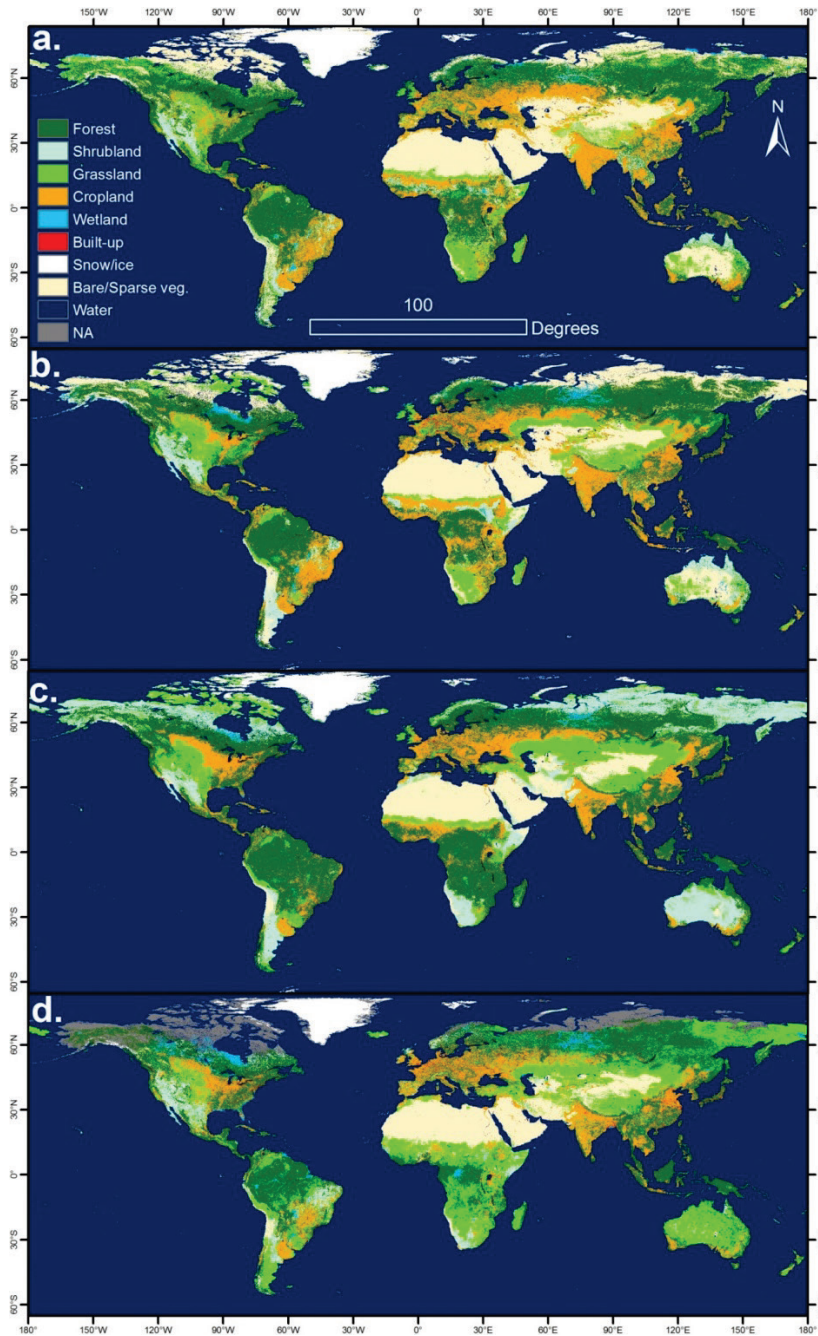


Figure 5.2. Input global land cover maps

5.2.2 Reference datasets

We used existing GLC reference datasets that are publicly accessible through the GOF-C-GOLD reference data portal, the Geo-Wiki platform and the International Steering Committee for Global Mapping (Fritz et al. 2009; GOF-C-GOLD 2014; Tateishi et al. 2011). The reference datasets used in this work are denoted as: GLC2000rd, GLCNMOrd, Geo-Wikird, MODIS/STEPrd, VIIRSrd and the Globcover-2005rd. The subscript “rd” is added here to avoid potential confusion between GLC maps and reference datasets. GLC2000rd concerns the reference dataset generated to validate the Global Land Cover 2000 map (Mayaux et al. 2006). The original version of this reference dataset with 22 LC classes was re-analyzed by experts to have a consolidated reference dataset with generalized 11 LC classes (Schultz et al. 2015). GLCNMOrd was created by Global Land Cover by National Mapping Organizations for calibration of a GLC map for 2003 (Tateishi et al. 2011). This reference dataset has 14 LC classes, which were assigned by international experts. MODIS/STEPrd was developed and updated by Boston University for calibration of the MODIS collection 4 and 5 GLC mapping (Friedl et al. 2010). It has 17 LC classes with the IGBP legend. Boston university also created VIIRSrd (Visible Infrared Imaging Radiometer Suite) for validating the VIIRS surface type products (GOF-C-GOLD 2014; Olofsson et al. 2012). This dataset has the same classes as MODIS/STEPrd and the reference LC labels were assigned by visual interpretation of very high resolution images. Geo-Wikird was developed through a volunteer based online platform. Volunteer interpretations of the reference LC were validated by a group of experts (Fritz et al. 2009). Globcover 2005rd was created to validate the Globcover 2005 GLC map (Bicheron et al. 2008). In the current study a re-interpreted version of the reference dataset by Tsendbazar et al. (2016) was used. Detailed information on the characteristics of the available reference datasets are provided in Tsendbazar et al. (2015a).

Since the used reference datasets have different sample site areas (Tsendbazar et al. 2015a), we assumed that the LC of the sample site refers to the LC of the centroid of that sample site (Tsendbazar et al. 2015b). The reference datasets were combined and the legends of available reference datasets were harmonized into the same nine general classes used in the harmonized GLC maps (Table 5.1).

The total number of reference sample sites used in this study was 24681 which comprises 2408 sites in Australia and Oceania, 10379 sites in Eurasia, 4715 sites in North America, 3298 sites in South America, and 3881 sites in Africa. Figure 5.3 shows the sample distribution of each reference dataset (a) and the reference LC classes (b).

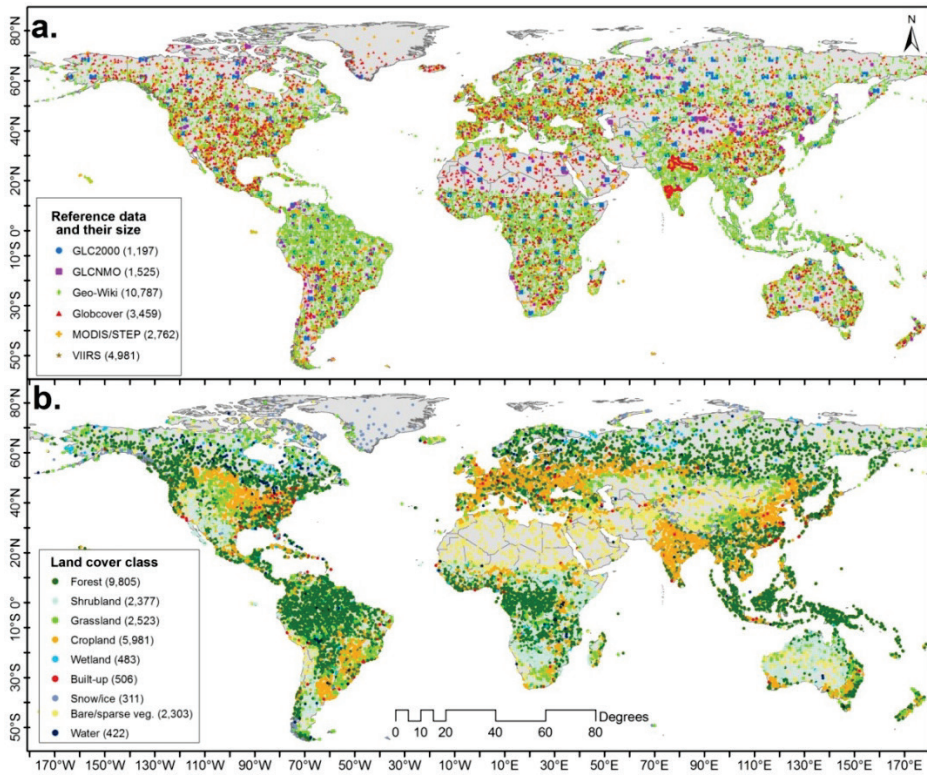


Figure 5.3. Spatial distribution of the reference datasets (a) and reference land cover classes (b)

Table 5.1. The generalised LC classes and their definitions

Code	Land cover class	Globcover	LC-CCI	Globeland 30	IGBP (MODIS, STEP& VIIRS)	GLC 2000	Geo-Wiki	GLCN MO	Definition
1	Forest	40-110, 170	160, 50-100, 160, 170	20	1-5, 8, 9	1-10	1	1-5	Lands where forests cover is more than 10% of the total land cover.
2	Shrubland	130	120	40	6,7	11, 12	2	7	Lands where shrubs cover is more than 10% of the total land cover.
3	Grassland	120, 140	110, 130, 140	30	10	13	3	8,9	Lands covered by natural grass with at least 10% of the total land cover
4	Cropland (incl. mixtures)	11-30	10-40	10	12, 14	16-18	4	11, 12,13	Lands where crops comprise at least 40% of the total land cover.
5	Wetland vegetation	180	180	50	11	15	6	15	Lands with a semi-permanent mosaic of water and herbaceous or woody vegetation.
6	Urban/built up	190	190	80	13	22	7	-	Land covered by buildings and other man-made structures and activities.
7	Snow/Ice	220	220	100	15	21	8,	-	Lands covered by permanent snow, glacier and icecap.
8	Bare/sparse vegetation	150, 200	150, 200	90	16	14, 19	9	10, 16, 17	Lands with vegetation cover lower than 10%.
9	Water	210	210	60	17	20	10	-	Water bodies in the land area, including river, lake, reservoir, pond, etc.

5.2.3 Integration of GLC datasets

We integrated the available GLC maps and reference datasets using the approach of Tsendbazar et al. (2015b). We used a regression kriging method that predicts local LC class presence probabilities since it produced the highest correspondence with reference data in comparison to voting, spatial correspondence, weighted voting and indicator kriging methods (Tsendbazar et al. 2015b). LC classes at the reference sample locations were extracted and used as calibration dataset in the regression analysis. The data were also used for estimating semivariograms of the regression residuals.

The presence probability of 9 LC classes were processed separately for five regions of the world (Figure 5.4) which were later merged to a global scale at 0.00278

degrees resolution. These regions were used to facilitate processing while avoiding boundary effects as the regions coincide with land and ocean borders.

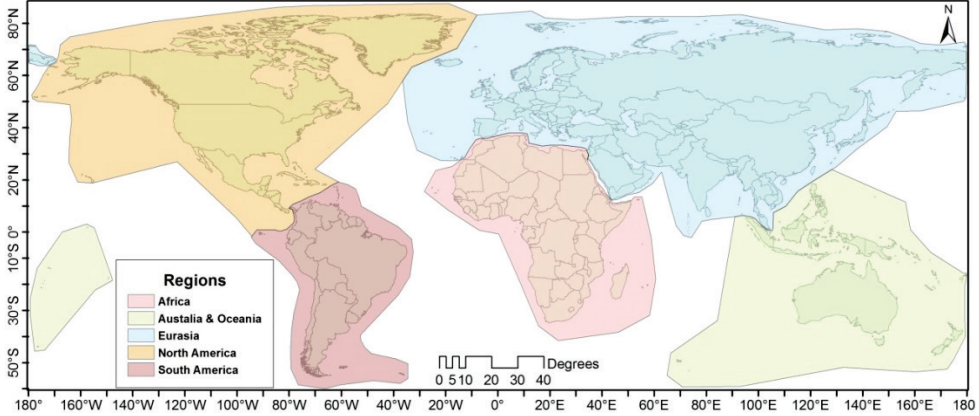


Figure 5.4. Boundary of the world regions used in the analysis

For each region, the general trend of presence probabilities of LC classes were predicted using a multinomial logistic (MNL) regression (Kempen et al. 2009). Next, the predicted probabilities were locally adjusted by interpolating regression residuals using simple kriging (Equation 5.1).

$$p_k(x) = \pi_k(x) + \varepsilon_k(x) \quad (5.1)$$

Here, $p_k(x)$ denotes the presence probability of LC class k at location x , $\pi_k(x)$ is a predicted probability trend of LC class k that was obtained by MNL regression and $\varepsilon_k(x)$ is the regression residual for that class. The latter has zero mean and is spatially correlated as modelled by a semivariogram; it is spatially interpolated using nearby reference data. A detailed description of this method can be found in Tsendbazar et al. (2015b). A semivariogram model was fitted for each class in each region following the method described in Tsendbazar et al. (2015b). For some LC classes, residual kriging was not used since no spatial correlation was observed in the regression residuals. The method was implemented using the `nnet` and `gstat` packages in R (Pebesma 2004; Ripley et al. 2014) and ran on the Amazon EC2 cloud computing facility (Amazon 2010).

After adjusting the predicted probabilities with residual kriging, any probability outside the interval $[0, 1]$ was set to the closest bound, zero or one. Subsequently, the estimates $p_k(x)$ $k=1, \dots, K$ were normalized by their sum to meet the condition $\sum_{k=1}^K p_k(x) = 1$ (de Bruin 2000a). In the integrated GLC map, pixels are assigned to the LC class having the highest probability .

The quality of the integrated GLC map was assessed by 10 fold cross-validation based on the reference dataset. For each region, the reference sample sites were partitioned randomly into 10 subsamples. Nine subsamples were used to train the integration method and one subsample was used to assess the integrated GLC map by calculating the overall correspondence between the reference LC and the integrated LC. This step was repeated 10 times so that each subsample was used for method training as well as validation while each sample site was only used once for validation. The median percentage of integrated LC classes locally corresponding with reference subsamples was calculated for each region. Note that these values should not be confused with the overall accuracy of LC maps since they are based on cross-validation using the calibration sample rather than comparison against an independent reference dataset obtained by probability sampling (Tsendbazar et al. 2015b).

5.2.4 GLC maps with user-specific legends

We proposed a concept of creating GLC maps with user-specific legends using LC class area fraction information. With this, GLC map users can create a flexible legend with class definitions (e.g. fraction thresholds) that fits their requirements. This concept was demonstrated from the perspectives of land system modelling and biodiversity assessments based on improved LC probability maps.

Land system models present land use systems describing both the heterogeneity of land cover within the landscape and land-use intensity highlighting human-environment interactions (Letourneau et al. 2012). Land use systems can include densely populated systems, cropland systems, grassland systems, forested systems and bare soils systems as well as mosaic systems which are combination of the main systems (Letourneau et al. 2012). For land system models, therefore, mosaic LC classes highlighting human influences to the environment can be useful in addition to the general LC classes. To create a GLC map from a perspective of land system models, we adopted several mosaic classes that were introduced in Asselen and Verburg (2012) and added these classes to the harmonized legend with nine general classes. The mosaic classes include peri-urban & villages (built-up:5-25%), mosaic cropland (20-50%) and grassland (20-50%), mosaic cropland (20-50%) and forest (15-40%), mosaic grassland (20-50%) and forest (15-40%), and mosaic grassland(20-50%) and bare (25-50%).

We used LC presence probability maps obtained by regression kriging at 0.00278 degrees resolution to create LC area fraction maps at 0.0834 degrees (5 arcminute) resolution that was used in land system model of Asselen and Verburg (2012). We selected the class probability maps instead of the integrated GLC map to avoid overestimation of frequent occurring classes (e.g. forests and cropland) and

underestimation of rare classes (e.g. urban and wetland). The expected fraction of LC classes in 5 arcminute resolution pixels ($E[A_k]$) was computed as follows (de Bruin 2000a):

$$E[A_k] = E[p_k] = \sum_{i=1}^n p_k(x_j) / n \quad (5.2)$$

where $p_k(x_j)$ is the probability of a class k for given the data at pixel x_j , n is the number of pixels at the original resolution within the 5 arc minute pixel. Equation 5.2 expresses that average class probabilities of pixels over a region (e.g. coarse resolution pixel) correspond to the expected class proportions in that region. Using the area fraction maps with 5 arcminute resolution, the mosaic classes were identified and the remaining general LC classes were assigned to the LC class having the maximum fraction.

We used a similar approach for creating a GLC map from a biodiversity perspective. Land use types that represents human impact such as cropland and urban have negative impact on species richness and species abundance (Newbold et al. 2015). Hence, characterization of LC classes such as cropland and built-up areas can be important for biodiversity assessments. Mosaic of these classes with classes that have less human influence such as natural vegetation could also be of interest. Therefore, we added four mosaic classes that could be of interest to biodiversity analysts on the general legend with nine LC classes. The mosaic classes include mosaic built-up(5-25 %) and cropland (20-70%), mosaic built-up (5-25 %) and nature vegetation (20-70%), mosaic cropland (20-70%) and grassland (20-70%), and mosaic cropland (20-70%) and forest/shrubland (20-60%). Note that the applied class fraction thresholds can be adjusted depending on user preferences. Using the LC area fraction maps at 5 arcminute resolution, the mosaic classes were first assigned while the remaining nine main LC classes were again assigned to the LC class having the maximum fraction.

5.3 Results

5.3.1 Integrated GLC map and LC probability maps

Each region was modeled by a distinct MNL regression model and a set of residual semivariograms. Figure 5.5 shows the nugget-to-sill ratios of the semivariogram models fitted to the regression residuals. The larger this ratio, the less the regression residuals are spatially correlated.

In general, the regression residuals tended to be spatially correlated in Africa and Australia and Oceania where nugget-to-sill ratios were smaller compared to other regions. In Australia and Oceania, classes such as wetland, built-up, water and

snow/ice that occupy small parts showed high spatial correlation (nugget-to-sill ratio=0) in the residuals. In contrast, in South America, the presence probability of more common classes such as forests, shrubland, grassland and cropland were entirely based on the MNL regression as the regression residuals were spatially uncorrelated (nugget-to-sill ratio =1) (Figure 5.5).

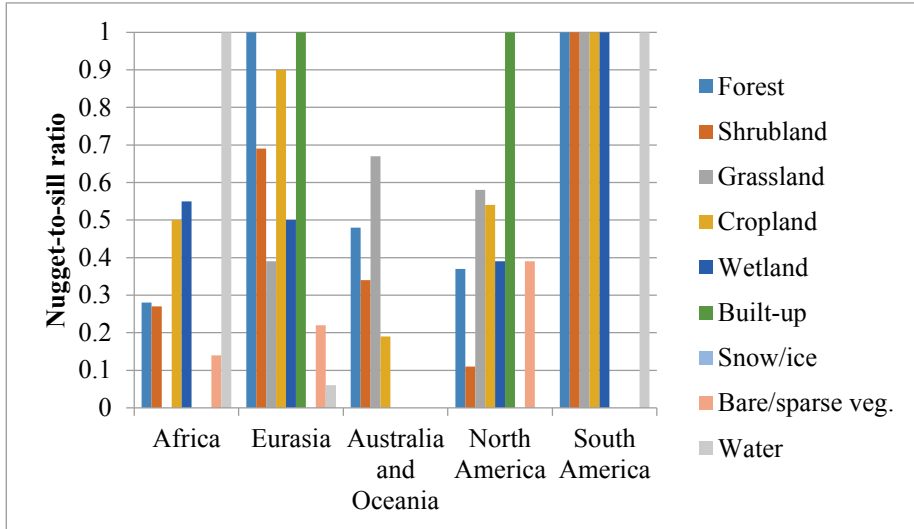


Figure 5.5. Nugget-to-sill ratio of the regression residuals

The integrated LC maps of the regions were merged to create the GLC map shown in Figure 5.6.

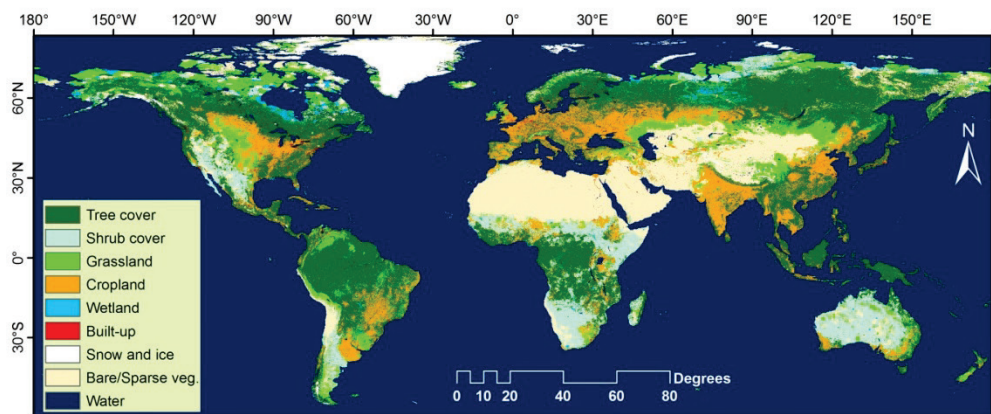


Figure 5.6. The integrated GLC map

The major patterns of homogeneous regions in the integrated GLC map are similar to those in the input GLC maps. These correspond to regions such as tropical and temperate forests, deserts across Asia and Africa and large cultivated areas in the different continents. The main differences between the integrated GLC map and the input maps are in areas where the input maps disagree amongst each other. In the integrated GLC map, these regions are characterized by combinations of LC classes that are present in one of the input GLC maps. For example, the western plateau and interior lowlands of Australia and tundra regions of Russia were mapped as shrubland, grassland and bare/sparse vegetation classes in the integrated GLC map (Figure 5.6), while these areas are mostly mapped as bare/sparse vegetation and grassland in the LC-CCI and Globcover maps, shrublands in the MODIS map, and grassland in the Globeland30 map (Figure 5.2). Compared to the input maps, the integrated GLC map presents more shrubland areas in African savannah regions and more grassland areas in Northern Canada. This effect is caused by the reference datasets.

LC class probability maps that were created from the integration of GLC datasets are presented in Figure 5.7. Forest, cropland, snow and ice, bare/sparse vegetation and water areas are spatially structured in zones with high probability and low probability. In contrast, transition areas of major biomes such as the Sahel and savannah region, tundra regions in Eurasia and North America and peripheries of Amazonian forests are characterized by approximately equal probabilities for multiple LC classes such as shrubland, grassland and bare/sparse vegetation.

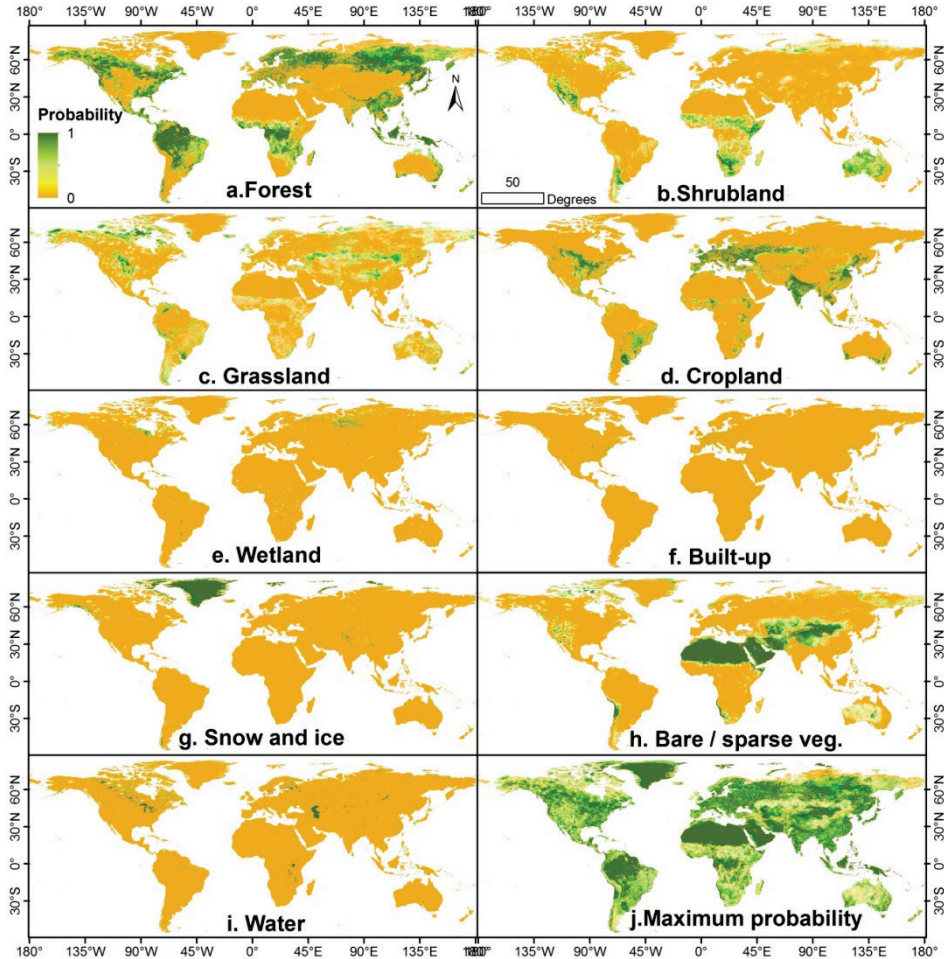


Figure 5.7. Probability maps of LC classes

Figure 5.7j shows the maximum probability values for different LC classes. Since the integrated GLC maps was based on a LC class that had the highest probability per pixel, this maximum probability map reveals information on the internal uncertainty of the integrated GLC map. The maximum probability value is lowest in areas such as transition areas of major biomes, tundra regions and central Australia, where multiple classes are possible based on the input GLC maps and reference data.

5.3.2 Cross validation of the integrated GLC map

Figure 5.8 presents the correspondence of the integrated GLC map with the reference dataset at region level. The correspondence of the integrated GLC map was highest in Eurasia (81%) followed by South America, North America, and Australia and Oceania (79%) and Africa (76%). When compared with the correspondence of the input GLC maps, these results are at least 12% higher in Australia and Oceania and Africa, 9 % higher in Eurasia and North America and 6% higher in South America. A plausible reason of the highest improvements in Australia and Oceania and Africa is that there was more room for improvements as the input maps had low correspondence with the reference data.

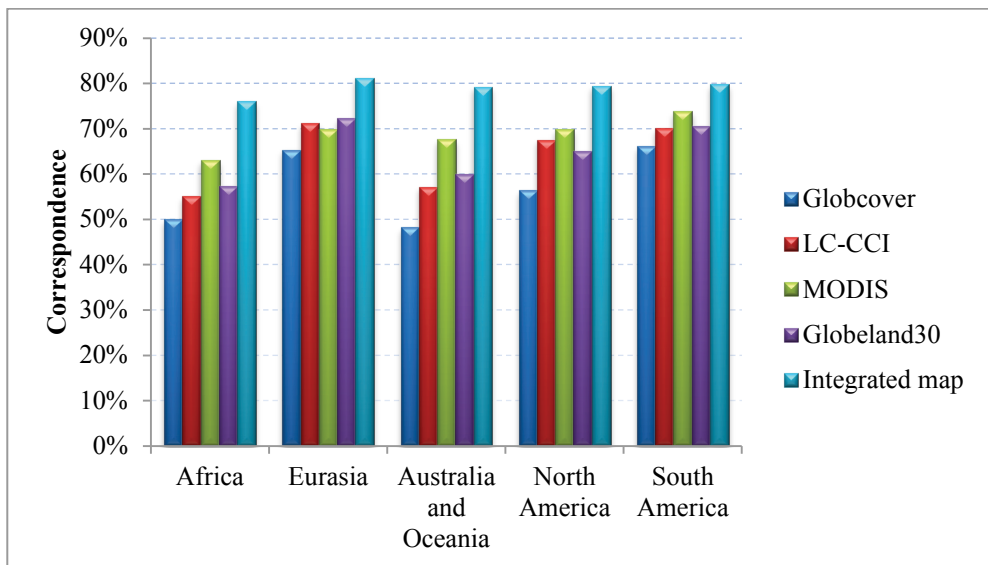


Figure 5.8. Correspondence of integrated GLC map with reference dataset.

Class-specific correspondence of the integrated GLC map with the reference dataset is provided in Table 5.2. The class-specific correspondence of the integrated GLC map was derived by cross-validation. Compared with the input maps, the integrated GLC map has a relatively high correspondence for the forest class in all regions except Africa. The correspondence of the shrubland class also improved except in South America. Mapping of grassland and cropland improved in regions other than Africa and South America. For rare classes such as wetland, built-up and water classes, the integrated map has the highest or one of the highest correspondences across all regions. Improved mapping of bare/sparse vegetation was achieved in all regions except Australia and Oceania. In this region, the Globcover and LC-CCI products had higher correspondence.

Table 5.2. Class-specific correspondence of the integrated GLC map compared with the input maps

	Globcover	LC-CCI	MODIS	Globeland30	Integrated
Africa					
Forest	71.1	67.3	90.2	63.7	84.9
Shrubland	11.9	21.3	26.9	17.3	70.8
Grassland	18.4	18.9	27.1	70.4	41.1
Cropland	57.7	79.2	66.7	76.0	75.0
Wetland	25.0	31.5	59.8	52.2	67.0
Built-up	74.5	91.5	78.7	91.5	89.4
Snow and ice	-	-	-	-	-
Bare/sparse vegetation	76.0	78.5	75.0	72.0	87.6
Water	80.0	80.0	70.0	78.0	86.7
Eurasia					
Forest	69.4	78.4	77.5	79.0	87.0
Shrubland	10.0	7.4	37.6	10.8	30.9
Grassland	9.3	31.5	45.9	47.1	59.2
Cropland	84.5	86.1	83.4	87.9	92.7
Wetland	19.0	36.8	31.3	41.1	41.1
Built-up	52.4	81.9	72.2	82.7	79.4
Snow and ice	33.1	33.1	34.6	27.6	44.9
Bare/sparse vegetation	77.4	70.0	51.2	58.6	81.3
Water	75.3	80.5	53.2	73.4	74.0
Australia and Oceania					
Forest	57.7	69.4	88.6	83.9	91.9
Shrubland	20.5	39.5	65.8	8.2	74.7
Grassland	14.2	21.1	22.2	53.1	57.5
Cropland	80.9	75.3	70.3	80.2	82.7
Wetland	0.0	4.0	44.0	36.0	44.0
Built-up	53.3	66.7	86.7	80.0	86.7
Snow and ice	37.5	37.5	25.0	100.0	62.5
Bare/sparse vegetation	73.9	67.4	8.7	6.5	23.9
Water	84.6	92.3	46.2	61.5	84.6
North America					
Forest	83.1	79.1	81.6	74.6	90.6
Shrubland	30.7	43.7	53.2	45.4	54.9
Grassland	20.9	46.8	45.1	26.5	52.8
Cropland	39.0	71.7	85.6	85.8	89.1
Wetland	3.6	25.7	37.9	35.7	45.7
Built-up	24.2	82.8	73.2	86.0	86.0
Snow and ice	60.1	63.5	61.5	52.7	70.9
Bare/sparse vegetation	41.6	36.8	9.5	6.5	50.6
Water	75.6	81.4	59.6	89.7	81.4
South America					
Forest	80.1	81.0	93.4	76.2	93.6
Shrubland	35.7	51.1	44.5	47.1	38.3
Grassland	6.9	15.2	31.4	56.6	44.1
Cropland	78.9	84.6	57.5	78.1	79.7
Wetland	35.9	29.7	65.6	46.9	54.7
Built-up	35.1	78.4	64.9	73.0	83.8
Snow and ice	59.3	66.7	40.7	51.9	66.7
Bare/sparse vegetation	48.9	48.9	34.1	34.1	55.7
Water	80.8	86.5	51.9	84.6	86.5

The overall correspondence of the integrated GLC map was 79.6%, which is at least 10% higher than the input GLC maps whose correspondence ranged between 59.6 and 69%.

5.3.3 GLC maps with user-specific legends

GLC maps with two user-specific legends are shown in Figure 5.9. Figure 5.9a demonstrates a GLC map with specific mosaic legends for land system models. Mosaic grassland and forest areas were mostly mapped in the tundra regions of Eurasia and North America while the mosaic grassland and bare class occur mostly in the northern tundra regions along the borders of deserts of Eurasia and Australia. Small patches of mosaic cropland and grassland areas can be observed along cultivated areas on the British Isles, in northeastern China and in eastern Brazil. Mosaic cropland and forest areas are identified along the borders of Amazonian forest and Savannah regions in Africa. Due to the coarse resolution of this map, large metropolitan areas in Japan, the USA, Benelux, western Germany and the UK are mapped as built-up areas. More dispersed urban areas in central and eastern Europe and eastern China are mapped as peri-urban and villages in this map.

Figure 5.9b presents a GLC map with specific mosaic classes for biodiversity assessments. Similar to Figure 5.9a, mosaic cropland and grassland class are mapped along the borders of main cultivated regions. However, owing to different area fraction thresholds, larger areas were assigned to this composite class. Substantial areas of mosaic cropland and forest/shrubland areas are mapped in the eastern USA, Brazil, central and eastern Europe, southwestern China and Savannah regions in Africa. Areas designated as peri-urban and villages in Figure 5.9a, are mapped mainly as mosaic of built-up and cropland areas in Figure 5.9b while mosaics of built-up and natural vegetation areas occur in small patches in western and eastern USA and east coast of Australia.

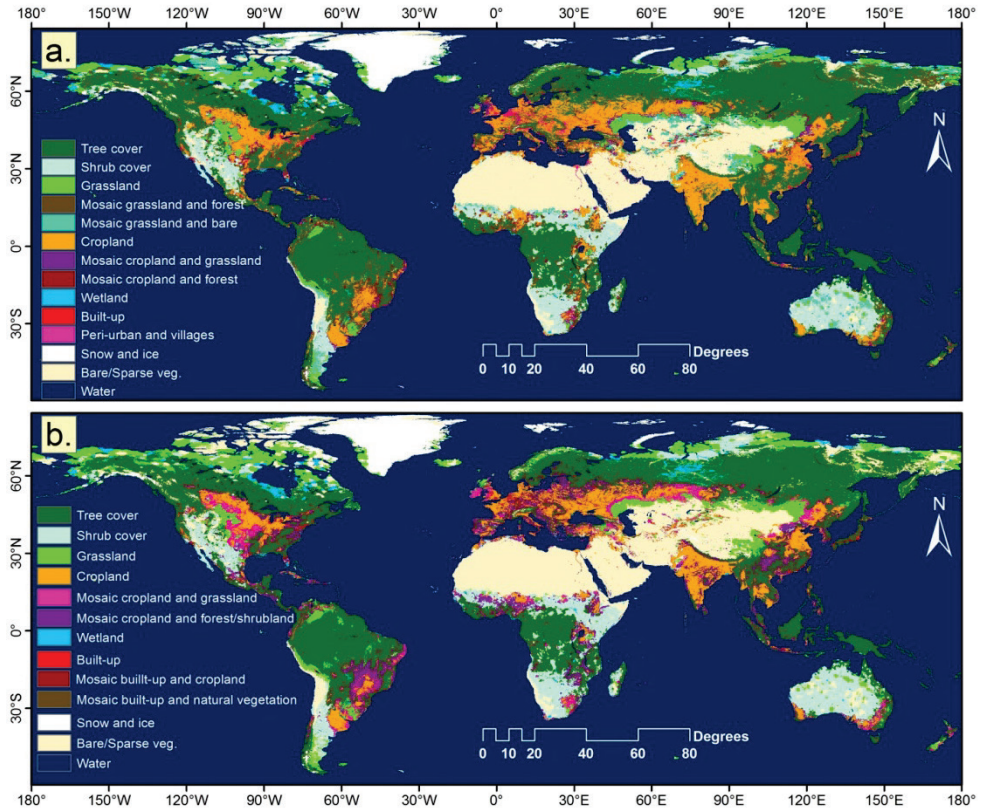


Figure 5.9. GLC maps with specific mosaic classes for land system models (a) and biodiversity assessments (b)

5.4 Discussion

5.4.1 Integration of GLC datasets

We created an integrated GLC map and LC class probability maps by predicting the probability of LC classes using MNL regression and modelling the spatial variation in the regression residuals. The regression residuals were spatially correlated to at least some extent for most LC classes (Figure 5.5). Particularly, correlation of the regression residuals were in general higher in Africa and Australia and Oceania compared to the other regions. Owing to this correlation, the correspondence of the integrated GLC map with the reference LC was at least 12% higher than that of the input maps in these regions (Figure 5.8). This shows the benefit of the use of geostatistical approach based on the available GLC maps as well as the reference datasets in case of spatial correlation in regression residuals (Tsendbazar et al. 2015b). On the other hand, the correspondence increase was the

least in South America (6%) where the regression residuals of the major LC classes were not spatially correlated.

Previous GLC map integration efforts achieved 3-4.5% accuracy improvements (Kinoshita et al. 2014; See et al. 2015). Based on the 10-fold cross-validation, the integrated GLC map had an overall 10% increase in the correspondence with respect to the input maps and regionally the improvement was up to 13%. Even though these figures are not directly comparable to accuracies provided in previous studies (Kinoshita et al. 2014; See et al. 2015), they are convincingly based on a large sample (24681 sites).

The maximum probability values that highlight the uncertainty of the integrated GLC map were low in areas such as transition areas of major biomes, tundra regions and central Australia, where the input maps disagreed (Figure 5.7j). These regions are known for being inconsistent amongst GLC maps (Fritz et al. 2011b; Herold et al. 2008) and having low spatial accuracies (Tsendbazar et al. 2015b). Such disagreement can be attributed to the complexity of heterogeneous landscapes (Herold et al. 2008). In the integrated map, these regions are mostly characterized by relatively small LC patches that emerge due to the influence of residual kriging using nearby reference data. This is another example demonstrating the role of reference datasets in map integration.

For integration, this study used publicly available reference datasets generated from various initiatives. As more reference datasets are becoming available to the public, GLC mapping can be further improved by using the pool of all available reference datasets. However, reference datasets typically differ in their design protocols (e.g. sample design, legend, sample site area) which has to be considered when using them for integration, as discussed by Tsendbazar et al. (2015b).

5.4.2 GLC maps with user-specific legends

GLC maps with user-specific legends (Figure 5.9) were created based on the integrated LC probability maps. The improved correspondence with reference LC (Figure 5.8) addresses the user requirement of reducing uncertainty about LC labelling. Furthermore, the maps in Figure 5.9 demonstrate the concept of producing GLC maps with user-specific legends based on area fraction estimates of LC classes. This approach makes it more flexible to address the different thematic requirements of multiple users. The new mosaic classes in Figure 5.9 occur mostly in heterogeneous landscapes where multiple LC classes have presence probability (Figure 5.7). This shows the importance of user-specific legends particularly in these regions.

Providing area fraction information of the LC classes can be useful not only to create user-specific mosaic LC classes, it can also be used for generalizing LC classes by summing class fractions for all composing classes. For example, the fraction for a class “non-vegetation” can be computed by summing the fractions of built-up, snow and ice and water. Moreover, users would be allowed to use their own area fraction thresholds for mapping mosaic classes. This can have implications for the area occupied by mosaic class as can be observed from Figure 5.9, where both GLC maps had the mosaic of cropland and grassland class while this class of Figure 5.9b was substantially larger than that of Figure 5.9a due to the threshold difference. Users can also adjust LC class fraction thresholds for mosaic or non-mosaic classes depending on the cost of GLC map misclassification for specific applications.

The expected fractions of LC classes at a coarser resolution were derived using the class probability maps following the method of de Bruin (2000a). Since the MNL regression used indicator data for class presence (zero or one), this integration method requires mutually exclusive LC classes. Therefore, class probability maps can be used to derive class fractions for coarser resolutions but not at the the pixel’s ground resolution (300 m resolution). The conversion method for the expected fractions of LC classes requires the LC regions (e.g., coarser resolution pixels) to be considerably larger than the pixels’ ground resolution (de Bruin 2000a). Otherwise, the statistical concept of expected area fractions still holds but there is a logical conflict with the requirement of mutually exclusive LC classes (De Bruin 2000b). In this study, the expected fraction of LC classes were calculated at 30x30 coarser resolution than the probability maps. As many GLC map users conduct applications at coarser resolution (Herold et al. 2011), the applied method could be used to create other user-specific legends from the integrated GLC map developed in this study. For users requiring medium resolution GLC maps, recently developed high resolution (30m) GLC maps such as the Globeland30 (Chen et al. 2015) could be used to create GLC maps with user-specific legends.

Internally, uncertainty about the LC fractions (i.e. uncertainty owing to the estimated probabilities) cannot be directly assessed from the LC fraction maps. Such assessment would require simulation, as demonstrated by (de Bruin 2000a). However, simulation at global scale is computationally demanding and it was not included in the current study. Alternatively, external assessment of the quality of the fractional LC maps would require independent reference data with fractional information. Existing reference datasets rarely provide fraction information of LC classes (Tsendbazar et al. 2015a). The reference dataset of Boston University, which is under development, could be used for this purpose. In this dataset, the

reference LC of 5x5km area is classified using object-based analyses of sub-meter resolution images (Olofsson et al. 2012). Hence, sample unit area can be reduced depending on the requirement and fraction of LC classes within sample unit area can be estimated and used as reference. Also comparisons with other LC type fraction datasets such as hybrid global forest cover map by Schepaschenko et al. (2015) are possible. However, the latter should not be confused with validation because that would imply that the map used for comparison is deemed error free.

Note that fractional forest cover as used in the legend of Figure 5.9, which follows the forest definition from Table 5.1, should not be confused with partial tree cover such as mapped by Hansen et al. (2013). This also applies to the fractions of other LC classes for which the original LC categories are maintained rather than element fractions which perhaps are easier to understand. In case a user requires element fractions rather than LC class fractions according to the legend definition in Table 1, different datasets such as Landsat-based tree cover by Hansen et al. (2013) and global cropland percentage map by Fritz et al. (2015) should be used.

5.5 Conclusions

This study applied regression kriging that makes use of both the GLC maps and the available reference datasets to develop an improved GLC map. The integrated GLC map showed 10% improvement in global correspondence and regionally up to 13% increase in correspondence with reference LC compared to the input GLC maps. Benefits of using geostatistical approach and available reference dataset were demonstrated as regions with more spatial correlation in the regression residuals resulted in with higher increase in the correspondence. This finding encourages the efforts of releasing available reference datasets to the public by international communities such as the GOFC-GOLD, Geo-Wiki portal and International Steering Committee for Global Mapping (Fritz et al. 2011a; GOFC-GOLD 2014; Tateishi et al. 2011). As more reference datasets are becoming available to the public, GLC mapping can be further improved by using the pool of all available reference datasets.

Along with the integrated GLC map, a probability map for each LC class was produced. Probability maps for LC classes are helpful in assessing the complexity of heterogeneous landscape and understanding the uncertainty of the integrated GLC map. Uncertainty in class assignment was higher in transition areas of major biomes and tundra regions where multiple LC classes had similar probabilities.

In this study, two GLC maps with user-specific legends for land system models and biodiversity assessments were created to demonstrate the concept of creating GLC

maps that meet the thematic requirement of different users. Area fraction maps of GLC classes are useful to create flexible user-specific legends particularly related to mosaic classes and thus address the thematic requirements of multiple users. Tuning LC maps to specific user needs should be taken into consideration for future GLC mapping efforts.

Acknowledgement

This study was supported by the ESA Land Cover CCI project. The authors would like to thank the ESA GOF-C-GOLD project office, Geo-Wiki platform and International Steering Committee for Global Mapping for providing the access to reference datasets on global land cover. We are grateful to the experts involved in the generation of the reference datasets for their valuable input.

Chapter

6

Synthesis

6.1 Main Results

Given the importance of GLC maps for a variety of user applications, this PhD research aimed to assess, compare and improve GLC maps while accounting for different user requirements. To address this objective, several research questions were answered in Chapters 2-5. This section summarizes the findings related to the research questions.

1. What is the suitability of GLC reference datasets from the perspectives of different user communities?

This research was formulated to assess the re-usability of GLC reference datasets in calibrating and validating GLC maps for different user communities. The research question is addressed in Chapter 2 in two steps.

Firstly, a literature analysis was conducted to obtain metadata information on available and forthcoming GLC reference datasets. The main characteristics of 12 GLC reference datasets in relation to the response design, sampling design protocols and current use of the datasets was reviewed. The results showed that the GLC reference datasets differ in aspects such as sample unit size and type, sample selection scheme as well as temporal coverage. The LCCS legend scheme and stratified random or systematic sample schemes were adopted in most datasets. In terms of response design, available satellite images, open source maps, geo-tagged photos and regional maps were used as source material for reference LC labelling, which was done by visual interpretation in most cases. Results showed that the existing GLC reference datasets have limitations to applications outside the scope of intended uses. Public accessibility of the datasets plays an important role for their efficient use.

Secondly, the user requirements regarding GLC reference datasets were reviewed and potential uses and limitations were assessed for four targeted GLC map user groups. These user groups included the climate modelling community (CMC), global forest change analysts, the GEO community of practice (CoP) for global agricultural monitoring and producers of improved GLC maps. Findings revealed GLC reference datasets have varying suitability levels depending on the reference data characteristics, user requirements and target maps. In general, LC-CCI, GOFD-GOLD, FAO-FRA and Geo-Wiki were identified as being the most suitable datasets for multiple user groups. Most datasets were identified to be at least marginally suitable for the GEO Global Agricultural Monitoring CoP and the improved GLC map production cases. On the other hand, many datasets were considered unsuitable for CMC and global forest change analysts groups since they

do not have fixed spatial sampling units that are visited repeatedly in time to facilitate change assessments.

The results of Chapter 2 stress the potentiality of GLC reference datasets for multiple uses and the need to provide public access to existing reference datasets which is likely to improve the usability of datasets outside their initially intended use.

2. Which GLC map has the highest accuracy considering the perspectives of specific applications?

This research question was answered in Chapter 3. Three GLC maps, i.e., Globcover, Land Cover-CCI and MODIS maps for the year 2005 were comparatively assessed. I re-interpreted and translated the existing reference dataset: Globcover-2005 into three different legends. This dataset was used as reference for accuracy assessment. The results showed that the LC-CCI and MODIS maps had similar overall accuracy ($70.8\pm 1.4\%$ and $71.4\pm 1.3\%$ respectively), while the Globcover map had the lowest accuracy ($61.3\pm 1.5\%$).

Since the impact of confusion errors varies for different applications, the maps were compared from the perspectives of five applications, i.e., general circulation models (GCM), dynamic global vegetation models (DGVM), agriculture assessments, carbon estimation and biodiversity assessments. Weighted accuracy assessment was done by deriving weights based on similarity of LC classes from the perspective of the applications. The corresponding overall weighted accuracy of the GLC maps varied between 80-92%. An accuracy increase was observed because confusions between some classes tended to be inconsequential for the applications considered. The MODIS map had the highest accuracy for all applications except DGVM, for which the LC-CCI map had slightly higher accuracy.

The impact of LC classes and confusions on the overall weighted accuracies were also assessed to identify priority classes and areas for improving GLC maps from the perspective of different applications.

The findings of this chapter suggest that the accuracy of GLC maps should be assessed per applications when the aim is to determine the fitness of GLC maps for an application; there is no single-figure accuracy estimate expressing map fitness for all purposes.

3. What are the local accuracies of the recent GLC maps?

Chapter 4 addressed this research question by assessing the spatial accuracy of the four recent GLC maps Globcover-2009, Land Cover-CCI-2010, MODIS-2010 and Globeland30. The assessment focused on Africa—a continent with complex heterogeneous landscapes that are known to cause inconsistencies among GLC maps. The study utilized GLC reference datasets that are publicly available from the GOF-C-GOLD Reference Data portal, the Geo-Wiki platform and the International Steering Community for Global Mapping.

It was found that the correspondence between GLC maps and the reference dataset varied spatially (Figure 4-7f). Regional differences in the quality of the GLC maps suggested the potential of creating an improved GLC map by integrating them.

The Sahel and dry savannah regions in Africa were identified as the areas with most uncertainty in LC class assignment.

4. Which integration method performs best for improving the recent GLC maps?

This research question was answered in Chapter 4. This research built on the answer to research question 3 that provided a motivation for integrating GLC maps based on their spatial variation in map correspondence. Geostatistical methods that are based on spatial variation in map accuracy and LC class probabilities were tested for improving the recent GLC maps. In addition, a conventional voting approach as commonly used for LC map integration (Iwao et al. 2011) was compared along with a geostatistical method that relies only on the reference data only. The performance of the methods was assessed by cross-validation.

Except for the voting method, all methods produced an increased correspondence with reference LC. The geostatistical integration methods using both the reference data and the existing maps resulted in 67.5%–76% correspondence with the reference LC, which is at least 4.5%–13% higher than that of the input maps. Regression kriging produced the highest correspondence (76%). This method ensures to reduce the classification errors by modelling the LC class probabilities using multinomial logistic regression and adjusting the predicted probabilities by kriging the local regression residuals. An improved LC map and LC class probability maps of Africa were created using this method.

The results of Chapter 4 highlight the benefits of using different reference datasets and geostatistical approaches for map integration.

5. How can improved GLC maps be produced to address the thematic requirements of different users?

Chapter 5 addressed this research question. An improved GLC map and LC class probability maps were created using the regression kriging method that was introduced in Chapter 4.

To answer the user aspect of research question 5, a concept of producing GLC maps with user-specific legends based on area fraction maps of LC classes was proposed. With LC class area fraction information, GLC map users can create a flexible legend that fits their requirements. Users can decide the number of LC classes as well as class definitions that involve fraction thresholds.

To demonstrate this concept, I created GLC maps with user-specific legends from the perspectives of land system modelling and biodiversity assessments. Based on LC class probability maps produced from the map integration, expected area fraction maps for LC classes at a coarser resolution were created. These were used to create additional mosaic classes that were considered useful for the user applications.

As results of this research, an integrated GLC map was created and this map showed a 10% increase in the global correspondence and regionally up to 13% increase in correspondence with reference LC compared to the input GLC maps. Based on the integrated probability maps, two GLC maps with 5 arcminute resolution for land system models and biodiversity assessments were created. Each map had additional mosaic LC classes customized to the considered user applications. Area fraction maps of LC classes allowed creating user-specific mosaic classes and thus can address the thematic requirements of multiple users. This should be taken into account in future GLC mapping efforts

6.2 Reflection and outlook

6.2.1 Revisiting GLC mapping and validation

This section revisits the developments in GLC mapping and validation including the assessments done in this thesis. The trends of GLC maps are discussed with respect to spatial, thematic and temporal properties along with their accuracy assessments and user considerations.

Interest in mapping global scale LC makes GLC mapping an actively progressing field. In the last 5 years, at least 14 GLC maps have been produced and this accounts for more than half of the currently available GLC products. Figure 6.1 summarizes the trends in GLC mapping including the assessments done in this thesis. Figure 6.1a shows the spatial resolution of GLC maps with respect to their published year. This figure shows a trend of GLC mapping progressing towards

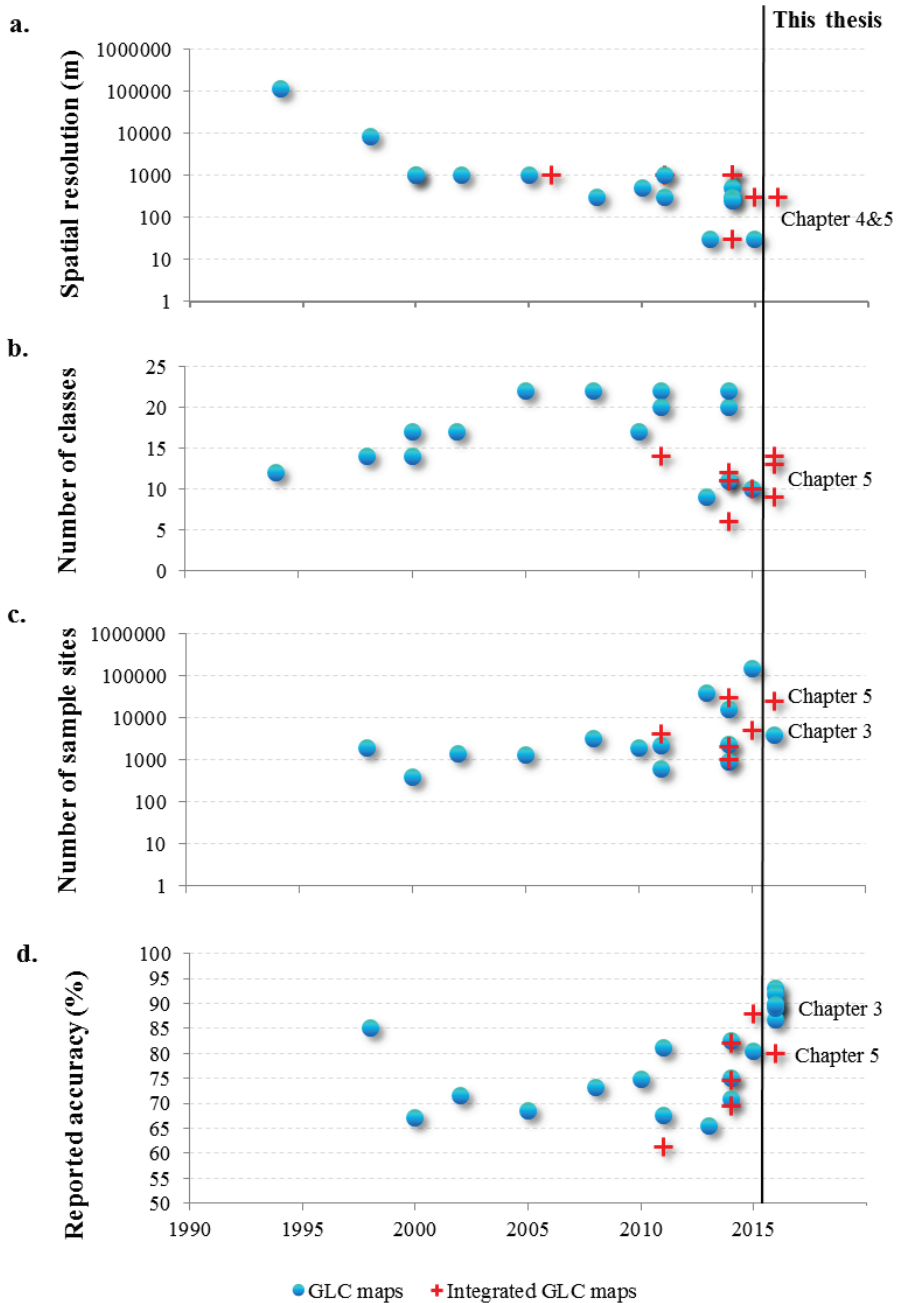


Figure 6.1. Trends in GLC mapping

higher resolution. In particular, with the release of the Landsat archive, high resolution (30m) GLC maps such as FROM-GLC and Globeland30 were developed (Chen et al. 2015; Gong et al. 2013). This is an important progress since high resolution satellite data can provide detail in heterogeneous landscapes (Mora et al. 2014). On the other hand, medium resolution times series data continue to have a significant role in GLC mapping due to their high temporal frequency and ability to characterize vegetation phenology (Friedl et al. 2002). The availability of multiple GLC maps led to the production of integrated maps such as presented in Chapter 5, which tend to have similar resolution as the maps used as input.

In terms of thematic detail, maps before 2010 showed an increase in the number of classes (Figure 6.1b). However, since 2011 two diverging trends in thematic detail are observed. GLC maps that are based on medium resolution satellite data continue to have more detailed thematic information (20-22 classes). On the other hand, Landsat based maps have fewer classes due to the lack of information on intra-annual land dynamics. Integrated maps also have fewer classes due to legend harmonization of several GLC maps that were used as inputs. This is also the case for the integrated GLC map presented in Chapter 5, although in Chapter 5 user-specific legends targeting specific user needs were also elaborated.

Note that Figure 6.1 refers to the published year of GLC maps rather than their target years. In terms of target year, apart from the maps generated for the year around 1993, GLC maps mostly concern LC in 2000, 2005 and 2010 (Table 1.1).

Figure 6.1c shows the number of reference sample sites used for assessing the accuracy of GLC maps. With increased availability of remote sensing data, more sample sites have been used for assessing map accuracies. In particular, Landsat based GLC maps were assessed using more than 30 000 sample sites. Other than data availability, crowd-based reference data collection efforts also enabled an increase in the number of sample sites used for accuracy assessment. This includes the Degree Confluence Project, which collects reference data based on photos and site descriptions gathered by volunteers who visited each confluence point of latitudes and longitudes (Iwao et al. 2006). Similarly, the Geo-Wiki platform collects global and regional scale reference data for LC through interpretation of satellite data or photo by volunteers (Fritz et al. 2009). This project collects reference data with a large number of sample sites that are being used to assess and calibrate multiple global scale maps (Fritz et al. 2015; Schepaschenko et al. 2015; See et al. 2015). Some reference datasets were made accessible to the public by international institutions such as the GOF-C-GOLD and International Steering Committee for Global Mapping. Reusability of such available reference dataset was assessed in Chapter 2 and accessible data were also used for accuracy

comparison (Chapter 3) as well as calibration and validation of the integrated GLC map (Chapter 5).

The accuracy of earlier GLC maps such as the map of the University of Maryland (Hansen et al. 2000) has not been assessed but over the last 15 years accuracy assessment has become commonplace. Except the 8-km resolution map of DeFries et al. (1998) who reported an overall accuracy range of 81.4 to 90.3%, maps produced before 2011 have a reported accuracy of around 70-75% (Figure 6.1d). Maps produced since 2011 have accuracies ranging from 61% to 87%. Including the integrated GLC map presented in Chapter 5, at least 6 maps were reported to have an overall accuracy of 80% or more. This could suggest a weak trend of increasing reported accuracy (Figure 6.1d). The assessments done in Chapter 3 showed higher accuracies when user's perspective was accounted for. Note that the reported accuracies of the GLC maps are not directly comparable due to the differences in thematic detail and validation strategies such as reference sample size and accuracy calculation e.g. area weighted accuracy.

As GLC maps are used for a variety of applications, requirements and perspectives of certain users have been considered in the map development procedure. The IGBP-DIScover and the LC-CCI maps were developed considering the requirements of modelling communities for climate, earth system and biogeochemical studies (CCI-LC 2014; Loveland et al. 2000). The MODIS collection 5 maps were developed with five different legends to address the thematic requirements of different users (Friedl et al. 2010). Map accuracy assessment as well as map integration have also been conducted by taking the perspective and need of specific users into account (DeFries and Los 1999; Jung et al. 2006; Tuanmu and Jetz 2014). The perspectives and requirements of certain users were addressed extensively in this thesis and this is discussed detail in the following section.

6.2.2 Added value of this research

This thesis aimed to address GLC map validation, comparison and integration from the perspective of GLC map users. In this scope:

- I. Chapter 2 assess the re-usability of current GLC reference datasets considering the requirements of specific user groups.
- II. Chapter 3 assess and compare the recent GLC maps from the perspective of several user applications.
- III. Spatial accuracy of several GLC maps are assessed and used for integration for an improved map in Chapter 4.

- IV. An integrated GLC map is created and the concept of GLC maps with user-specific legend is demonstrated in Chapter 5.

This PhD research assessed and demonstrated the importance of user requirements and perspectives in GLC mapping and validation. This was demonstrated in three different components related to reference datasets, accuracy assessment and map integration. Figure 6.2 provides an overview of user-oriented GLC mapping and validation addressed in this thesis.

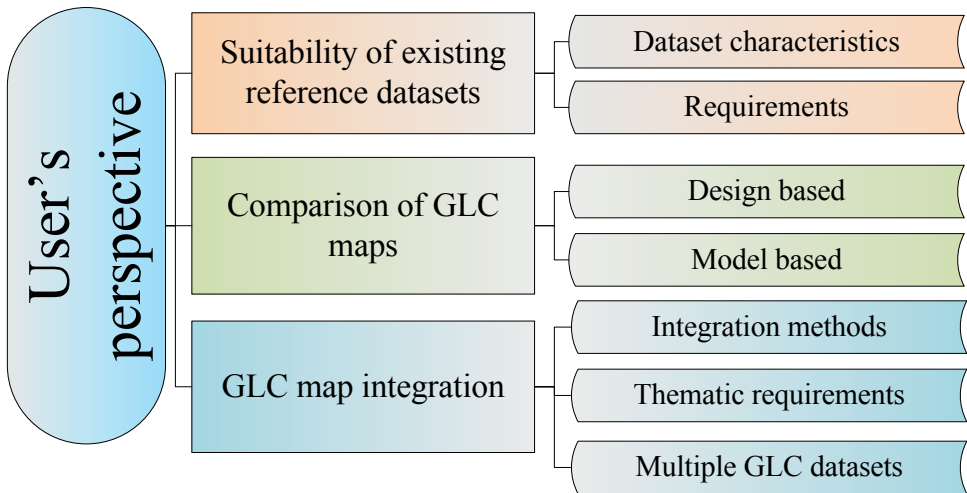


Figure 6.2. An overview of user-oriented GLC mapping and validation of this thesis

Firstly, given the need for reliable reference data, re-utilizing existing datasets for calibration and validation of GLC maps can be a useful approach. Several GLC reference dataset have become accessible for re-use by international communities and there are ongoing efforts to release new datasets (Fritz et al. 2011a; GOFCC-GOLD 2014). Owing to increased data availability, datasets with large numbers of sample sites have been produced recently (Figure 6.1). To make use of such datasets, it is important to know the dataset characteristics and suitability considering specific user requirements. Addressing this, Chapter 2 reviews systematically the metadata of several existing reference datasets and assessed their suitability in different applications. This assessment can be useful to re-use accessible datasets properly based on specific needs. Furthermore, this study is helpful for characterizing the specifications of new reference datasets that comply

with specific user requirements. This thesis also provides concrete examples on how the existing reference dataset can be used (Chapters 3-5).

Secondly, this thesis address comparison of GLC maps from the perspective of specific user applications. A number of GLC maps have been produced till date and more are coming with the developments of new satellites e.g. VIIRS, Sentinel and Proba-V. The accuracies of these maps have been assessed using different reference datasets having varying properties on sampling and response design (Herold et al. 2008). As a result, the reported accuracies of GLC maps (such as in Figure 6.1) are not directly comparable and this makes it challenging for map users who are interested in using the best available map fitting certain criteria. This issue is addressed in Chapter 3 where a design-based statistical accuracy assessments of recent GLC maps is described which considers different user perspectives.

Due to the lack of reference datasets that can be used to validate multiple maps, previous GLC map comparisons focused mostly on assessing agreement amongst several maps. Chapter 3 makes use of existing reference datasets with flexible thematic information to compare multiple maps. By providing comparative map accuracies, this assessment is useful for map users to know the best available map and to account user specific accuracies for their uncertainty assessments. An approach to objectively derive weights denoting the importance of misclassification errors is demonstrated and this can be used to consider other map user applications since the impact of misclassification errors varies depending on user applications (DeFries and Los 1999).

Furthermore, this thesis contributes scientifically to assessments of local accuracy of GLC maps. Since classification errors are not distributed evenly across a map (Carneiro and Pereira 2012), assessing map local accuracy complements to the commonly practiced statistical global accuracy assessments. A model-based geostatistical approach that models local accuracy of several GLC maps is provided in Chapter 4. Previous research focused on spatial accuracy assessment of LC maps in a relatively small regions (Carneiro and Pereira 2012; Kyriakidis and Dungan 2001). In contrast, this study demonstrates that spatial accuracy of GLC maps can be assessed for a large region (e.g. continent). This puts demands on computational power and requires large numbers of sample sites. Both these issues are being solved with current technological developments and progress in reference data generation and this opens a possibility to asses local accuracy of current and upcoming GLC maps.

Thirdly, this thesis also addresses requirements of improved map accuracy and flexible thematic information. Although there is a slight increase in reported

accuracies of GLC maps over recent years (Figure 6.1), these accuracies mostly do not meet the requirements of main map users (Kooistra et al. 2010). To improve accuracy, GLC maps have been integrated using a variety of methods (Jung et al. 2006; See et al. 2015; Tuanmu and Jetz 2014). Chapter 4 of this thesis compares multiple integration methods. The chapter accounts for local variation of map accuracy in map integration (See et al. 2015) by using geostatistics. Using cross-validation, regression kriging was shown to increase the correspondence of LC with reference data substantially. Based on this finding, an integrated GLC map is created in Chapter 5.

Chapter 5 also demonstrates the possibility to create GLC maps with user-specific legends. Although users of GLC maps often have interests in differing thematic information, current GLC maps are produced with a fixed legend. To tackle this limitation, GLC maps with user-specific legends for specific applications are demonstrated based on expected area fraction maps of LC classes. The approach allows the users to characterize additional mosaic classes depending on their requirements.

To conclude, this PhD research demonstrates the importance of requirements and perspectives of multiple users in GLC mapping and validation. Furthermore, benefits of using existing reference data in assessing, comparing and improving GLC maps are illustrated along with related advancing methods.

6.2.3 Limitations of the applied research methods and possible solutions

To better reflect the requirements and perspectives of different users for GLC mapping and validation, interaction with user groups is a key issue. This PhD research reflected the user's requirements and perspectives by reviewing published scientific papers and previous reports on user requirement analyses. This is a passive form of analysing user needs and it could limit better understanding of user's perspectives and needs. This can be improved by actively interacting with the users through a dedicated survey and facilitating user dialogs to engage users in mapping and improvement strategies. Furthermore, user engagement should not be a one way process only and map users should also proactively interact with the map producers to ensure the input GLC map for their analyses meet their requirements.

Most global models that use GLC maps—in particular climate models—aggregate the maps to a coarser resolution and use them in analyses e.g. by selecting the dominant LC class or area fraction information of LC classes within a grid (Nakaegawa 2011; Verburg et al. 2012). Assessments and comparisons of GLC

maps for different user perspectives in Chapter 3 and accuracy analyses of the integrated map in Chapter 5 were limited to the original resolution, not at the coarser resolution. Previous research has shown that the accuracy of GLC maps increased at coarser resolution due to spatial aggregation effect and area fraction at coarse resolution is strongly dependent on the original higher resolution data due to area ratio retention effect (Nakaegawa 2011; Yu et al. 2014). Accuracy assessment of GLC maps at coarser resolution is important for uncertainty propagation assessments of user applications. This can be addressed by assessing spatially aggregated GLC maps using an area-based reference dataset. With such dataset, the sample unit area can be reduced depending on the target map. Potential reference datasets for this can be the datasets by Boston University and LC-CCI projects of the ESA, which are under-development (Achard et al. 2011; Olofsson et al. 2012).

Existing GLC reference datasets were integrated and used for GLC map assessment and improvement using geostatistical approach (Chapter 4 and 5). This thesis does not deal with integration of GLC reference datasets for design-based accuracy assessments. Design-based statistical inference requires probability sampling. Reference datasets with probability sampling design could be integrated or re-used using methods introduced in Stehman et al. (2012) and (2014). These include methods to augment stratified random sampling (e.g. increase number of sample sites) and to re-use a stratified random sampling to maps with different strata.

Furthermore, other differences in reference datasets in terms of e.g. thematic legend and spatial support can be reduced by joint efforts of GLC mapping community and following internationally accepted map validation strategies (Herold et al. 2006; Mora et al. 2014).

6.2.4 Outlook

Current developments in GLC mapping and monitoring is expected to progress further towards more innovative and operational products in the near future. This is inevitable due to (1) the continued importance of GLC monitoring and (2) developments in technological innovations. Recently, the Conference of the parties (COP 21) of the UNFCCC came to an agreement to stabilize global temperature increase and to reduce global emissions (UNFCCC 2015). Since changes in LC are one of the main drivers contributing to global greenhouse-gas emissions and climate, the need to monitor and map GLC remains essential. Furthermore, innovative developments in satellite missions are contributing to improvements in remotely sensed data in terms of spatial, temporal and spectral domain. Recently launched ESA's Sentinel-1 and-2 missions can provide higher resolution (10-20m) optical and radar data with revisit times of a few days (ESA 2016). Sentinel-3 will

be launched to provide high temporal resolution optical and radar data with around 300m spatial resolution (ESA 2016). Ensuring data continuity between SPOT-VGT satellites and Sentinel 3, the Proba-V mission was also recently launched to provide daily observations of land surface and vegetation at 100m to 1km resolution (VITO 2013). Monitoring large regions using higher resolution (<10m) data is also foreseen with satellite missions such as SPOT-5 (Astrium 2012).

With the current data-rich era of earth observation, mapping LC at global scale based on a single sensor data is feasible to a certain extent. However, to improve GLC mapping in terms of spatial and temporal resolutions, fusing multi-sensor data is beneficial. For instance, high spatial and medium temporal resolution data (e.g. Sentinel-1 and-2) could be integrated with medium spatial and high temporal resolution data (e.g Sentinel-3 and Proba-V) to provide innovative high spatial and temporal resolution LC observations. While such innovate approach is useful to define the progresses of next generation GLC mapping, it is challenging to process global scale large volume of earth observation data and it calls for innovative methodical and technical solutions. Nevertheless, such efforts should be built upon existing two-decade experience and knowledge on GLC mapping. Focusing on uncertain areas and problematic classes of current GLC maps is an important step towards improvement in GLC mapping. In this respect, spatial and statistical accuracy assessments provided in Chapter 3 and 4 can be useful along with the integrated probability maps of LC classes presented in Chapter 5.

Assessments of GLC change are also progressing towards monitoring at finer scale at which most human impact occurs (GOF-C-GOLD 2011). Change in GLC can be assessed by post-classification comparison of Landsat based Globeland30 2000 and 2010 maps (Jun et al. 2014). Several products are also available for global forest change monitoring based on Landsat data (Hansen et al. 2013; Kim et al. 2014). GLC change monitoring can also benefit from integration of multi-sensor data as well as multiple LC change maps at regional scale. Integration of multi-sensor data to obtain high spatial and temporal resolution data can allow monitoring fine scale LC change with stable historical observation. Next generation products on global land monitoring are encouraged to include near real time monitoring of LC changes and research and developments of such products are planned by the ESA's Copernicus Land Monitoring Services (Copernicus-Programme 2015). Developments in forest change monitoring using time-series methods (Assunção et al. 2014; Verbesselt et al. 2010) can be useful for this. Furthermore, spatially combining national and regional scale LC change datasets to create LC change maps at global scale is also possible as demonstrated by Fuchs et al. (2015) who created decadal land cover/use change dataset for Europe.

Progresses in GLC map validation and reference data collection are emerging towards more international cooperation as well as to capitalize public knowledge. Many institutions are collaborating towards advancing methods and tools for operational map validation under the GEO framework (Chen et al. 2013). The role of international initiatives such as the CEOS Cal/Val working group and the GOFC-GOLD LC project office remains vital in supporting recommended practices for LC map validation, reference data sharing and jointly developing international knowledge on GLC map validation (Mora et al. 2014). In cooperation of the GOFC-GOLD, for example, Boston University is producing a reference dataset that can be used to assess multiple maps at different scales (Olofsson et al. 2012). In addition, the GOFC-GOLD LC project office continue to share reference datasets from different institutions. On the other hand, capitalizing the knowledge of the public in reference data collection for LC is progressing actively after the successful implementation of the Geo-Wiki platform (Fritz et al. 2009). This is an efficient way of collecting reference information for large number of locations although its quality can be questioned. After validation by experts, these crowdsourced data have been utilized for multiple calibration and validation purposes for improved GLC maps, forest masks and cropland maps (Fritz et al. 2015; Schepaschenko et al. 2015; See et al. 2015). While reference datasets with probability sampling are key for statistically rigorous accuracy assessments of GLC maps, crowdsourced data has value in calibration and integration of global scale LC products as proven in previous studies (Fritz et al. 2015; Schepaschenko et al. 2015; See et al. 2015).

Claims about advancements in GLC mapping and validation can only be made when the maps and their accuracies meet the requirements of the main map users. The ESA's LC-CCI project has twice organized user requirement assessments by means of surveys and active communications with users (LC-CCI 2014). Furthermore, the LC-CCI project continues to facilitate user interaction through workshops and conferences to reflect on their needs and provide feedback. These experiences along with the outcomes of this thesis stress the need to account for user requirements in upcoming GLC assessments.

References

- Achard, F., Defourny, P., Bontemps, S., Herold, M., & Mayaux, P. (2011). *Product validation plan: Land Cover CCI*. Louvain-la-Neuve, Belgium: Université catholique de Louvain (UCL)-Geomatics.
- Achard, F., Eva, H.D., Stibig, H.J., Mayaux, P., Gallejo, J., Richards, T., & Malingreau, J.P. (2002). Determination of deforestation rates of the world's humid tropical forests. *Science*, 297, 999-1002
- Aide, T.M., Clark, M.L., Grau, H.R., López-Carr, D., Levy, M.A., Redo, D., Bonilla-Moheno, M., Riner, G., Andrade-Núñez, M.J., & Muñiz, M. (2012). Deforestation and Reforestation of Latin America and the Caribbean (2001–2010). *Biotropica*
- Amazon, E. (2010, 2016). Amazon elastic compute cloud (Amazon EC2). Retrieved 15 Oct, 2015, from <https://aws.amazon.com/ec2/>
- Arino, O., Leroy, M., Ranera, F., Gross, D., Bicheron, P., Nino, F., Brockman, C., Defourny, P., Vancutsem, C., & Achard, F. (2007). GLOBCOVER-A Global Land Cover Service with MERIS. In, *Envisat Symposium 2007* (pp. 23-27). Montreux, Switzerland
- Asselen, S., & Verburg, P.H. (2012). A Land System representation for global assessments and land-use modeling. *Global Change Biology*, 18, 3125-3148
- Assunção, J., Gandour, C., & Rocha, R. (2014). DETERring deforestation in the Brazilian Amazon: environmental monitoring and law enforcement
- Avitabile, V., Herold, M., Heuvelink, G., Lewis, S., Phillips, O., Asner, G., Armston, J., Asthon, P., Banin, L., & Bayol, N. (2015). An integrated pan-tropical biomass map using multiple reference datasets. *Global Change Biology*
- Bartholomé, E., & Belward, A. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, 26, 1959-1977
- Becker-Reshef, I., Justice, C., Sullivan, M., Vermote, E., Tucker, C., Anyamba, A., Small, J., Pak, E., Masuoka, E., & Schmaltz, J. (2010). Monitoring global croplands with coarse resolution earth observations: The Global Agriculture Monitoring (GLAM) project. *Remote Sensing*, 2, 1589-1609
- Benítez, P., McCallum, I., Obersteiner, M., & Yamagata, Y. (2004). Global supply for carbon sequestration: Identifying least-cost afforestation sites under country risk consideration. *International Institute for Applied System Analysis*
- Béquignon, J., Caughey, J., Cramer, W., Fellous, J.L., Heip, C., Justice, C., Key, J.R., Koike, T., Lacaux, J.P., Lafaye, M., Lafeuille, J., Mathieu, P.P., Ranchin, T., Scholes,

- B., & Schroedter-Homscheidt, M. (2010). *GEO AND SCIENCE: A report prepared by the European Space Agency in the framework of the GEO Science and Technology Committee in support of the GEO Task ST-09-01 "Catalyzing Research and Development (R&D) Resources for GEOSS"*: European Space Agency.
- Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., Bontemps, S., Leroy, M., Achard, F., Herold, M., Ranera, F., & Arino, O. (2008). *GLOBCOVER: Products Description and Validation Report*. Toulouse, France: MEDIAS-France.
- Bierkens, M., & Burrough, P. (1993). The indicator approach to categorical soil data. *Journal of Soil Science*, 44, 361-368
- Bontemps, S., Defourny, P., Van Bogaert, E., Kalogirou, V., & Arino, O. (2011a). *GLOBCOVER 2009: Products Description and Validation Report*: UCLouvain and ESA.
- Bontemps, S., Herold, M., Kooistra, L., van Groenestijn, A., Hartley, A., Arino, O., Moreau, I., & Defourny, P. (2011b). Revisiting land cover observations to address the needs of the climate modelling community. *Biogeosciences Discussions*, 8, 7713-7740
- Brus, D., & De Gruijter, J. (1997). Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion)* 1. *Geoderma*, 80, 1-44
- Brus, D.J. (2011). Statistical sampling strategies for survey of soil contamination. *Dealing with Contaminated Sites*, 165-206
- Cai, X., Zhang, X., & Wang, D. (2011). Land availability for biofuel production. *Environmental Science and Technology-Columbus*, 45, 334
- Card, D.H. (1982). Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing*, 48, 431-439
- Carneiro, J.D., & Pereira, M.J. (2012). Geostatistical stochastic simulation for spatial accuracy assessment of land cover maps derived from remotely sensed data
- Cassman, K.G., Wood, S., Choo, P., Cooper, H., Devendra, C., Dixon, J., Gaskell, J., Kahn, S., Lal, R., & Lipper, L. (2005). Cultivated systems. *Ecosystems and human well-being: current state and trends*, 1, 745-794
- CCI-LC. (2014). *CCI-LC Product User Guide*. Belgium: UCL-Geomatics.

- Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., & Lu, M. (2015). Global land cover mapping at 30m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, 7-27
- Chen, L., Trinder, J., & Ban, Y. (2013). *Report of ISPRS/GEO Workshop on High Resolution Global Land Cover (GLC) Mapping, held in Beijing China on 27-28 April 2013*. Beijing.
- Clark, M., & Aide, T.M. (2011a). An analysis of decadal land change in Latin America and the Caribbean mapped from 250-m MODIS data. In, *ISPRS*. Melbourne, Australia
- Clark, M.L., & Aide, T.M. (2011b). Virtual Interpretation of Earth Web-Interface Tool (VIEW-IT) for Collecting Land-Use/Land-Cover Reference Data. *Remote Sensing*, 3, 601-620
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, 70, 213
- Comber, A., See, L., Fritz, S., Van der Velde, M., Perger, C., & Foody, G. (2013). Using control data to determine the reliability of volunteered geographic information about land cover. *International Journal of Applied Earth Observation and Geoinformation*, 23, 37-48
- Congalton, R.G. (1988). Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 54, 587-592
- Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37, 35-46
- Congalton, R.G., & Green, K. (2009). *Assessing the accuracy of remotely sensed data: principles and practices*. CRC
- Copernicus-Programme. (2015, 14 Dec 2015). Copernicus-The European Earth Observation Programme. Retrieved Jan 16, 2016
- Costa, M.H., Botta, A., & Cardille, J.A. (2003). Effects of large-scale changes in land cover on the discharge of the Tocantins River, Southeastern Amazonia. *Journal of Hydrology*, 283, 206-217
- de Bruin, S. (2000a). Predicting the areal extent of land-cover types using classified imagery and geostatistics. *Remote Sensing of Environment*, 74, 387-396

- De Bruin, S. (2000b). Querying probabilistic land cover data using fuzzy set theory. *International Journal of Geographical Information Science*, 14, 359-372
- De Fries, R., Hansen, M., Townshend, J., & Sohlberg, R. (1998). Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, 19, 3141-3168
- Defourny, P., Bontemps, S., Martin, B., Brockman, C., Fomferra, N., Grit, K., & Krüger, O. (2011a). *Landcover CCI: Product Specification Document: LandCover-CCI partnership*.
- Defourny, P., Bontemps, S., Obsomer, V., Van Bogaert, E., & Arino, O. (2010). Accuracy assessment of Global Land Cover Maps-lessons learnt from Globcover and Globcorine experiences. In
- Defourny, P., Bontemps, S., Schouten, L., Bartalev, S., Cacetta, P., de Wit, A., di Bella, C., Gérard, B., Giri, C., Gond, V., Hazeu, G., Heinimann, A., Herold, M., Jaffrain, G., Latifovic, R., Lin, H., Mayaux, P., Mûcher, S., Nonguierma, A., Stibig, H., Y. Shimabakuro, Van Bogaert, E., Vancutsem, C., Bicheron, P., Leroy, M., & Arino, O. (2011b). GLOBCOVER 2005 and GLOBCOVER 2009 validation : learnt lessons In, *GOFC-GOLD Global Land Cover & Change Validation Workshop*. Laxenburg, Austria
- Defourny, P., Mayaux, P., Herold, M., & Bontemps, S. (2012). Global Land-Cover Map Validation Experiences. *Remote Sensing of Land Use and Land Cover* (pp. 207-224): CRC Press
- Defourny, P., Vancutsem, C., Bicheron, P., Brockmann, C., Nino, F., Schouten, L., & Leroy, M. (2006). GLOBCOVER: a 300 m global land cover product for 2005 using Envisat MERIS time series. In, *ISPRS Commission VII Mid-Term Symposium: Remote Sensing: from Pixels to Processes* (pp. 8-11). Enschede, Netherlands: Citeseer
- DeFries, R., Hansen, M., Townshend, J., & Sohlberg, R. (1998). Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing*, 19, 3141-3168
- DeFries, R., & Los, S. (1999). Implications of land-cover misclassification for parameter estimates in global land-surface models: an example from the simple biosphere model (SiB2). *Photogrammetric Engineering and Remote Sensing*, 65, 1083-1088
- DeFries, R., & Townshend, J. (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: LCCS*. FAO
- Drobne, S., & Liseč, A. (2009). Multi-attribute decision analysis in GIS: weighted linear combination and ordered weighted averaging. *Informatica*, 33, 459-474
- ESA. (2016). Overview of Sentinel Satellites. Retrieved 22 Jan, 2016, from http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Overview4
- Eva, H., Carboni, S., Achard, F., Stach, N., Durieux, L., Faure, J.-F., & Mollicone, D. (2010). Monitoring forest areas from continental to territorial levels using a sample of medium spatial resolution satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, 191-197
- FAO. (2001). *FRA 2000: Global Forest Cover Mapping*. Rome: Forest Resources Assessment Programme, FAO.
- FAO. (2007). *Global Forest Resources Assessment 2010*. Rome: Forestry Department, FAO.
- FAO. (2008). *Terrestrial Essential Climate Variables: for climate change assessment, mitigation and adaptation*. Rome: FAO.
- FAO, & JRC. (2012). *Global forest land-use change 1990–2005*. Rome: FAO and JRC.
- Feddema, J.J., Oleson, K.W., Bonan, G.B., Mearns, L.O., Buja, L.E., Meehl, G.A., & Washington, W.M. (2005). The importance of land-cover change in simulating future climates. *Science*, 310, 1674
- Feldpausch, T.R., Lloyd, J., Lewis, S.L., Brienen, R.J.W., Gloor, M., Monteagudo Mendoza, A., Lopez-Gonzalez, G., Banin, L., Abu Salim, K., Affum-Baffoe, K., Alexiades, M., Almeida, S., Amaral, I., Andrade, A., Aragão, L.E.O.C., Araujo Murakami, A., Arets, E.J.M.M., Arroyo, L., Aymard, C., G.A., Baker, T.R., Bánki, O.S., Berry, N.J., Cardozo, N., Chave, J., Comiskey, J.A., Alvarez, E., de Oliveira, A., Di Fiore, A., Djagbletey, G., Domingues, T.F., Erwin, T.L., Fearnside, P.M., França, M.B., Freitas, M.A., Higuchi, N., C, E.H., Iida, Y., Jiménez, E., Kassim, A.R., Killeen, T.J., Laurance, W.F., Lovett, J.C., Malhi, Y., Marimon, B.S., Marimon-Junior, B.H., Lenza, E., Marshall, A.R., Mendoza, C., Metcalfe, D.J., Mitchard, E.T.A., Neill, D.A., Nelson, B.W., Nilus, R., Nogueira, E.M., Parada, A., Peh, K.S.H., Pena Cruz, A., Peñuela, M.C., Pitman, N.C.A., Prieto, A., Quesada, C.A., Ramirez, F., Ramirez-Angulo, H., Reitsma, J.M., Rudas, A., Saiz, G., Salomão, R.P., Schwarz, M., Silva, N., Silva-Espejo, J.E., Silveira, M., Sonké, B., Stropp, J., Taedoumg, H.E., Tan, S., ter Steege, H., Terborgh, J., Torello-Raventos, M., van der Heijden, G.M.F., Vásquez, R., Vilanova, E., Vos, V.A., White, L., Willcock, S., Woell, H., & Phillips, O.L. (2012). Tree height integrated into pantropical forest biomass estimates. *Biogeosciences*, 9, 3381-3403

- Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., & Gibbs, H.K. (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
- Foody, G.M. (2008). Harshness in image classification accuracy assessment. *International Journal of Remote Sensing*, 29, 3137-3158
- Foody, G.M. (2009). Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing*, 30, 5273-5291
- Foody, G.M. (2010). Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*, 114, 2271-2285
- 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, GB1016
- Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., & Cooper, A. (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., & Huang, X.M. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114, 168-182
- Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F., & Obersteiner, M. (2009). Geo-wiki. org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, 1, 345-354
- Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., van der Velde, M., Kraxner, F., & Obersteiner, M. (2011a). Geo-Wiki: An online platform for improving global land cover. *Environmental Modelling & Software*, 31, 110-123
- Fritz, S., & See, L. (2005). Comparison of land cover maps using fuzzy agreement. *International Journal of Geographical Information Science*, 19, 787-807
- Fritz, S., See, L., McCallum, I., Schill, C., Obersteiner, M., van der Velde, M., Boettcher, H., Havlík, P., & Achard, F. (2011b). Highlighting continued uncertainty in global land cover maps for the user community. *Environmental Research Letters*, 6, 044005

- Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., Duerauer, M., Albrecht, F., Schill, C., Perger, C., Havlik, P., Mosnier, A., Thornton, P., Wood-Sichra, U., Herrero, M., Becker-Reshef, I., Justice, C., Hansen, M., Gong, P., Abdel Aziz, S., Cipriani, A., Cumani, R., Cecchi, G., Conchedda, G., Ferreira, S., Gomez, A., Haffani, M., Kayitakire, F., Malanding, J., Mueller, R., Newby, T., Nonguierma, A., Olusegun, A., Ortner, S., Rajak, D.R., Rocha, J., Schepaschenko, D., Schepaschenko, M., Terekhov, A., Tiangwa, A., Vancutsem, C., Vintrou, E., Wenbin, W., van der Velde, M., Dunwoody, A., Kraxner, F., & Obersteiner, M. (2015). Mapping global cropland and field size. *Global Change Biology*, *21*, 1980-1992
- Fritz, S., See, L., You, L., Justice, C., Becker-Reshef, I., Bydekerke, L., Cumani, R., Defourny, P., Foley, J.A., Gilliams, S., Gong, P., Hansen, M., Hertel, T., Herold, M., Herrero, M., Kayitakire, F., Latham, J., Leo, O., McCallum, I., Obersteine, M., Ramankutty, N., Rocha, J., Tang, H., Thornton, P., Vancutsem, C., van der Velde, M., Wood, S., & Woodcock, C. (2013). The Need for Improved Maps of Global Cropland. *Eos, Transactions American Geophysical Union*, *94*, 2
- Fritz, S., You, L., Bun, A., See, L., McCallum, I., Schill, C., Perger, C., Liu, J., Hansen, M., & Obersteiner, M. (2011c). Cropland for sub-Saharan Africa: A synergistic approach using five land cover data sets. *Geophysical research letters*, *38*, L04404
- Fuchs, R., Herold, M., Verburg, P.H., Clevers, J.G., & Eberle, J. (2015). Gross changes in reconstructions of historic land cover/use for Europe between 1900 and 2010. *Global Change Biology*, *21*, 299-313
- Gallego, F.J., & Stibig, H.J. (2012). Area estimation from a sample of satellite images: The impact of stratification on the clustering efficiency. *International Journal of Applied Earth Observation and Geoinformation*, *22*, 139-149
- Gardin, S., Van Laere, S., Vancoillie, F., Anseel, F., Duyck, W., De Wulf, R., & Verbeke, L. (2011). Remote sensing meets psychology: a concept for operator performance assessment. *Remote Sensing Letters*, *2*, 251
- Ge, Y., Avitabile, V., Heuvelink, G.B., Wang, J., & Herold, M. (2014). Fusion of pan-tropical biomass maps using weighted averaging and regional calibration data. *International Journal of Applied Earth Observation and Geoinformation*, *31*, 13-24
- GEO. (2012). Group on Earth Observations. Retrieved 24 May, 2012, from http://www.earthobservations.org/about_geo.shtml
- Giri, C., Zhu, Z., & 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
- GOFC-GOLD. (2011). *A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals caused by*

- deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation*. Alberta, Canada: GOF-C-GOLD Project Office, Natural Resources Canada.
- GOF-C-GOLD. (2014). GOF-C-GOLD Reference Data Portal. Retrieved 22 Oct, 2014, from http://www.gofcgold.wur.nl/sites/gofcgold_refdataportal.php
- Göhmann, H., Herold, M., Jung, M., Schulz, M., & Schmullius, C. (2009). Prototyping a probability-based Best Map Approach for global land cover datasets at 1km resolution using MODIS, GLC2000, UMD and IGBP. In, *33rd ISRSE*. Stresa, Italy
- Goldewijk, K.K., Van Drecht, G., & Bouwman, A. (2007). Mapping contemporary global cropland and grassland distributions on a 5× 5 minute resolution. *Journal of Land Use Science*, 2, 167-190
- Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., & Chen, J. (2013). Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, 34, 2607-2654
- Goovaerts, P. (2010). Combining Areal and Point Data in Geostatistical Interpolation: Applications to Soil Science and Medical Geography. *Mathematical geosciences*, 42, 535-554
- Gower, J.C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 857-871
- Hagemann, S. (2002). *An improved land surface parameter dataset for global and regional climate models*. Max-Planck-Institut für Meteorologie
- Hansen, M., DeFries, R., Townshend, J., Carroll, M., Dimiceli, C., & Sohlberg, R. (2003). Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm. *Earth Interactions*, 7, 1-15
- Hansen, M., DeFries, R., Townshend, J., & 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., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., & Townshend, J.R.G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342, 850-853

- Hansen, M.C., & Reed, B. (2000). A comparison of the IGBP DISCover and University of Maryland 1km global land cover products. *International Journal of Remote Sensing*, 21, 1365-1373
- Hansen, M.C., Stehman, S.V., & Potapov, P.V. (2010). Quantification of global gross forest cover loss. *Proceedings of the National Academy of Sciences*, 107, 8650-8655
- Harris, N.L., Brown, S., Hagen, S.C., Saatchi, S.S., Petrova, S., Salas, W., Hansen, M.C., Potapov, P.V., & Lotsch, A. (2012). Baseline Map of Carbon Emissions from Deforestation in Tropical Regions. *Science*, 336, 1573-1576
- Hengl, T., Heuvelink, G.B.M., & Stein, A. (2004). A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*, 120, 75-93
- Herold, M., Hubald, R., & Di Gregorio, A. (2009a). *Translating and evaluating land cover legends using the UN Land Cover Classification System (LCCS)*. Jena, Germany: GOF-C-GOLD.
- 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., van Groenestijn, A., Kooistra, L., Kalogirou, V., & Arino, O. (2011). *User Requirements documents: Land Cover CCI*. Louvain-la-Neuve, Belgium: Université catholique de Louvain (UCL)-Geomatics.
- Herold, M., Woodcock, C.E., di Gregorio, A., Mayaux, P., Belward, A.S., Latham, J., & Schmullius, C.C. (2006). A joint initiative for harmonization and validation of land cover datasets. *Geoscience and Remote Sensing, IEEE Transactions on*, 44, 1719-1727
- Herold, M., Woodcock, C.E., Stehman, S.V., Nightingale, J., Friedl, M.A., & Schmullius, C. (2009b). The GOF-C-GOLD/CEOS land cover harmonization and validation initiative: technical design and implementation. In, *33rd ISRSE*. Stresa, Italy
- Hibbard, K., Janetos, A., van Vuuren, D.P., Pongratz, J., Rose, S.K., Betts, R., Herold, M., & Feddema, J.J. (2010). Research priorities in land use and land-cover change for the Earth system and integrated assessment modelling. *International Journal of Climatology*, 30, 2118-2128
- Huttich, C., Herold, M., Wegmann, M., Cord, A., Strohbach, B., Schmullius, C., & Dech, S. (2011). Assessing effects of temporal compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems: Implications for global monitoring. *Remote Sensing of Environment*, 115, 2445-2459

- Iwao, K., Nasahara, K.N., Kinoshita, T., Yamagata, Y., Patton, D., & Tsuchida, S. (2011). Creation of New Global Land Cover Map with Map Integration. *Journal of Geographic Information System*, 3, 160-165
- Iwao, K., Nishida, K., Kinoshita, T., & Yamagata, Y. (2006). Validating land cover maps with Degree Confluence Project information. *Geophysical research letters*, 33, L23404
- Janssen, L.L.F., & Van der Wel, F.J.M. (1994). Accuracy assessment of satellite derived land-cover data: a review. *Photogrammetric Engineering and Remote Sensing*, 60, 419-426
- JECAM. (2012). Joint Experiment of Crop Assessment and Monitoring Retrieved 22, November, 2012, from <http://www.jecam.org/>
- Jenkins, C.N., Pimm, S.L., & Joppa, L.N. (2013). Global patterns of terrestrial vertebrate diversity and conservation. *Proceedings of the National Academy of Sciences*, 110, E2602-E2610
- Joseph, S., Blackburn, G.A., Gharai, B., Sudhakar, S., Thomas, A., & Murthy, M.S.R. (2009). Monitoring conservation effectiveness in a global biodiversity hotspot: the contribution of land cover change assessment. *Environmental monitoring and assessment*, 158, 169-179
- Jun, C., Ban, Y., & Li, S. (2014). China: Open access to Earth land-cover map, 514, 434-434
- 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
- Kaptué Tchuenté, A.T., Roujean, J.L., & De Jong, S.M. (2011). Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale. *International Journal of Applied Earth Observation and Geoinformation*, 13, 207-219
- Kempen, B., Brus, D.J., Heuvelink, G.B., & Stoorvogel, J.J. (2009). Updating the 1: 50,000 Dutch soil map using legacy soil data: a multinomial logistic regression approach. *Geoderma*, 151, 311-326
- Kim, D.-H., Sexton, J.O., Noojipady, P., Huang, C., Anand, A., Channan, S., Feng, M., & Townshend, J.R. (2014). Global, Landsat-based forest-cover change from 1990 to 2000. *Remote Sensing of Environment*, 155, 178-193

- Kinoshita, T., Iwao, K., & Yamagata, Y. (2014). Creation of a global land cover and a probability map through a new map integration method. *International Journal of Applied Earth Observation and Geoinformation*, 28, 70-77
- Klein Goldewijk, K., Beusen, A., & Janssen, P. (2010). Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1. *The Holocene*
- Knight, J. (2002). Accuracy Assessment of Thematic Maps Using Inter-Class Spectral Distances. In *Forestry* (p. 127). Raleigh, The USA: North Carolina State University
- Kooistra, L., Groenestijn, A., Kalogirou, V., Arino, O., & Herold, M. (2010). User requirements from the climate modelling community for next generation global products from land cover CCI project. In *ESA-iLEAPS-EGU joint Conference 2010*. Frascati, Italy: ESA Special Publications SP-688
- Kyriakidis, P.C., & Dungan, J.L. (2001). A geostatistical approach for mapping thematic classification accuracy and evaluating the impact of inaccurate spatial data on ecological model predictions. *Environmental and ecological statistics*, 8, 311-330
- Kyriakidis, P.C., & Yoo, E.H. (2005). Geostatistical prediction and simulation of point values from areal data. *Geographical Analysis*, 37, 124-151
- Lapola, D.M., Oyama, M.D., Nobre, C.A., & Sampaio, G. (2008). A new world natural vegetation map for global change studies. *Anais da Academia Brasileira de Ciências*, 80, 397-408
- Latham, J., Cumani, R., Rosati, I., & Bloise, M. (2014). *FAO Global Land Cover (GLC-SHARE) Beta-Release 1.0 Database*. Rome, Italy: Land and Water Division, FAO
- LC-CCI. (2014). Overview of Phase 2: Land Cover CCI project. Retrieved Jan 22, 2016, from <http://www.esa-landcover-cci.org/?q=planification>
- Letourneau, A., Verburg, P.H., & Stehfest, E. (2012). A land-use systems approach to represent land-use dynamics at continental and global scales. *Environmental Modelling & Software*, 33, 61-79
- Li, W., & Zhang, C. (2011). A Markov chain geostatistical framework for land-cover classification with uncertainty assessment based on expert-interpreted pixels from remotely sensed imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 49, 2983-2992
- López-Carr, D., Davis, J., Jankowska, M.M., Grant, L., López-Carr, A.C., & Clark, M. (2011). Space versus place in complex human–natural systems: Spatial and multi-level

- models of tropical land use and cover change (LUCC) in Guatemala. *Ecological Modelling*, 229, 64-75
- Loveland, T., Reed, B., Brown, J., Ohlen, D., Zhu, Z., Yang, L., & Merchant, J. (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
- Magnussen, S., & de Bruin, S. (2003). Updating cover type maps using sequential indicator simulation. *Remote Sensing of Environment*, 87, 161-170
- Malczewski, J. (2000). On the use of weighted linear combination method in GIS: common and best practice approaches. *Transactions in GIS*, 4, 5-22
- Malczewski, J. (2006). Ordered weighted averaging with fuzzy quantifiers: GIS-based multicriteria evaluation for land-use suitability analysis. *International Journal of Applied Earth Observation and Geoinformation*, 8, 270-277
- Mayaux, P., Eva, H., Gallego, J., Strahler, A.H., Herold, M., Agrawal, S., Naumov, S., De Miranda, E.E., Di Bella, C.M., Ordoyne, C., Kopin, Y., & Roy, P.S. (2006). Validation of the global land cover 2000 map. *Geoscience and Remote Sensing, IEEE Transactions on*, 44, 1728-1739
- McCallum, I., Obersteiner, M., Nilsson, S., & Shvidenko, A. (2006). A spatial comparison of four satellite derived 1km global land cover datasets. *International Journal of Applied Earth Observation and Geoinformation*, 8, 246-255
- McCarthy, D.P., Donald, P.F., Scharlemann, J.P.W., Buchanan, G.M., Balmford, A., Green, J.M.H., Bennun, L.A., Burgess, N.D., Fishpool, L.D.C., Garnett, S.T., Leonard, D.L., Maloney, R.F., Morling, P., Schaefer, H.M., Symes, A., Wiedenfeld, D.A., & Butchart, S.H.M. (2012). Financial Costs of Meeting Global Biodiversity Conservation Targets: Current Spending and Unmet Needs. *Science*, 338, 946-949
- McRoberts, R.E. (2010). Probability-and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. *Remote Sensing of Environment*, 114, 1017-1025
- Milligan, G., & Cooper, M. (1988). A study of standardization of variables in cluster analysis. *5*, 181-204
- Monfreda, C., Ramankutty, N., & Foley, J.A. (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles*, 22, GB1022

- Mora, B., Tsendbazar, N.-E., Herold, M., & Arino, O. (2014). Global Land Cover Mapping: Current Status and Future Trends. *Land Use and Land Cover Mapping in Europe* (pp. 11-30): Springer Netherlands
- Nakaegawa, T. (2011). Uncertainty in land cover datasets for global land-surface models derived from 1-km global land cover datasets. *Hydrological Processes*, 25, 2703-2714
- NASA. (2013, 18 Feb 2013). Visible Infrared Imaging Radiometer Suite (VIIRS). Retrieved 18 Feb, 2013, from <http://npp.gsfc.nasa.gov/viirs.html>
- Newbold, T., Hudson, L.N., Hill, S.L., Contu, S., Lysenko, I., Senior, R.A., Börger, L., Bennett, D.J., Choimes, A., & Collen, B. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, 520, 45-50
- Olofsson, P., Foody, G.M., Stehman, S.V., & Woodcock, C.E. (2013). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*, 129, 122-131
- Olofsson, P., Stehman, S.V., Woodcock, C.E., Sulla-Menashe, D., Sibley, A.M., Newell, J.D., Friedl, M.A., & Herold, M. (2012). A global land-cover validation data set, part I: fundamental design principles. *International Journal of Remote Sensing*, 33, 5768-5788
- Pebesma, E.J. (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, 30, 683-691
- Pebesma, E.J., & Wesseling, C.G. (1998). Gstat: a program for geostatistical modelling, prediction and simulation. *Computers & Geosciences*, 24, 17-31
- Pérez-Hoyos, A., García-Haro, F.J., & San-Miguel-Ayanz, J. (2012). Conventional and fuzzy comparisons of large scale land cover products: Application to CORINE, GLC2000, MODIS and GlobCover in Europe. *ISPRS Journal of Photogrammetry and Remote Sensing*, 74, 185-201
- Perger, C., Fritz, S., See, L., Schill, C., Van Der Velde, M., McCallum, I., & Obersteiner, M. (2012). A Campaign to Collect Volunteered Geographic Information on Land Cover and Human Impact. In *GI-Forum 2012: Geovizualisation, Society and Learning* (pp. 83-91). Salzburg, Austria
- Pittman, K., Hansen, M.C., Becker-Reshef, I., Potapov, P.V., & Justice, C.O. (2010). Estimating global cropland extent with multi-year MODIS data. *Remote Sensing*, 2, 1844-1863
- Potapov, P., Hansen, M., Gerrand, A., Lindquist, E., Pittman, K., Turbanova, S., & Wilkie, M.L. (2011). The global Landsat imagery database for the FAO FRA remote sensing survey. *International Journal of Digital Earth*, 4, 2-21

- Poulter, B., Ciais, P., Hodson, E., Lischke, H., Maignan, F., Plummer, S., & Zimmermann, N. (2011). Plant functional type mapping for earth system models. *Geoscientific Model Development*, 4, 993-1010
- Poulter, B., MacBean, N., Hartley, A., Khlystova, I., Arino, O., Betts, R., Bontemps, S., Boettcher, M., Brockmann, C., Defourny, P., Hagemann, S., Herold, M., Kirches, G., Lamarche, C., Lederer, D., Ottlé, C., Peters, M., & Peylin, P. (2015). Plant functional type classification for Earth System Models: results from the European Space Agency's Land Cover Climate Change Initiative. *Geosci. Model Dev. Discuss.*, 8, 429-462
- Powell, R., Matzke, N., & De Souza, C. (2004). Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon. *Remote Sensing of Environment*, 90, 221-234
- Quaife, T., Quegan, S., Disney, M., Lewis, P., Lomas, M., & Woodward, F.I. (2008). Impact of land cover uncertainties on estimates of biospheric carbon fluxes. *Global Biogeochem. Cycles*, 22, GB4016
- Ramankutty, N., Evan, A.T., Monfreda, C., & Foley, J.A. (2008). Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles*, 22, GB1003
- Redo, D.J., Aide, T.M., Clark, M.L., & Andrade-NÚÑEZ, M.J. (2012). Impacts of internal and external policies on land change in Uruguay, 2001–2009. *Environmental Conservation*, 39, 122-131
- Ripley, B., Venables, W., & Ripley, M.B. (2014). Package ‘nnet’. In: March
- Rossiter, D. (2004). *Technical Note: Statistical methods for accuracy assesment of classified thematic maps*: International Institute for Geo-information Science & Earth Observation (ITC)
- Ruesch, A., & Gibbs, H.K. (2008). New IPCC Tier-1 global biomass carbon map for the year 2000
- Saaty, T. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15, 234-281
- Sachs, J., Remans, R., Smukler, S., Winowiecki, L., Anelman, S.J., Cassman, K.G., Castle, D., DeFries, R., Denning, G., & Fanzo, J. (2010). Monitoring the world's agriculture. *Nature*, 466, 558-560
- Scepan, J., Menz, G., & Hansen, M.C. (1999). The DISCover validation image interpretation process. *Photogrammetric Engineering and Remote Sensing*, 65, 1075-1081

- Schepaschenko, D., See, L., Lesiv, M., McCallum, I., Fritz, S., Salk, C., Moltchanova, E., Perger, C., Shchepashchenko, M., Shvidenko, A., Kovalevskiy, S., Gilitukha, D., Albrecht, F., Kraxner, F., Bun, A., Maksyutov, S., Sokolov, A., Dürauer, M., Obersteiner, M., Karminov, V., & Ontikov, P. (2015). Development of a global hybrid forest mask through the synergy of remote sensing, crowdsourcing and FAO statistics. *Remote Sensing of Environment*, 162, 208-220
- Schneider, A., Friedl, M., & Potere, D. (2009). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters*, 4, 044003
- Schultz, M., Tsendbazar, N., Herold, M., Jung, M., Mayaux, P., & Goehmann, H. (2015). Utilizing the Global Land Cover 2000 reference dataset for a comparative accuracy assessment of global 1km land cover maps MAPS. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 40-7/W3, 503-510
- See, L., Fritz, S., Thornton, P., You, L., Becker-Reshef, I., Justice, C., Leo, O., & Herrero, M. (2012). Building a Consolidated Community Global Cropland Map. Retrieved 22, November, 2012, from <http://www.earthzine.org/2012/01/24/building-a-consolidated-community-global-cropland-map/>
- See, L., Schepaschenko, D., Lesiv, M., McCallum, I., Fritz, S., Comber, A., Perger, C., Schill, C., Zhao, Y., Maus, V., Siraj, M.A., Albrecht, F., Cipriani, A., Vakolyuk, M.y., Garcia, A., Rabia, A.H., Singha, K., Marcarini, A.A., Kattenborn, T., Hazarika, R., Schepaschenko, M., van der Velde, M., Kraxner, F., & Obersteiner, M. (2015). Building a hybrid land cover map with crowdsourcing and geographically weighted regression. *ISPRS Journal of Photogrammetry and Remote Sensing*
- Sertel, E., Robock, A., & Ormeci, C. (2010). Impacts of land cover data quality on regional climate simulations. *International Journal of Climatology*, 30, 1942-1953
- Siebert, S., Henrich, V., Frenken, K., & Burke, J. (2013). *Update of the Digital Global Map of Irrigation Areas (GMIA) to Version 5* Bonn, Germany: Institute of Crop Science and Resource Conservation, Rheinische Friedrich-Wilhelms-Universität Bonn.
- Stehman, S.V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62, 77-89
- Stehman, S.V. (2000). Practical Implications of Design-Based Sampling Inference for Thematic Map Accuracy Assessment. *Remote Sensing of Environment*, 72, 35-45
- Stehman, S.V. (2009). Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, 30, 5243-5272

- Stehman, S.V. (2014). Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. *International Journal of Remote Sensing*, 35, 4923-4939
- Stehman, S.V., & Czaplewski, R.L. (1998). Design and Analysis for Thematic Map Accuracy Assessment:: Fundamental Principles. *Remote Sensing of Environment*, 64, 331-344
- Stehman, S.V., Olofsson, P., Woodcock, C.E., Herold, M., & Friedl, A. (2011). A global land cover validation dataset II: augmenting a stratified sampling design
- Stehman, S.V., Olofsson, P., Woodcock, C.E., Herold, M., & Friedl, M.A. (2012). A global land-cover validation data set, II: augmenting a stratified sampling design to estimate accuracy by region and land-cover class. *International Journal of Remote Sensing*, 33, 6975-6993
- Stehman, S.V., & Wickham, J.D. (2011). Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. *Remote Sensing of Environment*
- Strahler, A., Friedl, M., Schaaf, C., Zhang, X., & Hodges, J. (2003). *Validation of the consistent-year v003 MODIS land cover product*: Department of Geography, Boston University.
- Strahler, A.H., Boschetti, L., Foody, G.M., Friedl, M.A., Hansen, M.C., Herold, M., Mayaux, P., Morisette, J.T., Stehman, S.V., & Woodcock, C.E. (2006). Global land cover validation: Recommendations for evaluation and accuracy assessment of global land cover maps. *European Communities, Luxembourg*, 51
- Tateishi, R., Hoan, N.T., Kobayashi, T., Alsaaidh, B., Tana, G., & Phong, D.X. (2014). Production of Global Land Cover Data – GLCNMO2008. 2014, 6
- Tateishi, R., Uriyangqai, B., Al-Bilbisi, H., Ghar, M.A., Tsend-Ayush, J., Kobayashi, T., Kasimu, A., Hoan, N.T., Shalaby, A., Alsaaidh, B., Enkhzaya, T., Gegentana, & Sato, H.P. (2011). Production of global land cover data - GLCNMO. *International Journal of Digital Earth*, 4, 22-49
- Townshend, J.R., Masek, J.G., Huang, C., Vermote, E.F., Gao, F., Channan, S., Sexton, J.O., Feng, M., Narasimhan, R., & Kim, D. (2012). Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges. *International Journal of Digital Earth*, 5, 373-397
- Tsendbazar, N., De Bruin, S., & Herold, M. (2015a). Assessing global land cover reference datasets for different user communities. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, 93-114

- Tsendbazar, N.E., de Bruin, S., Fritz, S., & Herold, M. (2015b). Spatial Accuracy Assessment and Integration of Global Land Cover Datasets. *Remote Sensing*, 7, 15804
- Tsendbazar, N.E., de Bruin, S., Mora, B., Schouten, L., & Herold, M. (2016). Comparative assessment of thematic accuracy of GLC maps for specific applications using existing reference data. *International Journal of Applied Earth Observation and Geoinformation*, 44, 124-135
- Tuanmu, M.-N., & Jetz, W. (2014). A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography*, 23, 1031-1045
- UNFCCC. (2015). *Adoption of the Paris Agreement. Proposal by the President*. Geneva (Switzerland).
- Van Oort, P. (2005). Improving land cover change estimates by accounting for classification errors. *International Journal of Remote Sensing*, 26, 3009-3024
- Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, 114, 106-115
- Verburg, P.H., Asselen, S., Zanden, E.H., & Stehfest, E. (2012). The representation of landscapes in global scale assessments of environmental change. *Landscape Ecology*, 28, 1067-1080
- Verburg, P.H., Neumann, K., & Nol, L. (2011). Challenges in using land use and land cover data for global change studies. *Global Change Biology*, 17, 974-989
- VITO. (2013). The Proba-V instrument.
- Wang, J., Zhao, Y., Li, C., Yu, L., Liu, D., & Gong, P. (2015). Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250 m resolution. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, 38-47
- Wu, W., Shibasaki, R., Yang, P., Ongaro, L., Zhou, Q., & Tang, H. (2008). Validation and comparison of 1 km global land cover products in China. *International Journal of Remote Sensing*, 29, 3769-3785
- Wulder, M.A., Franklin, S.E., White, J.C., Linke, J., & Magnussen, S. (2006). An accuracy assessment framework for large-area land cover classification products derived from medium-resolution satellite data. *International Journal of Remote Sensing*, 27, 663-683
- Yager, R.R. (1988). On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *Systems, Man and Cybernetics, IEEE Transactions on*, 18, 183-190

- You, L., Wood, S., & Wood-Sichra, U. (2009). Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. *Agricultural Systems*, *99*, 126-140
- Yu, L., Wang, J., Li, X., Li, C., Zhao, Y., & Gong, P. (2014). A multi-resolution global land cover dataset through multisource data aggregation. *SCIENCE CHINA Earth Sciences*, *57*, 2317-2329

Summary

Observation of global-scale land cover is of importance to international initiatives, governments, and scientific communities that endeavour to understand and monitor changes affecting the environment. Various applications such as climate models, ecosystem modelling and hydrological models use a number of global land cover (GLC) maps that were produced from different initiatives. The users have different requirements regarding spatial, temporal and thematic aspects of GLC maps as well as their accuracy. For example, climate modellers typically use GLC maps at 1km spatial resolution or coarser whereas this resolution is too coarse for land change science studies to detect small-scale changes. Furthermore, to determine the fitness of GLC maps for certain applications, map accuracy assessments need to consider the perspectives of the users as confusion between certain classes can have a strong impact on specific applications whereas for other applications they are inconsequential. Therefore, generation and assessment of GLC maps needs to account for different user requirements and perspectives. This PhD research aimed to account for different user requirements in assessing, comparing and as well as improving GLC maps. In the Introduction chapter of this thesis, the rationale of this PhD research is given and five research questions are elaborated.

In Chapter 2, the characteristics of current GLC reference datasets that have been used for calibration and validation of GLC maps were reviewed and analysed. Twelve GLC reference datasets were further assessed in terms of their potential use in GLC map validation and calibration towards four GLC user groups, i.e., climate modellers, global forest change analysts, global agricultural monitoring community and map producers. Findings revealed varying GLC reference dataset suitability levels depending on the reference data characteristics, user requirements and target maps. Nonetheless, several datasets (LC-CCI, GOF-C-GOLD, FAO-FRA and Geo-Wiki) were identified as generally being suitable for re-use for multiple user groups. This chapter highlights the potentiality of GLC reference datasets for multiple uses and public access of existing reference datasets in improving the usability of the datasets outside their intended use.

Chapter 3 presents a comparative assessment of thematic accuracies of GLC maps based on an existing reference dataset. The Globcover-2005 reference dataset was processed to assess and compare Globcover, LC-CCI and MODIS maps for the year 2005. These maps were evaluated from the perspective of several user applications i.e., general circulation models, dynamic global vegetation models, agriculture assessments, carbon estimation and biodiversity assessments, using a weighted accuracy assessment procedure. Overall accuracies of the maps ranged between $61.3 \pm 1.5\%$ and $71.4 \pm 1.3\%$. The overall weighted accuracy varied between 80-92% for the considered applications. The latter accuracy is higher because confusions between some classes were deemed inconsequential for the

applications considered. To determine fitness of use of GLC maps, accuracy of GLC maps should be assessed per application; there is no single-figure accuracy estimate expressing map fitness for all purposes.

Chapter 4 assesses the spatial accuracy of Globcover-2009, Land Cover-CCI-2010, MODIS-2010 and Globeland30 in Africa using publicly available GLC reference datasets. Spatial accuracy was modelled by the spatial autocorrelation structures of the local correspondence between map and reference data. Created correspondence maps showed spatial patterns indicating zonal differences in the degree with which different GLC maps matched the reference data. The results showed the potentiality of integrating current GLC maps along with reference data to create an improved GLC map. Different integration methods including geostatistical approaches were tested and assessed by cross-validation. The integration methods based on geostatistical approach resulted in 4.5%–13% higher correspondence with the reference LC than any of the input GLC maps. An integrated LC map and LC class probability maps were created using regression kriging, which produced the highest correspondence (76%). This chapter demonstrates the added value of using reference datasets and geostatistics for improving GLC maps.

Chapter 5 follows up the findings of Chapter 4 and presents an improved GLC map and LC class probability maps. The integrated GLC map has 10% higher global correspondence and regionally up to 13% higher correspondence with reference LC than the individual input maps. To address the thematic requirements of different GLC map users, a concept of producing GLC maps with user-specific legends based on area fraction maps of LC classes is proposed. It is demonstrated by creating GLC maps with user-specific legends from the perspectives of land system modelling and biodiversity assessments. Based on LC class probability maps produced from map integration, expected area fraction maps for LC classes at coarser resolution were created. Based on these maps, two GLC maps at 5 arcminute resolution having additional mosaic LC classes which can be useful for the considered user applications were created. Area fraction maps of LC classes are useful for creating flexible user-specific legends particularly related to mosaic classes and hence they can address thematic requirements of multiple users.

This PhD research demonstrates the importance of accounting for the requirements and perspectives of user applications in validating, comparing and improving GLC maps. The work also includes improving the efficient use of existing GLC reference datasets, comparative accuracy assessment of GLC maps using both the design based and model based approaches as well as presenting an integration method to improve current GLC maps to better meet different application needs.

Acknowledgements

Looking back at my last 4 years of endeavour, I feel that the successful completion of my PhD was only possible with the support and contributions I received from many people who helped me make this journey, scientifically and personally, an interesting and pleasant experience. Herewith, I take this opportunity to thank all of you.

First of all, I would like to thank my parents and siblings for their love and support to make my life abroad an enjoyable experience. I will phrase it in Mongolian: “Хайрт ээж, аав, эгч, дүү нартаа надад гадаадад сурч амьдрахад минь үргэлж чин сэтгэлээсээ дэмжлэг үзүүлж байдагт тань талархаж явдагаа илэрхийлье. Та бүхнийхээ тусламжтайгаар Европд сурч докторын зэрэг хамгаалах гэж байгаадаа баяртай байна.” I was fortunate to start a family in the Netherlands; to my beloved husband Sukhad, your unconditional love, support and enthusiasm made my time as a PhD researcher and a mother a successful and happy experience. I am immensely grateful to have you in my life and thank you for always being by my side. To my little jewel, Misheel, who I gave birth to during this PhD period, your presence fills my life with happiness and joy. As the meaning of your name, the very thought of you makes me smile.

My sincere thanks to my promotor, Prof. Martin Herold. I would like to express my gratitude to you for providing me the great opportunity to undertake my PhD research under your guidance and support. You helped me learn from your scientific experiences and networks to enable me to develop greatly in the past four years. I greatly admire your genuine commitment to your students and team. It is my pleasure to continue working together with you here in Wageningen.

To my daily supervisor Sytze de Bruin, I am indebted to you for your commitment, patience and scientific expertise that helped me complete this PhD research successfully. I consider myself very fortunate to have you as my supervisor. Thank you for always being there to advise, comment and correct me.

I would like to thank my examining committee. Thank you for agreeing to be committee members and for taking the time to review my PhD thesis and attend my defense. To my paranympths; Sarah and Astrid, thank you for your support in organizing last wrap-ups of my PhD thesis and defense. Thank you for your company throughout my PhD period. I am thankful to Richard, Ben, Roberto, Sarah and Sukhad for their efforts in rigorously proofreading my papers and thesis chapters.

It was my pleasure to belong to this group of international students and researchers at the Laboratory of Geo-information Science and Remote Sensing. My gratitude to Arnold, Arend, Harm, Lammert, Ron, Willy, Jan Clevers, John, Lukasz, Jan Verbesselt, Corne and Aldo for their cooperation. I would also like to thank my officemates - Brice, Benjamin, Titia and Kim for the lively small-office conversations and your great company for activities in and outside the office. My special thanks goes to Truus and Antoinette for their endless administrative and personal support.

It was my pleasure to share these years with such a delightful circle of international colleagues - Niki, Erika, Noriko, María, Giulia, Laure, Arun, Loïc, Johannes, Michael, Yang, Peter, Valerio, Juha, Rosa, Hans, Valérie, Jessica, Kathleen, Jalal, José, Manos, Angela, Qijun, Eliakam, Eskender, Alvaro, Marcio, Saed, Jamshid, Simon, Mathieu, Kalkidan, Marston, Johannes Eberenz, Patric and Konstantin - thanks for your help and lovely company in research and social activities.

I had a great time being involved in social activities in Wageningen and around. I greatly enjoyed the good times shared with Zaya, Boloroo, Odko, the family of Buyanaa and Undrakh, Batkhuyag and Kadirbek. My heartfelt thanks to my neighbours and friends Roberto & Marcela (and now Javier), Yélica, Daniel & Arturo (and now Adel), Amy, Jiao & Xiang, Pranesh, Trupti & Pranjali, and of course Puja, for the great times we shared and also for taking care of Misheel, as if she were your own child, whenever we needed. Special thanks to Ben Wielstra and Yali, Christine, Loise and Pedro for your visits that I greatly enjoyed.

Thanks Shibani-di and Denny for making my escape and recharge times from the PhD research pleasant and comfortable. Hartstikke bedankt voor mijn Nederlandse taal mentor, Alida Maandag, voor de hulp en tijd die u besteed om mijn Nederlands te verbeteren. Bovendien wil ik Oyuna & Hans, Huib bedanken voor hun welwillendheid en steun.

Thank you, bedankt, bayarlalaa!

Nandin-Erdene Tsendbazar

Bennekom, 2016

List of publications

Peer reviewed journal publications

Tsendbazar, N.E., de Bruin, S., Mora, B., Schouten, L., & Herold, M. (2016). Comparative assessment of thematic accuracy of GLC maps for specific applications using existing reference data. *International Journal of Applied Earth Observation and Geoinformation*, 44, 124-135

Tsendbazar, N.E., de Bruin, S., Fritz, S., & Herold, M. (2015). Spatial Accuracy Assessment and Integration of Global Land Cover Datasets. *Remote Sensing*, 7, 15804

Tsendbazar, N.E., de Bruin, S., Herold, M., 2015. Assessing global land cover reference datasets for different user communities. *ISPRS Journal of Photogrammetry and Remote Sensing*. 103, 93-114.

Hussin, Y.A., Gilani, H., van Leeuwen, L., Murthy, M., Shah, R., Baral, S., **Tsendbazar, N.E.**, Shrestha, S., Shah, S.K., Qamer, F.M., 2014. Evaluation of object-based image analysis techniques on very high-resolution satellite image for biomass estimation in a watershed of hilly forest of Nepal. *Applied Geomatics* 6, 59-68.

Keshkamat, S., **Tsendbazar, N.**, Zuidgeest, M., Shiirev-Adiya, S., van der Veen, A., van Maarseveen, M., 2013. Understanding transportation-caused rangeland damage in Mongolia. *Journal of Environmental Management* 114, 433-444.

Keshkamat, S., **Tsendbazar, N.**, Zuidgeest, M.H., van der Veen, A., de Leeuw, J., 2012. The environmental impact of not having paved roads in arid regions: an example from Mongolia. *AMBIO: A Journal of the Human Environment* 41, 202-205.

Other scientific publications

Schultz, M., **Tsendbazar, N.E.**, Herold, M., Jung, M., Mayaux, P., & Goehman, H. (2015). Utilizing the Global Land Cover 2000 reference dataset for a comparative accuracy assessment of 1 km global land cover maps. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-7/W3, 503-510

Mora, B., **Tsendbazar, N.E.**, Herold, M., Arino, O., 2014. Global Land Cover Mapping: Current Status and Future Trends, Land Use and Land Cover Mapping in Europe. Springer Netherlands, pp. 11-30.

Tsendbazar N.E., Shiirev-Adiya. S et al., 2008. “Maps of Livestock Number and Pasture Resources of Mongolia (scale M1:5 000 000, M1:10 000 000). The National Atlas of Mongolia, Volume II, Ulaanbaatar, Mongolia.

Tserenchunt, B. **Tsendbazar. N.E.**, 2008. “The Results of Land Cover Change and Desertification Study” in MOGZA scientific magazine, Ulaanbaatar, Mongolia.

Short biography

Nandin-Erdene Tsendbazar was born in Bulgan sum, Dornod aimag in the eastern part of Mongolia on 18 September 1986. She did her primary and secondary schooling in Bulgan sum and in 2001, moved to Choibalsan city to attend her high school.



In 2003, she participated in the National Geographic Olympiad and won silver prize. Subsequently, she was invited to commence her undergraduate degree in geography from National University of Mongolia in Ulaanbaatar, Mongolia. Upon completing her degree in 2007, she joined the Institute of Geography of Mongolian Academy of Science as a research assistant. She worked there at the Laboratory of Geo-Information System and Remote Sensing and was involved in grassland degradation studies in central Mongolia and field surveys of nomadic livestock herders. She also worked actively in the project for renewal of the National Atlas of Mongolia particularly in the mapping and editing of national scale socio-economic maps.

In 2009, Nandika was awarded a scholarship under the Netherlands Fellowship Programme to study her Masters of Science degree at the Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands. She specialised in natural resources management and her MSc thesis focused on object based image analysis of Geo-Eye imagery for above ground carbon stock modelling. The field work was in the Himalayan mid-hills forests of Nepal in collaboration with ICIMOD, Katmandu. After her masters, Nandika moved to Wageningen in September 2011 to begin her PhD at the Laboratory of Geo-information Science and Remote Sensing. Her PhD research focused on user-oriented global land cover map validation, comparison and integration. She has also been actively involved with Land Cover-CCI and Geocarbon projects in global land cover map comparison studies.

Nandika's current research interests are related to mapping and validation of land cover with the use of data from the contemporary satellite sensors such as Proba-V and Sentinel. Her research interests also include integrating existing datasets and (geo)statistical assessments of map quality.

PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)



Review of literature (4.5 ECTS)

- Global land cover validation and user's requirements (2013)

Writing of project proposal (2 ECTS)

- Global land-high resolution hot spot monitoring
- Operation of the global land component, thematic domain vegetation and energy

Post-graduate courses (5.5 ECTS)

- Uncertainty propagation in spatial and environmental modelling; PE&RC (2011)
- Spatio-temporal analysis and big data processing using free and open source software; SENSE and PE&RC (2013, 2014)

Laboratory training and working visits (4 ECTS)

- Validation of land cover change assessment in Indonesia; EC-Joint Research Centre (2012)
- High resolution datasets for reference data generation; Boston University (2013)

Invited review of (unpublished) journal manuscript (2 ECTS)

- International Journal of Applied Earth Observation and Geoinformation: comparison and validation of global land cover products over China (2013)
- International Journal of Geographical Information Science: usage of VGI for validation of land cover maps (2014)

Deficiency, refresh, brush-up courses (3.9 ECTS)

- Spatial mixing analysis and image classification techniques to monitor forest change in the Brazilian Amazon; Laboratory of Geo-information Science and Remote Sensing, Wageningen University (2011)
- Spatial modelling and statistics; Laboratory of Geo-information Science and Remote Sensing, Wageningen University (2013)
- ArcGIS Story map training; Laboratory of Geo-information Science and Remote Sensing, Wageningen University (2015)

Competence strengthening / skills courses (5.8 ECTS)

- Project and time management; Wageningen Graduate Schools (2012)
- Social Dutch plus; Wageningen In'to Languages (2013)
- PhD Competence assessment; Wageningen Graduate Schools (2013)
- Scientific writing; Wageningen Graduate Schools (2014)

PE&RC Annual meetings, seminars and the PE&RC weekend (1.5 ECTS)

- PE&RC PhD weekend (2011)
- PE&RC PhD day (2012, 2015)

Discussion groups / local seminars / other scientific meetings (14.2 ECTS)

- Remote sensing thematic group meeting (2011-2016)
- Modelling and Statistics Network (MSN); PE&RC (2012-2013)
- 3rd GOF-C-GOLD Land monitoring symposium; Wageningen, the Netherlands (2013)
- R User meeting; WUR (2014)
- REDD Discussion group (2014-2015)

International symposia, workshops and conferences (5.5 ECTS)

- GEO-Carbon conference: carbon in a changing world; FAO, Rome, Italy (2011)
- ESA Living planet symposium; Edinburgh, United Kingdom (2013)
- GV2M: Global vegetation monitoring and modelling; Avignon, France (2014)
- 36th International Symposium on Remote Sensing of Environment (ISRSE) (2015)

Lecturing / supervision of practical's / tutorials (3 ECTS)

- Capita Selecta: global land cover validation; Wageningen, the Netherlands (2013)

Supervision of MSc students

- Methods for integrating global land cover datasets: a case study in Western Europe (2015)

Cover illustration is a RGB composite of three global land cover maps highlighting the differences between them.

Published by Uitgeverij BOXPress, Vianen