GIMA

Geographical Information Management and Applications

Methods for integrating global land cover datasets: a case study in Western Europe

Renske van Setten-Zaremba

November 2015

Methods for integrating global land cover datasets: a case study in Western Europe

GIMA Msc Thesis
Date: November 2015

Autor: Renske van Setten-Zaremba
Student number: s6010156

Contact information:
Kalanderplein 1-9
7511HX, Enschede
renskevansetten@gmail.com

First supervisor:
MSc Nandika (NE) Tsendbazar

Second supervisor:
dr.ir. Sytze de Bruin

Professor:
prof.dr. Martin Herold

Summary

Current global land cover (GLC) maps have an overall accuracy around 70%, varying from 67 to 81% (Mora et al. 2014; See et al. 2014). Map producers and users feel a need to improve the quality of GLC maps as errors add to the uncertainty of GLC applications (Herold et al. 2011; Mora et al. 2014; Tsendbazar et al. 2015; Verburg et al. 2011). Improved GLC maps can be achieved by integrating different land cover (LC) datasets (Herold et al. 2008). Before LC maps can be integrated, LC data needs to be harmonized to the same thematic legend and spatial extent. LC products have limitations due to product inconsistency (Tuanmu and Jetz 2014). The use of differing methodologies in LC mapping, integration, classification scheme and algorithms and data sources raises GLC mapping inconsistency issues (Mora et al. 2014). Inconsistencies between GLC dataset form an obstacle for map integration. Integration aims to label LC information to the most accurate LC class, but is dependent on the LC information from the LC maps used for integration.

There are different integration methodologies. Voting assigns a pixel to the LC class that occurs in the majority of the LC input datasets at the pixel's location. Voting is a widely accepted approach in data integration (Ge et al. 2014; Goovaerts. 1999; Iwao et al. 2011; Jung et al. 2006; Kinoshita et al. 2014; Tuanmu and Jetz. 2014). This research uses normal voting, weighted voting and probability voting for the map integration of: FROM-GLC hierarchy (2013), Globcover 2009, LC-CCI (2010) and MODIS5 (2010) LC maps. Normal voting is a new method that is purely map driven and uses a two-step approach: (1) in case the LC input map agree on a LC class, pixels were assigned to a LC class from simple majority voting. (2) In case the LC input maps disagreed on a LC class and formed a tie, pixels were assigned to a LC class based on LC class preferences calculated from step 1. In Weighted voting, pixels are assigned to the LC class that accumulates the highest weight that is derived from user's accuracy at that pixel's location. In case of probability voting this accounts for the probability of each LC class being the correct class. Weights and probabilities were derived from the published confusion matrices FROM-GLC (Yu et al. 2014) and Globcover 2009, LC-CCI (2010) and MODIS5 (2010) (Tsendbazar et al. 2016).

The integration methods were assessed on their overall and class specific accuracy in an external validation, by cross tabulating the assessed LC map against the reference dataset in a confusion matrix (Strahler et al. 2006; Foody 2005). The integration methods are evaluated on their improvement compared to each other and the LC input maps. As addition to the external validation, this research calculates the information entropy over the integration methods. Entropy is an internal measure of uncertainty and represents the amount of information necessary to require certainty (Shannon and Weaver. 1949). Based on the information entropy, probability voting was identified as the best integration method. A difference plot between the integration methods confirmed that normal voting and weighted voting achieved similar results. Normal voting, weighted voting and probability voting had an improved overall agreement with the reference dataset, respectively 70.85%, 71.72% and 71.40%, compared to the LC input maps. The improvement on class specific accuracies varied as LC-CCI (2010) often held the highest agreement metrics for LC classes. This can be explained; voting methods favor classes that have good probability or a high weight in the integration; therefore common classes are over-mapped and rare LC classes could have been under-mapped. The probability voting held the highest agreement metrics for LC classes among the voting methods.

Acknowledgements

I would like to thank a lot of people who made this research possible. First of all my supervisors; Nandika Tsendbazar and Sytze de Bruin for their helpful remarks, support and patience. I enjoyed learning from them about land cover maps, integration, R and critically evaluating validation results. I would like to thank my family for their support, enthusiasm for the research and understanding. Especially my partner Job van Setten and mother Mieke Zaremba. Also, I would like to thank my good friend Karel Epke for reading and discussing the research. Furthermore, I would like to thank the reader for his or her interest in my thesis. I hope you find the information provided in the research useful.

Keywords:

Land cover, map integration, class weight, class probability, disagreements land cover maps, validation, confusion matrices, agreement metrics and information entropy

Table of contents

List of Figures:

List of Tables:

1. Introduction

Global land cover (GLC) products often exhibit spatial disagreement and overall accuracy needs to be improved. Improved global land cover (GLC) maps can be achieved by integrating different land cover (LC) datasets (Herold et al. 2008). Some integration methods consider the agreements between two or more LC products (Jung et al. 2006), while others use map accuracy information to integrate LC products (Tuanmu and Jetz 2014). The purpose of these integration methods is to use available LC information to label the most accurate LC class.

This section will introduce the research context and background, problem description, and motivation of this research.

1.1. Context and Background

LC maps serve as critical baseline information for a wide variety of research purposes and environmental applications (Herold et al. 2008). These maps have been produced from remote sensing images and they characterize the world surface into different LC types. Many models use LC datasets as one of their data inputs, examples are: climate models, dynamic vegetation models, hydrological models and carbon stock models (Hibbard et al. 2010; Jung et al. 2006; Verburg et al. 2011). Modelling communities acknowledge the importance of an accurate representation of land use and land-cover change to understand and quantify the interactions and feedbacks with the climate and socio-economic systems (Hibbard et al. 2010). Errors in GLC datasets add to the modelling uncertainty (Nakaegawa et al. 2011). The selection of a specific LC dataset and its quality has an influence on the outcome of the respective model (Hibbard et al. 2010; Mora et al. 2014; Nakaegawa et al. 2011).

The accuracy of existing GLC maps is around 70%, varying from 67 to 81% (Mora et al. 2014; See et al. 2014), despite the significant developments in technology and methodology of GLC mapping (Fritz et al. 2011; Herold et al. 2011; Tsendbazar et al. 2015). Mora et al. (2014) mentions that there is a clear need to improve the current quality of GLC maps, with the example that datasets should have a maximum of 5-15% error as a target to be further used in modelling applications (Herold et al. 2011). There are also significant amounts of spatial disagreement across different LC maps, in particular in the cropland and forest domains (See et al. 2014). User communities have specific requirements, but generally require a higher spatial and thematic accuracy, interoperability and inter-comparability from LC maps for their applications (Herold et al. 2011; Mora et al. 2014; Tsendbazar et al. 2015; Verburg et al. 2011). There are multiple GLC maps and it is not readily apparent which is most useful for a particular application reflecting the user requirements or how to combine the different maps to provide an improved dataset (Herold et al. 2008). Map producers aim to improve the uncertainty components of GLC maps (Tsendbazar et al. 2015)

This research will focus on fusion of LC maps from the perspective of LC map producers, with the aim to improve the (overall) accuracy of LC maps. In order to make the research more feasible within the allocated time, this research focuses on Western European countries, namely: The United Kingdom, Ireland, The Netherlands, Belgium, France, Spain and Portugal as study area. This might enclose mapping matters in this research; however, findings are not limited to use within these countries. To date, there are several GLC maps, reference data sets, merges of existing products and LC hybrid maps. This is undertaken due to the significant spatial disagreement between LC products.

Voting is a procedure in which a pixel is assigned to the LC class that occurs in the majority of the LC datasets at the pixel's location. There are integration methods that use voting to assign a pixel to a LC class, or use voting combined with a other approach (Ge et al. 2014; Goovaerts. 1999; Iwao et al. 2011; Jung et al. 2006; Kinoshita et al. 2014; Tuanmu and Jetz. 2014). Jung et al. (2006) produced a SYNMAP merged existing LC products into a desired classification legend and exploring synergies of GLC products for carbon cycle modeling. Ge et al. (2014) proposed a fusion method to combine biomass maps linearly with weighted averaging and area bias corrections. Kinoshita et al. (2014) focused on an integration method to create a GLC map and a probability map, using probability of occurrence scores to perform a logistic regression analyses. Iwao et al. (2011) based their integration method on a voting approach, were a pixel is assigned to the LC class that occurs in the majority of the LC maps used for integration at that pixel's location. Iwao et al. (2011) uses map accuracy to decide on a LC class where maps disagree. Tuanmu and Jetz (2014) used a two-step integration approach to capture the heterogeneity of the finer LC maps and applied accuracy based weighting on class probabilities where the input maps disagreed. Previous work in the literature presents several geo-statistical approaches to integrate different LC products (Carneiro and Pereira 2012; Ge et al. 2014; See et al. 2014). See et al. (2014) introduced two hybrid LC map with crowdsourcing data and geographically weighted regression. Geographically weighted regression (GWR) weights observations and locally determines relationships between variables (Brunsdon et al. 1998; Fotheringham et al. 2003). Crowdsourcing data could improve the accuracy of LC maps where different LC products disagree (Comber et al. 2013). Carneiro and Pereira (2012) aimed to develop methodologies for mapping the spatial distribution of errors using indicator kriging.

The use of differing methodological approaches (e.g., classification scheme, data source and algorithms) for GLC mapping raises consistency issues and makes comparison difficult (Mora et al. 2014). Simultaneously, inconsistencies between LC products make integration of these products difficult. LC products have limitations due to product inconsistency (Tuanmu and Jetz 2014) and existing differences in LC legends are an obvious inconsistency that hinders the comparison of LC maps (Herold et al. 2008). This raises the issue of data harmonization before comparing or integrating LC products. Harmonization is done in order to be able to directly compare LC products. In addition to the differences in legends, there are spatial aspects that cause inconsistencies between LC datasets. For example, a different resolution between maps also makes integrating the available LC datasets difficult. This research is limited in the thematic and spatial harmonization. LC classes will be harmonized to a thematic legend with eight general LC classes. Spatial harmonization is based on the assumption that the sample unit area has homogenous LC type, so the datasets can be harmonized to have the same extent of sample units.

Inconsistencies between LC maps are not evenly distributed among all LC classes (Carneiro and Pereira 2012) and accuracy may vary locally within the map (Strahler et al. 2006; Foody 2005). Problematic classes are difficult to discriminate from other LC classes (Carneiro and Pereira 2012; Herold et al. 2008). Fragmented landscape, heterogeneous and transition areas generally have lower map accuracy (Carneiro and Pereira 2012; Herold et al. 2008; Jung et al. 2006; Tsendbazar et al. 2015). Class accuracies of problematic classes could be improved through integration of different LC data sets. This would also address the requirements of map producers because the approach improves problematic classes. LC map producers will benefit from an improved LC map with a higher overall accuracy and class accuracies, to better characterize GLC and problematic LC types. More accurate GLC map will also be beneficial to the users of GLC maps and provide a better basis for their applications.

1.2. Problem description

GLC map producers have stated that GLC maps only reached an overall accuracy around 70% despite the significant developments in technology and methodology in addition to the user requirements for improved accuracy (Fritz et al. 2011; Herold et al. 2011; Tsendbazar et al. 2015; Tsendbazar et al. 2016). There is a clear need to improve the current quality of LC maps from all perspectives and improved LC maps can be achieved by integrating different LC datasets. However, there are different integration methods and it is not known which is the most accurate integration method. Additionally, as the requirements on LC maps vary for different users, related research is often directed at a specific user community or problem.

Inconsistencies between GLC maps, caused by different methodologies, legends, source data, hinder comparison and integration between LC datasets (Mora et al. 2014). This will make integration even more difficult and raises the issue of data harmonization before comparing different integration methods (Herold et al. 2008). An integration method decides on a LC class for each pixel, based on LC labels of the input products. Integration aims to label LC information to the correct LC class, but is dependent on the LC information from input products. Inconsistencies among different GLC datasets are often attributed to landscape heterogeneity (Jung et al. 2006; Tsendbazar et al. 2015). Heterogeneous areas and transition areas of the main biomes are challenging to classify. These areas have fragmented landscape with different LC types and it is difficult to classify mixed LC information to a specific LC class (Carneiro and Pereira. 2012). Due to this difficulty, LC maps often disagree with one another in these regions. Carneiro and Pereira (2012) mention that inconsistencies are not evenly distributed among all LC classes and that some classes are easier to discriminate than other classes. These classes with low accuracies are known as problematic classes and difficult to separate spectrally. Herold et al. (2008) discusses that it is difficult to discriminate between classes in the cropland and forest domains, naming mixed trees, shrubs, cultivated and managed vegetation and barren LC types as examples of problematic classes. Problematic classes are often prominent in heterogeneous and transition areas (Herold et al. 2008). Class accuracies could be improved by dealing with problematic classes through integration and thereby also improved overall accuracy could be achieved.

This research aims for improved LC maps and therefore focuses on different integration methods that include: harmonization of LC classes, improved overall and class accuracy within the community of LC map producers.

1.3. Research objective and questions

This research focuses on LC map integration from the perspective of LC map producers and aims to improve (overall) accuracy of LC maps. Integration will be assessed with respect to the ability of improving overall and class specific accuracies. This research is limited to integration methods based on voting approaches, which are very commonly used for map integration.

Research objective:

The objective of this research is to generate improved LC products by integrating available LC products and reference datasets.

Research questions:

To achieve this research objective, the following research questions will be answered:

I. Can the selected LC map integration methods be applied to the study area considering data constraints and characteristics (e.g. inconsistent legends)?
II. How can LC datasets be integrated with the chosen integration methods and selected software?
III. Which is the most promising method based on internal validation?
IV. What is the agreement of the integrated LC maps with the reference dataset and how much has integration improved overall accuracy?

1.4. Outline

Chapter two describes the study area, study materials and available LC datasets and maps that were used for integration. Chapter three, the theoretical background describes the literature review for the integration methods, harmonization and validation. The fourth chapter reports on the methodology used in this research for the harmonization, integration methods and chapter five report the research results. Chapters six and seven discuss the results and conclude the research, respectively.

2. Study area, land cover data and materials

2.1. Study area

To make the research feasible within the allocated time, it focused on the following countries: United Kingdom, Ireland, The Netherlands, Belgium, France, Spain and Portugal as study area. The data such as LC maps and reference datasets were processed to match the study area. The analysis was done in the WGS84 geographic coordinate system and datum. Figure 1 presents the study area.

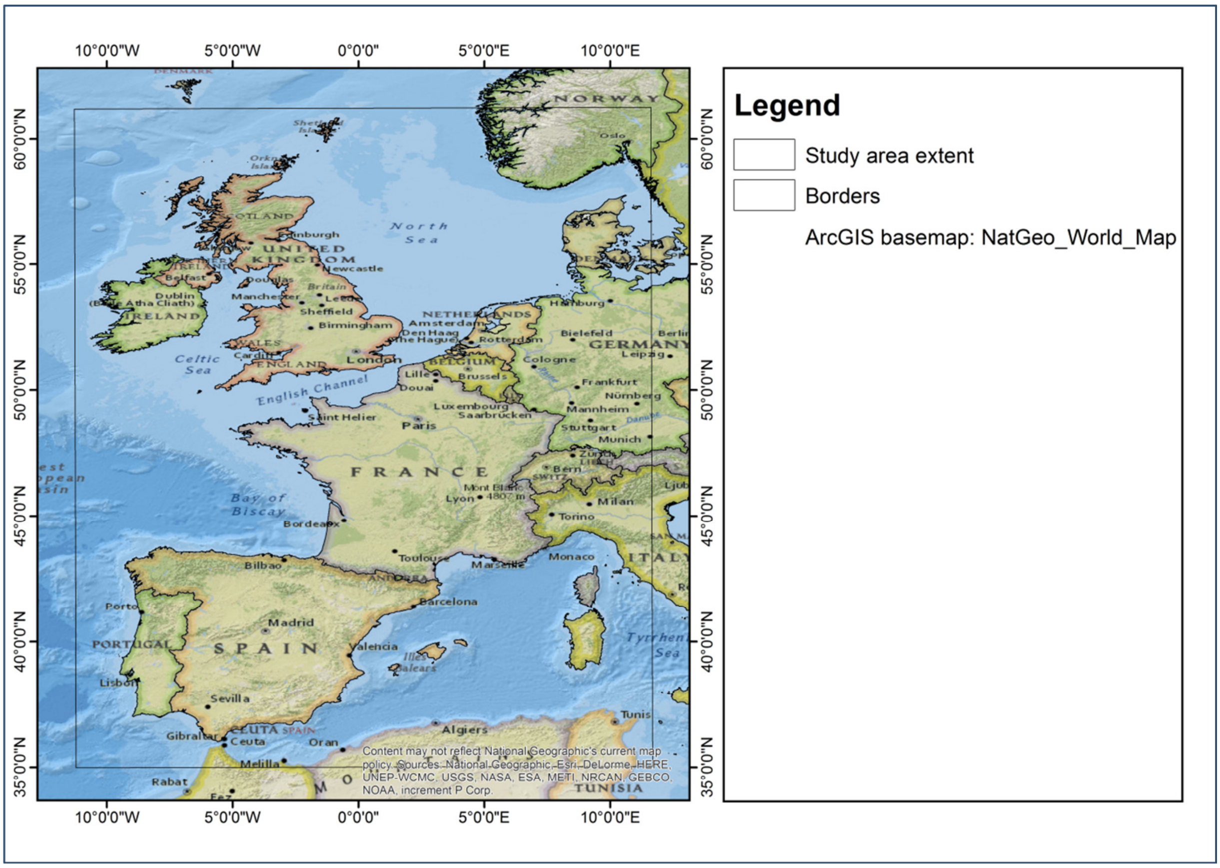

Figure 1: Study area

2.1.1. Area information

Western Europe has an ocean climate. The dominant LC are trees, shrubs, grass and croplands. Shrubs and sparse vegetation LC are more prominent in the southern part of Western Europe due to warmer temperatures. Norway's mountains area consisting of valleys and fjords, the French Alps and the Pyrenees on the border of France and Spain are areas with steep slopes, heterogeneous areas and fragmented landscapes.

Table 1 provides some meta data considering the target map of the study area. It should be noted that this information is acquired after harmonization is complete as the harmonized cell size influences these numbers. The decision for the chosen resolution is discussed in the methodology (section 4.1).

Table 1: Meta data target LC map

Reference system	Coordinate Reference system	WGS 1984
	Datum	WGS 1984
Resolution	Meters	300 (approximately)
	Decimal degrees	0.002778
Extent	x-minimum	-11.2936
	x-maximum	11.68418
	y-minimum	34.96944
	y-maximum	61.24722
Dimension	Rows	9460
	Columns	8272
	Cells	78253120

2.2. Land cover data

This research uses LC datasets for two purposes: (1) producing integrated LC maps and (2) assessing the performance of the integration methods by validating the produced LC maps. Four LC maps were used as input for the three chosen integration methods implemented in section 4.2. The LC reference data is used for the external validation of the produced LC maps in section 4.3.2.

2.2.1. Maps

FROM-GLC, Globcover 2009, LC-CCI (2010 and MODIS5 (2010) LC maps were used for integration. Table 2 and table 3 list all four LC maps and their metadata. Globcover 2009, LC-CCI (201) and MODIS5 (2010) were obtained from Wageningen University. FROM-GLC was downloaded from internet (Finer Resolution Observation and Monitoring Global Land Cover. 2015). The GLC maps were pre-processed in two steps: (1) the GLC maps were re-projected to the same geographic coordinate system and datum and (2) all GLC maps were clipped to the extent of the study area.

From the input maps, Globcover 2009, LC-CCI and MODIS5 had a WGS84 geographic coordinate system and datum. MODIS5 was re-projected to WGS84 by Tsendbazar et al. (2016) when acquired from Wageningen University. In QGIS all hierarchy tiles of the FROM-GLC hierarchy were merged to one raster dataset, re-projected to WGS84 and directly re-sampled to a 300 meter resolution. In the second step all maps were clipped to create a spatial subset of the raster within the extent of the study area with the clip (data management) tool in ArcGIS. For operations, like integration in R, it is required that all maps hold the same spatial extent and dimension (number of rows, columns and cells), therefore the environmental settings of snap raster were used.

FROM-GLC hierarchy 250 m is one of the FFROM-GLC family maps, which were produced using LandsatTM/ETM and MODIS EVI time series (Yu et al. 2014). The producers of FROM-GLC aimed for a LC data set for different user applications and therefore created FROM-GLC-hierarchy. The FROM-GLC-hierarchy was produced by an up-scaling / aggregation approach where a class type at coarser resolution is assigned based on the class type at a finer resolution (Yu et al. 2014). There are eight coarser resolutions in the hierarchy dataset, including the 250 meter used in this research, where the 30 meter map is the base map. The FROM-GLC-hierarchy 30 meter is produced from FROM-GLCagg, with a few improvements made on cell level at locations where there were misclassifications (Yu et al. 2014). These improvements were made by aggregating LC information from coarser resolution data as MODIS, Globcover 2009 and data to check among others global water and shorelines (Yu et al. 2014). Misclassifications that were improved were: (1) no data and cloud pixels that were replaced, (2) confusion between water bodies and shadows that were processed, (3) Bareland overestimation was reduced and (4) LC type confusion in shore areas was filtered (Yu et al. 2014). FROM-GLC-hierarchy is produced from FROM-GLCagg, which is based on FROM-GLC and FROM-GLCseg, for more information is referred to the research of Yu et al. (2014).

Globcover 2009 GLC map is produced from the automated classification of MERIS FR time series (Bontemps. 2011). Globcover 2009 has a LCCS classification with a legend of 22 LC classes, (Bontemps. 2011). The Globcover 2009 GLC map has a 300m resolution and is projected in Plate-Carrée (WGS84) (Bontemps. 2011). The reference database for the Globcover 2005 was used for validating Globcover 2009 LC map. LC-CCI (2010) and was produced from multi-year and multi-sensor strategy (Climate Change Initiative Land Cover project. 2015). The MERIS Full (2003-2012) and Reduced Resolution (FR and RR) archive were used as input to generate a 2003-2012 global land cover map, from which the 2010, 2005 and 2000 maps were produced. LC-CCI has a LCCS classification scheme with 22 LC classes. MODIS5 (2010) was- produced from: monthly EVI, LST, and 7 bands from 8 day composites (Friedl et al. 2010). MODIS5 has different LC classification system including IGBP (Mora et al. 2014) and has a legend with 17 classes.

Table 2: Meta data LC maps, production, classification and validation

Production		
LC map	Input data	Time of data collection
FROM-GLC-hierarchy (2013)	Landsat Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+). Improvements by aggregating other LC data sources.	2010
Globcover v2 (2009)	MERIS: Bi-monthly from 10 day composites	2009
LC-CCI (2010)	Unknown, made by Land Cover CCI project	Three maps for; 1998-2002, 2003-2007 and 2008-2012 epochs.
MODIS5 (2010)	MODIS: Monthly EVI, LST, and 7 bands from 8 day composites	2001-2008
Classification		
LC map	Classification method	Classification scheme
FROM-GLC-hierarchy (2013)	-	Two level classification system, with 10 classes
Globcover v2 (2009)	(Un)supervised spatio-temporal clustering	LCCS 22 class
LC-CCI (2010)	`-	-
MODIS5 (2010)	Supervised decision tree boosting	5 different LC classification system including IGBP, UMD
Validation		
LC map	Validation data	Absolute positional accuracy RMSE
FROM-GLC-hierarchy (2013)	38664 test samples collected from Landsat images, MODIS time series, high resolution images, field photos Google Earth	
Globcover v2 (2009)	Independent validation dataset from CHR satellite data and other datasets	77m
LC-CCI (2010)	-	
MODIS5 (2010)	Using HR satellite	50-100m

Source: Mora et al. 2014, Table 2.1; Climate Change Initiative Land Cover project. 2014; Yu et al. 2014

Table 3: Meta data LC maps, spatial reference and resolution

Meta data	LC maps			
	FROM-GLC hierarchy (2013)	Globcover v2 (2009)	LC-CCI (2010)	MODIS5 (2010)
Spatial reference				
Coordinate Reference System	Sinusoidal Datum: Sphere with radius 6371007.181 m.	WGS_1984	WGS_1984	Sinusoidal Datum: Sphere with radius 6371007.181 m
Datum		WGS 1984	WGS 1984	
Spatial resolution				
Meter in resolution at equator	250	300	300	500
Decimal degree	0.002661	0.002778	0.002778	0.004167

Source: Sinusoidal projection (2015); Tsendbazar et al (2016); Yu et al. 2014

2.2.2. Datasets

GLC2000, GLCNMO-tr, Geo-Wiki, Globcover 2005,MODIS-tr and VIIRS 3 are the reference datasets used for validating the produced LC maps. Figure 2 presents the distribution of the reference datasets over the study area. All reference datasets were obtained from Wageningen University (GOFC-GOLD reference data portal. 2015). The reference datasets were pre-processed with two steps: (1) the reference datasets re-projected to the right geographic coordinate system and datum and (2) all datasets were clipped to the extent of the study area.

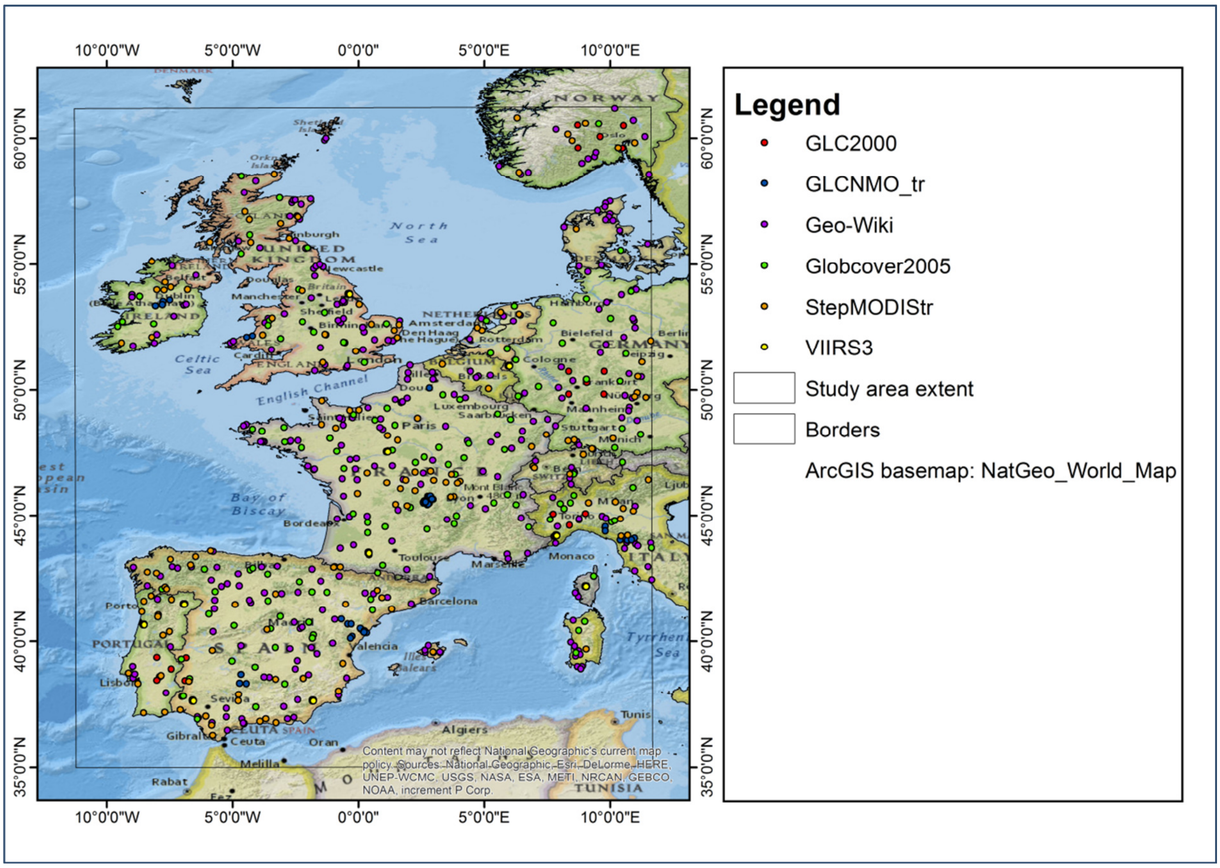

Figure 2: Reference datasets

Table 4 holds the distribution of sample sites over the individual reference datasets and table 5 lists the metadata information of the reference datasets (GOFC-GOLD reference data portal. 2015; Tsendbazar et al. 2015)

Table 4: Sample sites of reference datasets

Reference dataset	Number of Sample sites
GLC2000	19
GLCNMO-tr	50
Geo-Wiki	304
Globcover 2005	123
StepModus-tr	138
VIIRS 3	282
Total	916

The GLC2000 and Globcover 2005 reference dataset have a LCCS classification system. The GLC2000 reference dataset is the result of consolidation work, were data from the original GLC2000 dataset was randomly selected and reinterpreted using landsat scenes (GOFC-GOLD reference data portal 2015). Only samples that were identified as problematic were reinterpreted, this results in a final dataset with 1253 samples sites (GOFC-GOLD reference data portal 2015). GLC2000 has a block unit type sample design and 19 samples are within the study area and used for this research. GLC2000 reference dataset has a LCCS classification with 13 LC classes. The Globcover 2005 dataset is the result of consolidation work, were regional and national experts reinterpreted he original ESA-GlobCover 2005 dataset using Google Earth imagery (GOFC-GOLD reference data portal 2015). Globcover 2005 has a block unit type sample design and 123 samples are within the extent of the study area. Globcover 2005 has a LCCS classification system with 22 LC classes. GLCNMO-tr has a polygon unit type sample design and 50 samples are within the study area.

The StepMODIS-tr and Visible Infrared Imaging Radiometer Suite (VIIRS) reference dataset have an IGBP classification with 17 LC classes. 138 sample sites of stepMODIS-tr and 282 sample sites of VIIRS are within the extent of the study area (GOFC-GOLD reference data portal 2015). The STEP reference dataset is a model that derives LC parameters from remote sensing, collateral and field plot data that can be extracted with GIS to produce a GLC classification (GOFC-GOLD reference data portal 2015). StepMODIS-tr reference dataset has a polygon unit type sample design. VIIRS has an IGBP classification with 17 LC classes and clustered blocks of 5x5 km as its unit type sample design. VIIRS reference dataset is based on stratified random samples of 500 blocks (GOFC-GOLD reference data portal 2015), these clustered blocks are the unit type sample design. Geo-wiki reference dataset has a LCCS/IGBP classification with 10 LC classes and 304 sample sites within the study area. The Geo-wiki reference dataset is a Volunteered Geographical Information (VGI) reference dataset. Geo-wiki collects data samples from volunteers (which can be experts) in a web interface and asks these volunteers to interpret the reference LC from predefined sample locations (Comber et al. 2013; See et al. 2014). Comber et al. (2013) determined the reliability of volunteered geographical information (VGI) and shows that VGI can be used in LC mapping if it is reliable.

Table 5: Meta data reference datasets

Spatial reference			
Dataset	Coordinate Reference System	Datum	Projection
GLC2000	WGS 1984	WGS 1984	
GLCNMO-tr	WGS 1984	WGS 1984	Plate_Carree
Geo-Wiki	WGS 1984	WGS 1984	
Globcover 2005	WGS 1984	WGS 1984	
MODIS-tr	WGS 1984	WGS 1984	
VIIRS 3	WGS 1984	WGS 1984	
Production			
Dataset	Acquisition date	Classified by	Source
GLC2000	1999-2002	National/regional expert	Landsat TM, Aerial photographs, thematic maps, NDVI profile
GLCNMO-tr	2000-2003 for landsat and MODIS	National/regional expert	Google Earth image/photo, DCP photo, regional LC maps
Geo-Wiki	2000-2012	Volunteer	Google earth, Spot-NDVI, DCP, geo-tagged pictures
Globcover 2005	Circa 2005	International expert	SPOT VGT-NDVI profile, Google Earth
MODIS-tr	2001-2007	National/regional expert	Landsat 7 or higher resolution data, Google Earth
VIIRS 3			
Legend			
Dataset	Classification scheme	Number of classes	Hierarchical Classifier provided
GLC2000	LCCS	13	+
GLCNMO-tr	ST-LCCS	14	-
Geo-Wiki	LCCS/IGBP	10	-
Globcover 2005	LCCS	22	+
MODIS-tr	IGBP	17	+
VIIRS 3	IGBP	17	
Sample characteristics			
Dataset	Site	Unit type	Unit size
GLC2000	1265	Block	3x3km
GLCNMO-tr	1607	Polygon	>3x3km
Geo-Wiki	5608	Pixel	1 pixel for MODIS, Glob-Cover, GLC2000
Globcover 2005	4258; 3167 certain	Block	1.5x1.5km
MODIS-tr	1860	Polygon	1-376 pixel (0.5 km)
VIIRS 3	4500	Clustered blocks	1 km

Source: GOFC-GOLD reference data portal. 2015; Tsendbazar et al (2015), Table A2.

2.2.3. Confusion matrices of Global Land Cover maps

Accuracy information of the input maps was used for integration; this information was gained from the research of Tsendbazar et al. (2016) for the Globcover 2009, LC-CCI (2010) and MODIS5 (2010) input maps. For the input map, FROM-GLC-hierarchy, the confusion matrix of FROM-GLC agg 30 meter was used as accuracy information (Yu et al. 2014).

2.3. Software

The software used in this research includes ArcGIS version 10.2.2 (Esri ArcGIS software. 2015). ArcGIS is a Geographical Information System (GIS) used for spatial analysis and visualization. ArcGIS was mainly used for; harmonization, displaying LC data, results and finally the integrated LC maps. In the harmonization (section 4.1) ArcGIS was used to reclassify the LC maps and datasets. In the external validation ArcGIS was used to prepare (creating 3x3pixel blocks) the reference dataset to extract the LC class from the integrated maps which was used for the external validation (4.3.2). GGIS is free GIS software. In QGIS all tiles of the FROM-GLC-hierarchy were merged to one raster dataset, re-projected to WGS84 and re-sampled to 300m resolution. This was done in QGIS instead of ArcGIS because ArcGIS did not recognize the projection; during re-projection FROM-GLC was also resampled to a 300m resolution.

R is a programming language and environment for statistical computing (The R project for statistical computing. 2015). R was used to implement the integration methods, internal and external validation. Section 8.1 in the appendix holds all scripts that were used in R. In addition to R, Excel was used for harmonizing the legends of the LC maps and datasets to one legend and for assessing confusion matrices. The new harmonized legend was constructed in excel and implemented to the LC data in ArcGIS (section 4.1).

3. Theoretical background

3.1. Global land cover mapping

The current accuracy of GLC maps is around 70 % (Mora et al. 2014; See et al. 2014). Map producers and users feel a need to improve the quality of GLC maps as errors add to the uncertainty of GLC applications (Herold et al. 2011; Mora et al. 2014; Tsendbazar et al. 2015; Verburg et al. 2011).

Improved GLC maps can be achieved by integrating different LC datasets (Herold et al. 2008). The current GLC maps often do not agree with each other in certain areas and this results in spatial disagreements or inconsistencies between them. Inconsistencies between GLC maps come from many origins. Mora et al. (2014) mentions, that the use of differing methodological approaches (e.g., classification scheme, data source and algorithms) to produce GLC maps raises consistency issues. Herold & Schmullius (2004) mention inconsistencies that result from different standards being used to derive LC data, differences in: data model (vector/raster), cartographic standards, spatial aspects as spatial reference system, resolution, sample unit-type, unit-size and thematic (semantic) aspects as classification algorithms and LC legends. Jung et al. (2006) and Tsendbazar et al. (2015) mention that inconsistencies among different GLC datasets are often attributed to landscape heterogeneity. Due to fragmented landscapes, heterogeneous areas typically have many mixed cells and are therefore challenging to be classified. Mixed cells of LC information are difficult to label to a specific LC class (Carneiro and Pereira. 2012). Carneiro and Pereira. (2012) mention that transition areas have high inconsistency between LC products due to different classification algorithms in legends. Another problem in map inconsistencies are problematic classes, these classes are difficult to separate spectrally and therefore hold a lower accuracy (Carneiro and Pereira. 2012). It is difficult to discriminate between classes in the cropland and forest domains, Herold et al. (2008) identifies trees, shrubs, cultivated and managed vegetation and barren LC types as problematic classes. Tsendbazar et al. (2016) characterizes shrubs, grass and cropland classes, as LC classes with high confusion errors. Harmonization (section 4.1) needs to be done before being able to integrate GLC maps (Herold et al. 2008; Tsendbazar et al. 2016). Data harmonization aims to harmonize LC data to one standard (Herold et al. 2008). This makes it possible to compare and integrate the LC data in the same spatial dimension and thematic legend within this research (section 4.2). Harmonization has its limitations and will not ideally eliminate all spatial and thematic product inconsistencies. Harmonization is described in section 3.4 of this research.

An integration method decides on a LC class for each pixel in the new map based on LC input maps. Integration aims to label LC information to the most accurate LC class, but is dependent on the LC information from input maps. There are integration methods that consider the agreements between LC products (Jung et al. 2006), while others use map accuracy information (Iwao et al. 2011). Kinoshita et al. (2014) and Tuanmu and Jetz (2014) use probabilities in their integration approach which are based on map accuracies/agreements with the reference data. Kinoshita et al. (2014) mention that there are numerous studies that compare maps, but attempts at applying these result of for the creation of new integrated maps have been very limited. Section 3.2 of this chapter discusses integration methods with a voting approach in detail. Integration methods with a statistical approach, that take into account the location of LC classes, show promising results in accuracy improvements (Kinoshita et al. 2014; See et al. 2014). Geographically weighted regression (GWR) can be used for map integration and it weights observations and locally determines relationships between variables (Fotheringham et al. 2003). See et al. (2014) introduced two hybrid LC maps with crowdsourcing data and geographically weighted regression. In the approach, LC datasets were trained using volunteered geographical information (VGI) and a probability map was produced from each land cover dataset (See et al. 2014). A GWR was used to determine the best land cover product at each location by calculating the relationship between probability maps (See et al. 2014). A land cover hybrid map is produced from the probability maps, the LC class of a map with the highest probability of being the correct class was assigned to a pixel (See et al. 2014). The second hybrid map was better compared to GLC2000 and Globcover, but worse or similar in performance to MODIS map (See et al. 2014). Geostatistical approach can also be used for map indicator. For example, Indicator kriging includes indicators in the land cover datasets by transforming class occurrences to indicators: 1 if present, 0 if absent (Goovaerts. 1999). This approach could label indicators to each LC class present in the legend, maps would be made from each LC class indicator (Goovaerts. 1999). This approach could interpolate the indicators (co-variates or "soft" data) and reference data ("hard" data) with kriging and interpreted as probabilities. Section 3.3 discusses how the normal voting, weighted voting and probability voting methods are chosen for this research, considering the data constraints, characteristics and the study area.

The accuracy of a LC map is assessed through a map validation process. Strahler et al. (2006) describes validation as a term used for techniques that determine the quality of a particular map. It is important to assess overall map accuracy and specific class accuracies, but it should be recognized that accuracies varies locally within the map (Strahler et al. 2006; Foody 2005). Fragmented landscape, heterogeneous, transition areas, etc, are usually mapped with low accuracy and problematic classes have low class accuracies (Carneiro and Pereira 2012; Herold et al. 2008; Jung et al. 2006; Tsendbazar et al. 2015). Herold et al. (2006) mentioned that harmonization and validation complement and profit from each other in the overall quality of land cover products and boost interoperability and comparability. The theoretical background on validation is described in section 3.6.

3.2. LC map integration methods

Integration decides on a LC class for each pixel in the new map based on LC labels of the input products. This section studies different methods from literature on how LC datasets can be integrated to an improved LC product and which of these methods would be applicable to this research. The chosen integration methods are normal voting, weighted voting and probability voting.

3.2.1. Voting

Voting is a procedure in which a pixel is assigned to the LC class that occurs in the majority of the LC datasets at the pixel's location. Assignment is problematic or ambivalent in case of ties.

The method of Jung et al. (2006) is an example of using a voting approach for an integration map. Jung et al. (2006) merged existing LC products into a desired classification legend and exploited synergies of GLC products for carbon cycle modeling. The method of Jung et al. (2006) produced a map with a legend based on AVHRR, MODIS and vegetation satellite sensors using a fuzzy logic approach, legend harmonization and voting.

First the method defines a desired classification legend and secondly it links the defined legend with the legends of the original maps by assigning affinity scores between them (Jung et al. 2006). The LC class with the highest score is assigned to the map, which in principle can be seen as a voting procedure. Evaluation indicated a successful exploration of synergies between products and is therefore believed to be more accurate than existing products; however there is insufficient reference data to validate SYNMAP (Jung et al. 2006).

Gopal et al. (1999) used a voting approach to assign confidence estimates to conflicting predictions in classifications of GLC. In this voting strategy, volunteers voted on LC classes which resulted in that the predicted class received the largest number of votes (Gopal et al. 1999). Gopal et al. (1999) mentioned that a voting strategy improves classification accuracy and provides a way to evaluate map uncertainty, as voting results are contradictory in pixels with mixed LC classes.

3.2.2. Weighted voting

In weighted voting, a pixel is assigned to the class that accumulates the highest weight at the pixel's location.

Ge et al. (2014) proposed a fusion method to increase the accuracy of regional biomass estimates by using higher-quality calibration data and weighted averaging. Ge et al. (2014) combined biomass maps linearly using weights derived from the variance-covariance matrix associated with the accuracies of the source map and area bias corrections. This method uses minimization of the variance of the prediction error, where each map has an associated variance (Ge et al. 2014).

Iwao et al. (2011) presented an approach to create a new GLC map with map integration. Three existing GLC maps were combined (MOD12, GLC2000 and UMD) in six classes (forest, croplands, grasslands, wetlands, settlements and other land) of the Land Use, Land Use Change and Forestry classification scheme (LULUCF) (Iwao et al. 2011). The new map adopts the classification favoured by the majority of the input maps, where MOD12 with the highest accuracy was only replaced when GLC2000 and UMD agreed (Iwao et al. 2011). In principle this can be seen as weighted voting as a preference was given to all LC classes from MOD12, because MOD12 has a higher overall accuracy (OA). The new integration map, MOD12, GLC2000, UMD, each were validated with rates of agreement in a confusion matrix with reference points from the Degree Confluence Project (DCP). MOD12, GLC2000 and UMD have an overall agreement of respectively 60.4%, 85.9% and 55.5% with the DCP data (Iwao et al. 2011). The Results show that the overall accuracy in the new map was improved to 61.3% overall agreement, but one or more of the input maps show a higher accuracy for the Grasslands (UMD), Croplands (GLC2000), Settlement classes (MOD12) and the Arid and Polar climate zone (Iwao et al. 2011).

3.2.3. Probability voting

In probability voting, the voting process is applied on the probabilities of each class being the correct class. A pixel is assigned to the class that accumulates the highest probability at the pixel's location.

Kinoshita et al. (2014) created a GLC map and probability map through an integration method that uses probability voting. Kinoshita et al. (2014) used a harmonized legend with six classes to integrate six GLC maps: MODIS Land Cover Map Collection 4 (MOD12C4), MODIS Land Cover Map Collection 5 (MOD12C5), Global Land Cover 2000 (GLC2000), Globcover, the University of Maryland 1-km Global Land Cover products (UMD) and Global Land Cover by National Mapping Organizations (GLCNMO) (Kinoshita et al. 2014). The integration method consists of two steps: (1) calculating the probability of occurrence and (2) a logistic regression on the probability of occurrence. Kinoshita et al. (2014) first calculated the probability of occurrence of the six LC types from the reference dataset in each of the six GLC input maps, which can be described as class probabilities. Equation 1 describes the probability of occurrence where M indicates the map, n indicates the category class in a map, m indicates the category class in the reference data and N are the pixels located in M, n, m of the matrices (Kinoshita et al. 2014).

$$P_{M,n,m} = \frac{N_{M,n,m}}{\sum_{m=1}^{6} N_{M,n,m}} \qquad [1]$$

Source: Kinoshita et al. (2014)

In the second step Kinoshita et al. (2014) used the probability of occurrence scores to perform a logistic regression analyses to calculate the probability of LC class for each pixel in the map. This produces a probability map, from which an integrated map was produced with the LC class that holds the highest class probability in that pixel of the map (Kinoshita et al. 2014). The external validation reported an overall accuracy of 74.6% (Kinoshita et al. 2014). Kinoshita et al. (2014) found that map accuracy increased with the number of input maps but that the number of used training sites did not significantly affect accuracy.

Tuanmu and Jetz (2014) used an integration method which uses probability voting to generate a GLC product for biodiversity and ecosystem modeling. Tuanmu and Jetz (2014) used a two-step integration approach and the generalized scheme developed by Herold et al. (2008) for harmonization. First areal proportions from Globcover (2005) and MODIS2005 were integrated into an intermediate dataset with 500 meter resolution. In case of disagreement, a confusion matrix was used to calculate class probabilities for each class pair on which an accuracy based weighting was applied. Tuanmu and Jetz (2014) mentioned that accuracy based weighting was only applied because adding information from a coarser product would unnecessarily homogenize the heterogeneity captured by the finer product. In the second step, the generated intermediate dataset of 500 meter was integrated with the areal proportions of DISCover and GLC2000 of 1 kilometer. Information from the intermediate 500 meter dataset was kept if DISCover or GLC2000 agreed with the highest presented LC class. Priority was given to the intermediate dataset because some of its disagreement may result from LC change in the DISCover and GLC2000older products (Tuanmu and Jetz. 2014). In case of disagreement, the LC class was again calculated from accuracy based weighting applied on class probabilities from the intermediate 500 meter dataset, DISCover and GLC2000. The weights for the intermediate dataset were twice as high because the dataset was calculated from two products (Tuanmu and Jetz. 2014).

Tuanmu and Jetz (2014) calculate dissimilarities between the integrated map, Globcover (2005), MODIS2005 DISCover, GLC2000 and the randomly selected validation data. There was less dissimilarity in the integrated map compared to the input maps, which suggest the new map had improved (Tuanmu and Jetz. 2014). Classes were compared on their sensitivity, precision and F-score presented in graphs (Tuanmu and Jetz. 2014). Where sensitivity is the ratio of correctly identified pixels to the total number of pixels in the validation data and precision is the ratio of correctly identified pixels in an evaluated product (user accuracy) (Tuanmu and Jetz. 2014).

3.3. Applicability of integration methods in research

Normal voting, weighted voting and probability voting are the chosen integration methods in this research. This section discusses the applicability of integration methods within the data constraints of this research.

3.3.1. Study area and data constraints

The study area extent covers the United Kingdom, Ireland, The Netherlands, Belgium, France, Spain and Portugal presented in figure 1, section 2.1. FROM-GLC, Globcover (2009), LC-CCI (2010) and MODIS5 (2010) LC maps used as input maps for the integration methods hold LC information within the study area extent. Other areas from the GLC maps were excluded from this research. The LC input maps hold no constraints for the applicability of the integration methods but enclose mapping results within the study area. GLC2000, GLCNMO-tr, Geo-Wiki, Globcover 2005, MODIS-tr and VIIRS 3 are the reference datasets used for validating the produced LC maps. The reference datasets hold samples within the extent of the study area. Section 5.1 holds the results from harmonizing the dataset to the harmonized legend and one reference dataset. Table 6 holds the distribution of samples over the harmonized LC classes. The reference datasets hold constraints for the applicability of integration methods because LC classes (5) wetlands and (8) barren were not represented by enough samples, respectively four and nine reference points.

Table 6: Distribution of reference dataset over LC classes

LC Class	Description	Samples
1	Trees	224
2	Shrubs	65
3	Herbaceous vegetation	130
4	Cultivated and managed vegetation / agriculture (incl. mixtures)	426
5	Wetlands	4
6	Urban/built up	45
7	Water, permanent Snow and Ice	13
8	Barren	9
1 - 8	All classes	916

Weighted voting and probability voting however do require accuracy information to assign weights or calculate class probability; this information is often gained from a confusion matrix produced from reference data. The accuracy assessments of Tsendbazar et al. (2016) and Yu et al. (2014) were used and not an accuracy assessment produced from the reference data (section 2.2.3.). The accuracy assessment of FROM-GLC (Yu et al. 2014) differs from the accuracy assessments of the other input maps (Tsendbazar et al. 2016). The confusion matrix of FROM-GLC was actually produced from the stage of FROM-GLCagg and not FROM-GLC hierarchy and does not possess information on LC class (50) wetlands.

3.3.2. Selected methods for integration

Integration methods with a voting approach were the chosen integration methods in this research. The LC maps produced from normal voting, weighted voting and probability voting were evaluated by comparing the maps and an internal and external validation of the LC maps. This research evaluates which voting method achieves the highest results. Methods with a geo-statistical approach in their integration are promising and were considered but dropped because the harmonized LC classes (5) wetlands and (8) barren were not represented by enough samples. It was questionable if the reference data had enough reference points to produce a reliable result when used in an integration method that uses reference/training data at location. Additionally if such a method was chosen the reference dataset had to be divided in a training and validation dataset, which would have resulted in LC classes (5) wetlands and (8) barren having even less samples. Dividing the reference data into a training and validation dataset would have been required because a validation would reproduce measured training data when the same reference data is used for training and validation purposes, which would imply perfect predictions. Two of chosen methods require accuracy information. To keep all reference data samples for the validation, confusion matrices from the research of Tsendbazar et al. (2016) and Yu et al. (2014) were used instead of confusion matrices produced from the reference dataset. Weighted voting and probability voting were produced with weights and probabilities obtained from the research of Tsendbazar et al. (2016) and Yu et al. (2014). Normal voting, weighted voting and probability voting are implemented in the methodology (section 4.2) and their results are presented in section 5.2.

3.4. Harmonization

Harmonization of the available LC datasets and maps is necessary before comparing or integrating the datasets Harmonization between LC datasets can be understood as a process where similarities are emphasized and inconsistencies reduced (Herold et al. 2006). Harmonization should focus on making the datasets comparable and in case of inconsistencies, develop an understanding why and where LC products are not perfectly comparable.

The harmonization process searches for 'harmony' between datasets, a balance which involves inconsistencies opposite standardization (Herold et al. 2006). Harmonization is a 'bottom up' process and will not eliminate all differences, but should eliminate major inconsistencies between datasets (Herold et al. (2008). Standardization on the other hand, is a 'top down' process that eliminates all inconsistencies between datasets (Herold & Schmullius. 2004). One example on the thematic harmonization by Herold et al. (2006), is that the woodland of one dataset should not be the forest of another dataset, but is should be recognized that different LC products could characterize forest in different levels of detail.

Herold et al. (2006) and Herold & Schmullius (2004) describe joint initiatives for the harmonization and validation of LC datasets. Interoperability and comparability between LC products are an important aim of harmonization, as users benefit from these initiatives (Herold et al. 2011; Mora et al. 2014; Verburg et al. 2011). However, interoperability and comparability can be a challenge for data developers and map producers (Herold et al. 2006). Joint initiatives are only successful when strategic decisions are pushed through international bodies and agencies responsible for such tasks (Herold et al. 2006). In general mapping projects can profit from resources and harmonization experiences, especially in terms of identified problems and inconsistencies in existing legends (Herold et al. 2006). In later phases of map development, there is a higher chance of inconsistencies between products which cannot be solved by harmonization, like, for example, a different threshold defining a forest class in tree height (Herold et al. 2006).

Harmonization can be divided in spatial harmonization and the thematic harmonization.

3.4.1. Thematic harmonization

A thematic legend is developed from a classification system for a specific mapping purpose (Herold et al. 2006). LC classes are categorised with well-defined criteria to order spectral image data based on their characteristics, i.e. in terms of factors like percent cover and height (Herold et al. 2006). It is possible that LC legends that are derived from a different classification system, or even without an underlying classification system and therefore lack compatibility (Herold et al. 2006).

In thematic harmonization it is important to find a common language to describe LC and to translate between different legends (Herold et al. 2006). Harmonization is the translation of legends to one general legend (Herold et al. 2006). The Land Cover Classification System (LCCS) provides a valuable universal LC language for building and comparing LC legends. Herold and Schmullius (2004) identify LCCS as an appropriate classification system to provide a common language and translation device. Another common legend for GLC mapping would be the International Geosphere-Biosphere Programme (IGBP) classification scheme with 17 LC classes. Herold et al. (2006) describes that it is more important to standardize terminology than categories in legend harmonization. The harmonization process should first harmonize parameters used for the description of LC classes and thereafter focus on categories if this is necessary (Herold et al. 2006).

The process of legend translation highlights the similarities and differences between legends and shows which classes can be harmonized and where legends show inconsistencies (Herold et al. 2008). It should be noted that thematic harmonization does not resolve all inconsistencies between legends, for example, inconsistencies caused trough different thresholds will remain (Herold et al. 2006). In legend harmonization it is important to know where inconsistencies are not resolved, for further processes like integration and product evaluation.

Herold et al. (2008) uses LCCS to harmonize IGBP, GLC2000 and UMD classes into a general legend with 13 LC classes. Most LC maps and reference datasets available for integration have an LCCS or IGBP based legend.

3.4.2. Spatial harmonization

Spatial harmonization is concerned with harmonizing the different cell sizes of LC maps and the different sample units in reference datasets to one standard in order to compare and integrate them. This research will be limited in spatial harmonization and assume that LC type in a sample area is homogenous.

Yu et al. (2014) discussed the effects of up-scaling and down-scaling map resolution and mention that dominant LC types are overestimated when resolutions are coarsened in up-scaling. The majority LC class at a finer resolution is assigned to the coarser product which causes minority LC classes to be overruled (Yu et al. 2014).

3.5. Validation

An accuracy assessment derives a quantitative description of quality of GLC map (Strahler et al. 2006). There are several methods for performing an accuracy assessment; each has its own value and applicability in validating a given map (Strahler et al. 2006).

Strahler et al. (2006) listed the following priorities in an accuracy assessment: (1) an overall measure of map accuracy, (2) measures of class accuracy and (3) recognizing that accuracy may vary locally within the map. An overall measure of accuracy is needed to indicate quality of a thematic map. Measures of class accuracy are desired because users are interested in specific classes in a thematic map (Strahler et al. 2006). A basic approach in validation is the confusion matrix, were the given map is cross tabulated against a dataset which serves as reference in the analysis (Strahler et al. 2006). Thematic accuracy can be calculated from a confusion matrix (Strahler et al. 2006). These accuracy measures are based on the whole map and however, accuracy may vary locally within the map (Strahler et al. 2006). Foody (2005) states that one problem with the conventional approach, of thematic assessment with a confusion matrix, that it indicates a single, global estimate of thematic classification accuracy. There can be large spatial variation in map accuracy (Foody 2005). One example would be an overall accuracy of approximately 67% of International Geosphere Biosphere Programme's (IGBP) Data and Information System Cover (DISCover) GLC map with a spatial variation that differs by approximately 20% between continents (Foody 2005; Loveland et al. 1999).

The information entropy is a measure of per-pixel classification uncertainty (De Bruin and Gorte. 2000; Shannon and Weaver. 1949). De Bruin and Gorte (2000) used posterior probabilities and the information entropy to quantify uncertainty, when probabilities were concerned the information entropy was used as an internal measure of uncertainty. Entropy expresses uncertainty according to the vectors of posterior probabilities and does not involve uncertainty concerning the probabilities (De Bruin and Gorte. 2000). Gopal et al (1999) mentions that a voting strategy improves classification accuracy and provides a way to evaluate map uncertainty, as voting results are contradictory in cells with mixed LC classes.

Foody (2005) and Strahler et al. (2006) described various methods to indicate spatial variation in classification accuracy. Among others: confidence estimation for thematic class accuracies (Foody 2005; Loveland et al. 1999), calculating the accuracy for defined regions within the map (Foody 2005; Strahler et al. 2006), mapping misclassified areas (Foody 2005) and indication of classification uncertainty on a per pixel basis, which is discussed in the research of Foody (2005). Classification uncertainty can be a useful addition to an accuracy assessment but the relationship between uncertainty and accuracy may not be simple (Foody 2005). This is due to the possibility that LC classes can be allocated correctly but uncertain while other LC classes can be relatively certain but allocated incorrectly (Foody 2005). In mapping spatial accuracy locally, an appropriate testing set is required for an accuracy assessment within the area (Foody 2005).

3.5.1.Internal validation: Shannon information entropy

The information entropy represents the amount of information necessary to require certainty (Shannon and Weaver. 1949). Shannon and Weaver (1949) describe entropy with equation 2, the entropy of a set of probabilities where P_i is the probability being i.

$$H = -K \sum_{i=1}^{n} P_i * Log\ (P_i) \qquad [2]$$

Source: Shannon and Weaver (1949)

3.5.2. External validation: Confusion matrix

Strahler et al. (2006) suggest the confusion or error matrix as a basic approach for the validation of LC maps (Figure 3). In a confusion matrix, the classes of maps and reference data are cross tabulated against each other, from which metrics of classification accuracy or agreement is derived (Foody. 2005). Overall accuracy is an indication of the quality of the entire LC map that represents the percentage of correctly predicted samples over the total amount of samples from the reference dataset (Strahler et al. 2006). A confusion matrix is an accepted approach for map validation, specifically for overall accuracy and per-class accuracy (Foody. 2005; Jung et al. 2006; See et al. 2014; Strahler et al. 2006). The following researches use a confusion matrix to validate their LC maps: Herold et al. (2008), Iwao et al. (2011), Kinoshita et al. (2014), See et al. (2014). Tsendbazar et al. (2016). A confusion matrix is relatively easy to interpret and used by both map producer and map user communities (Strahler et al. 2006).

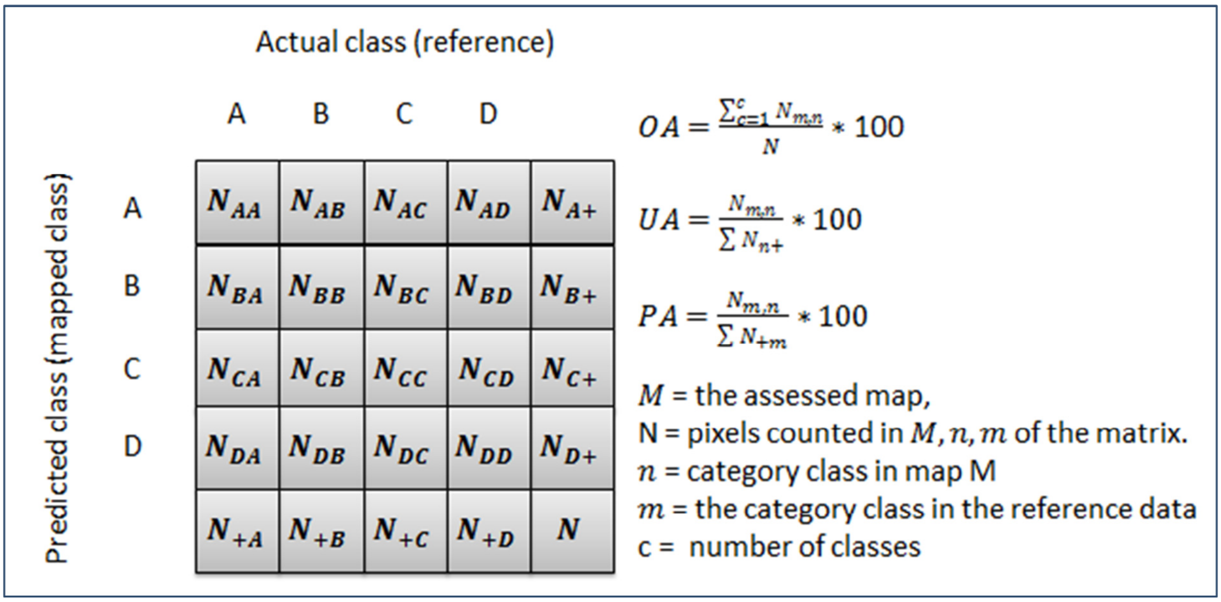

Figure 3: Confusion matrix, Strahler et al. (2006)

In a confusion matrix it is assumed that each pixel is allocated to a single class in the given map and reference dataset and that these datasets have the same spatial resolution (Strahler et al. 2006). In other words, the accuracy assessment considers a pixel as homogenous and not of mixed origin. Interpretation of the confusion matrix also requires consideration of the sample design from the reference dataset (Strahler et al. 2006).

4. Methods

The research methodology of this study consists of five phases: a literature review, legend harmonization, implementation of integration methods, the internal and external validation of the selected methods, and evaluation of the results and the improvements of the integrated LC maps (Figure 4).

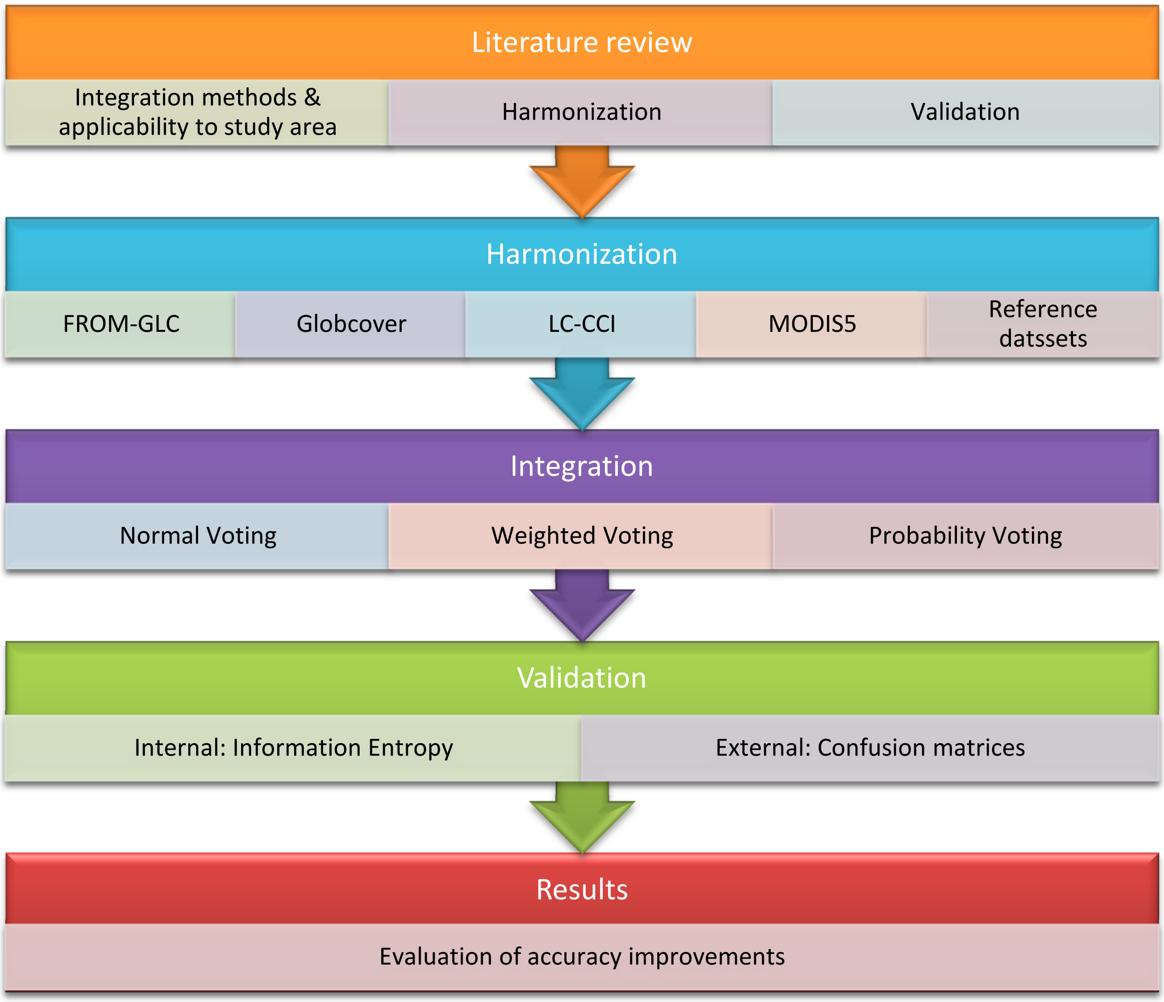

Figure 4: Overall research methodology

4.1. Data harmonization

Thematic legends of the LC data were harmonized to eight classes and the spatial harmonization differed for the (1) LC maps and (2) reference datasets: In the spatial harmonization of LC maps, each map is re-sampled to the chosen resolution. For the reference datasets the spatial harmonization is based on the assumption that the sample unit area has homogenous LC type, so the datasets can be harmonized to have the same extent of sample units.

Herold et al. (2008) used LCCS to harmonize IGBP, GLC2000 and UMD classes into a general legend with 13 LC classes. Most LC maps and reference datasets available for integration have an LCCS or IGBP based legend. From the LC maps: Globcover 2009 and LC-CCI have an LCCS based legend and MODIS5 (2010) is based on an IGBP legend. From the LC reference datasets: GLC2000, Gobcover 2005 and the GLCNMO-tr legends are based on LCCS and MODIS-tr, Geo-wiki and VIIRS are based on IGBP.

This research uses the LCCS based approach of Herold et al. (2008) and the research of Tsendbazar (2015) as a basis for the creation of the harmonized legend. The general legend of 13 LC classes used by Herold et al. (2008) and Tsendbazar (2015) were harmonized to eight classes by further combining all tree classes to one class and combining the classes water, permanent snow and ice.

Legends of the LC data were harmonized into one general legend with eight LC classes Table 27, table 28 and table 29 in the appendix describes the thematic harmonization in detail. Table 7 is a simplified version of these tables, which holds the harmonized legend of eight classes with the original LC classes that correspond to the original maps and reference datasets.

Table 7: Thematic harmonization

Harmonized legend		LC data						
LC Class	Description	FROM-GLC	GLC 2000	Geo-wiki	Globcover	GLCNMO	IGBP/ MODIS/ VIIRS	LC-CCI
1	Trees	20	1, 2, 3, 4, 5	1	40, 50, 60, 70, 90, 100, 110, 120, 160, 170	1, 2, 3, 4, 5	1, 2, 3, 4, 5, 8, 9	50, 60, 70, 80, 90, 100, 110, 160, 170
2	Shrubs	40	6	2	130	7	6, 7	120
3	Herbaceous vegetation	30	7	3	30, 140	8	10	40, 30, 140
4	Cultivated and managed vegetation / agriculture (incl. mixtures)	10	8	4	11, 14, 20	11, 13	12, 14	10, 20, 30
5	Wetlands	50, 70	9	6	180	-	11	180
6	Urban/built up	80	10	7	190	18	13	190
7	Water, permanent Snow and Ice	60, 100	11, 13	8, 10	210, 220	19, 20	15, 17	210, 220
8	Barren	90	12	9	150	10, 16, 17	16	15, 20

4.1.1. Land cover maps

Figure 5 describes in four steps the methodology used for harmonizing the FROM-GLC, Globcover, LC-CCI and MODIS LC maps. The harmonization was conducted in ArcGIS through a reclassification algorithm using the harmonized legend in excel (table 7).

Pre-processing steps were done in section 2.2.1. The spatial harmonization of the LC maps is done in step two of the methodology. The resolutions of the LC maps were re-sampled to a pixel size of 300 meter, which is 0.002778 by 0.002778 decimal degrees. A resolution of 300 meter was chosen because this is the same as the resolution of Globcover 2009 and LC-CCI 2010. The MODIS5 and FROM-GLC maps were resampled to have 300m resolution using nearest neighbor resampling. FROM-GLC was re-sampled during pre-processing in section 2.2.1. Table 2 and table 3 hold the metadata of the LC maps, among them the maps original spatial format.

The thematic harmonization was done in the third step of the methodology. FROM-GLC, Globcover 2009, LC-CCI and MODIS5 original legends from the LC maps were re-classified (spatial analyst) with the harmonized legend.

Figure 5: Harmonization method of LC maps

Figure 6 presents the FROM-GLC hierarchy, Globcover 2009, LC-CCI (2010) and MODIS5 (2010) LC maps that result from the harmonization, which are used as input maps for the integration methods to produce improved LC maps (section 4.2). Table 8 holds the user agreement (UA), producer agreement (PA) and overall agreement (OA) of all LC maps from the harmonized global confusion matrices of Tsendbazar et al. (2016) and Yu et al. (2014). The accuracy assessments of Globcover 2009, LC-CCI (2010) and MODIS5 were obtained from the research of Tsendbazar et al. (2016) and FROM-GLC was obtained from Yu et al. (2014). The confusion matrix from Yu et al. (2014) holds no information for LC class (5) wetlands and is produced from the stage of FROM-GLCagg and not FROM-GLC hierarchy (Yu et al. 2014) (section 3.3.2). Misclassifications from the stage of FROM-GLCagg were processed in FROM-GLC hierarchy (section 2.2.1) (Yu et al. 2014). The matrix of FROM-GLC holds a different accuracy assessment than the Globcover 2009, LC-CCI (2010) and MODIS5 (2010) LC input maps. This research uses the accuracy assessments for map integration with the chosen methods (section 4.2), but it is questionable if these accuracy assessments can be compared to each other (section 6.1.1) as is explained in section 3.3.2.

The FROM-GLC hierarchy LC map (figure 6a) shows a tiling effect which could, for example, be caused by differing acquisition times. LC-CCI (2010) (figure 6c) has the highest OA and UA for LC classes; (1) trees, (2) shrubs and (7) water, permanent snow and ice. MODIS5 (2010) (figure 6d) has the highest OA after LC-CCI (2010) and the highest UA for LC classes; (4) cultivated and managed vegetation/agriculture, (6) urban/built up and (8) barren. LC-CCI (2010) and MODIS5 (2010) class specific accuracies are generally high compared to the other input maps. Globcover 2009 (figure 6b) has a lower OA than LC-CCI (2010) and MODIS5 (2010) but a higher OA than FROM-GLC. Class specific accuracies of Globcover are generally a little lower or higher than the other input maps and only LC class (5) wetlands had the highest UA.

Table 8: Reported accuracies of input maps

	FROM-GLC		Globcover (2009)		LC-CCI (2010)		MODIS5 (2010)	
OA	66.10		67.81		74.70		73.92	
Class	UA	PA	UA	PA	UA	PA	UA	PA
1	80.33	79.92	73.92	80.43	90.39	81.10	87.25	79.23
2	48.32	38.06	45.49	36.34	61.67	55.39	46.01	66.91
3	53.22	34.62	33.25	25.19	47.02	45.84	42.63	55.63
4	57.78	66.86	76.83	75.83	79.23	82.20	84.60	84.83
5	0.00	0.00	68.57	39.34	46.88	49.18	52.38	32.84
6	41.57	25.37	74.07	40.00	66.04	70.00	85.71	60.00
7	81.64	90.53	82.39	89.31	85.33	94.81	82.96	82.96
8	62.64	90.67	68.05	85.83	72.83	83.64	98.28	62.09

Source: Tsendbazar et al. (2016) and Yu et al. (2014)

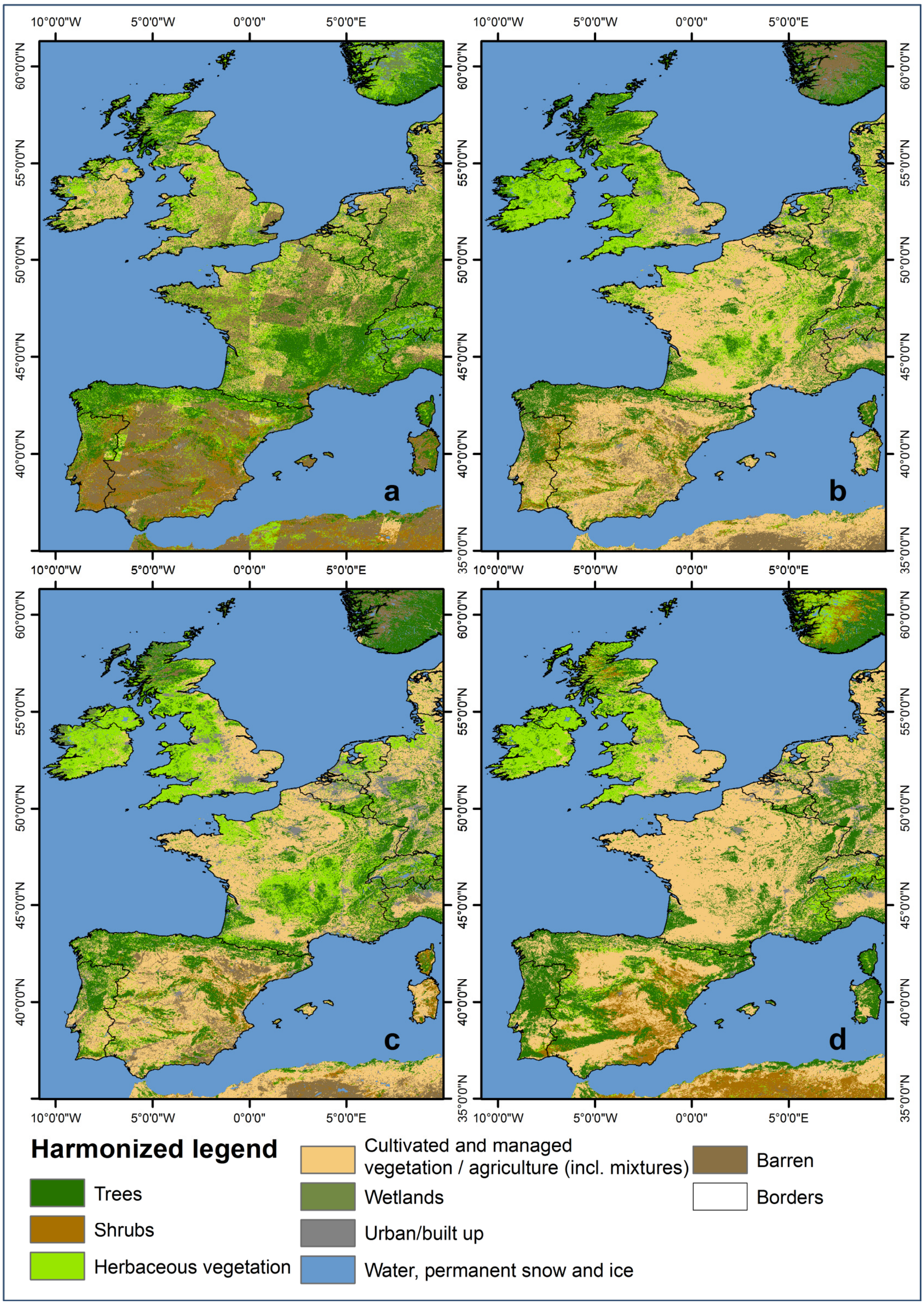

Figure 6: Harmonized LC maps

4.1.2. Reference datasets

Figure 6 describes in five steps the methodology used for harmonizing: GLC2000, GLCNMO-tr, Geo-wiki, Globcover 2005, MODIS-tr, and VIIRS to one reference dataset. The harmonization was performed in ArcGIS and used the harmonized legend table 7. Table 5 holds the metadata of the reference data, among them the data's original spatial format.

Pre-processing steps were done in section 2.2.2. The spatial harmonization of the reference datasets is done in step two of the methodology. All the reference datasets were harmonized to have a point as a sample unit under the assumption of LC homogeneity, with the feature to point (data management) tool in ArcGIS. The thematic harmonization of the reference datasets was done in the third step of the methodology. Each reference datasets was joined to the harmonized legend, based on a datasets original legend, by performing a join between the attribute table and the harmonized legend (table 7). Section 8.2 table 28 in the appendices holds the original legend of the reference datasets. The fourth step of the methodology completes the harmonization of the reference datasets. All point features of each reference dataset were merged to one reference dataset with the merge (data management) tool in ArcGIS. The reference dataset was used for the external validation, section 4.3.2 further explains the external validation methodology.

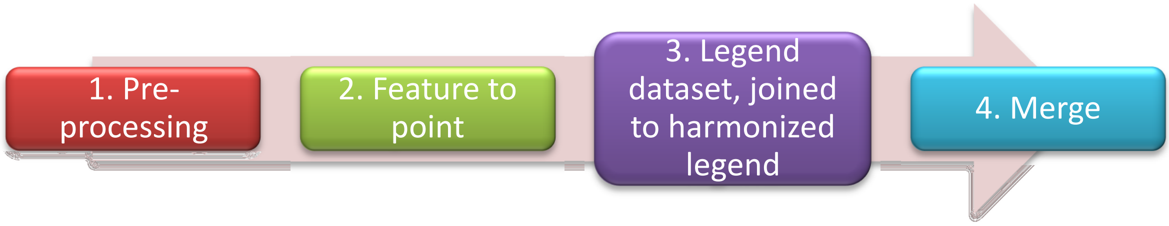

Figure 7: Harmonization method of LC reference datasets

Figure 8 presents the results from the harmonization of GLC2000, GLCNMO-tr, Geo-Wiki, Globcover 2005,MODIS-tr and VIIRS 3 to one reference dataset. The reference dataset has 916 samples. Table 6 (section 3.3.1) presents the distribution of the samples over the harmonized LC classes and discusses that LC classes (5) wetlands and (8) barren were not represented by enough samples (section 3.3.3).

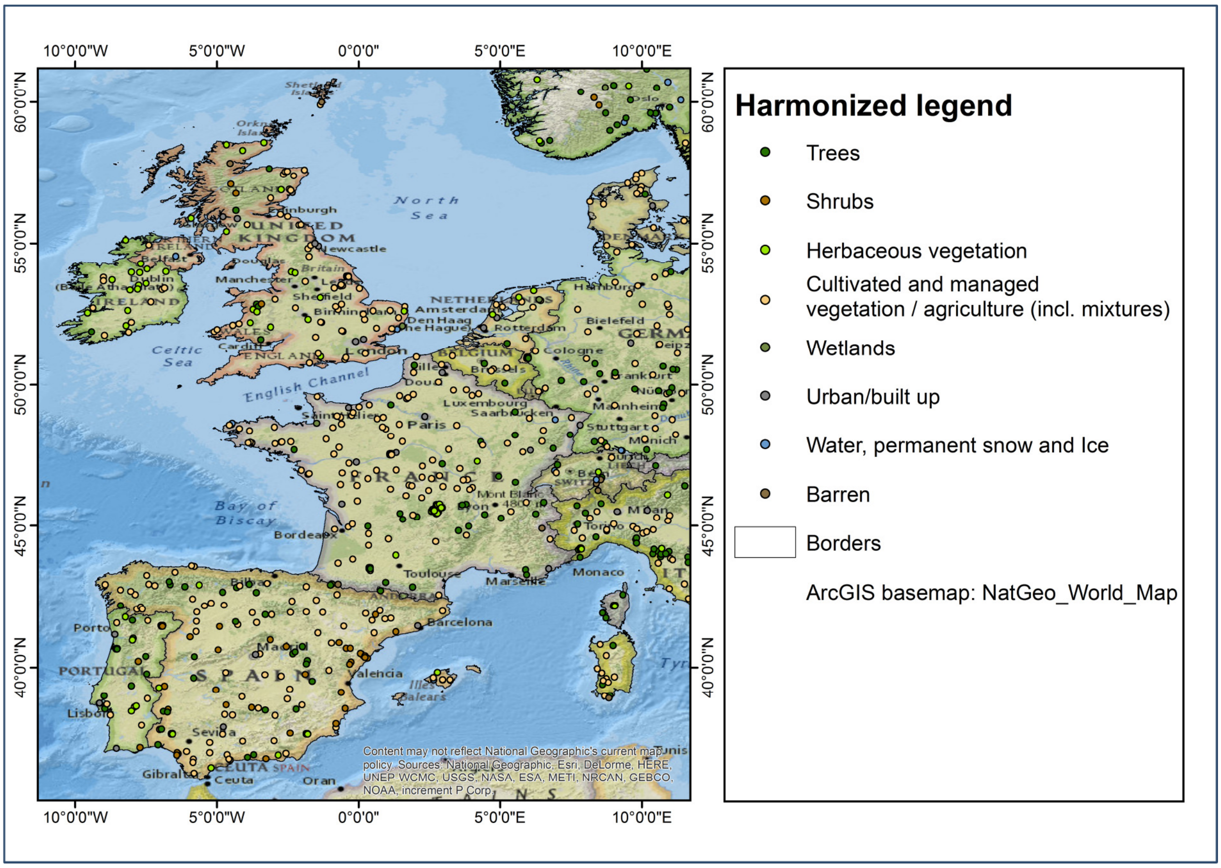

Figure 8: Harmonized reference dataset

4.2. Integration methods

An integration method decides on a LC class for each pixel in the new map based on the LC information from the input products. Section 3.2 discusses the theoretical background on the chosen integration methods: normal voting, weighted voting and probability voting. This section covers the implementation of the integration methods with the harmonized FROM-GLC, Globcover 2009, LC-CCI and MODIS5 LC maps. The integration methods were implemented in R. Two of the chosen integration methods require accuracy information from each input map. The accuracy assessments of Globcover 2009, LC-CCI (2010) and MODIS5 were obtained from the research of Tsendbazar et al. (2016) and FROM-GLC was obtained from Yu et al. (2014). Section 3.3.2 explains the use of accuracy information from external research instead of an accuracy assessment with own reference data.

4.2.1. Normal Voting

The methodology for "Normal voting" was based on a common voting procedure, known as majority voting. A pixel was assigned to the class that occurs in the majority of the LC maps at that pixel's location. As mentioned in the theoretical background, a class is easily assigned where all LC datasets agree, but it is more difficult when LC datasets disagree and form a tie. There are five conditions possible in the voting procedure based on the preconditions: (1) there are four input maps and (2) the voting procedure is performed on the classes that are present. The voting procedure decides on a LC class in case all maps agree, three maps agree or two maps agree while the remaining two disagree. In case of a tie, the voting procedure remains undecided on a LC class. Table 9 shows the five possible conditions in the initial voting procedure.

Table 9: Initial voting, from agreement to disagreement conditions

Indicator	Condition
10	All agree
20	Three maps agree, one disagrees
30	Two maps agree, remaining disagree
40	Tie: two maps agree, remaining also agree
50	Tie: all maps disagree

Part of this methodology was to decide on how to deal with ties through which a complete voting result can be achieved. This research uses a new approach in solving ties based on class preferences that are calculated from class occurrences of each input map. The methodology of the normal voting was built from three phases (figure 9): (1) the initial voting map, (2) indicators of conditions to deal with ties and the (3) final voting map.

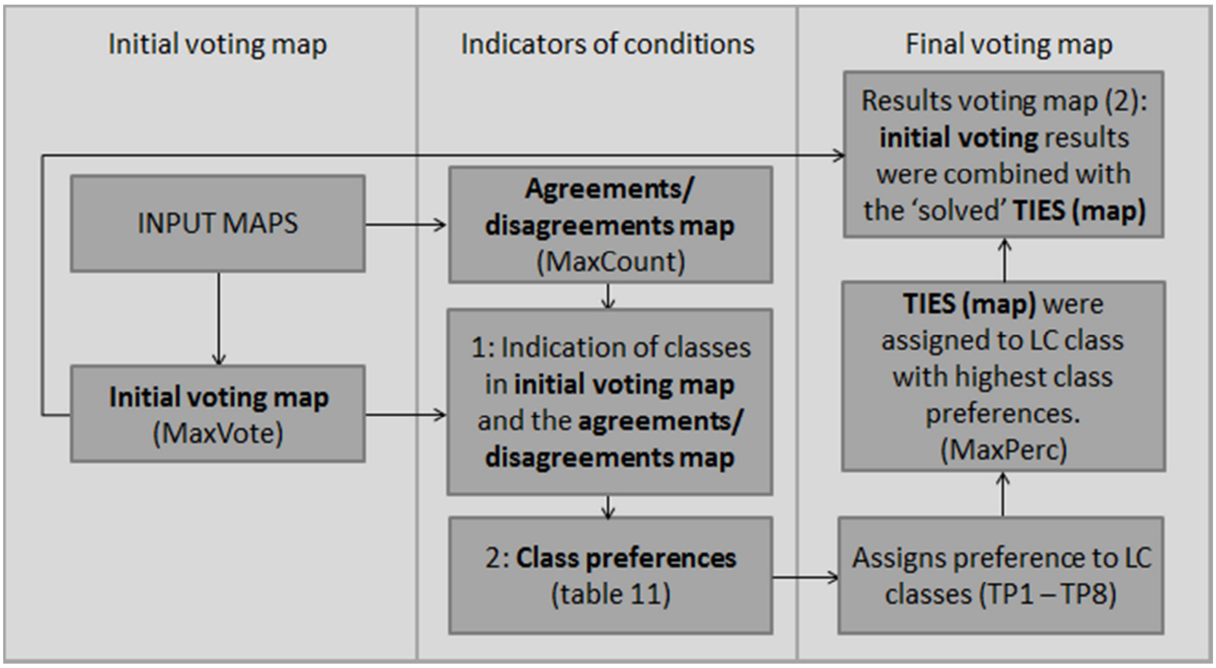

Figure 9: Methodology normal voting

Initial voting map

The initial phase is the majority voting performed on the class presence of the four input maps and produced an initial voting map (section 5.1, figure 13a). In the initial voting map, pixels were assigned to a LC class that occurs most often at a pixels location, where the four LC input maps agree on a LC class. In case of a tie, the voting procedure remains undecided on a LC class, represented by a deviating value in the initial voting map (initial voting script: MaxVote, section 8.1.1).

Indicators of conditions to deal with ties

This phase deals with the ties from the initial voting process by: (1) producing an agreement/disagreement map (2) calculating occurrences within the initial voting map and agreement/disagreement map for each of the LC input maps and (3) processing this information to class occurrences in excel (table 30 and table 31, 8.2.2), which were simplified to class preferences (table 10). The final voting map uses this information to decide on a LC class at a tie location.

The agreement/disagreement map (section 5.1 figure 13b) indicated the five conditions (table 9) from the initial voting map produced by (initial voting: VoteCount, section 8.1.1). The agreement/disagreement map represents conditions where: (10) all maps agree, (20) three maps agree and (30) two maps agree while the remaining two disagree. In these conditions the voting procedure from the initial voting map decides on a LC class. The voting procedure remains undecided on a LC class in case of a tie. Ties are represented by the conditions where: (40) two maps agree while the remaining two also agree on a LC class and (50) where all maps disagree. In phase two is calculated for each LC input map when the eight classes occur in the initial voting map and in each of the five conditions (script: indicators in condition script, section 8.1.2). In phase three this information is exported to excel (table 30 and table 31, section 8.2.2): Counted cells from classes that concur in the initial voting, its input map and each of the five conditions were converted to percentages set against the counted cells from classes of the initial voting results. Tie locations were excluded, as the eight classes do not occur in the tie locations of the initial voting map.

This information represents in percentages how much an input map with its classes concurs with the initial voting. Additionally the information indicates the amount of agreement with the other maps by being distributed over the three agreeing conditions.

The idea is to "solve" tie locations based on the occurrences of LC classes (from each input map) in the initial voting results and assumes that using a class with a high occurrence could improve map accuracy. Table 10 class preferences were calculated from these class occurrences (table 30 and table 31) by summing the class occurrence over the three agreeing conditions. Class preferences of each LC input map, represents in percentages the contribution of a LC class to the initial voting map. Classes with a high preference will hold a higher ranking in the voting than classes with a lower preference.

Table 10: Class preferences

Class preferences (percentages of class occurrences in initial voting)						
		Initial voting	FROM-GLC	Globcover	LC_CCI	MODIS 5
All classes 1 - 8		100	74.28	85.81	86.73	84.79
Class description	No	Initial voting	FROM-GLC	Globcover	LC_CCI	MODIS 5
Trees	1	100	82.99	78.47	84.16	83.96
Shrubs	2	100	80.90	46.35	51,51	68.04
Herbaceous vegetation	3	100	45.74	72.85	82.39	67.89
Cultivated and managed vegetation / agriculture	4	100	30.08	88.27	84.39	87.95
Wetlands	5	100	37.45	65.51	83.15	18.77
Urban/built up	6	100	31.95	62.73	98.08	77.44
Water, Snow and Ice	7	100	99.90	99.73	99.93	99.00
Barren	8	100	60.00	92.48	82.26	6.46

Locations with ties were solved by using the class preferences to decide on a LC class. Four input maps give four possible LC classes, in case of a tie; the class with the highest preference will be assigned to that location. This is a new approach for solving ties that uses the agreements between the initial voting procedure and its input maps.

Final voting map

The final voting map was produced by: (1) using the class preferences to assign a LC class to each tie location and (2) produce a voting result map by combining the initial voting from agreeing conditions with the ties that have been solved. The script used to produce a complete voting result is in section 8.1.3 of the appendix.

In step one, in case of a tie, LC class one to eight of each input map were replaced with the corresponding class preference. Table 11 is an example of this approach, for each tie location there are four inputs of class preferences (table 10) that correspond to the actual class of the input map. The resulting class preferences of each map were summed from TP1 to TP8 and normalized by dividing each TP trough the sum of TP1 to TP8. The script calculated which TP had the highest preference and assigned that LC class to the ties location. The output is a raster with a voting outcome for each tie location.

Table 11: Example for solving ties in normal voting

Normal voting, assigning a class to a tie location										
INPUT MAP	Actual Class	Class preferences								
		1	2	3	4	5	6	7	8	
FROM-GLC	1	82.99	0	0	0	0	0	0	0	
Globcover	7	0	0	0	0	0	0	99.73	0	
LC-CCI	4	0	0	0	84.39	0	0	0	0	
MODIS	3	0	0	67.89	0	0	0	0	0	
		TP1	TP2	TP3	TP4	TP5	TP6	TP7	TP8	TP-SUM
TP		82.99	0	67.89	84.39	0	0	99.73	0	335.00
Normalized		0.25	0	0.20	0.25	0	0	0.30	0	1
Assigned class								7		

The voting results were produced by joining the initial voting with the solved ties by a " if else statement ". The normal voting map holds the values of the "solved ties" in case of a tie and the values of the initial voting in case of agreeing conditions where there were no ties.

4.2.2. Weighted Voting

Weighted voting is based on a methodology that applies weights into the voting procedure. A pixel is assigned to the class that accumulates the highest weight at that pixel's location. This research's "weighted voting" derives weights from the user accuracy (UA) of each LC class which were obtained from the global confusion matrices of FROM-GLC (Yu et al. 2014) Globcover 2009, LC-CCI (2010) and MODIS5 (Tsendbazar et al. 2016). This research bases the weights on the user accuracies, as user accuracies represent the agreement of the LC map with the reference data. The weights used for weighted voting are based on the UA in table 12.

The methodology of weighted voting is presented in figure 10 and bears similarity to the methodology used within tie locations from the normal voting method. The script used to produce the weighted voting map is in section 8.1.4 of the appendix. The methodology of weighted voting was applied on each location in the study area.

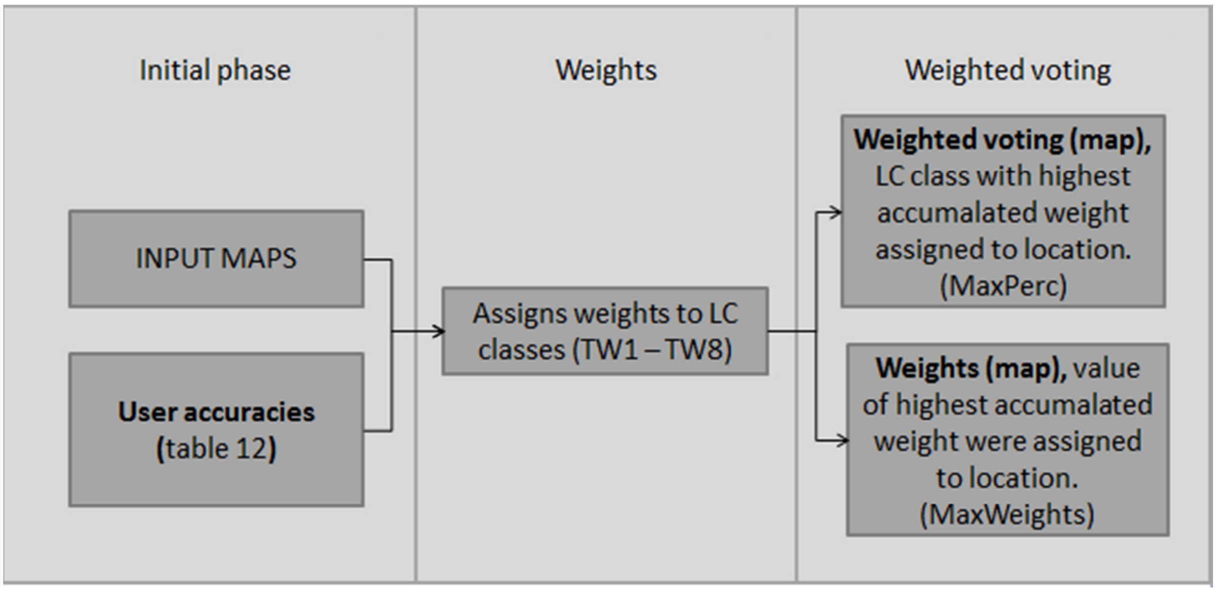

Figure 10: Methodology weighted voting

The weights used for weighted voting are based on the UA in table12.

Table 12: User accuracies weighted voting

UAwv.csv				
Class	FROM-GLC	Globcover	LC_CCI	MODIS 5
1	80.33	73.92	90.39	87.25
2	48.32	45.49	61.67	46.01
3	53.22	33.25	47.02	42.63
4	57.78	76.83	79.23	84.60
5	0.00	68.57	46.88	52.38
6	41.57	74.07	66.04	85.71
7	81.64	82.39	85.33	82.96
8	62.64	68.05	72.83	98.28

Table 13 is an example of the weighted voting approach. Classes one to eight of each input map were replaced with the corresponding user accuracies from table 12.in the initial phase of the methodology. Each location has four inputs of user accuracies from the actual class of the input map. These user accuracies were converted to weights when the accuracies of each map were summed in TW1 to TW8 and normalized by dividing each TW trough the sum of TW1 to TW8. The weighted voting map was produced when the script calculated which TW has the highest accumulated weight and assigns the corresponding class to that location. Additionally, a map was made from weights used in the weighted voting by returning the weights value instead of LC class.

Table 13: Example weighted voting

Weighted voting										
INPUT MAP	Actual Class	User accuracies, used as weights								
		1	2	3	4	5	6	7	8	
FROM-GLC	4	0	0	0	57.78	0	0	0	0	
Globcover	4	0	0	0	76.83	0	0	0	0	
LC-CCI	5	0	0	0	0	46.88	0	0	0	
MODIS	3	0	0	42.63	0	0	0	0	0	
Weights		WP1	WP2	WP3	WP4	WP5	WP6	WP7	WP8	WP-SUM
		0	0	42.63	134.62	46.88	0	0	0	224.12
Normalized weights		0	0	0.19	0.60	0.21	0	0	0	1
Assigned class					4					

4.2.3. Probability Voting

In probability voting, a voting procedure is applied on the probabilities of each class being the correct class. A pixel is assigned to the class that accumulates the highest probability at the pixel's location. Section 3.2.2 from the theoretical background studies the researches of Kinoshita et al. (2014) and Tuanmu and Jetz (2014) on how probability voting has been used in previous studies. In this research, the four input maps were integrated simultaneously with probability voting because FROM-GLC and MODIS respectively have a finer and coarser resolution than the resolution of 300 meter from Globcover 2009 and LC-CCI. The methodology of probability voting is presented in figure 11. The script used to produce the probability voting map is in section 8.1.5 of the appendix.

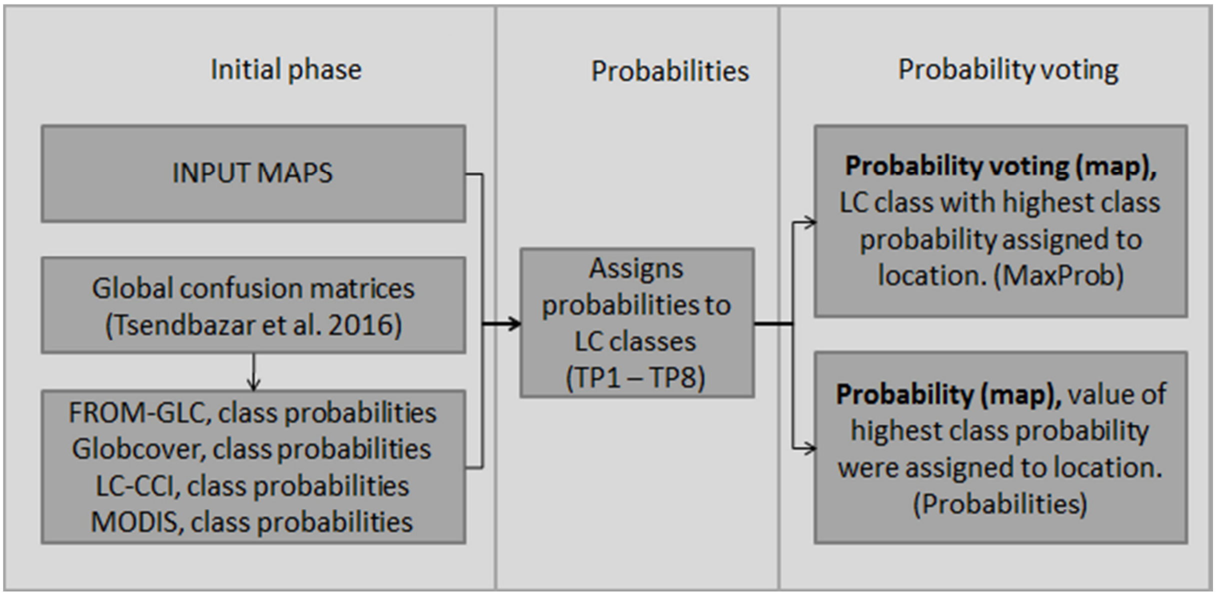

Figure 11: Methodology probability voting

Class probabilities

The probabilities used in the methodology were obtained from converting the harmonized global confusion matrices from the research of Tsendbazar et al. (2016) and Yu et al. (2014) to class probabilities with equation 1 (section 3.2.3) from the research of Kinoshita et al. (2014). Probabilities were calculated from the user perspective in this research, as this represents the agreement of the LC input maps with the reference data. Kinoshita et al. (2014) calculated class probabilities for six LC classes from the original legends of the input maps. This research calculated class probabilities for the harmonized legend with eight LC classes. The matrices were converted to probabilities by dividing the category class of the mapped class n, set against the category class of the reference class m, through the sum of the mapped class described in equation 1. Table 14 on the next page holds the class probabilities calculated with equation 1. Probabilities with value zero were given a value close to infinity for the multiplication in the methodology, described by equation 5 and table 15.

Table 14: Class probabilities of LC input maps

FROM-GLC		Reference classes								SUM
		1	2	3	4	5	6	7	8	
Mapped Classes	1	0.80	0.05	0.10	0.03	0	0.00	0.00	0.00	1
	2	0.19	0.48	0.20	0.05	0	0.00	0.00	0.07	1
	3	0.17	0.17	0.53	0.08	0	0.01	0.01	0.03	1
	4	0.10	0.06	0.22	0.58	0	0.02	0.00	0.02	1
	5	0	0	0	0	0	0	0	0	0
	6	0.05	0.05	0.08	0.13	0	0.42	0.03	0.25	1
	7	0.07	0.02	0.05	0.02	0	0.00	0.82	0.02	1
	8	0.02	0.11	0.20	0.03	0	0.00	0.01	0.63	1

Globcover		Reference classes								SUM
		1	2	3	4	5	6	7	8	
Mapped Classes	1	0.74	0.05	0.08	0.10	0.02	0.00	0.01	0.01	1
	2	0.21	0.45	0.15	0.14	0	0.01	0	0.04	1
	3	0.10	0.21	0.33	0.21	0.01	0.01	0.00	0.12	1
	4	0.06	0.04	0.10	0.77	0.00	0.01	0.00	0.01	1
	5	0.11	0	0.20	0	0.69	0	0	0	1
	6	0	0	0	0.19	0	0.74	0.07	0	1
	7	0.06	0.01	0.03	0.01	0.01	0.01	0.82	0.05	1
	8	0.03	0.04	0.18	0.05	0.01	0.01	0.00	0.68	1

LC-CCI		Reference classes								SUM
		1	2	3	4	5	6	7	8	
Mapped Classes	1	0.90	0.02	0.03	0.04	0.00	0	0.00	0.00	1
	2	0.12	0.62	0.13	0.07	0.00	0	0	0.06	1
	3	0.06	0.10	0.47	0.23	0.00	0.01	0.00	0.12	1
	4	0.07	0.05	0.07	0.79	0.01	0.01	0.00	0.01	1
	5	0.28	0.03	0.13	0.05	0.47	0	0.02	0.03	1
	6	0.04	0	0.02	0.28	0	0.66	0	0	1
	7	0.05	0	0.01	0.01	0.03	0	0.85	0.05	1
	8	0.03	0.04	0.17	0.01	0.01	0.00	0.00	0.73	1

MODIS		Reference classes								SUM
		1	2	3	4	5	6	7	8	
Mapped Classes	1	0.87	0.03	0.02	0.06	0.01	0.00	0.00	0.00	1
	2	0.18	0.46	0.22	0.03	0.02	0.00	0.01	0.08	1
	3	0.09	0.13	0.43	0.14	0.01	0.00	0.01	0.19	1
	4	0.07	0.03	0.03	0.85	0.01	0.01	0	0.00	1
	5	0.26	0	0.07	0	0.52	0.02	0.10	0.02	1
	6	0.03	0.03	0	0.06	0	0.86	0.03	0	1
	7	0.04	0	0.01	0	0	0	0.83	0.13	1
	8	0	0.00	0.00	0.00	0	0	0.01	0.98	1

Source: Yu et al. (2014) and Tsendbazar et al. (2016).

Multiple dependent events in conditional probability

A location with the probability of being $Class_n$ depends on the class probability of each LC input map: $M1_{class_n}$, $M2_{class_n}$, $M3_{class_n}$ and $M4_{class_n}$. These are multiple dependent events in joint probability. Joint probability is expressed as $P(A, B)$ which means that both events occur and conditional probability is expressed as $P(A|B)$ which means that event A occurs given event B occurs (UMICH, electronic textbook; probability. 2015). Equation 3 describes multiple events in conditional probability with Baye's rule's (UMICH, electronic textbook; probability. 2015).

$$P(A|B) = \frac{P(B|A)*P(A)}{P(B)} \qquad [3]$$

Source: Probabilities, Bayes Rule (2015)

Equation 3 can be rewritten as equation 4, to represent $P(C|M1 \dots M4)$, the probability of being $Class_n$ given the class probabilities of the four LC input maps. $P(C)$ is the probability of being $Class_n$ in the new integrated map and $P(M1..M4)$ are the class probabilities of the LC input maps.

$$P(C|M1 \dots M4) = \frac{P(M1..M4|C)*P(C)}{P(M1..M4)} \qquad [4]$$

The methodology of probability voting from this research is based on two assumptions:

1. The input maps are independent events
2. The LC classes have equal priors

Based on assumption two, $P(C)$ in the numerator can be removed from equation4, because the probabilities of the LC input classes ($P(M1)$ to $P(M4)$) have equal prior. Based on assumption one, the LC input maps can be seen as independent events and $P(M1..M4|C)$ can be solved by multiplying the LC classes from the input maps. In this research, the new assigned class is calculated by multiplying the LC class probabilities from the input maps divided trough the sum of all class probabilities to normalize the probabilities (equation 5).

$$P(C|M1 \dots M4) = \frac{P(M1..M4|C)}{P(M1..M4)} \qquad [5]$$

Implementation

Table 15 is an example of probability voting. Probability voting takes into account the probability of the mapped class and the probability of being another class, which makes that each input class holds eight class probabilities. On a specific location, there are four classes from the input maps that together hold 32 class probabilities for the integrated class. For each map, eight probabilities that correspond to the input class are imported from table 14. All probabilities being $class_n$ from each map are multiplied to TP1 to TP8 and normalized by dividing each TP trough the sum of all TP's (equation 5). The probability voting map is produced from assigning the LC class with the highest TP/class probability to that location. Additionally, a probability map, from which the LC map was produced, is made by returning the probabilities value instead of LC class.

Table 15: Example probability voting

Probability voting										
INPUT MAP	Actual Class	Class probabilities								
		1	2	3	4	5	6	7	8	SUM
FROM-GLC	4	0.10	0.06	0.22	0.58	0.00	0.02	0.00	0.02	1
Globcover	4	0.06	0.04	0.10	0.77	0.00	0.01	0.00	0.01	1
LC-CCI	5	0.28	0.03	0.13	0.05	0.47	0.00	0.02	0.03	1
MODIS	3	0.09	0.13	0.43	0.14	0.01	0.00	0.01	0.19	
(multiplied)		TP1	TP2	TP3	TP4	TP5	TP6	TP7	TP8	TP-SUM
Probabilities		1.55 E-04	9.01 E-06	1.18 E-03	2.88 E-03	2.39 E-25	8.85 E-27	5.00 E-10	1.61 E-06	4.23 E-03
Normalized P		0.04	0.00	0.28	0.68	0.00	0.00	0.00	0.00	1
Assigned class					4					

4.3. Validation of methods

This research used confusion matrices to validate the integrated LC maps and the information entropy as a measure of uncertainty in the classification of the integrated methods: Normal voting, weighted voting and probability voting. The validation of LC maps from this research can be divided in an internal and external validation. The confusion matrices validate LC maps on their agreement with the reference dataset. The information entropy is an internal measure of uncertainty and independent of the external validation.

4.3.1. Internal validation

The information entropy is an internal measure of uncertainty and represents the amount of information necessary to require certainty (Shannon and Weaver. 1949). Next to the external validation, the information entropy was used as addition to evaluate the integration methods by measuring their uncertainty. A high entropy value stands for a high uncertainty in the LC classification and a low entropy value represents certainty. The methodology computed the information entropy of LC classes for each method accumulated over the LC map. The best integration method could be chosen as the one having least entropy aggregated over the study area. The information entropy over the LC maps was calculated by equation 6, similar to equation 2 in section 3.5.1:

$$H = -K \sum_{i=1}^{n} P_i * Log\ (P_i) \qquad [6]$$

$i = possibilities\ of\ LC\ class\ (1\ to\ 8), probabilities\ in\ probability\ voting\ and\ weights\ in\ weighted\ voting$

Source: Shannon and Weaver. 1949

Normal voting is based on the presence of a LC class from each input map and not class probabilities or weights. Therefore the information entropy (equation 6) is implemented on the presence of LC classes from the LC input maps, each map has a possibility of 0.25% to be correct, when maps agree possibilities are summed. The Shannon function from the normal voting script is in section 8.1.3 of the appendix and implements the information entropy on normal voting.

Weighted voting and probability voting respectively hold "weights" and "probabilities" that represent possibilities. These possibilities were implemented as $p(c_n)$ in equation 6 and hold the TW1 to TW8 (or TP1 to TP8) values from the methodology of weighted voting (figure 10) and probability voting (figure 11). The same Shannon function was used for weighted voting and probability voting. The weighted voting script (section 8.1.4) and probability voting script (section 8.1.5) in the appendix implement the information entropy on weighted voting and probability voting.

Table 16 and presents an example for calculating the information entropy for one situation in weighted voting. $p(c_n)$ in equation 6 and TW1 to TW8 do not hold a value for each LC class in weighted voting. In case all maps agree, there is one value for $p(c_n)$ and the TP's. In case all maps disagree, there are four values for $p(c_n)$ and the TW's. The function LogSpecial inside the Shannon function gives $Log(pc_n)$ the value 0 in case $p(c_n)$ is 0, and the function gives $Log(pc_n)$ the value of $Log(pc_n)$ in case $p(c_n)$ is not 0.

Table 16: Example for calculating information entropy over weighted voting

Weighted voting										
INPUT MAP	Actual Class	User accuracies, used as weights								
		1	2	3	4	5	6	7	8	
FROM-GLC	4	0	0	0	57.78	0	0	0	0	
Globcover	4	0	0	0	76.83	0	0	0	0	
LC-CCI	5	0	0	0	0	46.88	0	0	0	
MODIS	3	0	0	42.63	0	0	0	0	0	
		WP1	WP2	WP3	WP4	WP5	WP6	WP7	WP8	WP-SUM
Weights		0	0	42.63	134.62	46.88	0	0	0	224.12
Normalized weights		0	0	0.19	0.60	0.21	0	0	0	1
Assigned class					4					
Calculation of information entropy										
		1	2	3	4	5	6	7	8	
$p(c_n)$		0.00	0.00	0.19	0.60	0.21	0.00	0.00	0.00	
$Log(pc_n)$		0.00	0.00	-2.39	-0.74	-2.26	0.00	0.00	0.00	
$p(c_n) * Log(pc_n)$		0.00	0.00	-0.46	-0.44	-0.47	0.00	0.00	0.00	
$* -1$		0.00	0.00	0.46	0.44	0.47	0.00	0.00	0.00	
$\sum$ =		1.37	H							

Table 17 presents an example for calculating the information entropy for one situation in probability voting. (c_n) in equation 6 and TP1 to TP8 hold a value for each LC class in probability voting because probability voting accounts for the probability of a LC class being another class.

Table 17: Example for calculating information entropy over probability voting

Probability voting										
INPUT MAP	Actual Class	Class probabilities								
		1	2	3	4	5	6	7	8	SUM
FROM-GLC	4	0.10	0.06	0.22	0.58	0.00	0.02	0.00	0.02	1
Globcover	4	0.06	0.04	0.10	0.77	0.00	0.01	0.00	0.01	1
LC-CCI	5	0.28	0.03	0.13	0.05	0.47	0.00	0.02	0.03	1
MODIS	3	0.09	0.13	0.43	0.14	0.01	0.00	0.01	0.19	1
(multiplied)		TP1	TP2	TP3	TP4	TP5	TP6	TP7	TP8	TP-SUM
Probabilities		1.55 E-04	9.01 E-06	1.18 E-03	2.88 E-03	2.39 E-25	8.85 E-27	5.00 E-10	1.61 E-06	4,23 E-03
Normalized P		0.04	0.00	0.28	0.68	0.00	0.00	0.00	0.00	1
Assigned class					4					
Calculation of information entropy										
		1	2	3	4	5	6	7	8	
$p(c_n)$		0.04	0.00	0.28	0.68	0.00	0.00	0.00	0.00	
$Log(pc_n)$		-4.77	-8.87	-1.84	-0.55	-73.91	-78.66	-23.01	-11.36	
$p(c_n) * Log(pc_n)$		-0.17	-0.02	-0.51	-0.38	0.00	0.00	0.00	0.00	
$* -1$		0.17	0.02	0.51	0.38	0.00	0.00	0.00	0.00	
$\sum$ =		1.09	H							

4.3.2. External validation

Confusion matrices were used for the external validation of LC maps, which is explained in the theoretical background section 3.5.2. The LC maps were assessed on their agreement with the reference dataset and the integration methods were evaluated on their improvement compared to the other maps (section 5.3). Section 5.3 presents the accuracy assessment of seven LC maps: normal voting, weighted voting, probability voting, FROM-GLC, Globcover (2009), LC-CCI (2010) and MODIS5 (2010). In a confusion matrix, the assessed map was cross tabulated against the reference dataset described in section 3.5.2. The reference dataset was produced from harmonizing GLC2000, GLCNMO-tr, Geo-wiki, Globcover 2005, MODIS-tr, and VIIRS to one reference dataset (section 4.1.2). The theoretical background section 3.5.2 mentions that a reference dataset is used as reference, but may contain misclassification (Strahler et al. 2006). It is questionable if the reference data is validate as the reference data comes from multiple sources. Therefore it would be more correct to refer to the derived metrics of the confusion matrix as agreement instead of accuracy. Figure 12 presents the methodology of the external validation:

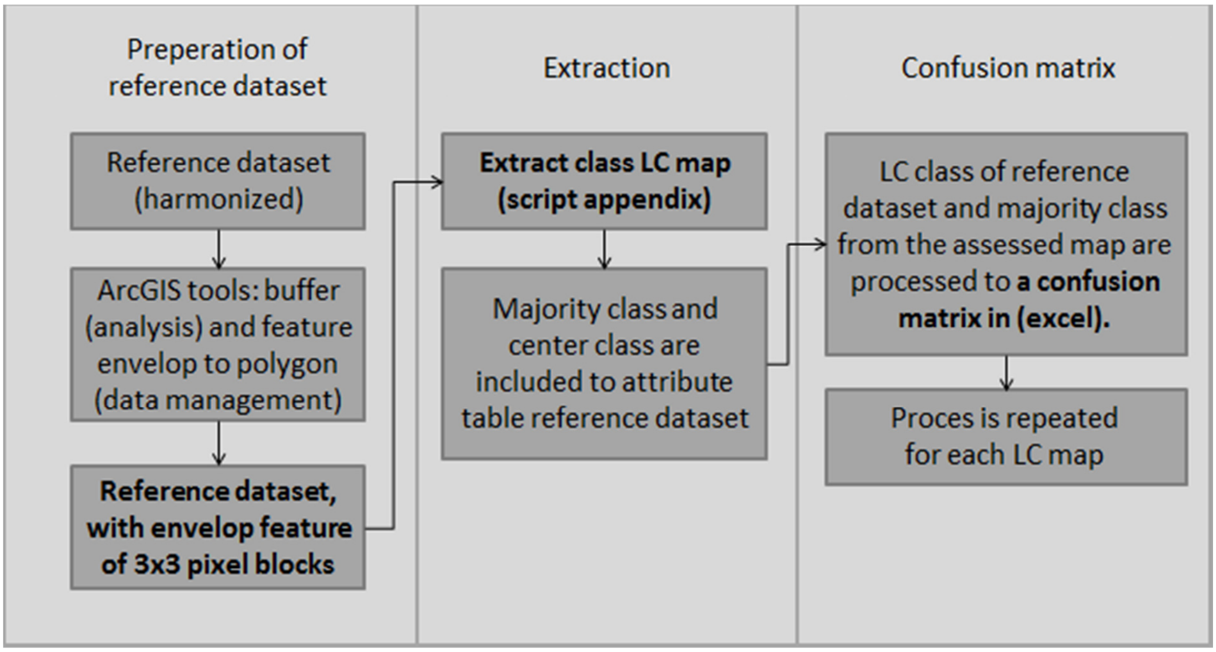

Figure 12: Methodology external validation

The reference dataset (section 4.1.2) was prepared for the external validation by using the buffer (analysis) and feature envelop to polygon (data management) tool to create an envelope around the point features of the reference datasets. The envelops were set to have a size of 3x3 pixel blocks, 0.00833x0.00833 decimal degree, with a point feature from the reference dataset as center point. Each envelop around a point feature was used to extract the majority class from the assessed LC maps with the external validation script in section 8.1.6 the appendix. The extracted majority class and center class from the assessed map were added to the attribute table of the reference dataset as a result (section 8.1.6 appendix). The original LC class of the reference dataset and the extracted majority class from the assessed map were processed to the format of a confusion matrix in excel. This was repeated for the normal voting, weighted voting, probability voting, FROM-GLC, Globcover (2009), LC-CCI (2010) and MODIS5 (2010) LC maps. The confusion matrices of the integration methods are presented in table 21 section 5.3.1. For the LC input maps the confusion matrices are presented in table 22 and table 23 in section 5.3.2.

An example of a confusion matrix is presented in table 18 where M indicates the assesed map, n indicates the category class in a map, m indicates the category class in the reference data and N are the pixels located in M, n, m of the matrix. Overall agreement (OA), producer agreement (PA) and user agreement (UA) were derived from the confusion matrices and calculated by equation 7.

$$OA = \frac{\sum_{c=1}^{c} N_{m,n}}{N} * 100$$

$$UA = \frac{N_{m,n}}{\sum N_{n+}} * 100 \qquad [7]$$

$$PA = \frac{N_{m,n}}{\sum N_{+m}} * 100$$

Source: Strahler et al. (2006)

Table 18: Example of a confusion matrix

Map: M		Reference classes C_m								Correct	$\sum C_n$	UA
		1	2	3	4	5	6	7	8			
Mapped classes C_n	1	$\boldsymbol{N_{11}}$	N_{12}	N_{13}	N_{14}	N_{15}	N_{16}	N_{17}	N_{18}	N_{11}	N_{1+}	…
	2	N_{21}	$\boldsymbol{N_{22}}$	N_{23}	N_{24}	N_{15}	N_{26}	N_{27}	N_{28}	N_{22}	N_{2+}	…
	3	N_{31}	N_{32}	$\boldsymbol{N_{33}}$	N_{34}	N_{35}	N_{36}	N_{37}	N_{38}	N_{33}	N_{3+}	…
	4	N_{41}	N_{42}	N_{43}	$\boldsymbol{N_{44}}$	N_{45}	N_{46}	N_{47}	N_{48}	N_{44}	N_{4+}	…
	5	N_{51}	N_{52}	N_{53}	N_{54}	$\boldsymbol{N_{55}}$	N_{56}	N_{57}	N_{58}	N_{55}	N_{5+}	…
	6	N_{61}	N_{62}	N_{63}	N_{64}	N_{65}	$\boldsymbol{N_{66}}$	N_{67}	N_{68}	N_{66}	N_{6+}	…
	7	N_{71}	N_{72}	N_{73}	N_{74}	N_{75}	N_{76}	$\boldsymbol{N_{77}}$	N_{78}	N_{77}	N_{7+}	…
	8	N_{81}	N_{82}	N_{83}	N_{84}	N_{85}	N_{86}	N_{87}	$\boldsymbol{N_{88}}$	N_{88}	N_{8+}	…
correct		N_{11}	N_{22}	N_{33}	N_{44}	N_{55}	N_{66}	N_{77}	N_{88}	$N_{m,n}$		
$\sum N_m$		N_{+1}	N_{+2}	N_{+3}	N_{+4}	N_{+5}	N_{+6}	N_{+7}	N_{+8}		N	OA
PA		…	…	…	…	…	…	…	…		OA	…

5. Results

This section holds the results from the integration methods, the internal validation and the external validation. The harmonized LC maps we used as input for the normal voting, weighted voting and probability voting integration methods. A difference plot shows the similarities and dissimilarities between the integrated LC maps. The internal and external validations were used to evaluate the integration methods.

5.1. Integration methods

This section holds the LC results from the normal voting, weighted voting and probability voting integration methods. Normal voting makes use of internal information from a voting procedure Weighted voting and probability voting respectively base their weights and probabilities on the global confusion matrices of Tsendbazar et al. (2016) and Yu et al. (2014). Accuracy assessments of external research were used because it was questionable if the reference data from this research was sufficient for producing reliable results, since certain LC classes were not represented by enough samples

5.1.1. Normal voting

Figure 13 presents the initial voting map (a) and the agreement/disagreement map (b) which indicates the five conditions (table 9) from the initial voting procedure. In the initial voting map, pixels were assigned to a LC class where the four input maps agree on a LC class. In case of a tie, the voting procedure remained undecided on a LC class. There are 5338552 pixels of the 78252697 pixels in the LC map where the voting procedure remains undecided which is 6.82% of the map.

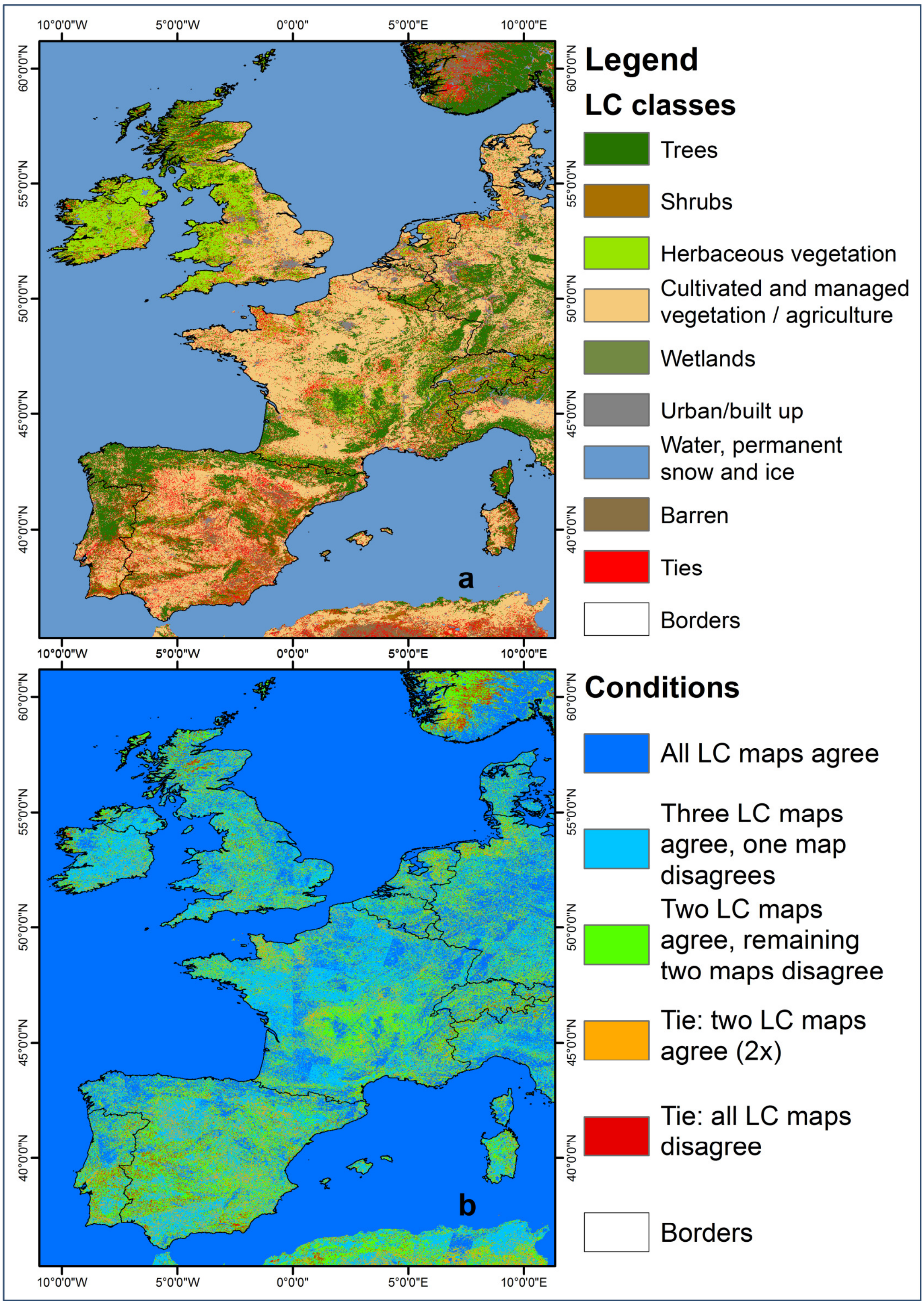

Figure 13: Initial voting map and its agreement/disagreement conditions from input maps

Ties were solved by using class preferences to decide on a LC class at that location, class preferences are presented in table 10 section 4.2.3. These class preferences were calculated from class occurrences in table 30 and table 31 of the appendix. Map occurrences (table 19) were calculated to further analyze the results of the initial voting. Map occurrences were calculated from table 30 and table 31 of the appendix by summing the class occurrence over all LC classes and identifying how much each input map has contributed to the initial voting map.

In 63.51% all input map agree on a LC class in the initial voting. LC-CCI contributed the most to the initial voting map. LC-CCI has the highest map occurrence where three maps agree and where two maps agree. Globcover and MODIS map occurrences are close to LC-CCI. FROM-GLC contributed the least to the initial voting map, where FROM-GLC disagrees with 25.72 % of the initial voting map.

Table 19: Map occurrence

LC classes 1 -8	Conditions			
	All agree	3 agree	2 agree	Disagree, ties
FROM-GLC	63.51	7.14	3.62	25.72
Globcover	63.51	15.54	6.76	14.19
LC-CCI	63.51	16,38	6.84	13.27
MODIS5	63.51	15.62	5.66	15.21

This section uses the observations from on class preferences (table 10, section 4.2.3) to describe the initial voting results. A high class preference demonstrates that the corresponding LC class is more often assigned to the normal voting map.LC-CCI holds for most of the LC classes the highest class preferences: (1) trees, (3) herbaceous vegetation, (5) wetlands, (6) urban/built up and (7) water, permanent snow and ice respectively 84.16%, 82.39%, 83.15%, 98.08% and 99.93%. Class preferences of Globcover and MODIS come close to the class preferences of LC-CCI. MODIS does not hold the highest class preference for any of the LC classes and has a very low class preference for (5) wetlands and (8) barren respectively 18.77% and 6.46%. Globcover holds the highest class preference for LC classes (4) cultivated and managed vegetation/agriculture and (8) barren respectively 88.27% and 92.48%. FROM-GLC holds the highest class preference for LC class (2) shrubs among the LC input maps. FROM-GLC generally holds the lowest class preferences over most of the LC classes, but holds a high preference for LC classes (1) trees, (2) shrubs and (7) water, snow and ice respectively 82.99%, 80.90% and 99.90%. FROM-GLC holds relative low class preferences because FROM-GLC has a low occurrence in the initial voting map.

The normal voting map is presented in figure 14 and was produced by joining the initial voting results with the "solved ties" in case of disagreeing conditions from the LC input maps.

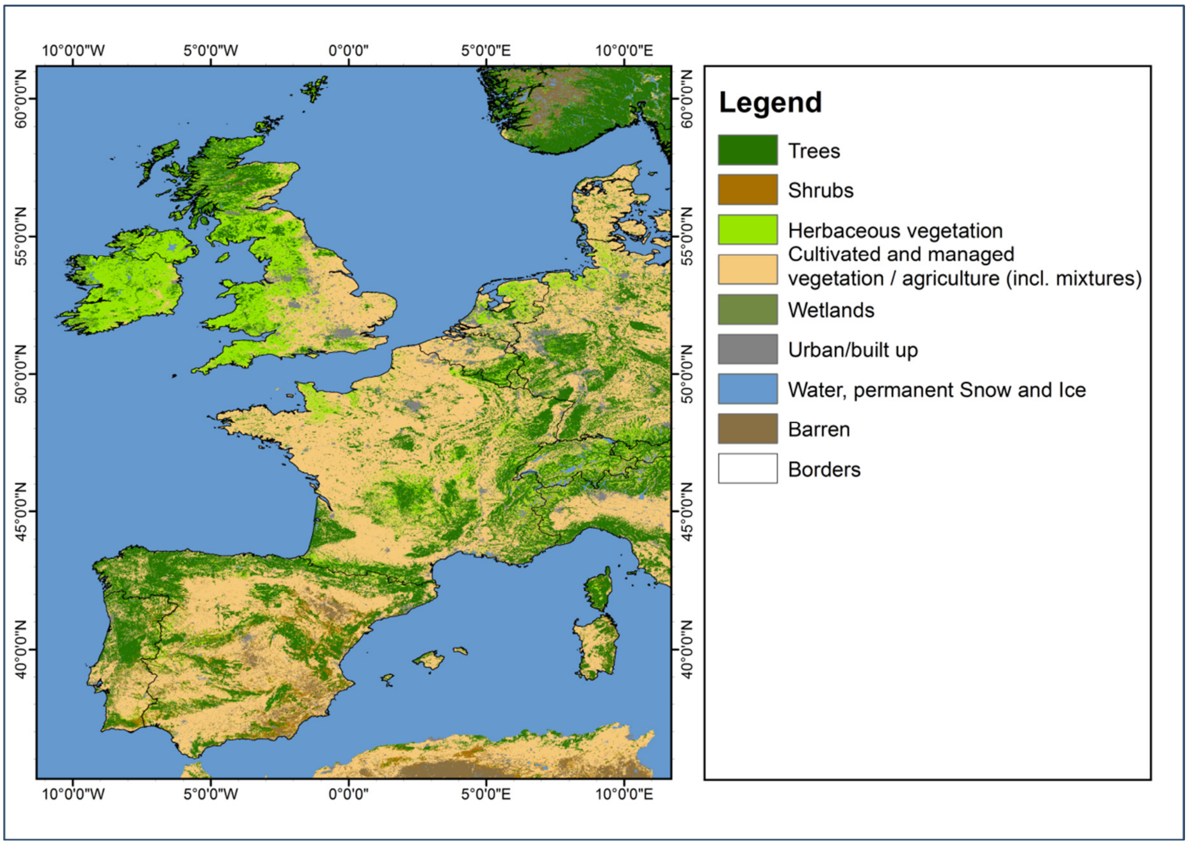

Figure 14: Integrated map obtained by normal voting

5.1.2. Weighted voting

Figure 15 presents the resulted weighted voting LC map (a) and the highest accumulated weight (b) over eight LC classes from which the integrated map was generated. The weights are based on the UA of each LC class of the input maps. Presented in table 8 section 4.1.1 and table12 section 4.2.2, both tables holds the UA of each LC input map obtained from the research of Tsendbazar et al. (2016) and Yu et al. (2014). Class maximum weights were generally high in the weighted map (figure 15b), but holded less contrast than the maximum probabilities of the probability voting method (figure 17b). Class weights are low in area where the LC input maps disagree (figure 13b).

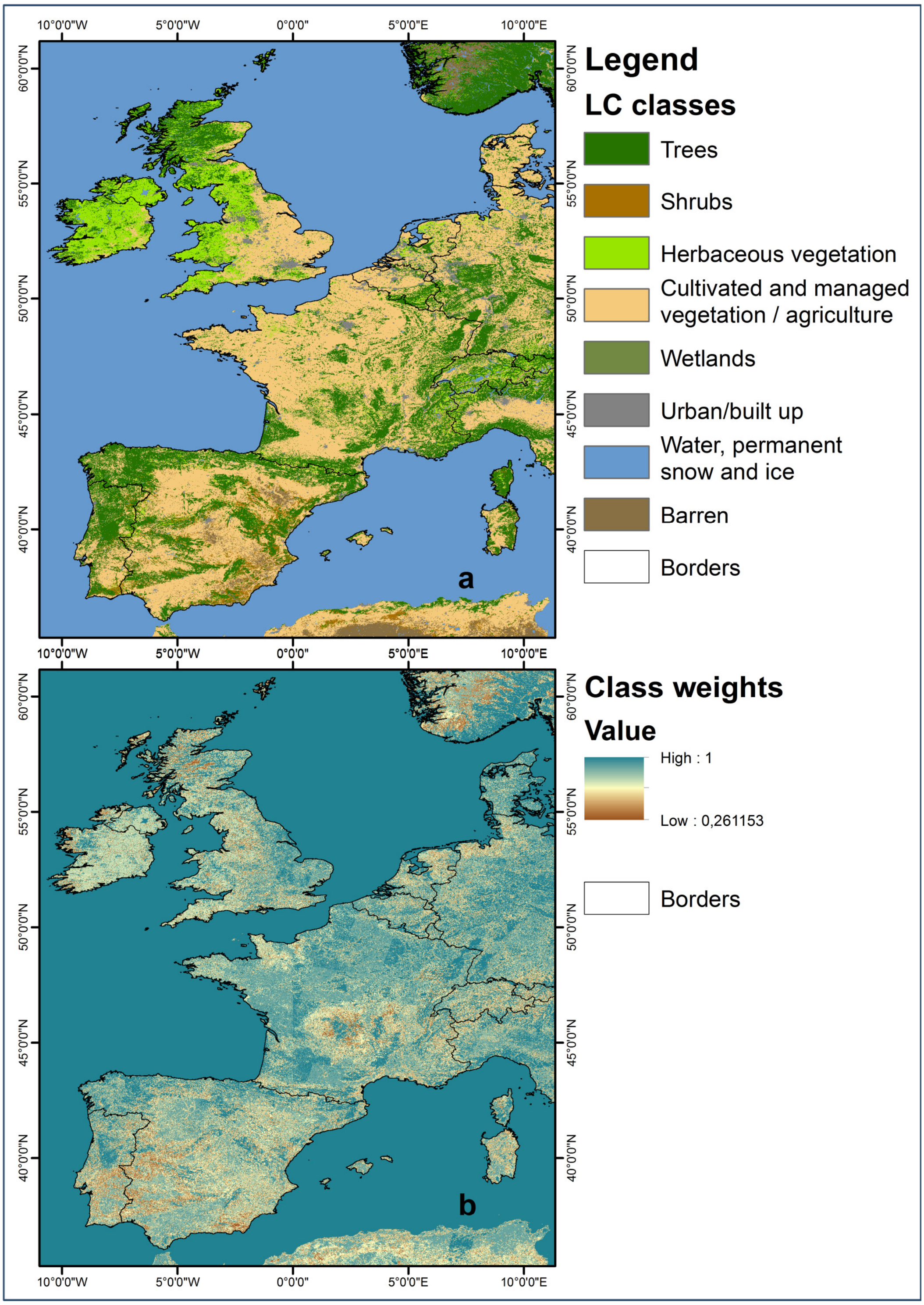

Figure 15: Integrated map obtained by weighted voting and its class weights

Figure 16 presents the weights for each of the eight LC classes. LC classes (1) trees, (4) cultivated and managed vegetation/agriculture and (7) water, permanent snow and ice hold high weights to be assigned at locations within the integrated map. Common LC classes with high weights were favored in the weighted voting process. There is a high weight for a LC class when LC input maps agree on a LC class. Weighted voting has less contrast between its weights than the probability voting maps (figure 18).

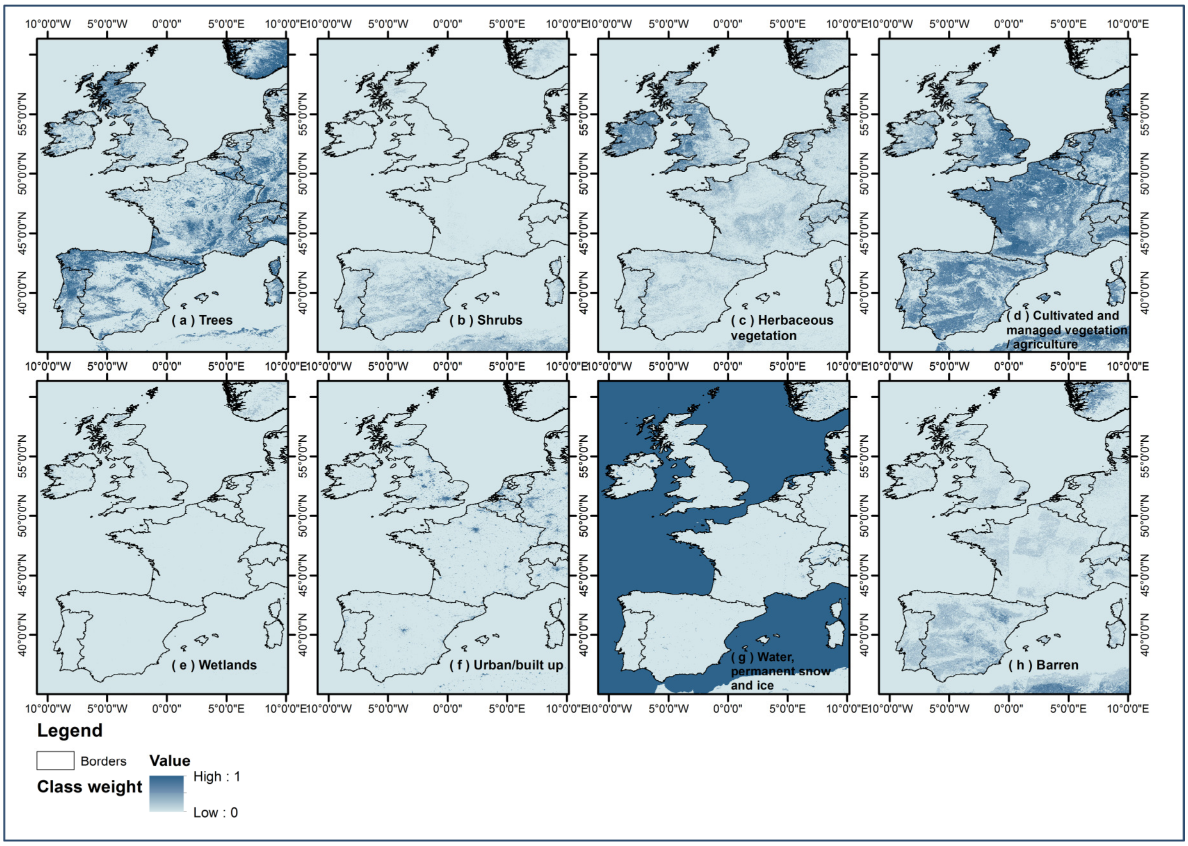

Figure 16: Class specific weights LC classes

This section uses the observations from table 8 (section 4.1.1) that hold the OA, UA and PA of the LC input maps to describe the weighted voting results. A high UA means a high weight input for the corresponding LC class of the input map, since the weights are based on the UA.

LC-CCI (2010) has the highest OA of 74. 70% from the input maps. LC-CCI and MODIS each hold the highest UA for three LC classes amongst the eight LC classes of each input map. LC-CCI for LC classes:(1) trees (2) shrubs and (7) water, permanent snow and ice with respectively weights of 90.39%, 61.67% and 85.33%. MODIS for LC classes: (4) cultivated and managed vegetation/agriculture, (6) urban/built up and (8) barren with respectively weights of 84.60%, 85.71% and 98.28%. FROM-GLC has the highest UA for LC class (3) herbaceous vegetation and Globcover has the highest UA for LC class (5) wetlands among the input maps. The high presence of trees in FROM-GLC LC map with the relative high UA for LC class (1) trees, could cause an overestimation of trees in the weighted voting map. The UA for LC class (5) wetlands from FROM-GLC is unknown because the confusion matrix from Yu et al. (2014) holds no information for wetlands, which is class 50 in the FROM-GLC classification system. There is an average difference 28.94% in the UA of a specific LC class, there are no significant differences with the exception of LC class (5) wetlands.

5.1.3. Probability voting

In each location of the probability voting map, a pixel is assigned to the LC class that holds the highest probability at the pixel's location. Figure 17 presents the resulted probability voting LC map (a) and the highest class probabilities (b) from which the LC map is generated. Maximum class probabilities are generally high in the map and low in area where the LC input maps disagree (figure 13b). The areas with low probability are typically located in areas with fragmented landscape, high slope or transition areas, for example: the mountain areas in Norway or the French Alps.

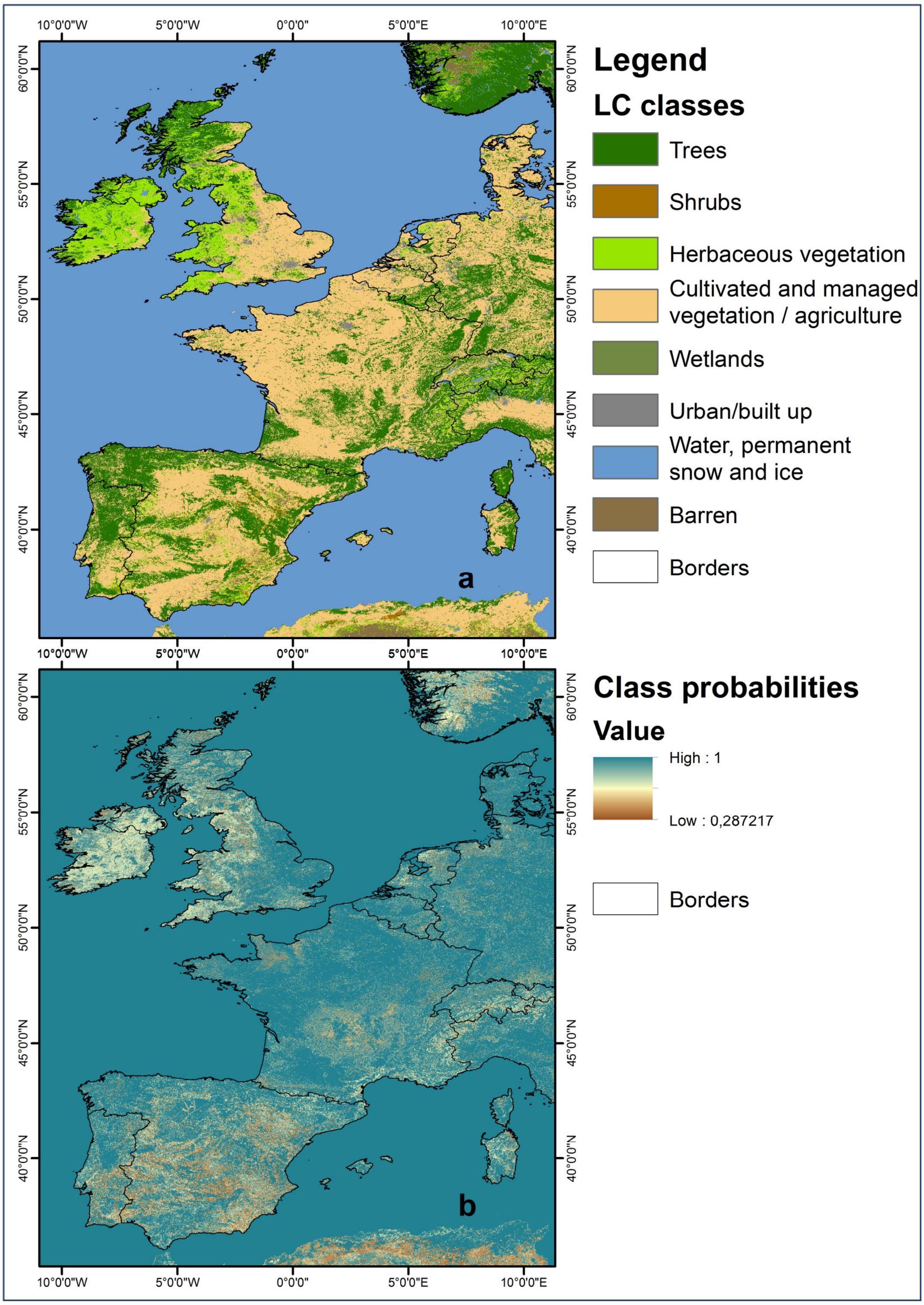

Figure 17: Integrated map obtained by probability voting and its class probability

Figure 18 presents the probability for each of the eight LC classes. LC classes (1) trees, (3) herbaceous vegetation, (4) cultivated and managed vegetation/agriculture and (7) water, permanent snow and ice have a high probability to be assigned at locations within the integrated map. LC class (7) water, permanent snow and ice had a high class probability and a relative low probability to be seen for another class (table 14). Additionally, in most locations, all input maps agree on LC class (7) water, permanent snow and ice, which confirm a high probability. Like weighted voting, common LC classes with good probability are favored in the voting process. Probability voting has more contrast between high and low class probability since class probabilities of the input maps are multiplied before normalization (table 15, section 4.2.3). Weighted voting methodology summed class weights before normalization (table 13, section 4.2.2) and therefore holds less contrast between high and low weights (figure 15 and figure 16).

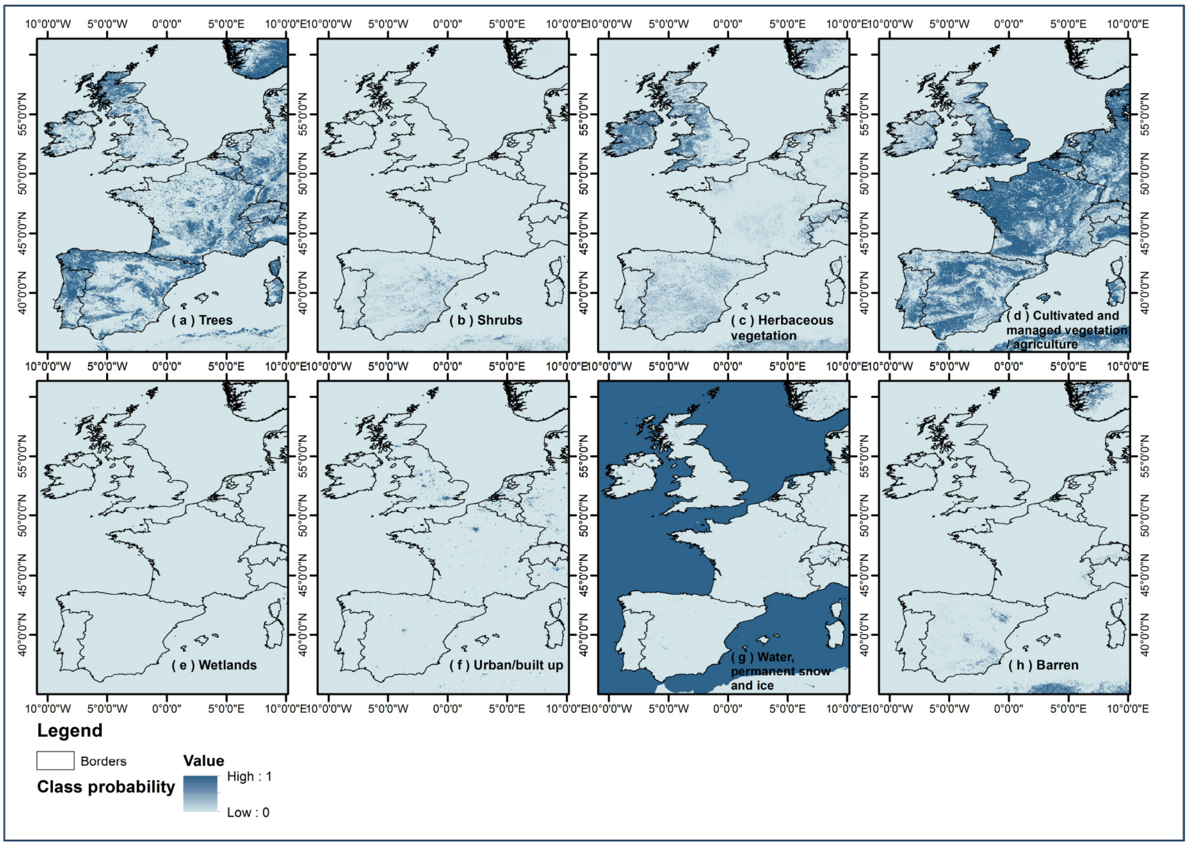

Figure 18: Class specific probability of LC classes

This section further describes class probabilities and the probability of being another LC class; these are observed from table 14 (section 4.2.3) and figure 18. Table 14 holds the probabilities of each LC class from FROM-GLC, Globcover 2009, LC-CCI (2010) and MODIS5 (2010), based on the research of Tsendbazar et al. (2016) and Yu et al. (2014).

Generally LC classes (2) shrubs, (3) herbaceous vegetation and (4) cultivated and managed vegetation/agriculture have a relative high probability to be LC classes (1) trees, 2, 3 and 4. This is the case for FROM-GLC, in Globcover this only accounts for LC classes (2) shrubs and (3) herbaceous vegetation. LC class (2) shrubs from LC-CCI holds the probability of being (1) trees and (3) herbaceous vegetation where (3) holds the probability to be seen as (4) cultivated and managed vegetation/agriculture. In MODIS, LC class (2) shrubs holds the probability to be LC class (1) trees and (3) herbaceous vegetation where (3) holds the probability to be seen as (2) shrubs, (4) cultivated and managed vegetation/agriculture and (8) barren.

The input maps hold more differences in the class probabilities of LC classes: (5) wetlands, (6) urban/built up, (7) water, permanent snow and ice and (8) barren. LC classes (6) urban/built up of FROM-GLC holds the probability to be LC class (4) cultivated and managed vegetation/agriculture and (8) barren where (8) holds the probability to be seen for LC class (3) herbaceous vegetation. In Globcover, LC classes (5) wetlands and (8) barren hold the probability to be LC class (3) herbaceous vegetation and LC class (6) urban/built up holds the probability to be LC class (4) cultivated and managed vegetation/agriculture. In LC-CCI LC classes (5) wetlands, (6) urban/built up and (8) barren respectively hold the probability to be mistaken for LC class (1) trees, (4) cultivated and managed vegetation/agriculture and (3) herbaceous vegetation. LC classes (5) wetlands and (7) water, permanent snow and ice from MODIS respectively hold the probability to be mistaken for LC class (1) trees and (8) barren.

5.1.4. Difference plots

Different plots were used to view the similarities and dissimilarities between the voting methods. Different plots do not assess the integration methods but give an overview where the methods disagree on LC classes. Figure 19 shows a difference plot between the integrated map and table 20 holds the number of pixels from the difference plot. Differences between the integration methods occur, where, respectively weights from weighted voting (figure 15) and class probabilities from probability voting (figure 17) and where low and LC input maps disagreed in the initial voting (figure 13b). Meaning that, integration methods disagreed where methods are less certain in assigning a pixel to a specific LC class. Differences between the integration methods typically occur in heterogeneous areas, fragmented landscape, and transition areas. For example, the mountain areas in Norway or the French Alps have fragmented landscape due to high slope. Bretagne in north-east France and the Netherlands hold differences between the integration methods, but this cannot be explained by heterogeneity, fragmentation of the landscape. Table 20 holds information on these results, where specific methods disagree in the disagreement areas plotted in figure 19.

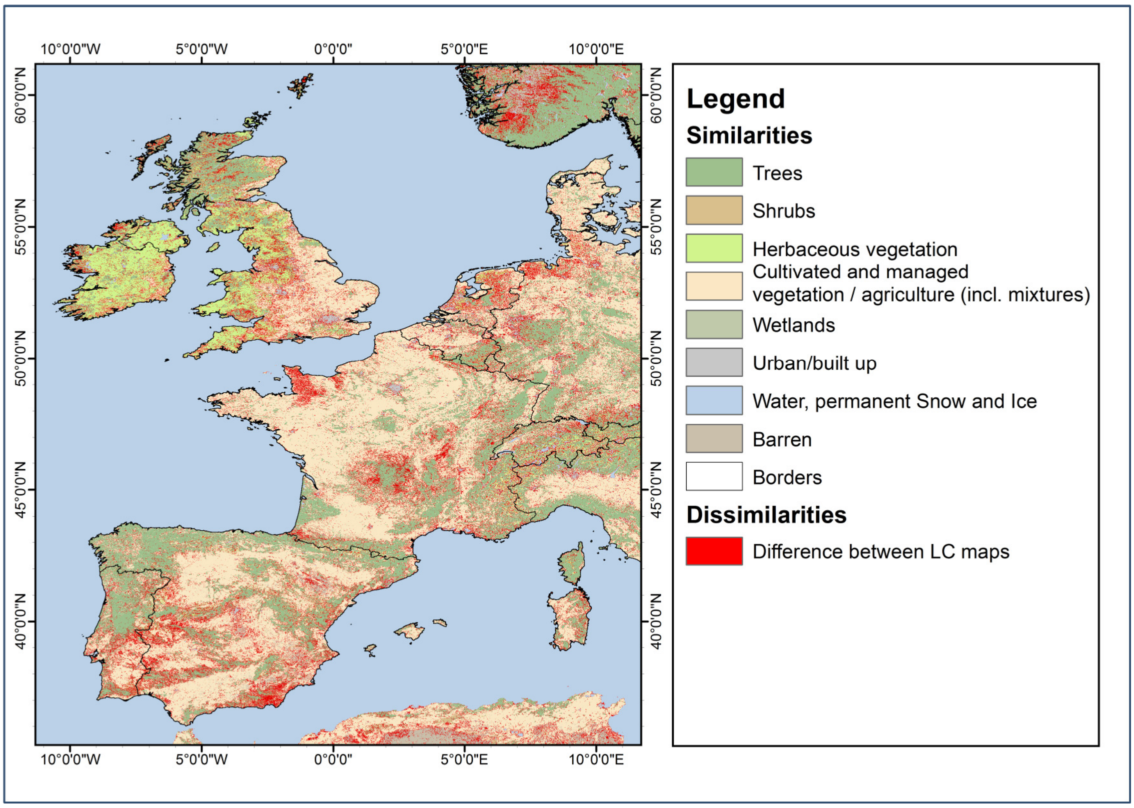

Figure 19: Difference plot, similarities and dissimilarities between LC maps

5610888 pixels of the 78252697 pixels disagree in the difference plot, this is a 7.17% difference between the LC maps. When excluding the LC class (7) water, permanent snow and ice there is still a 15.76% difference between the land classes of the LC maps. From the disagreement areas, 96.74% of the pixels differ between normal voting and probability voting. Weighted voting has less difference with the other maps in the disagreement area, a 45.54% pixel difference with normal voting and a 62.65% pixel difference with probability voting. From the 5610888 pixels that disagree between the integration methods, there is only a 4.93% pixel difference where all three maps disagree and 47.61% of those pixels are located in tie locations from the initial voting.

Table 20: Information on difference plot

Agreement / disagreement integration methods		
Agreement between maps	92.83	%
Disagreement between maps	7.17	%
Description of disagreement inside the difference plot		
Disagreement between normal voting and weighted voting	45.54	%
Disagreement between normal voting and probability voting	96.74	%
Disagreement between weighted voting and probability voting	62.65	%
Where all integrated maps disagree	4.93	%
Where maps disagree in tie locations from initial voting	47.61	%

5.2. Internal validation

This section presents the results of the internal validation. Figure 20 holds the information entropy calculated over the integration methods LC maps: normal voting (a), weighted voting (b) and probability voting (c). A high entropy value stands for a high uncertainty in the LC classification and a low entropy value represents certainty. The information entropy of normal voting and weighted voting is similar, as both maps have the same patterns over the study area. Weighted voting seems had slightly lower values in areas with high uncertainty than normal voting, for example; the high uncertainty in the mountain areas of Norway. The information entropy calculated over probability voting was dissimilar from the other integration methods. Probability voting had the lowest uncertainty in comparison to the other methods and provides the best classification based on the information entropy. Most areas have classification certainty, a low uncertainty; these areas have a high contract with areas of high uncertainty.

Areas with high uncertainty seem to be roughly highest in areas where the normal voting, weighted voting and probability voting LC maps disagree (figure 19). Similarly, areas with high uncertainty (figure 20) concur partly with the tie locations (figure 13) from the initial process of normal voting. This is logical for normal voting as the information entropy is implemented on class presence of the LC input maps (section 4.3.1), the five conditions from the normal voting methodology in table 9 section 4.2.1. Section 5.2.4 reports with table 20 that 47.61% of the locations where the methods disagree (figure 20), concur with tie locations (figure 13) from the initial voting.

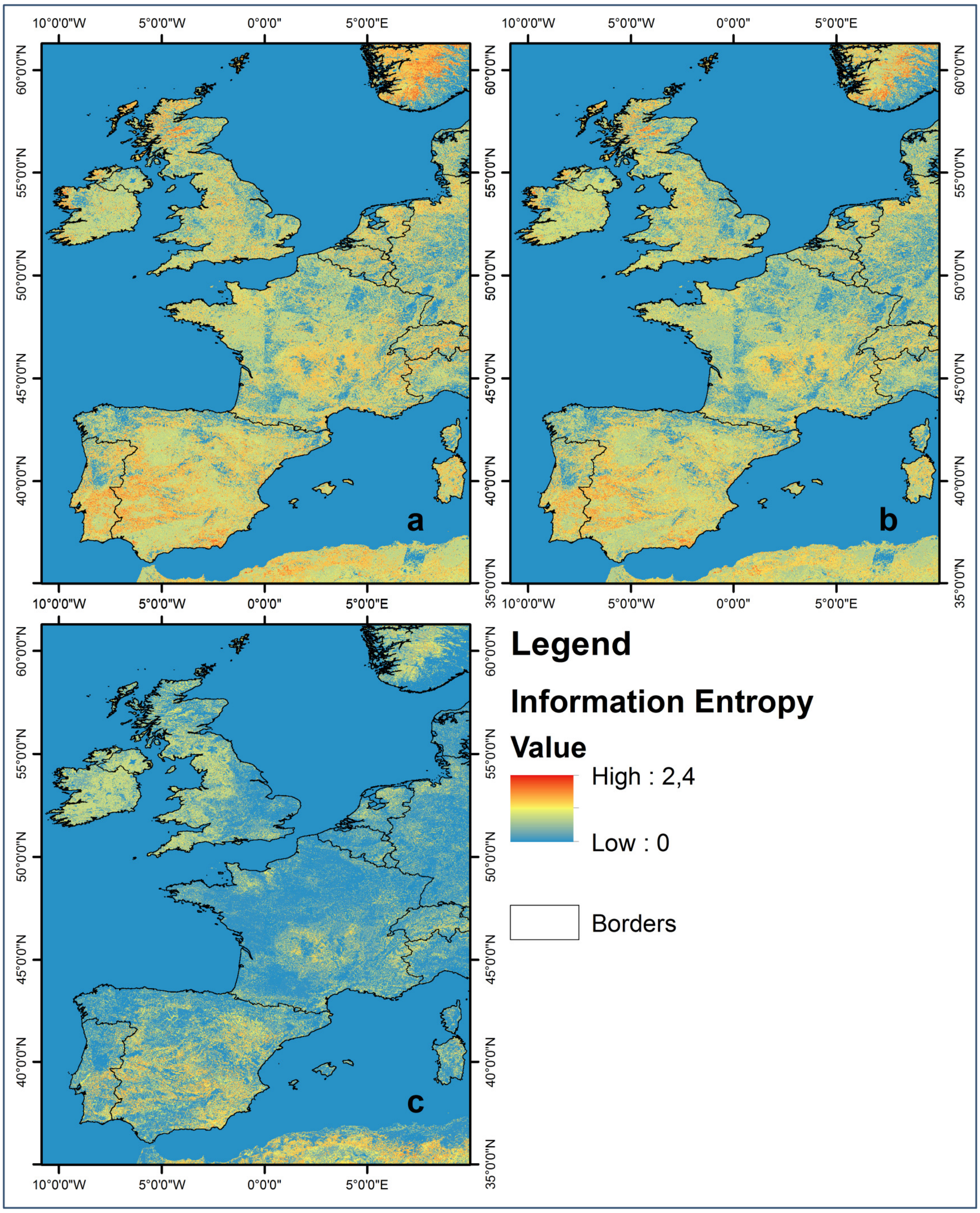

Figure 20: Information entropy over LC maps of integration methods

5.3. External validation

Confusion matrices were used for the external validation of the integrated LC maps: normal voting, weighted voting and probability voting. This section presents the results from the confusion matrices of the input maps and integrated LC maps. The integration methods were evaluated on their improvement, compared to the assessment of the other methods and input maps. Figure 27, figure 28 and figure 30 present the derived metrics in a graph for the discussion of the results in section 6.1.4.

5.3.1. Integration methods

Table 21 presents the confusion matrices of the integration methods. Table 24, table 25 and table 26 hold an overview of the metrics derived from the confusion matrices: overall agreement, producer agreement and user agreement of the input map with reference dataset.

Table 21: Confusion matrices of normal voting, weighted voting and probability voting

Normal Voting		Reference classes*								Correct	Total	UA
		1	2	3	4	5	6	7	8			
Mapped classes	1	169	22	19	27	0	2	0	3	169	242	69.83
	2	10	16	16	1	0	0	0	0	16	43	37.21
	3	6	8	64	42	1	1	0	1	64	123	52.03
	4	31	17	26	347	1	5	0	2	347	429	80.89
	5	0	0	0	0	1	0	0	0	1	1	100.00
	6	5	0	0	4	0	36	0	0	36	45	80.00
	7	0	0	0	1	1	1	13	0	13	16	81.25
	8	3	2	5	4	0	0	0	3	3	17	17.65
Correct		169	16	64	347	1	36	13	3	649		
Total		224	65	130	426	4	45	13	9		916	OA
PA		75.45	24.62	49.23	81.46	25.00	80.00	100.0	33.33		OA	70.85
Weighted Voting		Reference classes*								Correct	Total	UA
		1	2	3	4	5	6	7	8			
Mapped classes	1	185	26	24	31	1	4	0	3	185	274	67.52
	2	8	16	17	1	0	0	0	0	16	42	38.10
	3	3	8	39	24	1	0	0	2	39	77	50.65
	4	23	14	46	363	1	3	0	1	363	451	80.49
	5	0	0	0	0	1	0	0	0	1	1	100.00
	6	4	0	0	3	0	37	0	0	37	44	84.09
	7	0	0	0	1	0	1	13	0	13	15	86.67
	8	1	1	4	3	0	0	0	3	3	12	25.00
Correct		185	16	39	363	1	37	13	3	657		
Total		224	65	130	426	4	45	13	9		916	OA
PA		82.59	24.62	30.00	85.21	25.00	82.22	100,0	33.33		OA	71.72
Probability Voting		Reference classes*								Correct	Total	UA
		1	2	3	4	5	6	7	8			
Mapped classes	1	187	37	28	35	2	2	3	3	187	297	62.96
	2	2	8	3	0	0	0	0	0	8	13	61.54
	3	5	6	55	22	1	0	0	4	55	93	59.14
	4	28	13	44	368	1	15	1	1	368	471	78.13
	5	0	0	0	0	0	0	0	0	0	0	-
	6	2	0	0	0	0	26	0	0	26	28	92.86
	7	0	0	0	1	0	2	9	0	9	12	75.00
	8	0	1	0	0	0	0	0	1	1	2	50.00
Correct		187	8	55	368	0	26	9	1	654		
Total		224	65	130	426	4	45	13	9		916	OA
PA		83.48	12.31	42.31	86.38	0.00	57.78	69.23	11.11		OA	71.40

5.3.2. Land cover input maps

Table 22 and table 23 present the confusion matrices of the LC input maps. The agreement assessments of the LC input maps with reference datasets were used as basis to evaluate the improvement of the integration methods compared to their input maps. Table 24, table 25 and table 26 hold an overview of the metrics derived from the confusion matrices: overall agreement, producer agreement and user agreement of the input map with reference dataset.

Table 22: Confusion matrices of FROM-GLC and Globcover (2009)

FROM-GLC		Reference classes*								Correct	Total	UA
		1	2	3	4	5	6	7	8			
Mapped classes	1	166	12	32	57	0	9	0	2	166	278	59.71
	2	21	21	21	13	0	1	0	1	21	78	26.92
	3	20	14	39	69	2	14	1	3	39	162	24.07
	4	6	4	20	141	0	4	0	0	141	175	80.57
	5	1	1	0	0	0	0	0	2	0	4	0.00
	6	1	0	1	7	0	15	0	0	15	24	62.50
	7	0	0	0	1	2	1	12	1	12	17	70.59
	8	9	13	17	138	0	1	0	0	0	178	0.00
Correct		166	21	39	141	0	15	12	0	394		
Total		224	65	130	426	4	45	13	9		916	OA
PA		74.11	32.31	30.00	33.10	0.00	33.33	92.31	0.00		OA	43.01
Globcover (2009)		Reference classes*								Correct	Total	UA
		1	2	3	4	5	6	7	8			
Mapped classes	1	138	15	21	47	0	4	1	2	138	228	60.53
	2	11	20	8	2	0	0	0	1	20	42	47.62
	3	19	3	53	64	0	1	0	1	53	141	37.59
	4	49	25	44	302	2	15	1	2	302	440	68.64
	5	1	0	0	0	1	0	0	0	1	2	50.00
	6	3	0	0	4	0	25	0	0	25	32	78.13
	7	0	0	0	1	0	0	11	0	11	12	91.67
	8	3	2	4	6	1	0	0	3	3	19	15.79
Correct		138	20	53	302	1	25	11	3	553		
Total		224	65	130	426	4	45	13	9		916	OA
PA		61.61	30.77	40.77	70.89	25.00	55.56	84.62	33.33		OA	60.37

Table 23: Confusion matrices of LC-CCI (2010) and MODIS5 (2010)

LC-CCI (2010)		Reference classes*										
		1	2	3	4	5	6	7	8	Correct	Total	UA
Mapped classes	1	167	17	16	18	0	0	0	2	167	220	75.91
	2	3	17	2	0	0	0	0	1	17	23	73.91
	3	9	4	66	68	1	1	0	1	66	150	44.00
	4	32	13	17	293	0	2	0	0	293	357	82.07
	5	0	3	3	3	3	0	0	0	3	12	25.00
	6	9	1	1	14	0	42	0	0	42	67	62.69
	7	1	0	0	1	0	0	13	0	13	15	86.67
	8	3	10	25	29	0	0	0	5	5	72	6.94
Correct		167	17	66	293	3	42	13	5	606		
Total		224	65	130	426	4	45	13	9		916	OA
PA		74.55	26.15	50.77	68.78	75.00	93.33	100.00	55.56		OA	66.16
MODIS5 (2010)		Reference classes*										
		1	2	3	4	5	6	7	8	Correct	Total	UA
Mapped classes	1	157	31	20	31	0	4	4	3	157	250	62.80
	2	11	16	25	8	1	0	0	0	16	61	26.23
	3	6	12	46	27	1	1	0	6	46	99	46.46
	4	45	5	39	355	2	4	1	0	355	451	78.71
	5	0	0	0	0	0	0	0	0	0	0	-
	6	5	1	0	5	0	35	1	0	35	47	74.47
	7	0	0	0	0	0	1	7	0	7	8	87.50
	8	0	0	0	0	0	0	0	0	0	0	-
Correct		157	16	46	355	0	35	7	0	616		
Total		224	65	130	426	4	45	13	9		916	OA
PA		70.09	24.62	35.38	83.33	0.00	77.78	53.85	0.00		OA	67.25

5.3.3. Overview of agreement metrics

This section gives an overview of the agreement metrics from normal voting, weighted voting and probability voting integration methods and the FROM-GLC hierarchy, Globcover 2009, LC-CCI (2010) and MODIS5 (2010) LC maps. OA, UA and PA are presented in tables: 24, 25 and 26.

Integrated LC maps:

The overall agreement of weighted voting was the highest among the integration methods with an agreement of 71.72% (table 26). The results present no significant difference between the overall agreement of the three integration methods, with a difference of 0.87% agreement between the highest (weighted voting) and lowest (normal voting) overall agreement.

Table 24: Overall agreement of integrated methods and input maps

LC map	OA
Normal voting	70.85
Weighted voting	71.72
Probability voting	71.40
FROM-GLC-hierarchy	43.01
Globcover (2009)	60.37
LC-CCI (2010)	66.16
MODIS5(2010)	67.25

Observations of class agreements between the methods (table 25 and table 26):

1. Normal voting and weighted voting seem to achieve similar results in the PA and UA of LC class (5) wetlands and in the PA of LC class (2) shrubs and (7) water, permanent snow and ice and (8) barren. This with just slight differences between their confusion matrices for these classes.
2. There is no information on metrics for LC class (5) wetlands in probability voting, because none of the samples from the reference dataset agree with the LC map. This could be caused by the low number of four samples in LC class (5) wetlands from the reference dataset (table 6 section 3.3.1).
3. Probability voting holds low agreement metrics for LC class (7) water, permanent snow and ice compared to the other methods, a 69.23% producer agreement and 75.00% user agreement.
4. Probability voting achieved the most LC class improvement among the integration methods and could possibly achieve higher results when classes like LC class (5) wetlands and LC class (5) wetlands have more samples.

Table 25 present the user agreement of the LC maps produced by the integration methods and their input LC maps.

Table 25: User agreement of integrated methods and input maps

LC map		Normal voting	Weighted voting	Probability voting	FROM-GLC	Globcover (2009)	LC-CCI (2010)	MODIS5 (2010)
		UA	UA	UA	UA	UA	UA	UA
Class agreement	1	69.83	67.52	62.96	59.71	60.53	75.91	62.80
	2	37.21	38.10	61.54	26.92	47.62	73.91	26.23
	3	52.03	50.65	59.14	24.07	37.59	44.00	46.46
	4	80.89	80.49	78.13	80.57	68.64	82.07	78.71
	5	100.00	100.00	-	0.00	50.00	25.00	-
	6	80.00	84.09	92.86	62.50	78.13	62.69	74.47
	7	81.25	86.67	75.00	70.59	91.67	86.67	87.50
	8	17.65	25.00	50.00	0.00	15.79	6.94	-

Table 26 present the producer agreements of the LC maps produced by the integration methods and their input LC maps.

Table 26: Producer agreement of integrated methods and input maps

LC map		Normal voting	Weighted voting	Probability voting	FROM-GLC	Globcover (2009)	LC-CCI (2010)	MODIS5 (2010)
		PA	PA	PA	PA	PA	PA	PA
Class agreement	1	75.45	82.59	83.48	74.11	61.61	74.55	70.09
	2	24.62	24.62	12.31	32.31	30.77	26.15	24.62
	3	49.23	30.00	42.31	0.00	40.77	50.77	35.38
	4	81.46	85.21	86.38	33.10	70.89	68.78	83.33
	5	25.00	25.00	0.00	0.00	25.00	75.00	0.00
	6	80.00	82.22	57.78	33.33	55.56	93.33	77.78
	7	100.00	100.00	69.23	92.31	84.62	100.00	53.85
	8	33.33	33.33	11.11	0.00	33.33	55.56	0.00

LC input maps:

All integrated maps have an improved overall agreement compared to the input maps (table 24). MODIS5 has the highest overall agreement between the input maps with 67.25%, slightly higher than LC-CCI with an overall agreement of 66.16%. The reported accuracies of the input maps table 8 (Tsendbazar et al. 2016; Yu et al. 2014) in section 4.1.1 present LC-CCI with an overall accuracy of 74.70% and MODIS5 with an overall accuracy of 73.92% when their accuracy assessment is harmonized to the eight LC classes. Tsendbazar et al. (2016) research presents the confusion matrices of Globcover, LC-CCI and MODIS5 with 13 generalized LC classes were Globcover, LC-CCI and MODIS5 have an overall accuracy reported of respectively 61.3%, 70.8% and 71.4%.

LC-CCI holds the highest producer accuracy over the integration methods for LC classes: (3) herbaceous vegetation, (5) wetlands, (6) urban/built up, (7) water, permanent snow and ice and (8) barren. LC-CCI, normal voting and weighted voting LC maps hold a 100% producer for LC class (7) water, permanent snow and ice. LC-CCI further holds a higher user agreement over the integration methods for LC classes (1) trees, (2) shrubs and (4) cultivated and managed vegetation/agriculture. FROM-GLC holds the highest producer agreement for LC class (2) shrubs and Globcover holds the highest user agreement for LC class (7) water, permanent snow and ice. In other cases, one or all of the integration methods hold higher agreement in the metrics.

Observations of class agreement between the methods and input map (table 25 and table 26):

1. Probability voting has improved class agreements, but LC-CCI more often hold the highest agreement metrics for a LC class compared to the methods and input maps.
2. The highest producer agreement for LC class (2) shrubs from the integration methods is similar to the lowest producer agreement for LC class (2) shrubs from the input maps (MODIS5).
3. Probability voting, FROM-GLC and MODIS5 hold no information for the metrics of LC class (5) wetlands, this could be caused by the low number of four samples for LC class (5) barren from the reference dataset (table 6 section 3.3.1).
4. FROM-GLC and MODIS5 hold no information for the metrics of LC class (8) barren, like LC class (5) wetlands, LC class (8) barren has low number of nine samples (table 6 section 3.3.1).

6. Discussion, conclusion and recommendations

This chapter gives an overview of the research by; discussing the results and methodology of this research, drawing conclusions, restrictions of this research and recommendations.

6.1. Discussion

This section discusses the methodology and results of the research: harmonization, integration method, the most accurate method based on the information entropy and the amount of improvement in the integrated LC maps.

6.1.1. LC data

FROM-GLC seems to have a lower accuracy than the other input maps for the integration. Yu et al. (2014) reports an overall accuracy of 66.10 % for FROM-GLCagg, were Tsendbazar et al. (2016) reports on a Globcover (2009), LC-CCI (2010) and MODIS5 (2010) overall accuracy of respectively: 67.81%, 74.70% and 73.92% (table 8 section 4.1.1). But it should be noted that these accuracy assessments come from different sources and therefore cannot be compared directly (section 4.1.1). The accuracy assessment used for FROM-GLC-hierarchy does belong to the 30m base map of FROM-GLCagg and might not accurately represent FROM-GLC-hierarchy, but it is unknown how much the pre-processing from FROM-GLCagg to FROM-GLC hierarchy influences the maps accuracy (section 2.2.1). The confusion matric of FROM-GLCagg holds no information for LC class (5) wetlands, in the original legend class 50 which also stand for wetlands (section 2.2.1). Additionally FROM-GLC-hierarchy has a tilling effect over the LC map (figure 7a section 4.1.1). This all suggests that FROM-GLC-hierarchy holds a low accuracy in Western Europe. LC classes (5) wetlands and (8) barren are represented by a low number of sample sites, respectively four and nine samples (section 3.3.2). A higher number of sample sites could give a better result for LC classes (5) wetlands and (8) barren in the external validation. Table 6 (section 3.3.1) presents the distribution of the samples over the harmonized LC classes.

6.1.2. Harmonization

Herold et al. (2006) mention that one universal legend would provide too much standardization and reduce the relevance and applicability for many applications. Harmonization to eight general LC classes reduces the applicability to discriminate between problematic classes. For example the harmonization of all tree classes to one LC class makes it impossible to characterize forest in detail. The spatial harmonization of the reference dataset is based on the assumption that the sample unit area has homogenous LC type, so the datasets were harmonized to have the same extent of sample units. In reality this is not always true; some areas have heterogeneous LC types.

6.1.3. Integration methods

Normal voting was produces from a common voting procedure, were a pixel was assigned to the class that occurs in the majority of the LC maps at that pixel's location. In the methodology (section 4.2.1) is explained that ties occur with this procedure were the four LC input maps disagree on a LC class. Ties are assigned to a LC class by giving a high class preference to LC classes that occur more often in the initial voting: these class preferences were presented in table 10. The normal voting LC map was produced by merging the initial voting results with the "solved ties" in case the LC input maps disagree. This is a new approach used in this research, solving ties based on class preferences (table 10) that are calculated from class occurrences of each input map in the initial voting (table 30 and table 31). The advantage of this method is that it's purely map driven and rejects input maps with a low contribution in the initial process in case maps disagree and form a tie. The disadvantage of this method would be that it does not use accuracy information which could reject inaccurate classes in the beginning of the voting process.

In weighted voting, a pixel is assigned to the class that accumulates the highest weight at that pixel's location. Weighted voting derives weights from the user accuracy of each LC class which were obtained from the published confusion matrices (Yu et al. 2014) (Tsendbazar et al. 2016). Weights based on class accuracies could improve the integrated map by giving preference to accurate classes and reject inaccurate classes of the input maps. In weighted voting studies the research of Ge et al. (2014) and Iwao et al. (2011), weights were applied in a different manner. Ge et al. (2014) assigns weights to LC classes according to the accuracy of the mapped class and Iwao et al. (2011) gives preference to each LC class from the map with the highest overall accuracy (OA). Ge et al. (2014) looks at accuracies from the perspective of the source map and Iwao et al. (2011) uses overall accuracies but assesses the maps from the user perspective. The advantage of weighted voting is that the method uses accuracy information for map integration, which could reject inaccurate classes. A disadvantage would be that such an integration method favors LC classes with a high class weight and therefore tends to over-map these LC classes, for example: LC class (1) trees and (7) water, permanent snow and ice.

In probability voting, a voting procedure is applied on the probabilities of each class being the correct class. A pixel is assigned to the LC class that accumulates the highest probability at the pixel's location. Kinoshita et al. (2014) and Tuanmu and Jetz (2014) used probability voting in previous studies. Kinoshita et al. (2014) achieves improved map accuracy by performing a logistic regression analysis with class probabilities, but the regression analysis is not used in this research as the scope focuses on integration methods with a voting approach. Probabilities of harmonized LC classes were calculated with the same equation Kinoshita et al. (2014) used to calculate class probabilities. Probabilities were calculated from the user perspective in this research, because this represents the agreement of the LC input maps with the reference data. Tuanmu and Jetz (2014) apply probability voting in case products disagree and use a workflow that keeps the heterogeneity captured from the product with a finer resolution. In this research, the four input maps are integrated simultaneously with probability voting because FROM-GLC and MODIS respectively have a finer and coarser resolution than the resolution of 300 meter from Globcover 2009 and LC-CCI. Probability voting has the same advantages and disadvantages as weighted voting. Probability voting rejects LC classes with a low class probability but favors LC classes that have a good probability in the integration. Smaller LC classes like LC class (2) shrubs tend to be under-mapped.

Figure 19 presents a difference plot where the normal voting, weighed voting and probability voting maps are similar and dissimilar. Integration methods disagreed where methods were less certain in assigning a pixel to a specific LC class. Generally on locations where maximum weights from weighted voting (figure 15) and maximum class probabilities from probability voting (figure 17) were low and and LC input maps disagreed in the initial voting (figure 13b). Differences between the integration methods typically occur in heterogeneous areas, fragmented landscape, and transition areas. 5610888 pixels disagree in the difference plot; this is a 7.17% difference between the LC maps (table 20). 54.49% of the difference plot is where the methods agree on LC class (7) water, permanent snow and ice. Even when excluding the LC class (7) water, permanent snow and ice there is a 15.76% difference between the land classes of the LC maps. Normal voting, weighed voting and probability voting are quite similar. From table 20 can be seen when methods disagree in the disagreeing areas of the difference plot: (1) normal voting and weighed voting disagree in 45.54%, (2) normal voting and probability voting disagree in 96.74%, (3) weighted voting and probability voting disagree in 62.65%, (4) all maps disagree in 4.93% and (5) 47.61% where the methods disagree are tie locations from the initial voting of normal voting. Normal voting and probability voting seem to disagree most, were normal voting and weighted voting seem more similar compared to the other methods.

Weighted voting and probability voting base their weights and probabilities on the accuracy assessment of Ye et al. (2014) and Tsendbazar et al. (2016). The overall accuracies of FROM-GLCagg, Globcover (2009), LC-CCI (2010) and MODIS5 (2010) of respectively: 66.10% , 67.81%, 74.70% and 73.92% presented in table 8 section 4.1.1 (Ye et al. 2014; Tsendbazar et al. 2016). The discussion on the LC data mentions that it is questionable if the accuracy assessments can be compared to each other, since they come from different sources. This accounts for the results of weighted voting and probability voting, as respectively their weights and probabilities are based on these accuracy assessments (Ye et al. 2014; Tsendbazar et al. 2016). Additionally the accuracy assessment of FROM-GLCagg is used for FROM-GLC-hierarchy which does not hold information on LC class (5) wetlands and FROM-GLC holds a tilling effect over its LC map (figure 8a). The discussion on the improvement of the integrated maps reports the overall agreement from the external validation of FROM-GLCagg, Globcover (2009), LC-CCI (2010) and MODIS5 (2010) of respectively: 43.01%, 60.37%, 66.16% and 67.25%.

6.1.4. Internal validation

The internal validation uses entropy as an internal measure of uncertainty, which represents the amount of information necessary to require certainty (Shannon and Weaver. 1949). Accuracy measures by a confusion matrix are based on the whole map, but it is known accuracy may vary locally within the map (Foody 2005; Strahler et al. 2006). This research used the information entropy as an addition to the external validation, to calculate the information entropy as an internal measure of uncertainty.

Probability voting had the lowest uncertainty and the highest contrast between certainty and uncertainty in comparison to the other methods and therefore provides the best classification based on the information entropy. Like the difference plot between the integration methods, the information entropy calculated over normal voting and weighted voting suggest that the methods are similar (figure 20a and b). Both methods seem to have similar patterns in the information entropy aggregated over the study area, were normal voting seems to have slightly higher values of uncertainty. In each integration methods, uncertainty seems to be highest in areas were: (1) tie locations (figure 13) from the initial voting and were (2) the integration maps disagree in the difference plot (figure 19). In normal voting it is logical that areas with high uncertainty are located in tie locations as the information entropy is implemented on the five conditions from table 9 (section 4.2.1, 4.3.1 and 5.2.1). The difference plot (figure 20) reports in section 5.1.4 that 47.61% of the locations where the methods disagree concur with tie locations from the initial voting (table 20).

The information entropy ranks probability voting as the best integration method and normal voting and weighted voting as similar integration methods. The external validation suggests weighted voting achieved the best overall results, but probability voting achieved better results for LC classes and only has a slightly lower overall agreement. The difference plot does confirm that normal voting and weighted voting achieved similar results.

6.1.5. Improvements of integrated LC maps

This section evaluates the improvements of the integrated LC maps: Normal voting, weighted voting and probability voting based on the overall agreement, user agreement and producer agreement metrics derived from their external validation. The integration methods assessments are compared to each other and the LC input maps. The external validation uses a basic approach by cross tabulating the assessed LC map against the reference dataset. Section 5.3 holds the confusion matrices and their agreement metrics of integrated LC maps and input maps.

Section 3.5.2 in the theoretical background mentions: a reference dataset is used as reference in an external validation, but may contain misclassification (Strahler et al. 2006). Section 4.3.2 describes that it would be more correct to refer to the derived metrics of the confusion matrix as agreement instead of accuracy, because it is questionable if the reference data is valid as the reference data comes from multiple sources.

The reference dataset used for the validation uses 3x3 pixel envelop to extract he majority class from the assessed map, which is processed to confusion matrices in excel table.., table.. and table.. Strahler et al. (2006) mentioned that interpretation of the confusion matrix requires consideration of the sample design from the reference dataset (Strahler et al. 2006). Since the sample design is assumed to be homogenous LC type, this research does not consider the sample design of: GLC2000, GLCNMO-tr, Geo-Wiki, Globcover 2005,MODIS-tr and VIIRS 3 datasets. This research does not calculate spatial variation from the confusion matrices as is suggested by Strahler et al. 2006), but uses the information entropy as a measure of internal classification uncertainty.

Integration has improved overall accuracy in the normal voting, weighted voting and probability voting LC maps compared to the input maps (figure 21). Weighted voting has the highest overall agreement among the integration methods with an agreement of 71.72%, but differs slightly from normal voting and weighted voting (figure 21). Table 24 (section 5.3.3) and figure 21 show that FROM-GLC-hierarchy has an extreme low agreement of 43.01% with the reference dataset compared to Globcover (2009), LC-CCI (2010) and MODIS5 (2010) with respectively 60.37%, 66.16% and 67.25%.

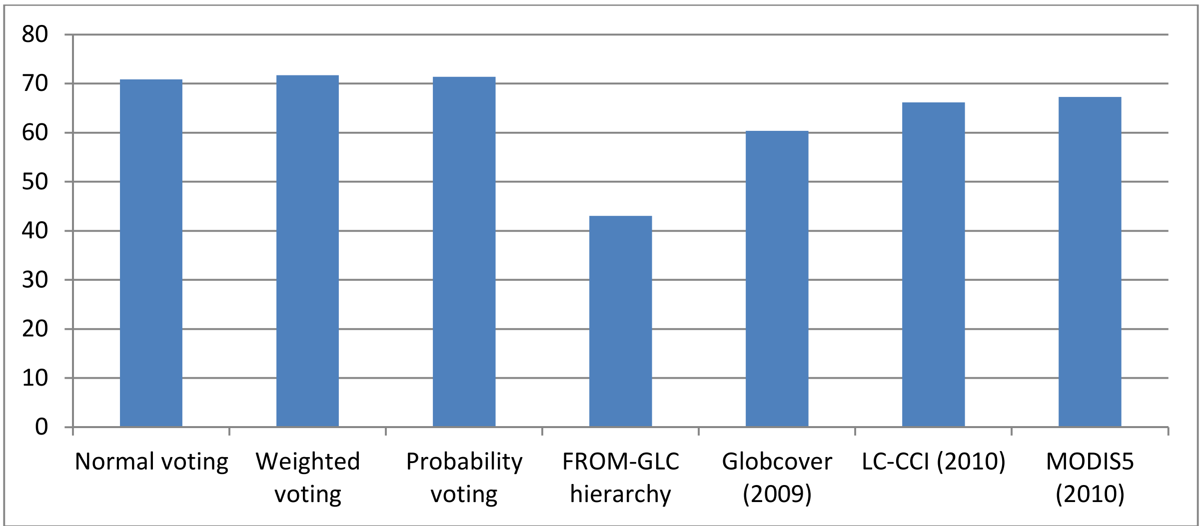

Figure 21: Overall agreement LC maps

Table 6 (section 3.3.1) presents the distribution of the samples over the harmonized LC classes and describes that LC classes (5) wetlands and (8) barren were not represented by enough samples (section 3.3.3). This causes that FROM-GLC-hierarchy and MODIS5 hold no information for the metrics of LC classes (5) wetlands and (8) barren, were probability voting has no information on the user agreement for LC class (5) wetlands. According to the external validation, probability voting results in the best class improvement among the integration method. Probability voting more often holds improved agreement in LC classes compared to the other integration methods and input maps (table 25 and table 26). Probability voting does hold low agreement metrics for LC class (7) water, permanent snow and ice compared to the other methods and input maps, a 69.23% producer agreement and 75.00% user agreement. LC-CCI often holds higher agreement metrics for the eight LC classes compared to the integration methods. According to the internal validation, probability voting is also the most promising method as it has the least entropy aggregated over the study area.

Tsendbazar et al. (2016) characterizes shrubs, grass and cropland classes, respectively LC class (2) shrubs, (3) herbaceous vegetation, and (4) cultivated and managed vegetation/agriculture of this research, as LC classes with high confusion errors. This is generally also the case for this research, especially for LC class (2) shrubs; class agreements are lower for each integration method than the LC input maps.

LC classes: (3) herbaceous vegetation, (5) wetlands, (6) urban/built up and (8) barren from the integration methods show a significant improvement in user agreement.

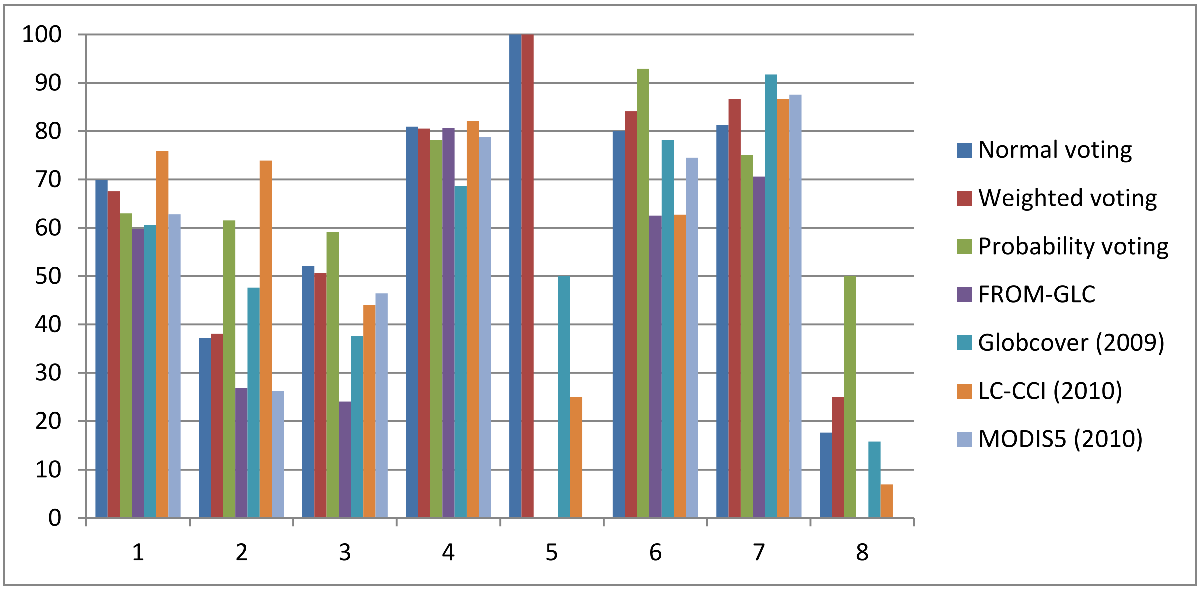

Figure 22: User agreement of LC maps

LC classes (1) trees and (4) cultivated and managed vegetation/agriculture from the integration methods show a significant improvement in producer agreement.

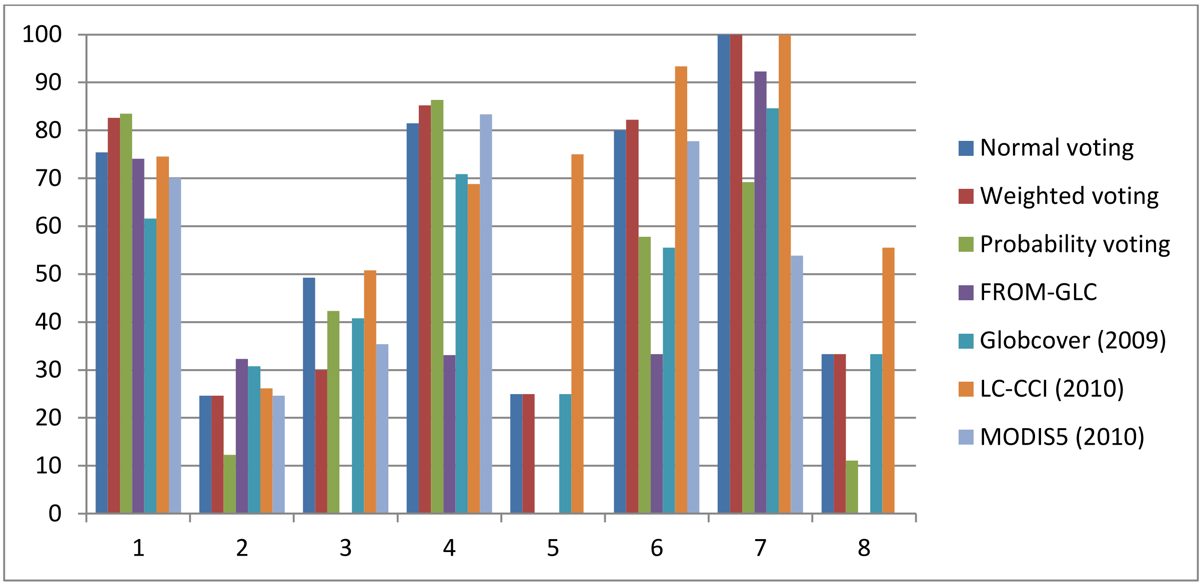

Figure 23: Producer agreement of LC maps

6.2. Conclusions

This section gives answers to the research question which were used to reach the objective and conclude the research.

6.2.1. Research questions

This section answers the research questions:

I. *Can the selected LC map integration methods be applied to the study area considering data constraints and characteristics?*

Normal voting, weighted voting and probability voting integration methods have been applied to the study area, considering the data constraints and characteristics. The reference dataset constraints and characteristics influenced the research discussed in section 3.3. Table 6 presents the distribution of the samples sites over the harmonized LC classes, LC classes (5) wetlands and (8) barren were not represented by enough samples (section 3.3.3).

II. *How can LC datasets be integrated with the chosen integration methods and selected software?*

All integration methods are implemented in R and the scripts used for implementation are in the appendix 8.1. Normal voting uses a common voting procedures and assigns ties to LC classes based on class preferences (table 10) calculated from the occurrences of LC classes from each input map in the initial voting (table 30 and tale 31). In weighed voting and probability voting a pixel is assigned to the class that accumulates respectively the highest weight/highest probability at that pixel's location. The weights and probabilities are based on the confusion matrices of FROM-GLC (Yu et al. 2014) Globcover 2009, LC-CCI (2010) and MODIS5 (Tsendbazar et al. 2016).

III. *Which is the most promising method based on internal validation?*

The most promising integration method based on internal validation is probability voting, because probability voting has the least entropy aggregated over the study area. Probability voting has the lowest uncertainty (figure 20c) compared to normal voting (figure 20a) and weighted voting (figure 20b). Generally, probability voting has low uncertainty in the information entropy map, there is high contract with the areas of high uncertainty.

IV. *What is the agreement of the integrated LC maps with the reference dataset and how much has integration improved overall accuracy?*

Normal voting, weighted voting and probability voting respectively have an overall agreement with the reference dataset of 70.85%, 71.72% and 71.40%. Integration has improved overall accuracy compared to the overall agreement of FROM-GLC-hierarchy, Globcover (2009), LC-CCI (2010) and MODIS5 (2010) with respectively: 43.01% , 60.37%, 66.16% and 67.25%. There is no clear improvement in the agreement metrics of LC-classes. Some LC classes of the input maps often had a higher agreement metric for LC-classes compared to one or more integration methods, particularly for LC class (2) shrub.

6.2.2. Research conclusion

Compared to their input LC maps: FROM-GLC-hierarchy, Globcover (2009), LC-CCI (2010) and MODIS5 (2010), normal voting, weighted voting and probability voting have an improved overall agreement. It is difficult to decide which integration method is most accurate among normal voting, weighted voting and probability voting. There is a 7.17% pixel difference between the LC maps (table 24) and there are no significant differences between the overall agreement metrics of the integration methods (figure 21). There is less improvements in class agreements of the integration methods, compared to the LC input maps (figure 22 and figure 23). Voting methods favour classes that have good probability or a high weight in the integration; therefore common LC classes are over-mapped and rare LC classes could have been under-mapped. Probability voting has the least entropy aggregated over the study area, the most improvements in class agreements, but the external validation shows a poor result for LC classes: (5) wetlands, (7) water, permanent snow and ice due to limited sample sites. Normal voting and weighted voting seem to achieve similar results because (1) these methods achieve similar patterns in the information entropy and (2) that these methods have the least disagreement over the study area.

6.3. Recommendation

Each voting integration method is suited for GLC map integration to improve overall accuracy. It is recommended that these methods are compared to other integration methods with a statistical approach as GWR and kriging, which takes into account the location of LC classes, to select which method is most accurate. Weighted voting and probability voting could possibly achieve better results when the weighs/probabilities are based on the same accuracy assessment. The methodology used for normal voting seems to provide reliable results, without the use of weights and class probabilities in the voting procedure, since the achieved results are similar to weighted voting and the overall agreement is close to weighted voting and probability voting.

Recommendations for further research:

- Comparing voting integration methods to other statistical integration methods
- Integration of LC maps with their accuracy assessment should have a trusted high level of accuracy
- Reference dataset should be represented by enough samples for each LC class and consideration of the sample design.
- Harmonization to more than eight LC classes to be able to discriminate between other problematic classes, for example, characterizing different forest classes

7. References

- Bontemps, S., et al. (2011). "GLOBCOVER 2009-Products description and validation report."
- Brunsdon, C., et al. (1996). "Geographically weighted regression: a method for exploring spatial nonstationarity." Geographical analysis 28(4): 281-298.
- Brunsdon, C., et al. (1998). "Geographically weighted regression." Journal of the Royal Statistical Society: Series D (The Statistician) 47(3): 431-443.
- Carneiro, J. D. and M. J. Pereira (2012). "Geostatistical stochastic simulation for spatial accuracy assessment of land cover maps derived from remotely sensed data."
- Charlton, M., et al. (2009). "Geographically weighted regression." White paper. National Centre for Geocomputation. National University of Ireland Maynooth.
- Climate Change Initiative Land Cover project (2014), overview. WWW site http://www.esa-landcover-cci.org/?q=overview [accessed: 10-12-2014].
- Climate Change Initiative Land Cover project (2014), map viewer. WWW site http://maps.elie.ucl.ac.be/CCI/viewer/index.php [accessed: 10-12-2014].
- Climate Change Initiative Land Cover project (2015), Land cover maps. WWW site http://www.esa-landcover-cci.org/ [accessed: 02-11-2015].
- Comber, A., et al. (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.
- De Bruin, S. and B. Gorte (2000). "Probabilistic image classification using geological map units applied to land-cover change detection." International Journal of Remote Sensing 21(12): 2389-2402.
- Di Gregorio, A. (2005). Land cover classification system: classification concepts and user manual: LCCS, Food & Agriculture Org.
- Esri ArcGIS software (2015). WWW site http://www.esri.com/software/arcgis [accessed: 30-09-2015].
- Finer Resolution Observation and Monitoring Global Land Covert (2015), overview. WWW site http://data.ess.tsinghua.edu.cn/ [accessed: 21-01-2015].
- Foody, G. (2005). "Local characterization of thematic classification accuracy through spatially constrained confusion matrices." International Journal of Remote Sensing 26(6): 1217-1228.
- Foody, G. M. (2002). "Status of land cover classification accuracy assessment." Remote Sensing of Environment 80(1): 185-201.
- Fotheringham, A. S., et al. (2003). Geographically weighted regression: the analysis of spatially varying relationships, John Wiley & Sons.
- Friedl, M. A., et al. (2010). "MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets." Remote Sensing of Environment 114(1): 168-182.
- Fritz, S., et al. (2011). "Highlighting continued uncertainty in global land cover maps for the user community." Environmental Research Letters 6(4): 044005.
- Ge, Y., et al. (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.
- GOFC-GOLD reference data portal (2015). WWW site http://www.gofcgold.wur.nl/sites/gofcgold_refdataportal.php#glc2000 [accessed: 23-09-2015]
- Goovaerts, P. (1999). "Geostatistics in soil science: state-of-the-art and perspectives." Geoderma 89(1): 1-45.

- Gopal, S., et al. (1999). "Fuzzy neural network classification of global land cover from a 1 AVHRR data set." Remote Sensing of Environment 67(2): 230-243.
- Herold, M., et al. (2008). "Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets." Remote Sensing of Environment 112(5): 2538-2556.
- Herold, M. and C. Schmullius (2004). Report on the harmonization of global and regional land cover products. Workshop report at FAO, Rome, Italy
- Herold, M., et al. (2011). "CCI Land Cover project–User Requirements Document (version 2.2)." Project report 22.
- Herold, M., et al. (2006). "A joint initiative for harmonization and validation of land cover datasets." Geoscience and Remote Sensing, IEEE Transactions on 44(7): 1719-1727.
- Hibbard, K., et al. (2010). "Research priorities in land use and land-cover change for the Earth system and integrated assessment modelling." International Journal of Climatology 30(13): 2118-2128.
- Iwao, K., et al. (2011). "Creation of new global land cover map with map integration." Journal of Geographic Information System 3(02): 160.
- Jung, M., et al. (2006). "Exploiting synergies of global land cover products for carbon cycle modeling." Remote Sensing of Environment 101(4): 534-553.
- Loveland, T. R., et al. (1999). "An analysis of the IGBP global land-cover characterization process." Photogrammetric Engineering and Remote Sensing 65(9): 1021-1032.
- Kinoshita, T., et al. (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.
- Mora, B., Tsendbazar, N.E., Herold, M., Arino, O., (2014). "Global land cover mapping: Current status and future trends. In: Manakos, I., Braun, M. (Eds.), Land Use and Land Cover Mapping in Europe: Practices and Trends. " Springer, Netherlands, p.441.
- Nakaegawa, T. T., Maki ; Yamashiki, Yosuke (2011). "Uncertainty in land cover datasets for global land-surface models derived from 1-km global land cover datasets." Hydrological Processes, 2011, Vol.25(17), pp.2703-2714 [Peer Reviewed Journal].
- QGIS (2015). WWW site http://www.qgis.org/nl/site/forusers/download.html [accessed: 28-01-2015].
- See, L., et al. (2014). "Building a hybrid land cover map with crowdsourcing and geographically weighted regression." ISPRS Journal of Photogrammetry and Remote Sensing.
- Shannon, C. E. and W. Weaver (1949). The mathematical theory of communication (Urbana, IL, University of Illinois Press IL.
- Sinusoidal projection (2015) WWW site https://code.env.duke.edu/projects/mget/wiki/SinusoidalMODIS [accessed: 02-11-2015].
- Stehman, S. V., et al. (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(22): 6975-6993.
- Strahler, A. H., et al. (2006). "Global land cover validation: Recommendations for evaluation and accuracy assessment of global land cover maps." European Communities, Luxembourg: 51.
- The R project for statistical computing (2015). WWW site http://www.R-project.org
- Tsendbazar, N., et al. (2015). "Assessing global land cover reference datasets for different user communities." ISPRS Journal of Photogrammetry and Remote Sensing.

- Tsendbazar, N., et al. (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. and W. Jetz (2014). "A global 1-km consensus land-cover product for biodiversity and ecosystem modelling." Global Ecology and Biogeography.
- UMICH, electronic textbook; probability (2015) WWW site https://controls.engin.umich.edu/wiki/index.php/Bayes_Rule,_conditional_probability,_independence [accessed: 28-08-2015].
- Verburg, P. H., et al. (2011). "Challenges in using land use and land cover data for global change studies." Global Change Biology 17(2): 974-989.
- Yu, L., et al. (2014). "A multi-resolution global land cover dataset through multisource data aggregation." Science China Earth Sciences 57(10): 2317-2329.

8. Appendices

This section holds the appendixes of this research.

8.1. Scripts for implementing voting methods

This section holds the scripts used for implementing the integration methods, internal validation and external validation in R.

8.1.1. Initial voting

```
F <- raster("C:\\...\\FROMGLCNG8.tif")
G <- raster("C:\\...\\GlobcoverG8.tif")
L <- raster("C:\\...\\LCCCIG8.tif")
M <- raster("C:\\...\\MODISG8.tif")

STNV <- stack(F, G, L, M)
plot(STNV)

MaxVote <- function(x, ...){
  # here I assume 8 classes
  if(any(is.na(x))) return(NA) # return NA if all(!) maps have NA
  vec1 <- lapply(1:8, FUN = function(y) sum(x==y))
  if(length(which(vec1==1))==4) return(-2) # tie with all different
  if(length(which(vec1==2))==2) return(-1) # tie with two different
  which.max(vec1)
}

V <- calc(STNV, MaxVote)
plot(V)
writeRaster(V, filename = "NormalVoting_SA.tif", format = "GTiff", overwrite = TRUE)

VoteCount <- function(x, ...){
  # here I assume 8 classes
  if(any(is.na(x))) return(NA) # return NA if all(!) maps have NA
  vec1 <- lapply(1:8, FUN = function(y) sum(x==y))
  if(length(which(vec1==1))==4) return(50) # tie with all different
  if(length(which(vec1==2))==2) return(40) # tie with two different
  if(length(which(vec1==2))==1) return(30) # Two are equal, remaining two are contrary
  if(length(which(vec1==3))==1) return(20) # Three are equal
  if(length(which(vec1==4))==1) return(10) # All are equal
}

I <- calc(STNV, VoteCount)
plot(I)
writeRaster(I, filename = "NormalVotingI_SA.tif", format = "GTiff", overwrite = TRUE)
```

8.1.2. Indicators in conditions

```
F <- raster("C:\\...\\FROMGLCNG8.tif")
G <- raster("C:\\...\\GlobcoverG8.tif")
L <- raster("C:\\...\\LCCCIG8.tif")
M <- raster("C:\\...\\MODISG8.tif")
```

```
V <- raster("C:\\...\\NormalVoting_SA.tif")
I <- raster("C:\\...\\NormalVotingI_SA.tif")

B <- stack(F, G, L, M, V, I)

# F, calculating votes from F in indicators (10,20,30,40,50):
for(i in 1:8){
  assign(paste('FI', i, sep="), calc(B, fun=function(x) ifelse((x[1]==i&&x[1]==x[5]&&x[6]==10),
10+i, 0)) +
         calc(B, fun=function(x) ifelse((x[1]==i&&x[1]==x[5]&&x[6]==20), 20+i, 0)) +
         calc(B, fun=function(x) ifelse((x[1]==i&&x[1]==x[5]&&x[6]==30), 30+i, 0)) +
         calc(B, fun=function(x) ifelse((x[1]==i&&x[1]==x[5]&&x[6]==40), 40+i, 0)) +
         calc(B, fun=function(x) ifelse((x[1]==i&&x[1]==x[5]&&x[6]==50), 50+i, 0)))
}
FIS <-sum(FI1,FI2,FI3,FI4,FI5,FI6,FI7,FI8)
writeRaster(FIS, filename = "Fi_SA.tif", format = "GTiff", overwrite = TRUE)

# G, calculating votes from G in indicators (10,20,30,40,50):
for(i in 1:8){
  assign(paste('GI', i, sep="), calc(B, fun=function(x) ifelse((x[2]==i&&x[2]==x[5]&&x[6]==10),
10+i, 0)) +
         calc(B, fun=function(x) ifelse((x[2]==i&&x[2]==x[5]&&x[6]==20), 20+i, 0)) +
         calc(B, fun=function(x) ifelse((x[2]==i&&x[2]==x[5]&&x[6]==30), 30+i, 0)) +
         calc(B, fun=function(x) ifelse((x[2]==i&&x[2]==x[5]&&x[6]==40), 40+i, 0)) +
         calc(B, fun=function(x) ifelse((x[2]==i&&x[2]==x[5]&&x[6]==50), 50+i, 0)))
}
GIS <-sum(GI1,GI2,GI3,GI4,GI5,GI6,GI7,GI8)
writeRaster(GIS, filename = "Gi_SA.tif", format = "GTiff", overwrite = TRUE)

# L, calculating votes from L in indicators (10,20,30,40,50):
for(i in 1:8){
  assign(paste('LI', i, sep="), calc(B, fun=function(x) ifelse((x[3]==i&&x[3]==x[5]&&x[6]==10),
10+i, 0)) +
         calc(B, fun=function(x) ifelse((x[3]==i&&x[3]==x[5]&&x[6]==20), 20+i, 0)) +
         calc(B, fun=function(x) ifelse((x[3]==i&&x[3]==x[5]&&x[6]==30), 30+i, 0)) +
         calc(B, fun=function(x) ifelse((x[3]==i&&x[3]==x[5]&&x[6]==40), 40+i, 0)) +
         calc(B, fun=function(x) ifelse((x[3]==i&&x[3]==x[5]&&x[6]==50), 50+i, 0)))
}
LIS <-sum(LI1,LI2,LI3,LI4,LI5,LI6,LI7,LI8)
writeRaster(LIS, filename = "Li_SA.tif", format = "GTiff", overwrite = TRUE)

# M, calculating votes from M in indicators (10,20,30,40,50):
for(i in 1:8){
  assign(paste('MI', i, sep="), calc(B, fun=function(x) ifelse((x[4]==i&&x[4]==x[5]&&x[6]==10),
10+i, 0)) +
         calc(B, fun=function(x) ifelse((x[4]==i&&x[4]==x[5]&&x[6]==20), 20+i, 0)) +
         calc(B, fun=function(x) ifelse((x[4]==i&&x[4]==x[5]&&x[6]==30), 30+i, 0)) +
```

```
        calc(B, fun=function(x) ifelse((x[4]==i&&x[4]==x[5]&&x[6]==40), 40+i, 0)) +
        calc(B, fun=function(x) ifelse((x[4]==i&&x[4]==x[5]&&x[6]==50), 50+i, 0)))
}
MIS <-sum(MI1,MI2,MI3,MI4,MI5,MI6,MI7,MI8)
writeRaster(MIS, filename = "Mi_SA.tif", format = "GTiff", overwrite = TRUE)
```

8.1.3. Normal voting

```
F <- raster("C:\\...\\FROMGLCNG8.tif")
G <- raster("C:\\...\\GlobcoverG8.tif")
L <- raster("C:\\...\\LCCCIG8.tif")
M <- raster("C:\\...\\MODISG8.tif")

#### Normal Voting
STNV <- stack(F, G, L, M)
MaxVote <- function(x, ...){
  # here I assume 8 classes
  if(any(is.na(x))) return(NA) # return NA if all(!) maps have NA
  vec1 <- lapply(1:8, FUN = function(y) sum(x==y))
  if(length(which(vec1==1))==4) return(-2) # tie with all different
  if(length(which(vec1==2))==2) return(-1) # tie with two differen
  which.max(vec1)
}
V <- calc(STNV, MaxVote)
writeRaster(V, filename = "NV_SA.tif", format = "GTiff", overwrite = TRUE)

VoteCount <- function(x, ...){
  # here I assume 8 classes
  if(any(is.na(x))) return(NA) # return NA if Find= maps have NA
  vec1 <- lapply(1:8, FUN = function(y) sum(x==y))
  if(length(which(vec1==1))==4) return(50) # tie with all different
  if(length(which(vec1==2))==2) return(40) # tie with two different
  if(length(which(vec1==2))==1) return(30) # Two are equal, remaining two are contrary
  if(length(which(vec1==3))==1) return(20) # Three are equal
  if(length(which(vec1==4))==1) return(10) # All are equal
}
I <- calc(STNV, VoteCount)
writeRaster(I, filename = "NVI_SA.tif", format = "GTiff", overwrite = TRUE)

####  Stack with initial results: inputmaps, Voting & Indicators ####
STIV<-stack(F,G,L,M,V,I)
CSV <- read.csv("Class Preferences.csv")          ( table 10 section 4.2.1)
for(i in 1:8){
  assign(paste('TP', i, sep=''), calc(STIV, fun=function(x) ifelse((x[5]==-2&&x[1]==i||x[5]==-
1&&x[1]==i), CSV[i,2], 0)) +
        calc(STIV, fun=function(x) ifelse((x[5]==-2&&x[2]==i||x[5]==-1&&x[2]==i), CSV[i,3], 0)) +
        calc(STIV, fun=function(x) ifelse((x[5]==-2&&x[3]==i||x[5]==-1&&x[3]==i), CSV[i,4], 0)) +
        calc(STIV, fun=function(x) ifelse((x[5]==-2&&x[4]==i||x[5]==-1&&x[4]==i), CSV[i,5], 0)))
}
```

```
####Stack 8 resulting maps####
TPNV<-stack(TP1,TP2,TP3,TP4,TP5,TP6,TP7,TP8)
####Total percentages summed for normalizing TP's####
TPT <- sum(TP1,TP2,TP3,TP4,TP5,TP6,TP7,TP8)
TP1N <- TP1/TPT
TP2N <- TP2/TPT
TP3N <- TP3/TPT
TP4N <- TP4/TPT
TP5N <- TP5/TPT
TP6N <- TP6/TPT
TP7N <- TP7/TPT
TP8N <- TP8/TPT
####Stack 8 normalised resulting maps ####
NTPNV<-stack(TP1N,TP2N,TP3N,TP4N,TP5N,TP6N,TP7N,TP8N)

MaxPerc <- function(x, ...){
  if(any(is.na(x))) return(NA) # return NA if any(!) maps have NA
  if(all(x==0)) return (0)  # if all layers are 0 return 0, this is to ensure only the cells with -2 and -1
cases will get a result
  which.max(x)
}

TIES <- calc(NTPNV, MaxPerc)
plot(TIES)
writeRaster(TIES, filename = "TIESNV_SA.tif", format = "GTiff", overwrite = TRUE)

####Stacks all maps####
STNV2<-stack(F,G,L,M,V,I,TIES)
MapNV<-calc(STNV2, fun=function(x) ifelse((x[5]==-2||x[5]==-1), x[7], x[5]))
####Normal Voting map:####
writeRaster(MapNV, filename = "MapNV_SA.tif", format = "GTiff", overwrite = TRUE)

# Shannon Entropy in bits (base 2 log) for voting
logSpecial <- function(x) ifelse(x==0, 0, log(x, 2))  # dealing with log(0)
Shannon <- function(x, ...){
  if (is.na(x[1])) return(NA)
  # convert to probs for each class (8 classes)
  probs <- lapply(1:8, FUN = function(y) 0.25*sum(x==y))  # assuming 4 maps
  vec2 <- lapply(probs, FUN=function(x) logSpecial(x) * -x)
  sum(unlist(vec2))
}
InformationEntropyNV <- calc(STNV, Shannon)
plot(InformationEntropyNV)
writeRaster(InformationEntropyNV, filename = "InformationEntropyNV_SA.tif", format = "GTiff",
overwrite = TRUE)
```

8.1.4. Weighted voting

```
F <- raster("C:\\...\\FROMGLCNG8.tif")
G <- raster("C:\\...\\GlobcoverG8.tif")
L <- raster("C:\\...\\LCCCIG8.tif")
M <- raster("C:\\...\\MODISG8.tif")

ST <- stack(F, G, L, M)
CSV <- read.csv("UAwv.csv")            (Table 12 section 4.2.2)

####  Weights ####
for(i in 1:8){
  assign(paste('TW', i, sep=''),
         calc(ST, fun=function(x) ifelse((x[1]==i), CSV[i,1], 0)) +
         calc(ST, fun=function(x) ifelse((x[2]==i), CSV[i,2], 0)) +
         calc(ST, fun=function(x) ifelse((x[3]==i), CSV[i,3], 0)) +
         calc(ST, fun=function(x) ifelse((x[4]==i), CSV[i,4], 0)))
}
####Stack 8 resulting maps####
TWS<-stack(TW1,TW2,TW3,TW4,TW5,TW6,TW7,TW8)
####Total weights summed for normalizing TW's####
TWT <- sum(TW1,TW2,TW3,TW4,TW5,TW6,TW7,TW8)
TW1N <- TW1/TWT
TW2N <- TW2/TWT
TW3N <- TW3/TWT
TW4N <- TW4/TWT
TW5N <- TW5/TWT
TW6N <- TW6/TWT
TW7N <- TW7/TWT
TW8N <- TW8/TWT
####Stack 8 normalised resulting maps####
TWSN<-stack(TW1N,TW2N,TW3N,TW4N,TW5N,TW6N,TW7N,TW8N)

MaxPerc <- function(x, ...){
  if(any(is.na(x))) return(NA) # return NA if any(!) maps have NA
  if(all(x==0)) return (0)  # if all layers are 0 return 0,
  which.max(x)
}

####  Weighted Voting map:####
WV <- calc(TWSN, MaxPerc)
writeRaster(WV, filename = "MapWV_SA.tif", format = "GTiff", overwrite = TRUE)

####returns instead of class value, the weights used in weighted voting####
MaxWeights <- function(x, ...){
  if(any(is.na(x))) return(NA) # return NA if any(!) maps have NA
  max(x) # check which layer has the highest value, then returns value (which=probability)
}
```

```
WN <- calc(TWSN, MAxWeights)
writeRaster(WN, filename = "WeightsWV_SA.tif", format = "GTiff", overwrite = TRUE)

# Shannon Entropy in bits (base 2 log) for voting
logSpecial <- function(x) ifelse(x==0, 0, log(x, 2))  # dealing with log(0)
Shannon <- function(x, ...){
  vec2 <- lapply(x, FUN=function(x) sum((logSpecial(x) * -x)) )
  sum(unlist(vec2))
}
InformationEntropyWV <- calc(TWSN, Shannon)
writeRaster(InformationEntropyWV, filename = "InformationEntropyWV_SA.tif", format = "GTiff",
overwrite = TRUE)
```

8.1.5. Probability voting

```
F <- raster("C:\\...\\FROMGLCNG8.tif")
G <- raster("C:\\...\\GlobcoverG8.tif")
L <- raster("C:\\...\\LCCCIG8.tif")
M <- raster("C:\\...\\MODISG8.tif")

ST <- stack(F, G, L, M)
CSVF <- read.csv("Fp.csv")              (Table 14 section 4.2.3)
CSVG <- read.csv("Gp.csv")              (Table 14 section 4.2.3)
CSVL <- read.csv("Lp.csv")              (Table 14 section 4.2.3)
CSVM <- read.csv("Mp.csv")              (Table 14 section 4.2.3)

##Probability voting##
for(i in 1:8){
  assign(paste('TP', i, sep=''),
        calc(ST, fun=function(x) CSVF[x[1],i]) *
        calc(ST, fun=function(x) CSVG[x[2],i]) *
        calc(ST, fun=function(x) CSVL[x[3],i]) *
        calc(ST, fun=function(x) CSVM[x[4],i]))
}

####Stack 8 resulting maps####
TPS<-stack(TP1,TP2,TP3,TP4,TP5,TP6,TP7,TP8)
TPT <- sum(TP1,TP2,TP3,TP4,TP5,TP6,TP7,TP8)
TP1N <- TP1/TPT
TP2N <- TP2/TPT
TP3N <- TP3/TPT
TP4N <- TP4/TPT
TP5N <- TP5/TPT
TP6N <- TP6/TPT
TP7N <- TP7/TPT
TP8N <- TP8/TPT
####Stack 8 normalised resulting maps####
TPSN<-stack(TP1N,TP2N,TP3N,TP4N,TP5N,TP6N,TP7N,TP8N)
```

```
MaxProb <- function(x, ...){
  if(any(is.na(x))) return(NA) # return NA if any(!) maps have NA
  which.max(x) # check which layer has the highest value, then return layer nr (which=landclass)
  }
PV <- calc(TPSN, MaxProb)
writeRaster(PV, filename = "MapPV_SA.tif", format = "GTiff", overwrite = TRUE)

####returns instead of class, the probabilities used in probability voting####
#Probabilities <- function(x, ...){
  #if(any(is.na(x))) return(NA) # return NA if any(!) maps have NA
    #max(x) # check which layer has the highest value, then returns value (which=probability)
}
#PN <- calc(TPSN, Probabilities)
#writeRaster(PN, filename = "ProbabilitiesPV_SA.tif", format = "GTiff", overwrite = TRUE)

# Shannon Entropy in bits (base 2 log) for voting
logSpecial <- function(x) ifelse(x==0, 0, log(x, 2))  # dealing with log(0)
Shannon <- function(x, ...){
  vec2 <- lapply(x, FUN=function(x) sum((logSpecial(x) * -x)) )
  sum(unlist(vec2))
}
InformationEntropyPV <- calc(TPSN, Shannon)
writeRaster(InformationEntropyPV, filename = "InformationEntropyPV_SA.tif", format = "GTiff",
overwrite = TRUE)
```

8.1.6. External validation

```
polygons <- readShapePoly("Reference dataset.shp")
#querying
fullimage <- raster("ASSESSED LCMAP.tif")
fullimage@data
#project polygon
polygons@proj4string <- fullimage@crs

#start preparing query
extr_data <- extract(fullimage, polygons)
extr_data_center<-extract(fullimage, coordinates(polygons))#in case you want the landcover of the
point locations

#majority, count, fraction, anyclass
majority <- unlist(lapply(extr_data, modal))

## add to the reference  polygons
polygons$gl_maj<-majority
polygons$gl_cent<-extr_data_center

##write everything back to the shapefile.
writeOGR(polygons, ".", "FROMGLCNG8_E_extract", driver="ESRI Shapefile")
```

8.2. Original legends, harmonization and LC occurrences in initial voting

This section presents the tables that hold the thematic harmonization of LC data and the LC occurrences in initial voting.

8.2.1. Thematic harmonization

Table 27: Thematic harmonization of LC classes 1, and 2

Harmonized LC data			FROM-GLC		GLC2000		Geo-wiki		Globcover		GLCNMO		IGBP/ MODIS/ VIIRS		LC-CCI	
Harmonized legend																
no	LC class	LCCS	no	LC class	no	LC class	no	LC class	no	LC class	no	LC class	no	D LC class	no	LC class
1	Trees (all sort of trees)	A12-A3 and A24-A3.A20.B2	20	20 Forest (20/21 Broadleave, 20/22 Needleleave, 20/23 Mixed, 20/24 Orchard)	1, 2, 3, 4, 5	Evergreen Needleleaf, Evergreen broadleaf, Deciduous Needleleaf Trees, Deciduous Broadleaf Trees, Mixed/Other Trees	1	1 Tree cover	40, 50, 60, 70, 90, 100, 110, 120, 160, 170	40 Closed to open (>15%) broadleaved evergreen or semi-deciduous forest (>5m); 50 Closed (>40%) broadleaved deciduous forest (>5m); 60 Open (15-40%) broadleaved; deciduous forest/woodland (>5m); 70 Closed (>40%) needleleaved evergreen forest (>5m); 90 Open (15-40%) needleleaved deciduous or evergreen forest (>5m); 100 Closed to open (>15%) mixed broadleaved and needleleaved forest (>5m); 110 Mosaic forest or shrubland (50-70%) / grassland (20-50%); 120 Mosaic grassland (50-70%) / forest or shrubland (20-50%); 160 Closed to open (>15%) broadleaved forest regularly flooded (semi-permanently or temporarily) - Fresh or brackish water; 170 Closed (>40%) broadleaved forest or shrubland permanently flooded - Saline or brackish water	1, 2, 3, 4, 5	1 Broadleaf Evergreen Forest; 2 Broadleaf Deciduous Forest; 3 Needleleaf Evergreen Forest; 4 Needleleaf Deciduous Forest; 5 Mixed Forest	1, 2, 3, 4, 5, 8, 9	1 Evergreen Needleleaf forest; 2 Evergreen Broadleaf Forest; 3 Deciduous Needleleaf forest; 4 Deciduous Broadleaf Forest; 5 Mixed Forest; 8 Woody Savanna; 9 Savanna	50, 60, 70, 80, 90, 100, 110, 160, 170	50 Tree cover, broadleaved, evergreen, closed to open (>15%); 60 Tree cover, broadleaved, deciduous, closed to open (>15%); 70 Tree cover, needleleaved, evergreen, closed to open (>15%); 80 Tree cover, needleleaved, deciduous, closed to open (>15%); 90 Tree cover, mixed leaf type (broadleaved and needleleaved); Mosaic tree and shrub (>50%) / herbaceous cover (<50%); 110 Mosaic herbaceous cover (>50%) / tree and shrub (<50%); 160; Tree cover, flooded, fresh or brakish water; 170 Tree cover, flooded, saline water
2	Shrubs	A12-A4.A20.B3	40	40 Shrub	6	Shrubs	2	2 Shrub cover	130	130 Closed to open (>15%) (broadleaved or needleleaved, evergreen or deciduous) shrubland (<5m)	7	7 Shrub	6, 7	6 Closed Shrubland; 7 Open Shrubland	120	120 Shrubland

Table 28: Thematic harmonization of LC classes 3,4 and 5

Harmonized LC data			FROM-GLC		GLC2000		Geo-wiki		Globcover		GLCNMO		IGBP/MODIS/ VIIRS		LC-CCI	
Harmonized legend																
no	LC class	LCCS	no	LC class	no	LC class	no	LC class	no	LC class	no	LC class	no	D LC class	no	LC class
3	Herbaceous vegetation	A12-A2.A 20.B 4	30	30 Grass (30/31 Managed & 30/32 Nature)	7	Herbaceous vegetation	3	3 Herbaceous vegetation / Grassland	30, 14 0	30 Mosaic vegetation (grassland/shrubland/f orest) (50-70%) / cropland (20-50%); 140 Closed to open (>15%) herbaceous vegetation (grassland, savannas or lichens/mosses)	8	8 Herbaceous	10	10 Grasslands	40, 130, 140	40 Tree cover, broadleaved, evergreen, closed to open (>15%); 130 Grassland; 140 Lichens and mosses
4	Cultivated and managed vegetation / agriculture (incl. mixtures)	A11 and A23	10	10 Cropland (10/11 Rice, 10/12 Greenhouse, 10/13 Other)	8	Cultivated and managed vegetation / agriculture (incl. mixtures)	4	4 Cultivated and managed / Cropland	11, 14, 20	11 Post-flooding or irrigated croplands (or aquatic); 14 Rainfed croplands; 20 Mosaic cropland (50-70%) / vegetation (grassland/shrubland/f orest) (20-50%)	11, 13	11 Cropland; 13 Cropland / Other Vegetation Mosaic	12, 14	12 Cropland; 14 Cropland /Natural vegetation	10, 20, 30	10 Cropland, rainfed; 20 Cropland, irrigated or post-flooding; 30 Mosaic cropland (>50%) / natural vegetation (tree, shrub, herbaceous cover) (<50%)
5	Other shrub/herba ceous vegetation	A24-A2 and A24-A4	50, 70	50 Wetlands (30/51 Grass & 90/52 Silt) ; 70 Tundra (40/71 Shrub & 30/72 Grass)	9	Other shrub/herba ceous vegetation	6	6 Regularly flooded / wetland	18 0	180 Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil - Fresh, brackish or saline water			11	11 Permanent Wetlands	180	180 Shrub or herbaceous cover, flooded, fresh/saline/brakish water

Table 29: Thematic harmonization of LC classes 6, 7 and 8

Harmonized LC data			FROM-GLC		GLC2000		Geo-wiki		Globcover		GLCNMO		IGBP/MODIS/ VIIRS		LC-CCI	
Harmonized legend																
no	LC class	LCCS	no	LC class	no	LC class	no	LC class	no	LC class	no	LC class	no	D LC class	no	LC class
6	Urban/built up	B15	80	80 Impervious (80/81 High albedo, 80/82 Low albedo)	10	Urban/built up	7	7 Urban / built up	190	190 Artificial surfaces and associated areas (Urban areas >50%)	18	18 Urban	13	13 Urban and Built Up	190	190 Urban areas
7	Water, permanent Snow and Ice	B27-A1, B27-A2, B27-A3, B28-A1, B28-A2, B28-A3	60, 100	60 Water (60/61 Lake, 60/62 Pond, 60/63 River, 60/64 Sea) ; 100 Snow/Ice (100/101 Snow & 100/102 Ice)	11, 13	Open water; Snow and Ice	8, 10	8 Snow and ice; 10 Open water	210, 220	210 Water bodies; 220 Permanent snow and ice	19, 20	19 Snow / Ice; 20 Water bodies	15, 17	15 Snow and Ice; 17 Water bodies	210, 220	210 Water bodies; 220 Permanent snow and ice
8	Barren	B16 and A12-A1A 14, A12-A2A 14	90	90 Bareland (90/91 Saline-Alkali, 90/92 Sand, 90/93 Gravel, 10/94 Bare Cropland, 90/95 Dry river/lake bed, 90/96 Other)	12	Barren	9	9 Barren	150	150 Sparse (<15%) vegetation; 200 Bare areas	10, 16, 17	10 Sparse vegetation; 16 Bare area, consolidated (gravel,rock); 17 Bare area, unconsolidated (sand)	16	16 Barren or Sparseley Vegetated	15, 20	150 Sparse vegetation (tree, shrub, herbaceous cover) (<15%); 200 Bare areas

8.2.2. LC classes occurrences of LC input maps in initial voting

Table 30: Class occurrences of FROM-GLC and Globcover (2009) in initial voting its conditions

Pixels		Initial Voting	FROM-GLC in Votes				Globcover (2009) in Votes			
			Agrees	3 agree	2 agree	Disagree	Agrees	3 agree	2 agree	Disagree
			10	20	30	0	10	20	30	0
All 1/tm 8		78252697	49700895	5590350	2830974	20130478	49700895	12158805	5291816	11101181
Separate classes:										
Trees	1	9163819	4421809	2288646	894363	-	4421809	1947641	821846	-
Shrubs	2	956572	54273	313220	406378	-	54273	202386	186705	-
Herbaceous vegetation	3	3907952	369064	689987	728303	-	369064	1690876	787148	-
Cultivated and managed vegetation / agriculture	4	13852954	2440157	1363298	363108	-	2440157	7179440	2608809	-
Other shrub/herbaceous vegetation	5	49649	54	1580	16962	-	54	1321	31151	-
Urban/built up	6	983353	149761	110215	54222	-	149761	326616	140472	-
Water, Snow and Ice	7	42744471	42237333	397812	68457	-	42237333	350822	42851	-
Barren	8	1255375	28444	425592	299181	-	28444	459703	672834	-
Tie, two maps disagree (2x)	-1	4158335	-	-	-		-	-	-	-
Tie: all maps disagree	-2	1180217	-	-	-		-	-	-	-
Percentages		Total Voting	FROM-GLC in Votes				Globcover (2009) in Votes			
			10	20	30	0	10	20	30	0
All 1/tm 8		100	63,51	7,14	3,62	25,72	63,51	15,54	6,76	14,19
Separate classes:										
Trees	1	100	48,25	24,97	9,76	-	48,25	21,25	8,97	-
Shrubs	2	100	5,67	32,74	42,48	-	5,67	21,16	19,52	-
Herbaceous vegetation	3	100	9,44	17,66	18,64	-	9,44	43,27	20,14	-
Cultivated and managed vegetation / agriculture	4	100	17,61	9,84	2,62	-	17,61	51,83	18,83	-
Other shrub/herbaceous vegetation	5	100	0,11	3,18	34,16	-	0,11	2,66	62,74	-
Urban/built up	6	100	15,23	11,21	5,51	-	15,23	33,21	14,29	-
Water, Snow and Ice	7	100	98,81	0,93	0,16	-	98,81	0,82	0,10	-
Barren	8	100	2,27	33,90	23,83	-	2,27	36,62	53,60	-

Table 31: Class occurences of LC-CCI (2010) and MODIS5 (2010) in initial voting and its conditions

Pixels		Initial Voting	LC-CCI (2010) in Votes				MODIS5 (2010) in Votes			
			Agrees	3 agree	2 agree	Disagree	Agrees	3 agree	2 agree	Disagree
			10	20	30	0	10	20	30	0
All 1/tm 8		78252697	49700895	12818931	5349880	10382991	49700895	12220296	4428242	11903264
Separate classes:										
Trees	1	9163819	4421809	2492586	797873	-	4421809	2363422	908408	-
Shrubs	2	956572	54273	252152	186271	-	54273	249623	346990	-
Herbaceous vegetation	3	3907952	369064	1861235	989499	-	369064	1617286	666570	-
Cultivated and managed vegetation / agriculture	4	13852954	2440157	6975178	2275234	-	2440157	7531066	2212455	-
Other shrub/herbaceous vegetation	5	49649	54	2256	38973	-	54	1791	7472	-
Urban/built up	6	983353	149761	389150	425589	-	149761	346515	265237	-
Water, Snow and Ice	7	42744471	42237333	396919	81687	-	42237333	64398	14647	-
Barren	8	1255375	28444	449455	554754	-	28444	46195	6463	-
Tie, two maps disagree (2x)	-1	4158335	-	-	-		-	-	-	-
Tie: all maps disagree	-2	1180217	-	-	-		-	-	-	-
Percentages		Initial Voting	LC-CCI (2010) in Votes				MODIS5 (2010)in Votes			
			10	20	30	0	10	20	30	0
All 1/tm 8		100	63,51	16,38	6,84	13,27	63,51	15,62	5,66	15,21
Separate classes:										
Trees	1	100	48,25	27,20	8,71	-	48,25	25,79	9,91	
Shrubs	2	100	5,67	26,36	19,47	-	5,67	26,10	36,27	
Herbaceous vegetation	3	100	9,44	47,63	25,32	-	9,44	41,38	17,06	
Cultivated and managed vegetation / agriculture	4	100	17,61	50,35	16,42	-	17,61	54,36	15,97	
Other shrub/herbaceous vegetation	5	100	0,11	4,54	78,50	-	0,11	3,61	15,05	
Urban/built up	6	100	15,23	39,57	43,28	-	15,23	35,24	26,97	
Water, Snow and Ice	7	100	98,81	0,93	0,19	-	98,81	0,15	0,03	
Barren	8	100	2,27	35,80	44,19	-	2,27	3,68	0,51	