

## Production of Global Land Cover Data – GLCNMO2008

Ryutaro Tateishi<sup>1</sup>, Nguyen Thanh Hoan<sup>2</sup>, Toshiyuki Kobayashi<sup>1</sup>, Bayan Alsaadeh<sup>1</sup>, Gegen Tana<sup>1</sup> & Dong Xuan Phong<sup>1</sup>

<sup>1</sup> Center for Environmental Remote Sensing, Chiba University, Chiba, Japan

<sup>2</sup> Institute of Geography, Vietnam Academy of Science and Technology (VAST), Hanoi, Vietnam

Correspondence: Ryutaro Tateishi, Center for Environmental Remote Sensing, Chiba University, 1-33 Yayoi-cho Inage-ku Chiba, 263-8522, Japan. Tel: 81-43-290-3850. E-mail: [tateishi@faculty.chiba-u.jp](mailto:tateishi@faculty.chiba-u.jp)

Received: June 2, 2014 Accepted: June 19, 2014 Online Published: July 21, 2014

doi:10.5539/jgg.v6n3p99

URL: <http://dx.doi.org/10.5539/jgg.v6n3p99>

### Abstract

A fifteen-second global land cover dataset — GLCNMO2008 (or GLCNMO version 2) was produced by the authors in the Global Mapping Project coordinated by the International Steering Committee for Global Mapping (ISCGM). The primary source data of this land cover mapping were 23-period, 16-day composite, 7-band, 500-m MODIS data of 2008. GLCNMO2008 has 20 land cover classes, within which 14 classes were mapped by supervised classification. Training data for supervised classification consisting of about 2,000 polygons were collected globally using Google Earth and regional existing maps with reference of this study's original potential land cover map created by existing six global land cover products. The remaining six land cover classes were classified independently: Urban, Tree Open, Mangrove, Wetland, Snow/Ice, and Water. They were mapped by improved methods from GLCNMO version 1. The overall accuracy of GLCNMO2008 is 77.9% by 904 validation points and the overall accuracy with the weight of the mapped area coverage is 82.6%. The GLCNMO2008 product, land cover training data, and reference regional maps are available through the internet.

**Keywords:** land cover, MODIS, decision tree method, Global Mapping Project

### 1. Introduction

Land cover is a basic environmental parameter. Since the first global land cover dataset by satellite data, GLCC, was produced using AVHRR data of 1992 (Loveland & Belward, 1997; Loveland et al., 1999) [Web references: w1, hereafter simply describe 'w1' for a website reference], several products have been released by different initiatives. The University of Maryland produced global land cover data using AVHRR data (Hansen et al., 2000). Boston University has also produced MODIS 1-km land cover data (Friedl et al., 2002). The Joint Research Centre (JRC) of the European Commission produced GLC2000 using SPOT/VEGETATION data of 2000 (Bartholome & Belward, 2005) [w2]. The European Space Agency (ESA) published higher resolution global land cover data, GlobCover, using 300 m resolution ENVISAT/MERIS of 2005-2006 and 2009 [w3]. The Global Mapping project by national mapping organizations published 1-km global land cover data, GLCMO version 1 using MODIS of 2003 (Tateishi et al., 2011).

The current trends of global land cover mapping are: (i) to produce a new map with higher accuracy, (ii) to produce a new map with better resolution, (iii) to use multiple existing maps for the improvement (Giri et al., 2005; Herold et al., 2008), (iv) to publish and share the training data (Tateishi et al., 2011).

In response to "Agenda21" adopted at the "Earth Summit" in 1992, the "Global Mapping" project was proposed for contribution to global environmental conservation from National Mapping Organizations. Later in 1996, the International Steering Committee for Global Mapping (ISCGM) was established [w4]. The Global Mapping Project is an international project to develop eight basic global datasets including land cover through cooperation of National Mapping Organizations from about 180 countries/regions. The product of land cover is called the Global Land Cover by National Mapping Organizations (GLCNMO) (Tateishi et al., 2011).

This paper describes how authors produced a new global land cover data, GLCNMO version 2 under the Global Mapping Project. The features of this mapping are (a) 500m resolution, (b) use of six existing global land cover data products to derive 'potential land cover map', (c) improvement of the method of individual mapping for six classes, and (d) release training data used in this mapping project.

## 2. Data Used

### 2.1 MODIS Data

The main data used in the GLCNMO2008 mapping project are MODIS data observed in 2008. The source MODIS data are the MODIS/Terra+Aqua Nadir BRDF-Adjusted Reflectance 16-day L3 Global 500m SIN Grid V005 (MCD43A4). These are 16-day composite, 7-band, 500-m, 10-degree tile data. The source MODIS data were reprojected into a latitude/longitude coordinate system, and mosaicked to produce the data of five continents (Eurasia, North America, South America, Africa, and Oceania) and one ocean region, South Pacific Ocean Islands, as shown in Figure 1. Mosaicking and reprojection were done using MODIS Reprojection Tool (MRT) software distributed from U.S. Geological Survey. 15 arc-second resampled pixels are nominally called 500 m resolution because it is approximately 500 m at the equator. Geometric accuracies after mosaicking ranges from 96 m to 200 m in RMSE for the east–west direction or north–south direction in four continents except Oceania when compared with Landsat images from GLCF from the University of Maryland as a correct location. The RMSEs of Oceania are 264 m and 344 m for the east–west direction or north–south direction, respectively.

Cloud-contaminated pixels were replaced through linear interpolation of two cloud-free pixels before and after the cloud pixel in cases where the cloud-contaminated period is equal or less than six 16-day periods. If the cloud-contaminated period is more than six 16-day periods, the cloud-contaminated pixel was replaced by the average of 2007 and 2009 MODIS data at the same location and the same time of the year. Preprocessed MODIS data consist of 23 periods of 16-day composite in 2008. More detail explanation of the preprocessing of MODIS data was published by the authors (Hoan et al., 2013).

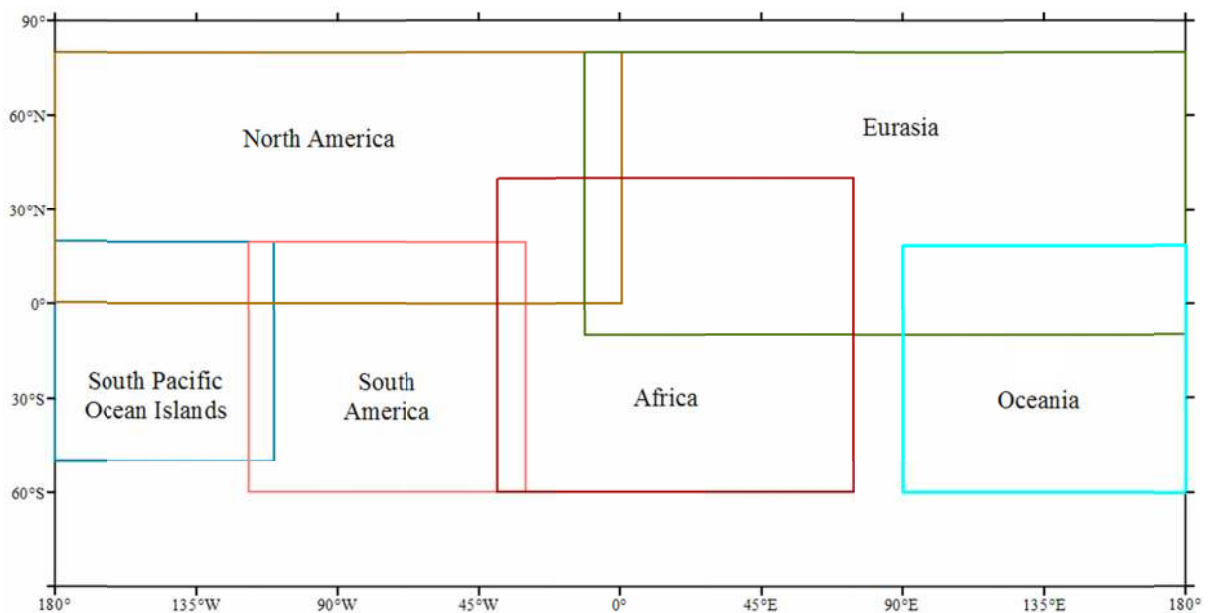


Figure 1. MODIS data regions, after mosaicking and reprojection, of each of the five continents and south pacific ocean islands (North America, South America, Africa, Eurasia, Oceania and South Pacific Ocean Islands)

### 2.2 Other Data

Other data used in this study are as follows.

#### (1) Satellite data

Two types of satellite data were used except MODIS data.

- Landsat ETM+ data: (used for wetland mapping, mangrove mapping)
- DMSP-OLS data (Version 4 DMSP-OLS Nighttime Lights Time Series, 2008): (used for urban mapping)

The data are cloud-free composites made using all the available archived DMSP-OLS smooth resolution data for calendar years. The products are 30 arc second grids, spanning -180 to 180 degrees longitude and -65 to 65 degrees latitude [w5].

#### (2) Global land cover data

The following six existing global land cover products were used to select candidates of land cover training data by producing 'potential land cover maps' (see the section 4 in this paper).

- GLCC v2.0: 1 km (resolution), derived from AVHRR 1992 -1993, by USGS [w1]
- UMD LC: 1 km, derived from AVHRR 1992-1993, by University of Maryland [w6]
- GLC2000: 1 km, derived from SPOT-VEGETATION from November 1999 to December 2000, by Joint Research Centre [w2]
- MOD12Q1 V004: 1 km, derived from MODIS 2001, by Boston University
- GLCNMO v1: 1 km, derived from MODIS 2003, by ISCGM [w7]
- GlobCover v.2.2: 300m, derived from MERIS from December 2004 to June 2006, by ESA [w3]

#### (3) Regional land cover maps

In the mapping process of GLCNMO version 1 (or GLCNMO2003), 164 existing regional maps were used to check intermediate classification results and to select candidates of land cover training data. The same map images were used in the process to select candidates of land cover training data. The complete list of 164 maps is available from the CEReS website [w8]. These maps can be displayed and downloaded by the data sharing/overlay system, CEReS Gaia [w9] (Tateishi et al., 2012; Tateishi & Sri Sumantyo, 2013).

In addition to the above data, the following four land cover data were also used for Europe and North America.

- -CORINE Land Cover 2006, raster data, 100 m resolution, version 16 (04/2012) [w10]
- 2005 Land Cover of North America at 250 meters, Edition: 1.0 [w11]
- Land Cover map of Canada, resolution 30 m [w12]
- NLCD 2006 Land Cover, Edition: 1.0, by USGS, resolution 30 m [w13]

#### (4) Other data

Many other data were used for individual mapping as described in section "5.3 Individual mapping" (see Appendix A).

### 3. Legend

The land cover legend for GLCNMO2008 is the same as the first version of GLCNMO. The legend consists of 20 land cover classes as shown in Table 1. The detail definition of the legend based on Land Cover Classification System (LCCS) [w14] was given by Tateishi et al. (2011).

Table 1. Land cover legend for GLCNMO

| Code | GLCNMO land cover class                | LCC Label  |
|------|--|--|
| 1    | Broadleaf Evergreen Forest             | Broadleaved Evergreen Closed to Open (100-40)% Trees   |
| 2    | Broadleaf Deciduous Forest             | Broadleaved Deciduous Closed to Open (100-40)% Trees   |
| 3    | Needleleaf Evergreen Forest            | Needleleaved Evergreen Closed to Open (100-40)% Trees  |
| 4    | Needleleaf Deciduous Forest            | Needleleaved Deciduous Closed to Open (100-40)% Trees  |
| 5    | Mixed Forest                           | Broadleaved Closed to Open Trees /Needleleaved Closed to Open Trees  |
| 6    | Tree Open                              | Open (40 - (20-10)%) Trees (Woodland)  |
| 7    | Shrub                                  | Closed to Open Shrubland (Thicket)   |
| 8    | Herbaceous                             | Closed to Open Herbaceous Vegetation, Single Layer   |
| 9    | Herbaceous with Sparse Tree / Shrub    | Closed to Open Herbaceous Vegetation with Trees and Shrubs   |
| 10   | Sparse Vegetation                      | Sparse Herbaceous Vegetation //Sparse Woody Vegetation   |
| 11   | Cropland                               | Herbaceous Crop(s)   |
| 12   | Paddy field                            | Graminoid Crops //Non-Graminoid Crops  |
| 13   | Cropland /Other Vegetation Mosaic      | Cultivated and Managed Terrestrial Area(s) /Natural And Semi-Natural Primarily Terrestrial Vegetation //Cultivated Aquatic or Regularly Flooded Area(s) /Natural And Semi-Natural Primarily Terrestrial Vegetation |
| 14   | Mangrove                               | Closed to Open Woody Vegetation Water Quality: Saline Water  |
| 15   | Wetland                                | Closed to Open Woody Vegetation Water Quality: Fresh Water // Closed to Open Woody Vegetation Water Quality: Brackish Water // Closed to Open Herbaceous Vegetation  |
| 16   | Bare area, consolidated (gravel, rock) | Consolidated Material(s)   |
| 17   | Bare area, unconsolidated (sand)       | Unconsolidated Material(s)   |
| 18   | Urban                                  | Artificial Surfaces and Associated Area(s)   |
| 19   | Snow / Ice                             | Perennial Snow // Perennial Ice  |
| 20   | Water Bodies                           | Artificial Waterbodies // Natural Waterbodies  |

#### 4. Use of Existing Global Land Cover Products

In order to use the existing global land cover products effectively for a new mapping attempt, six products (GLCC, UMD LC, GLC2000, MOD12Q1, GLCNMO, GlobCover; see section 2.2 (2)) were overlaid to produce a new map showing reliability of the mapped result for each class. This map is called 'potential land cover map' in this study. Figure 2 shows an example of 'potential land cover map' for cropland. For example, red color shows the area of cropland in all six existing products which means this area is most probably cropland. On the other hand, green color shows less reliable cropland areas because this areas are cropland by only three products out of six. 'Potential land cover maps' for all 20 classes were produced as reference information for the selection of training data or validation data. Its land cover type was identified by using other materials stated in the section 5.2.1.

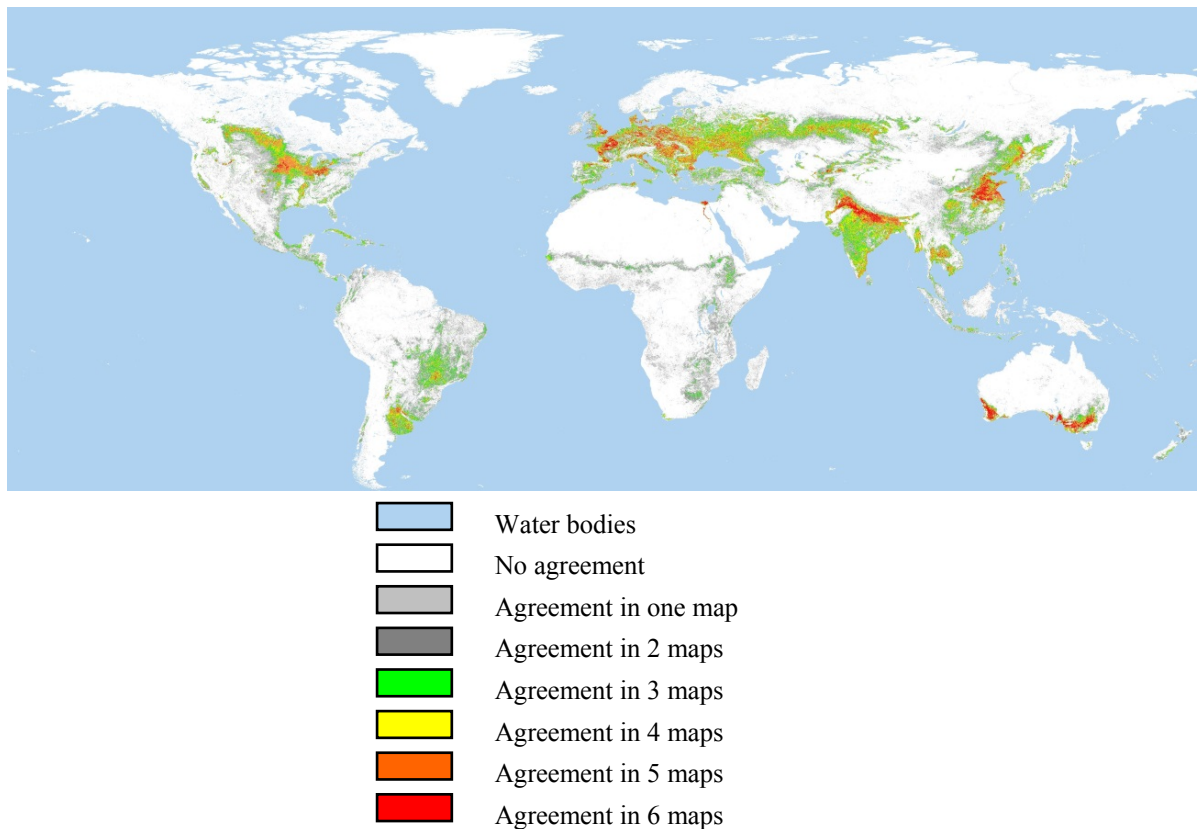


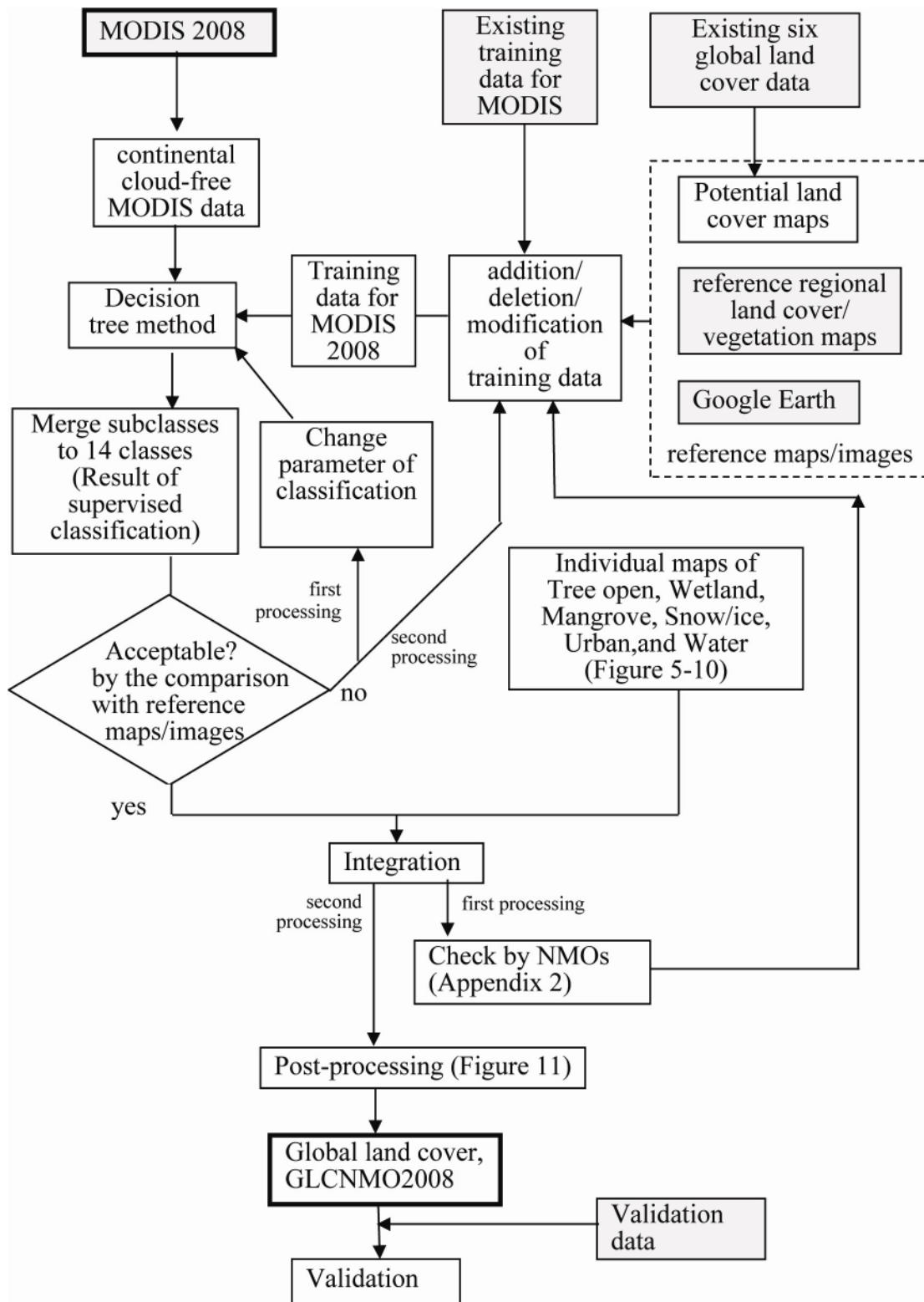
Figure 2. Potential map of cropland from six existing global land cover map

## 5. Land Cover Mapping

### 5.1 Method of Classification

The basic method to produce GLCNMO2008 is the combination of supervised classification and individual mapping for six classes (Tree open, Wetland, Mangrove, Snow/Ice, Urban, and Water). This method is an improved method of the previous mapping of GLCNMO2003 (Tateishi et al., 2011). Figure 3 shows the whole flow of the mapping for GLCNMO2008. The main features of the method are as follows.

- i) Reuse of training data for GLCNMO2003 after verification and addition, for GLCNMO2008 (see the section 5.2.1)
- ii) Use of existing global land cover products as ‘potential land cover maps’ (see the section 4)
- iii) Comparison of two supervised classification methods: decision tree method and maximum likelihood method (see the section 5.2.2)
- iv) Improvement of individual mapping for six classes from the method of GLCNMO2003 (see the section 5.3)



A shaded box means the data used.

Figure 3. Flow of GLCNMO2008 mapping

## 5.2 Supervised Classification

### 5.2.1 Training Data

The existing training data for the previous version of GLCNMO (Tateishi et al., 2011) which is available from the CERES website, were used as the initial land cover training data. These data were updated as follows.

- a) The existing training data were examined whether they were acceptable to use for the mapping of 2008 or not, using the following references.
  - 'Potential land cover maps'
  - Google Earth images around 2008
  - Existing regional land cover maps (see section 2.2 (3))
  - NDVI phenological curves of MODIS2008
  - Color composite images of MODIS for four seasons (January, April, July, and November)
- b) Candidate areas of additional training data for a specific class were selected with the reference of 'potential land cover maps.' This is a good example of effective use of existing global land cover maps. The candidate areas were examined using the same references of the step a), and identified their land cover types.

The above training data preparation was done by experts carefully. Through this process, 14 land cover classes were divided into further detail subclasses based on NDVI phenological curves and geographic locations. The numbers of subclasses are 1000, 525, 297, 210, 18 for Eurasia, North America, Africa, South America, and Oceania, respectively. Large number of subclass is herbaceous and cropland in Eurasia. These two classes have totally 509 subclasses. Finally, 312,753 pixels from 2,080 polygons for all subclasses were collected as globally distributed training data. Figure 4 shows distribution of 2,080 training polygons.

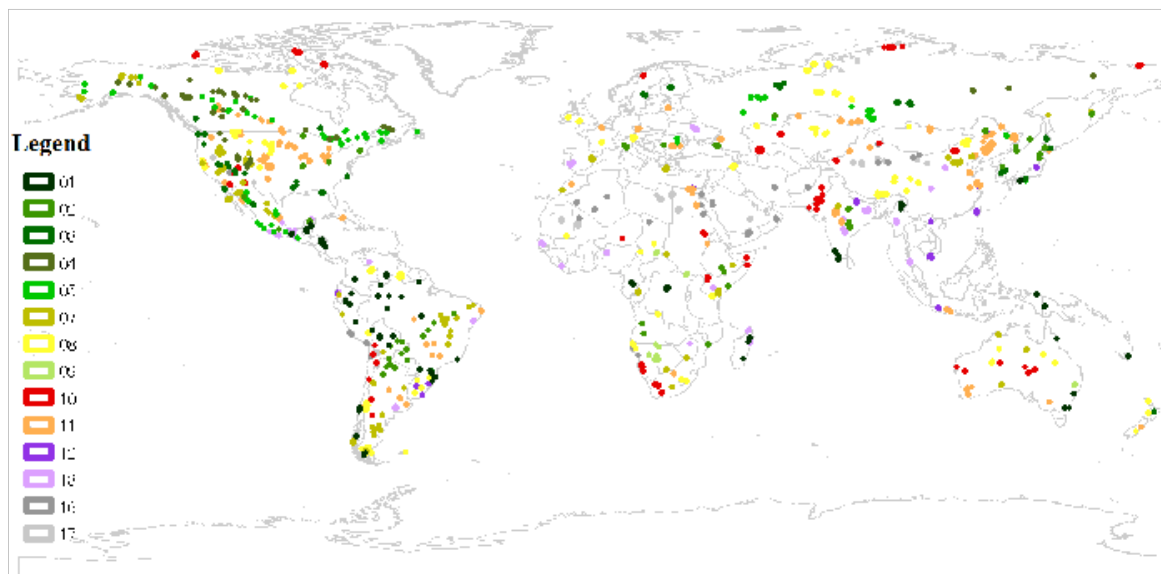


Figure 4. Distribution of 2080 training polygons

The number in the legend corresponds to the code in Table 1. Multiple dots at near locations are represented as one dot.

### 5.2.2 Decision Tree Method

Variables used for supervised classification are MODIS five bands (band 1, 2, 4, 6, and 7) and NDVI of 12 periods (period 1, 3, 5, ---, 23) in 2008, totally  $(5+1) \times 12 = 72$  variables. Decision tree method (commercial software, See5) and maximum likelihood method (by commercial software, ENVI) were applied to continental MODIS data using training data described above. Many subclasses classified were merged to 14 classes and compared with 'potential land cover maps' and reference regional maps visually. Then, decision tree method was selected because of less apparent errors. Another reason to select decision tree is that it is easier to modified classified result using a parameter than maximum likelihood method.

As a result, 14 classes were classified by decision tree method for five continents. The classification was done by continental basis. MODIS data of “South Pacific Ocean Islands” was merged to that of “South America,” and classified together. After decision rules were obtained by See5 software, classification using these rules for each pixel was performed by the software, National Land Cover Development (NLCD) distributed by U.S. Geological Survey.

Classified results of 14 classes were compared with reference maps/images as shown in Figure 3 visually. If the result is not acceptable, first, a parameter of decision tree method is changed, and secondly, training data are modified.

After the integration of 14 classes with other six classes, the result was verified by National Mapping Organizations (NMOs) (Appendix B). Based on the feedback from NMOs, some training data were modified/deleted and some were added. The number of training pixels and polygons mentioned at the end of 5.2.1 are the final one.

### 5.3 Individual Mapping

Six land cover classes—Tree open, Wetland, Mangrove, Snow/ice, Urban, Water—were extracted independently as follows.

#### 5.3.1 Tree open

Global percent tree cover data were produced from MODIS data of 2008 by the authors using decision tree method. The method to produce percent tree cover is being submitted as another paper (Kobayashi and Tateishi 2013). This product indicates a percentage of area coverage of tree canopy per a unit area (pixel). Pixels with 15–40% were assigned as “Tree Open” Because forest class was defined as 40-100 % tree coverage in this project. The threshold value of 40% was selected based on Forest Resources Assessment (FRA) program by the Food and Agriculture Organization (FAO) of the United Nations (Zhu & Waller, 2003).

#### 5.3.2 Wetland

The wetland mapping except North America was carried out by the following steps (see Figure 5).

Step 1: Tasseled cap transformation is applied to MODIS seven bands data for each period of 2008. Three indices of Brightness, Greenness, and Wetness were obtained for each period. Coefficients of Tasseled cap transformation are based on Lobser and Cohen (2007).

Step 2: In order to reduce much manual operation, wetlands to be mapped were limited to the ones larger than 1000 km<sup>2</sup> by Ramsar Sites Database (Appendix A, 1). Wetlands in 90 selected regions were extracted.

Step 3: For each wetland, the best index and period to separate wetland and non-wetland are selected manually with the reference of Google Earth images.

Step 4: Threshold value is decided for the selected index/period image with the reference of Google Earth image.

Step 5: Wetland map is obtained by integrating extracted wetlands by the previous step.

The detail methodology of wetland mapping was published by GegenTana and Tateishi (2013).



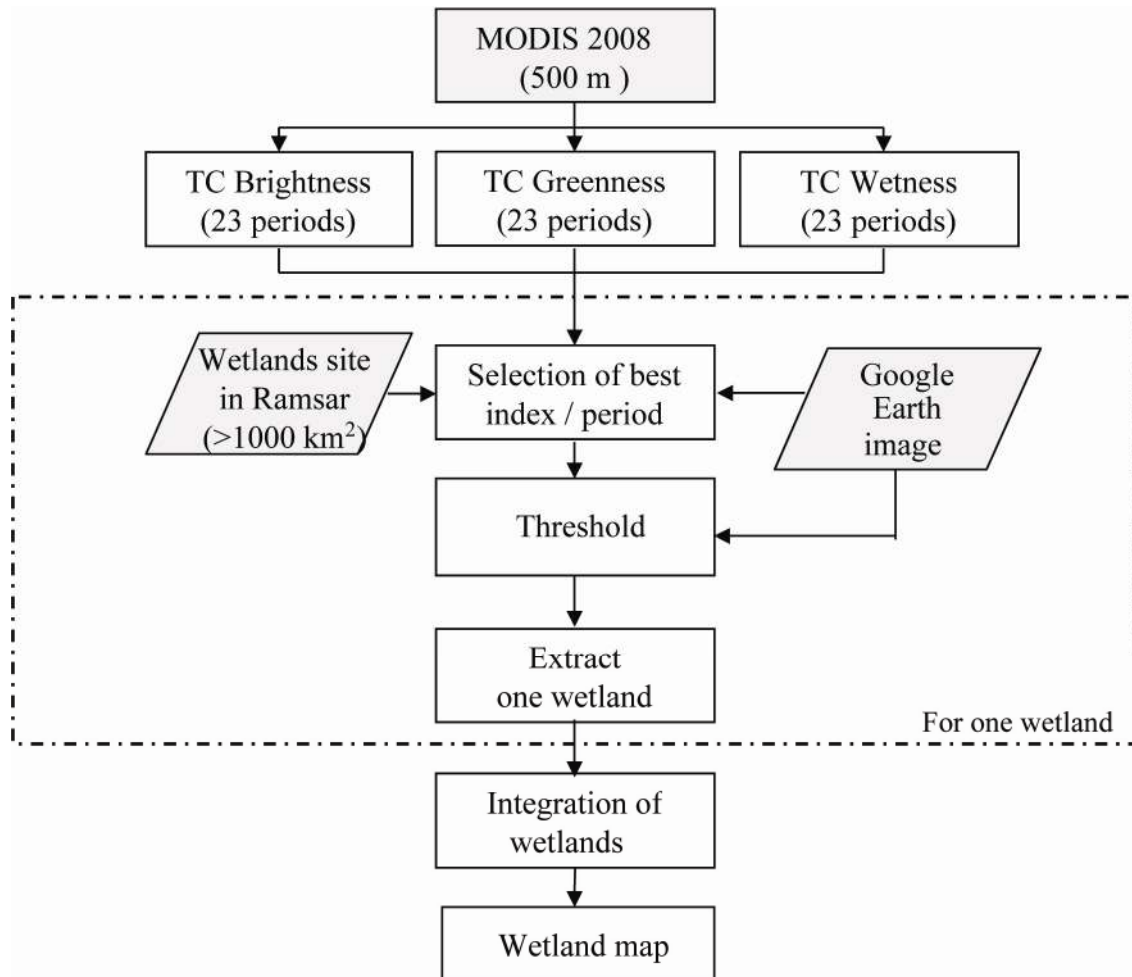


Figure 5. Flow of wetland mapping except North America

In North America, reliable wetland data are available; U.S. National Wetlands Inventory (Appendix A, 4), Canadian Wetland Inventory (Appendix A, 5), and Canadian land cover data (Appendix A, 6). Therefore, in North America, more sophisticated mapping method than the other continents was applied as the following steps (see Figure 6).

Step 1: Same as the above Step 1.

Step 2: Training data of six classes, wetland, forest, shrub, herbaceous, cropland, and sparse vegetation were collected with the reference of Ramsar Sites Database, U.S. National Wetlands Inventory and Canadian Wetland Inventory, and 'potential land cover maps' described in 5.2.1 a).

Step3: Decision tree method was applied to 69 variables consisting of 23 periods and three indices (Brightness, Greenness, and Wetness) from MODIS with the training data derived in the above step. This step was carried out independently in five sub-regions of Köppen-Geiger climate map (Appendix A, 2).

Step 4: Four masks of water, snow/ice, urban/bare, and elevation were created in order to exclude wetland areas in these regions. Water mask was produced by the threshold of Tasseled cap Brightness index of the period 13. Snow/ice mask was produced by the threshold of the mean value of Tasseled cap Brightness. Urban and bare area mask was obtained from GLCNMO version 1. The threshold of elevation mask was determined based on the highest elevation of the wetland training data, which was 1236 m. The elevation data used is GTOPO30 (Appendix A, 3).

Step 5: The final wetland map was produced by integrating wetlands classified in the step 3 and excluding wetlands in the masked regions in the step 4.

The detail methodology of wetland mapping in North America was published by GegenTana et al. (2013).

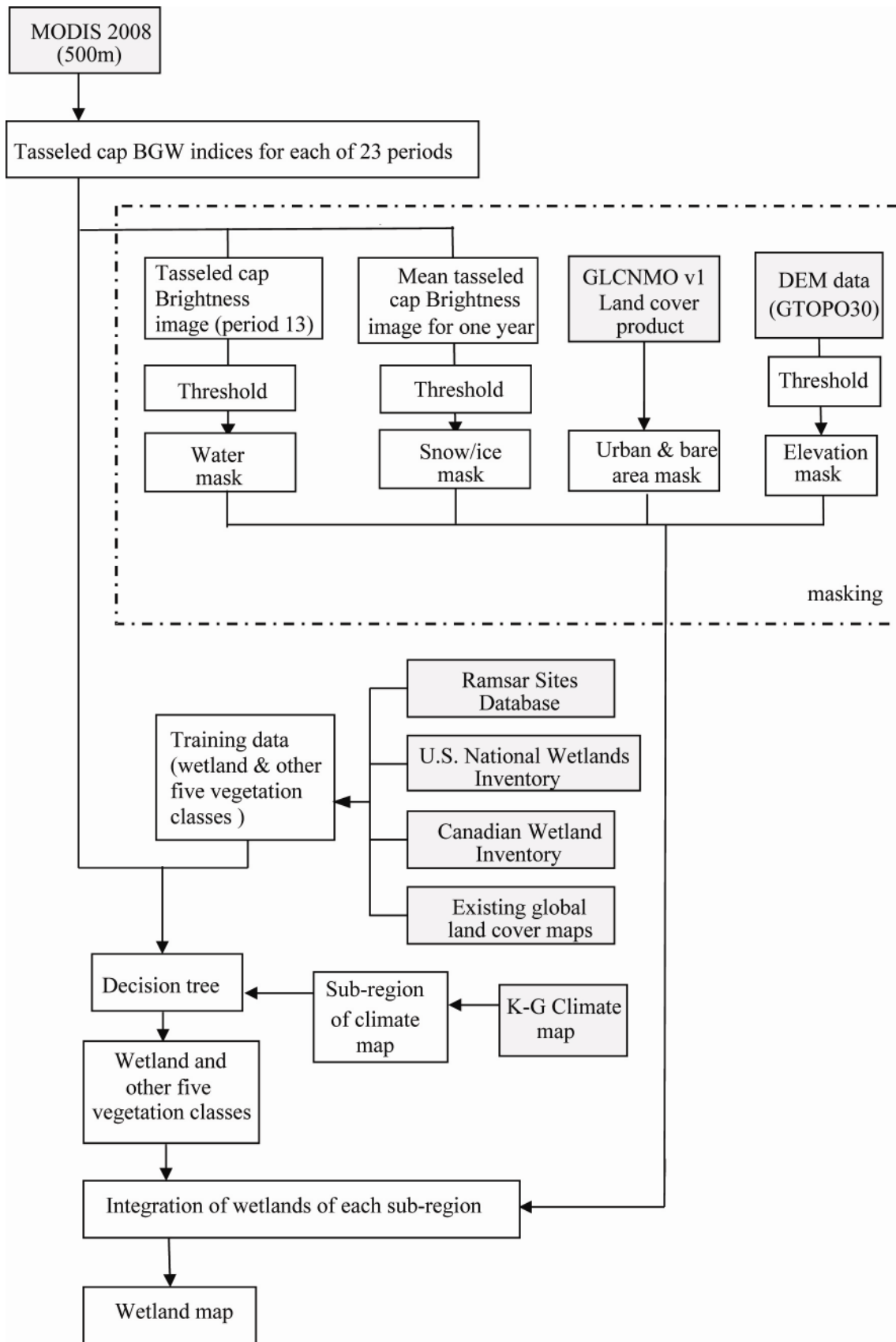


Figure 6. Flow of wetland mapping in North America

### 5.3.3 Mangrove

The mangrove mapping was carried out by the following steps (see Figure 7).

Step 1: Color composite of MODIS band 6 for red, MODIS band 6/band 2 for green, and the difference of normalized difference water index (NDWI) and normalized difference vegetation index (NDVI) for blue, was produced for the collection of training data for each continent. Here, NDWI is defined by  $(NIR-SWIR)/(NIR+SWIR) = (band2-band6)/(band2+band6)$  (Gao 1996). MODIS band 6 was selected because of lower DN value for mangrove forest. The selection of a ratio, band 6/band 2, was based on the study by Green et al. (1998), and the selection of NDWI-NDVI was based on Tong et al. (2004).

Step 2: Training data of mangrove and other land cover classes (ex. forest, cropland, shrub, fish pond, urban) were collected by the visual interpretation of color composite images produced in the step 1. In this process, higher elevation than 20 m was excluded from training data of mangrove.

Step 3: Maximum likelihood method was applied to four variables of MODIS band 6, band 6/band 2, NDWI-NDVI, and elevation data (Appendix A, 7). The classification was applied for each continent.

Step 4: Within the classified mangrove regions, a single isolated pixel was removed. Then, the final mangrove map was produced.

There are detail reference materials available in Japan, Cambodia such as aerial photographs in Japan, vegetation maps at scales 1:50,000 and 1:25,000 in Japan, ground survey information in Cambodia, and Cambodia Atlas (Appendix A, 8-11). Therefore, in these regions, Landsat ETM+ 30m data were used instead of 500m MODIS. The methodology to map mangrove is the same as the above steps. The detail method was published by the author (Alsaideh et al., 2013a, 2013b).

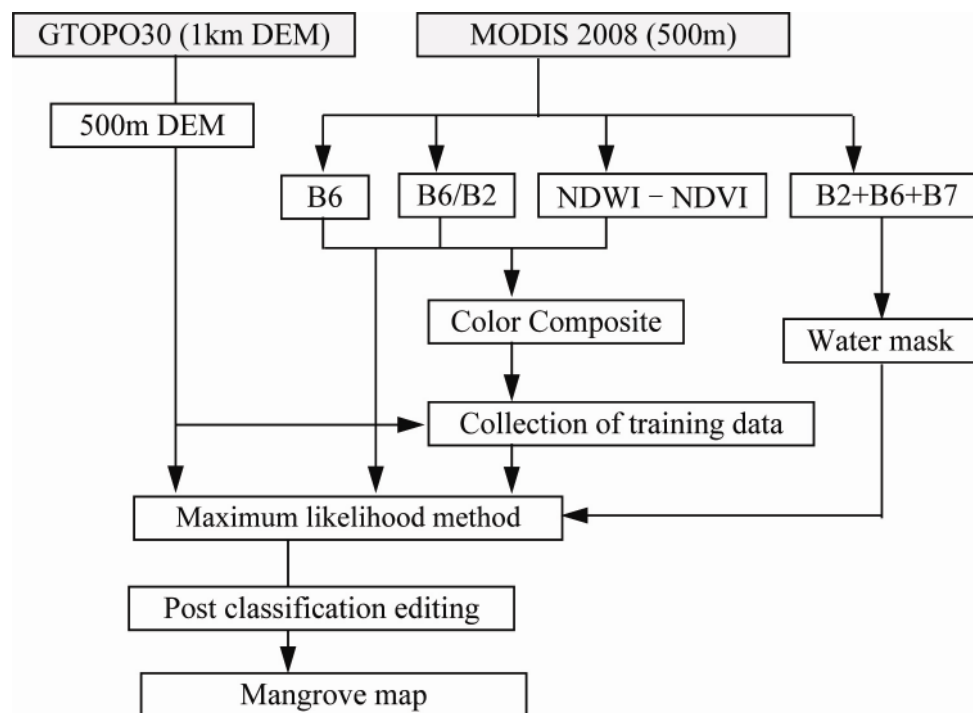


Figure 7. Flow of mangrove mapping

### 5.3.4 Snow/Ice

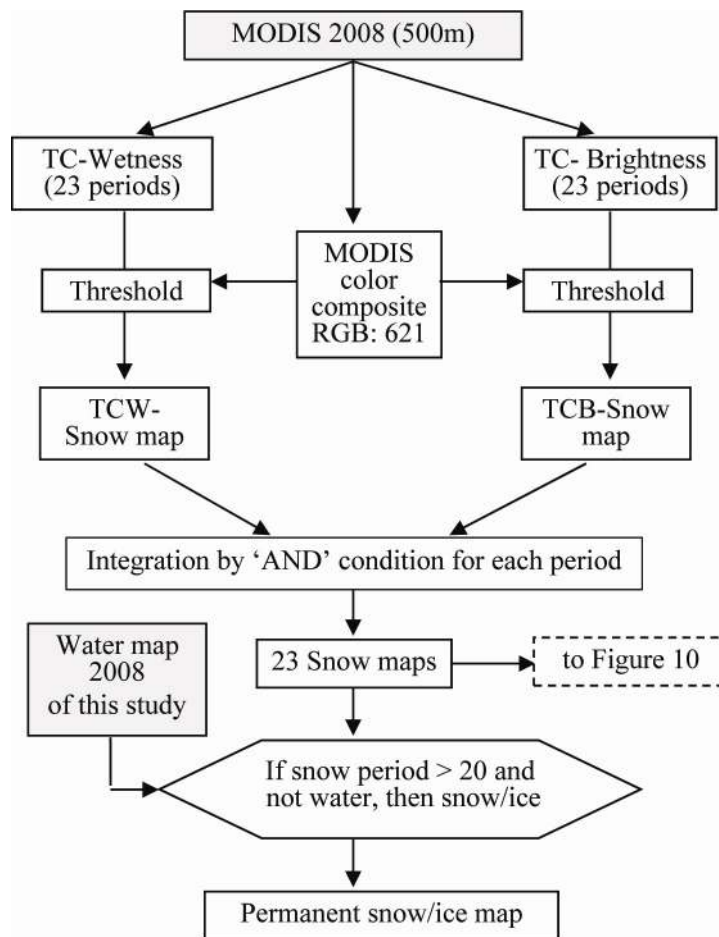
The permanent snow/ice mapping was carried out by the following steps (see Figure 8).

Step 1: Tasseled cap transformation is applied to MODIS seven bands data for each period of 2008. Two indices of Brightness and Wetness were obtained for each period. And color composite of band 6 (red), band 2 (green), and band 1 (blue) of MODIS were produced for each period of 2008.

Step 2: Snow areas were mapped by the threshold of Wetness data for each period with the reference of visual interpretation of color composite of MODIS. This is TCW-Snow map in Figure 8. Similarly, TCB-Snow map were produced using Brightness data for each period.

Step 3: Twenty-three snow maps were derived by the “AND” operation of TCW-Snow map and TCB-Snow map.

Step 4: The final permanent snow/ice map was produced by the logical operation using 23 snow maps and a water map of 2008. If number of snow periods is more than 20 and the other periods are not water, this pixel is decided as permanent snow/ice. Here, the water map is the same one as described in section 5.3.6 of this paper, that is a part of GLCNMO version 2.



TCW: Tasseled Cap Wetness,  
TCB: Tasseled Cap Brightness

Figure 8. Flow of snow/ice mapping

### 5.3.5 Urban

The urban mapping was carried out using mainly population data, DMSP-OLS night-time light data, constructed impervious surface area data ISA (Appendix A, 12-14), and MODIS NDVI data by the following steps (see Figure 9). Resolutions of 1 km and 500m in Figure 9 are actually 30 arc-second and 15 arc-second respectively. This is the same for other individual mappings.

Step 1: Population density for 1-km (30 arc-second) pixel is derived from population data (Appendix A, 12) and area of a pixel. 1-km (30 arc-second) population density data are resampled to 500 m (15 arc-second) by generating 2 by 2 same value pixels.

Step 2: Candidate of urban area is derived from population density data by the threshold which is decided by visual interpretation of Landsat ETM+ images and Google Earth images. Four threshold values are decided based on Gross Domestic Product (GDP) per capita. That is, higher GDP (Appendix A, 14) countries have higher threshold value to extract candidate of urban areas.

Step 3: In order to exclude less night-time light area from urban, DMSP-OLS data were used. The excluded areas were decided by the threshold which is decided by the same way as the step 2. The higher GDP countries have higher threshold value.

Step 4: In order to exclude less impervious surface areas (ISA) from urban, ISA data (Appendix A, 13) were used. As no ISA data is available for 2008, the data for 2010 were used. The excluded areas were decided by the threshold which is decided by the same way as the step 2. The higher GDP countries have higher threshold value.

Step 5: In order to exclude more vegetation areas from urban, MODIS NDVI data were used. The excluded areas were decided by the threshold which is decided by the same way as the step 2. The higher GDP countries have higher threshold value because there are more trees on streets and residential areas in developed countries.

Step 6: The final urban map was produced by excluding areas of less night-time light (step 3), areas of less impervious surface (step 4), and more vegetated areas (step 5) from candidate of urban areas (step 2).

The detail methodology of urban mapping was published by Phong et al. (2013).

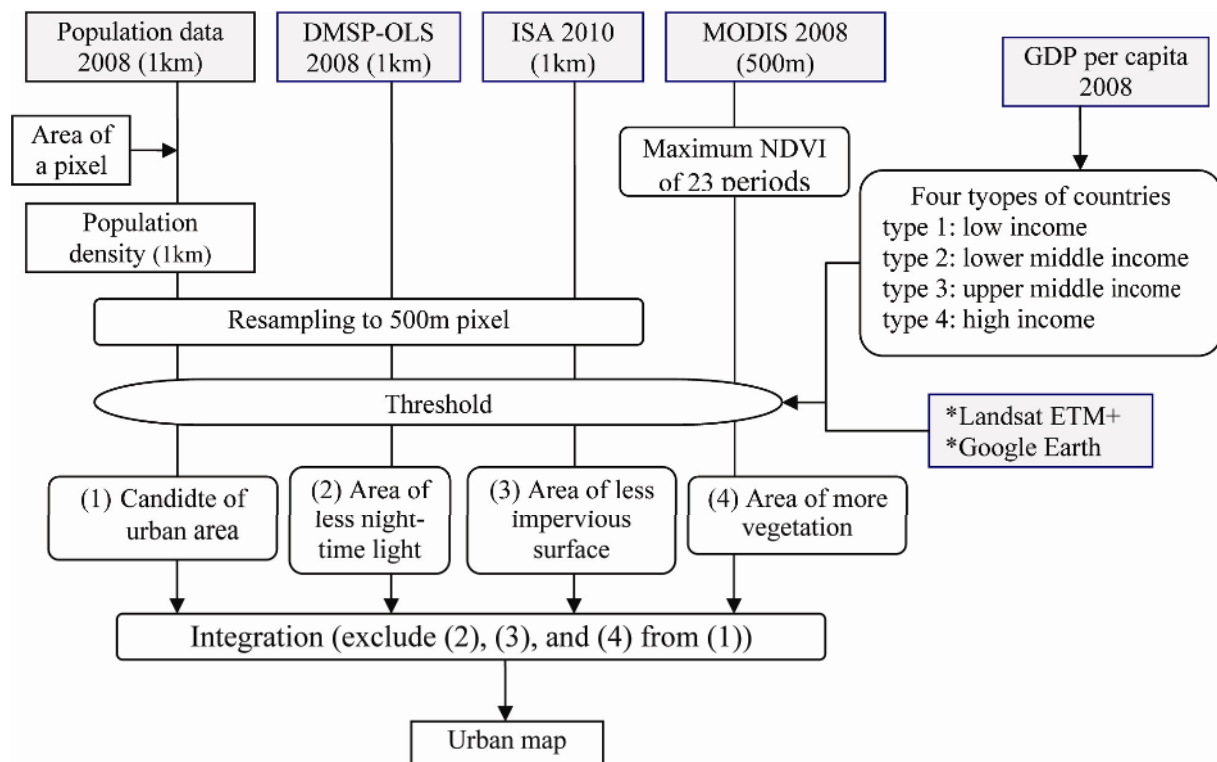


Figure 9. Flow of urban mapping

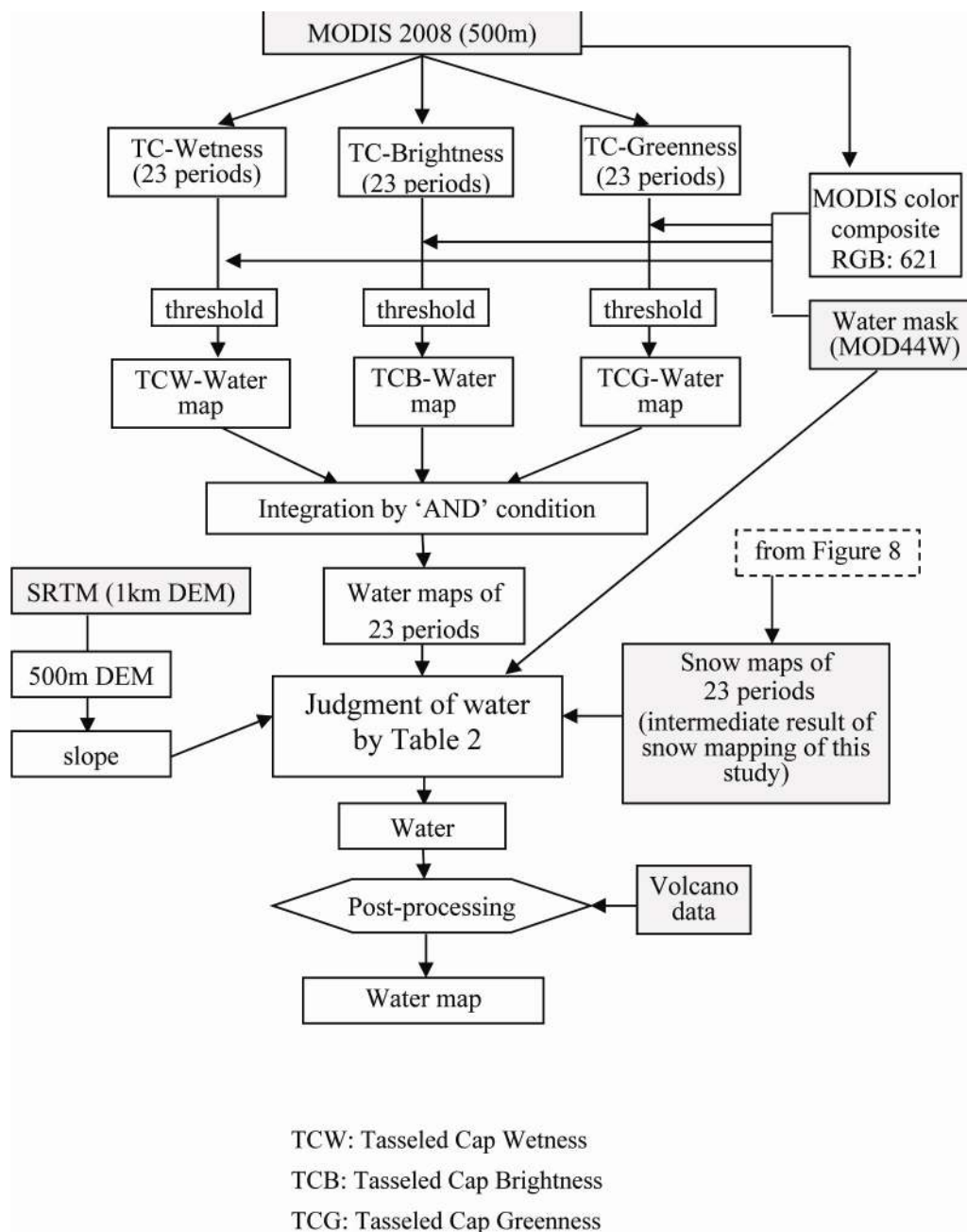


Figure 10. Flow of inland water mapping

### 5.3.6 Water

The water mapping was carried out by the following steps (see Figure 10).

- Step 1: Tasseled cap transformation is applied to MODIS 7 bands data for each period of 2008. Three indices of Brightness, Greenness, and Wetness were obtained for each period.
- Step 2: Color composite of band 6 (red), band 2 (green), and band 1 (blue) of MODIS were produced for each period of 2008.
- Step 3: For each of 23 periods, a water map is produced from Tasseled cap Wetness by selecting a threshold. This water map is TCW-Water map in Figure 10. The selection of the threshold is done by the visual comparison with MODIS color composite image. Similarly, water maps were produced using TC Brightness and TC Greenness independently. They are called TCB-water map and TCG-water map, respectively in Figure 10. The threshold values for water pixel were selected as  $TCW > -800$ ,  $TCB < 4300$ ,

TCG < 1400 in the case the pixel is water in the existing water mask data MOD44W (Appendix A, 15). When the pixel is not water in MOD44W, threshold values are TCW > -70, TCB < 3500, TCG < 120.

Step 4: For each of 23 periods, three water maps from TCW, TCB, and TCG were integrated by logical “AND” operation. As a result, 23 water maps were produced.

Step 5: 23 water maps were integrated using slope data from SRTM DEM, water mask data MOD44W, and 23 snow maps which are the intermediate result of snow data in this study (Step 3 of 5.3.4 snow/ice). The judgment of water pixel is based on threshold values of Table 2.

Step 6: As some volcanos were misclassified to water, the existing volcano data (Appendix A, 16) were used for post-processing to avoid this error. Then the final water map was obtained.

The detail methodology of water mapping was published by Hoan et al. (2012).

Table 2. Judgment of water area

| Latitude (L)           | MOD44W (water mask data) | Condition to be mapped as water  |
|------------------------|--------------------------|--|
| L >= 70 deg.           | Water                    | at least 20 periods of Water or Snow, and slope < 7 degrees                                  |
|                        | Not water                | at least 2 periods of Water, and at least 20 periods of Water or Snow, and slope < 5 degrees |
| 70 deg. > L >= 60 deg. | Water                    | at least 2 periods of Water, and at least 20 periods of Water or Snow, and slope < 7 degrees |
|                        | Not water                | at least 3 periods of Water, and at least 20 periods of Water or Snow, and slope < 5 degrees |
| L < 60 deg.            | Water                    | at least 6 periods of Water, and slope < 10 degrees  |
|                        | Not water                | at least 12 periods of Water, and slope < 7 degrees  |

#### 5.4 Final Mapping by Integration and Post-Processing

Classified result by supervised method and individually extracted six classes were integrated by the order shown in Figure 11. Six classes were overlaid in this order which means that “Water” has the highest priority. One exception is that the class ‘Cropland/Other vegetation Mosaic’ has higher priority than the class “Tree Open”.

As the post-processing, the following two steps were processed.

- 1) Since MODIS data covers up to 80 degree in north latitude, the area of 80-90 degree north was added from GlobCover data. By this step, continental land cover data were completed.
- 2) Five continental data were combined to make global data. For Antarctica, ‘land’ and ‘lake’ of SCAR Antarctic Digital Database was added as ‘Snow/ice’ class and ‘Water’ class, respectively. By this step, the global land cover data, GLCNMO2008, was completed. Figure 12 shows the final map product, GLCNMO2008.

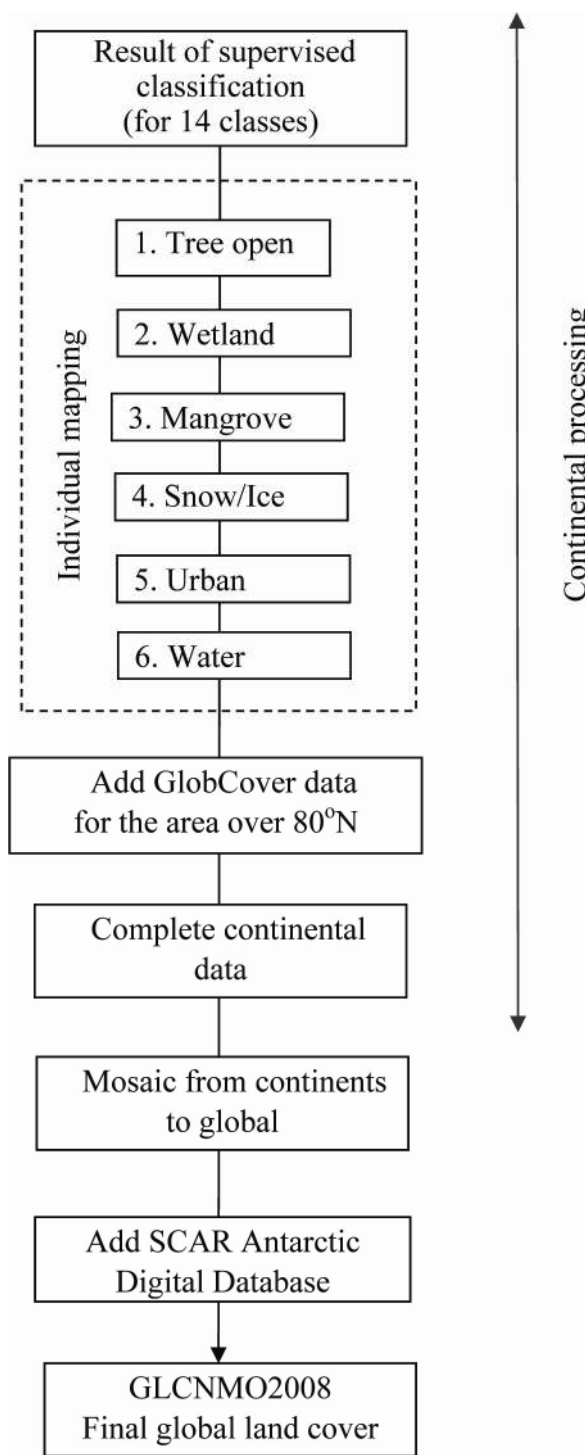


Figure 11. Flow of integration and post-processing



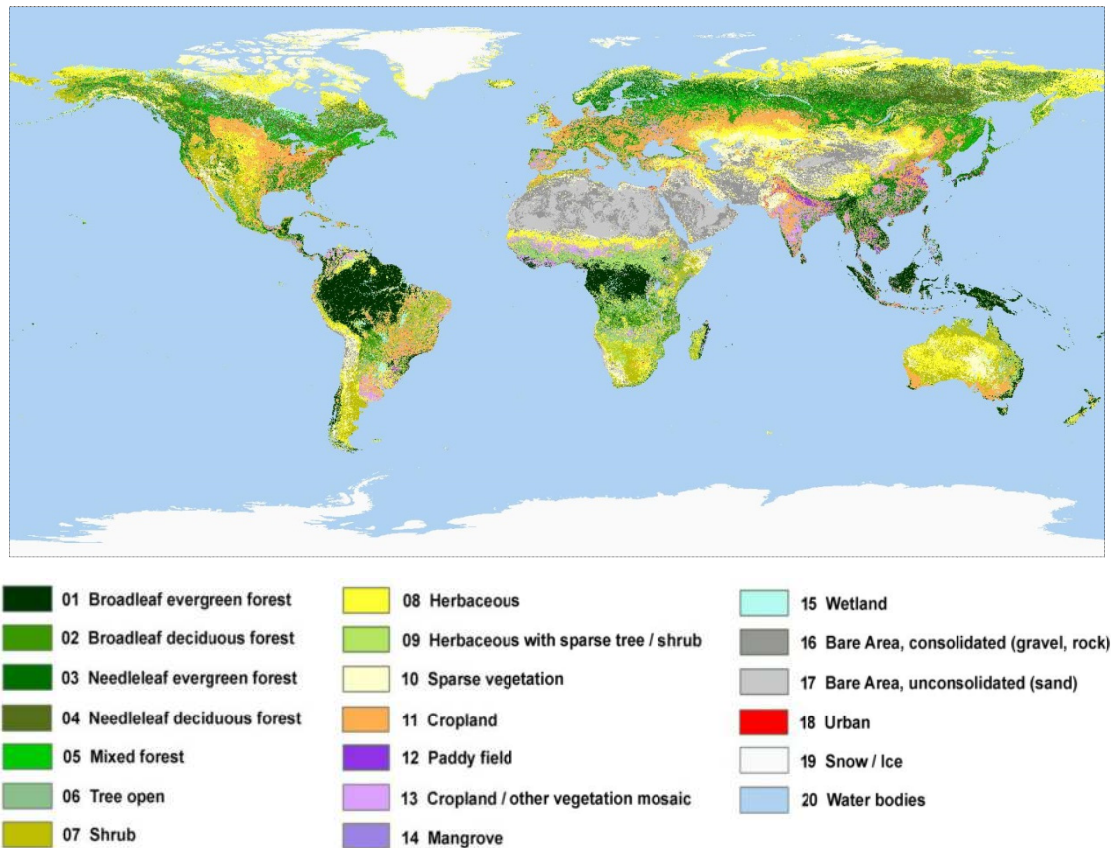


Figure 12. GLCNMO2008 (GLCNMO version2)

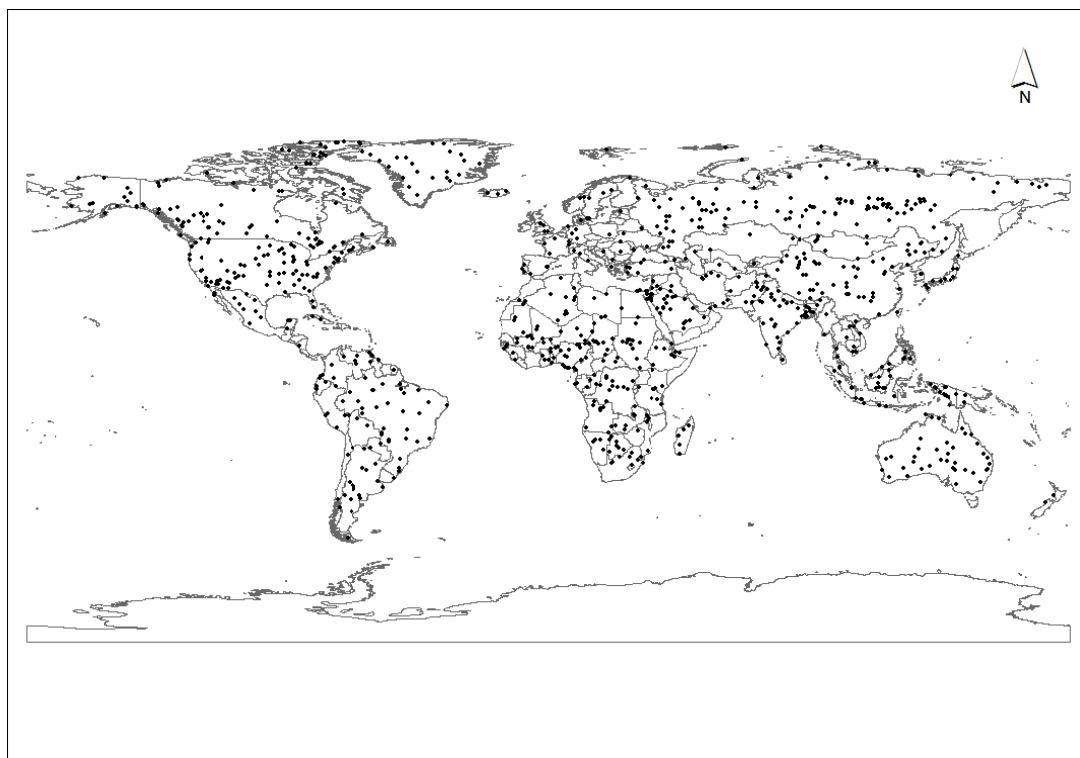


Figure 13. Distribution of 904 validation points for GLCNMO2008

## 6. Validation

Validation points (pixels) were selected by stratified random sampling. “at least 40 points per class” was a target. Therefore approximately 100 points per class were sampled with the assumption of half of them cannot be surely identified its land cover class. Identification of land cover class was made by combined method of i) visual interpretation of Google Earth, ii) referencing regional maps and ‘potential land cover maps’, and iii) MODIS NDVI phenological curves. Finally 904 points were selected for 20 classes globally after the careful check by experts (Figure 13). Table 3 shows a confusion matrix of GLCNMO2008.

Table 3 shows a confusion matrix. The overall accuracy is 77.9%; the average of user’s accuracy is also 78.9%, and the average of the producer’s accuracy is 76.8%. Similar classes were aggregated to produce eight classes: Forest, Other natural vegetation, Cropland, Wetland, Bare area/ sparse vegetation, Urban, Snow/ice, and Water. A confusion matrix for eight aggregated classes is presented in Table 4. The aggregated overall accuracy is 91.4%, the average of user’s accuracy is 94.0%, and the average of producer’s accuracy is 92.0%. The accuracy percentage value increases 13.5% by aggregation of classes from 20 to 8. Another mode of expressing accuracy is the area-weighted accuracy, by which accuracy is calculated with weight values of the classified land area of each class. Table 5 presents both overall accuracies and area-weighted overall accuracy. The area-weighted overall accuracy for 20 classes is increased to 82.6%, and the aggregated case is increased slightly to 92.5%.

Table 3. Confusion matrix of GLCNMO2008

| Class name                                  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20  | Total | User's accuracy (%) |
|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-----|-------|---------------------|
| 1. Broadleaf evergreen forest               | 50 | 1  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 2  | 0  | 0  | 0  | 0  | 0  | 0   | 57    | 88                  |
| 2. Broadleaf deciduous forest               | 1  | 22 | 2  | 4  | 3  | 1  | 1  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 1  | 0  | 0  | 0  | 0  | 0   | 37    | 59                  |
| 3. Needleleaf evergreen forest              | 2  | 4  | 32 | 3  | 6  | 1  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 2  | 0  | 0  | 0  | 0  | 0   | 52    | 62                  |
| 4. Needleleaf deciduous forest              | 0  | 2  | 1  | 30 | 1  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 36    | 83                  |
| 5. Mixed forest                             | 0  | 4  | 2  | 7  | 26 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 39    | 67                  |
| 6. Tree open                                | 1  | 5  | 3  | 7  | 3  | 29 | 6  | 0  | 1  | 0  | 4  | 1  | 5  | 0  | 3  | 0  | 0  | 1  | 0  | 0   | 69    | 42                  |
| 7. Shrub                                    | 0  | 0  | 0  | 0  | 0  | 3  | 26 | 2  | 3  | 2  | 1  | 0  | 1  | 0  | 1  | 0  | 0  | 0  | 0  | 0   | 39    | 67                  |
| 8. Herbaceous                               | 1  | 1  | 0  | 1  | 0  | 0  | 5  | 35 | 3  | 3  | 0  | 1  | 2  | 0  | 3  | 1  | 0  | 0  | 2  | 0   | 58    | 60                  |
| 9. Herbaceous with sparse tree/shrub        | 0  | 1  | 0  | 0  | 0  | 1  | 0  | 0  | 33 | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 36    | 92                  |
| 10. Sparse vegetation                       | 0  | 1  | 0  | 0  | 0  | 0  | 3  | 0  | 1  | 30 | 1  | 0  | 0  | 0  | 0  | 7  | 1  | 0  | 0  | 0   | 44    | 68                  |
| 11. Cropland                                | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 2  | 2  | 0  | 37 | 4  | 6  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 53    | 70                  |
| 12. Paddy field                             | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 2  | 30 | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 33    | 91                  |
| 13. Cropland / other vegetation mosaic      | 0  | 0  | 0  | 0  | 0  | 2  | 0  | 3  | 0  | 0  | 2  | 2  | 21 | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 30    | 70                  |
| 14. Mangrove                                | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 43 | 1  | 0  | 0  | 0  | 0  | 0   | 44    | 98                  |
| 15. Wetland                                 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 30 | 0  | 0  | 0  | 0  | 0   | 30    | 100                 |
| 16. Bare Area, consoli- dated (gravel,rock) | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 3  | 0  | 0  | 0  | 0  | 0  | 30 | 7  | 0  | 0  | 0   | 41    | 73                  |
| 17. Bare Area, uncon- solidated (sand)      | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 3  | 0  | 0  | 0  | 0  | 0  | 2  | 39 | 0  | 0  | 0   | 44    | 89                  |
| 18. Urban                                   | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 53 | 0  | 0   | 53    | 100                 |
| 19. Snow / ice                              | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 53 | 0   | 54    | 98                  |
| 20. Water bodies                            | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 55  | 55    | 100                 |
| Total                                       | 55 | 41 | 40 | 52 | 40 | 41 | 43 | 42 | 43 | 41 | 47 | 41 | 40 | 45 | 41 | 41 | 47 | 54 | 55 | 55  | 904   |                     |
| Producer's accuracy (%)                     | 91 | 54 | 80 | 58 | 65 | 71 | 60 | 83 | 77 | 73 | 79 | 73 | 53 | 96 | 73 | 73 | 83 | 98 | 96 | 100 |       |                     |

Overall accuracy = 77.9%.

Average of user's accuracy = 78.9%.

Average of producer's accuracy = 76.8%.

Table 4. Confusion matrix of aggregated classes of GLCNMO2008

| Class name                             | Corresponding classes before aggregation  | 1. Forest | 2. Onv | 3. Crop | 4. Wet | 5. Bare | 6. Urban | 7. Snow | 8. Water | Total | User's accuracy (%) |
|--|---|-----------|--------|---------|--------|---------|----------|---------|----------|-------|---------------------|
| <b>1. Forest</b>                       | 1. Broadleaf evergreen forest<br>2. Broadleaf deciduous forest<br>3. Needleleaf evergreen forest<br>4. Needleleaf deciduous forest<br>5. Mixed forest<br>6. Tree open | 257       | 8      | 16      | 8      | 0       | 1        | 0       | 0        | 290   | 88.6                |
| <b>2. Other natural vegetation</b>     | 7. Shrub<br>8. Herbaceous<br>9. Herbaceous with sparse tree/shrub   | 8         | 107    | 6       | 4      | 6       | 0        | 2       | 0        | 133   | 80.5                |
| <b>3. Cropland</b>                     | 11. Cropland<br>12. Paddy field<br>13. Cropland/other vegetation mosaic   | 3         | 8      | 105     | 0      | 0       | 0        | 0       | 0        | 116   | 90.5                |
| <b>4. Wetland</b>                      | 14. Mangrove<br>15. Wetland   | 0         | 0      | 0       | 74     | 0       | 0        | 0       | 0        | 74    | 100                 |
| <b>5. Bare area/ Sparse vegetation</b> | 16. Bare Area,consolidated (gravel, rock)<br>17. Bare Area, unconsolidated (sand)<br>10. Sparse vegetation  | 1         | 5      | 1       | 0      | 122     | 0        | 0       | 0        | 129   | 94.6                |
| <b>6. Urban</b>                        | 18. urban   | 0         | 0      | 0       | 0      | 0       | 53       | 0       | 0        | 53    | 100                 |
| <b>7. Snow/ice</b>                     | 19. Snow/Ice  | 0         | 0      | 0       | 0      | 1       | 0        | 53      | 0        | 54    | 98.1                |
| <b>8. Water</b>                        | 20. Water   | 0         | 0      | 0       | 0      | 0       | 0        | 0       | 55       | 55    | 100                 |
| <b>Total</b>                           |   | 269       | 128    | 128     | 86     | 129     | 54       | 55      | 55       | 904   |                     |
| <b>Producer's accuracy (%)</b>         |   | 95.5      | 83.6   | 82.0    | 86.0   | 94.6    | 98.1     | 96.4    | 100      |       |                     |

Overall accuracy = 91.4%.

Average of user's accuracy = 94.0%.

Average of producer's accuracy = 92.0%.

Table 5. Overall accuracy and area-weighted overall accuracy of GLCNMO 2008

|  | GLCNMO 2008 (20 classes)<br>Table 3 | GLCNMO 2008<br>(8 aggregated classes)<br>Table4 |
|--|-------------------------------------|---|
| Overall accuracy                                   | 77.9%                               | 91.4%   |
| Overall accuracy<br>(Weighted by classified areas) | 82.6%                               | 92.5%   |

## 7. Conclusions and Recommendations

Fifteen arc second global land cover, GLCNMO2008 (or GLCNMO version 2) has been produced with the overall accuracy of 77.9 % by 904 validation points sampled by stratified random method. The product is available from ISCGM website [w4]. MODIS 500 m data of 2008 used for classification, globally distributed 2080 training polygons with shape format, and reference maps are also published from the data sharing system developed by the author, CEReS Gaia [w9]. All published data produced by this study are listed in Appendix C.

Through the experience of this mapping project, the following recommendations were obtained for the future global land cover mapping.

### (1) Accumulation of training data

Different land cover projects usually prepare their training data independently. In case of global land cover

mapping, training data by different mapping projects must be shared and accumulated for more accurate mapping in the future. The first step of sharing training data is to publish training data when land cover product is published. For this purpose authors publish training data of this study from CEReS website [w8] and CEReS Gaia [w9].

- (2) In addition to the sharing of training data, reference maps should be also shared. Sharing and accumulation of reference maps will increase reliability of training data and validation data. For this purpose, authors publish reference map images from CEReS Gaia [w9].
- (3) Land cover scientists have produced several global land cover products since the first production of GLCC. Comparison or integration of them is important task for future better mapping. The data sharing system, CEReS Gaia, provides opportunities to compare/overlay these existing global land cover images.

### Acknowledgments

This global land cover mapping was conducted as an activity of the Global Mapping Project coordinated by the International Steering Committee for Global Mapping (ISCGM). The authors express their sincere gratitude to the collaborating National Geospatial Information Authorities (or National Mapping Organizations) (Appendix B) for their contribution of intermediate validation, and thank the Secretariat of ISCGM for their assistance to cooperate with NMOs. Authors also appreciate the U.S. Geological Survey for providing MODIS data.

This work was supported by a Japan Society for the Promotion of Science (JSPS) research grant, KAKENHI (22220011).

### References

- Alsaaidh, B., Al-Hanbali, A., Tateishi, R., Kobayashi, T., & Hoan, N. T. (2013a). Mangrove forests mapping in the southern part of Japan using Landsat ETM+ with DEM. *Journal of Geographic Information System*, 5(4), 369-377. <http://dx.doi.org/10.4236/jgis.2013.54035>
- Alsaaidh, B., Al-Hanbali, A., Tateishi, R., & Kobayashi, T. (2013b). A method for integrating MODIS data and DEM for detecting mangrove forests in Kalimantan. *Proceedings of the International Symposium on Remote Sensing (ISRS)*, 111-114, Chiba, Japan, 15-17 May.
- Bartholome, E., & Belward, A. S. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. *Int. J. of Remote Sensing*, 26(9 & 10), 1959-1977. <http://dx.doi.org/10.1080/01431160412331291297>
- Carroll, M. L., Townshend, J. R., Dimiceli, C. M., Noojipady, P., & Sohlberg, R. A., 2009. A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, 2(4), 291-308. <http://dx.doi.org/10.1080/17538940902951401>
- Centre for Topographic Information, Earth Sciences Sector, Natural Resources. (2009). *Canada Land Cover, circa 2000-Vector: Data Product Specifications – Edition 1.0*. Retrieved from <http://www.geobase.ca/geobase/en/data/landcover/index.html>
- Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., ... Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83, 287-302. [http://dx.doi.org/10.1016/S0034-4257\(02\)00078-0](http://dx.doi.org/10.1016/S0034-4257(02)00078-0)
- Gao, B. C. (1996). NDWI - a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257-266. [http://dx.doi.org/10.1016/S0034-4257\(96\)00067-3](http://dx.doi.org/10.1016/S0034-4257(96)00067-3)
- GegenTana, & Tateishi, R. (2013). Mapping wetlands in Africa using MODIS tasseled cap indices. *Proceedings of the International Symposium on Remote Sensing (ISRS)*, 654-657, Chiba, Japan, 15-17 May.
- GegenTana, Letu, H., Cheng, Z. K., & Tateishi, R. (2013). Wetlands mapping in North America by decision rule classification using MODIS and ancillary data, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. <http://dx.doi.org/10.1109/JSTARS.2013.2249499>
- Giri, C., Zhu, Z., & Reed, B. (2005). A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets, *Remote Sensing of Environment*, 94, 123-132. <http://dx.doi.org/10.1016/j.rse.2004.09.005>
- Green, E. P., Clark, C. D., Mumby, P. J., Edwards, A. J., & Ellis, A. C. (1998). Remote sensing techniques for mangrove mapping. *International Journal of Remote Sensing*, 19(5), 935-956. <http://dx.doi.org/10.1080/014311698215801>

- Hansen, M. C., Defries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing*, 21(6 & 7), 1331-1364. <http://dx.doi.org/10.1080/014311600210209>
- Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., & Schullius, C. (2008). Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets, *Remote Sensing of Environment*, 112, 2538-2556. <http://dx.doi.org/10.1016/j.rse.2007.11.013>
- Hoan, N. T., Tateishi, R., Phong, D. X., & Johnson, B. (2012). Global Water Mapping using MODIS Tasseled Cap Indexes. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 7161-7164, Munich, Germany, 22-27 July.
- Hoan, N.T., Kobayashi, T., Gegentana, Alsaaidh, B., Zhang, N., Phong, D. X., & Tateishi, R. (2013). A new global land cover map 2008 - GLCNMO ver. 2, *Proceedings of the International Symposium on Remote Sensing (ISRS)*, 107-110, Chiba, Japan, 15-17 May.
- Kobayashi, T., & Tateishi, R. (2013). Comparison of a new percent tree cover dataset with existing one and categorical land cover datasets in Eurasia, *Advances in Remote Sensing* (pp. 345-357). <http://dx.doi.org/10.4236/ars.2013.24037>
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorol. Z.*, 15, 259-263. <http://dx.doi.org/10.1127/0941-2948/2006/0130>
- Lobser, S. E., & Cohen, W. B. (2007). MODIS tasseled cap: land cover characteristics expressed through transformed MODIS data. *International Journal of Remote Sensing*, 28(22), 5079-5101. <http://dx.doi.org/10.1080/01431160701253303>
- Loveland, T. R., & Belward, A. S. (1997). The IGBP-DIS global 1 km land cover data set, DISCover: first results. *International Journal of Remote Sensing*, 18, 3289-3295. <http://dx.doi.org/10.1080/014311697217099>
- Loveland, T. R., Zhu, Z., Ohlen, D. O., Brown, J. F., Reed, B. C., & Yang, L. (1999). An analysis of the IGBP Global Land Cover Characterization Process. *Photogrammetric Engineering and Remote Sensing*, 65, 1021-1032.
- Phong, D. X., Nguyen, H. T., Kobayashi, T., & Tateishi, R. (2013). A global 500-m urban map for GLCNMO version 2. *Proceedings of the International Symposium on Remote Sensing (ISRS)*, 666-669, Chiba, Japan, 15-17 May.
- Tateishi, R., Bayaer, U., Al-Bilbisi, H., Aboel Ghar, M., Tsend-Ayush, J., Kobayashi, T., ... Gegentana, Sato, H. P. (2011). Production of global land cover data – GLCNMO, *International Journal of Digital Earth*, 4(1), 22-49. <http://dx.doi.org/10.1080/17538941003777521>
- Tateishi, R., & Sri Sumantyo, J. T. (2012). Development of geospatial data sharing/ overlay system - CEReS Gaia -, 2012. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 558-561, Munich, 22-26 July.
- Tateishi, R., & Sri Sumantyo, J. T. (2013). Geospatial data sharing/overlay system: CEReS Gaia. *Proc. 34th ACRS*, SC09/421-428, Bali Indonesia, 20-24 October.
- Tiner, R. W. (1997). NWI maps: What they tell us. *Natl. Wetlands Newslett.*, 19(12), 7-12.
- Tong, P. H. S., Auda, Y., Populus, J., Aizpuru, M., Al Habshi, A., & Blasco, F. (2004). Assessment from space of mangroves evolution in the Mekong Delta, in relation to extensive shrimp farming. *International Journal of Remote Sensing*, 25(21), 4795-4812. <http://dx.doi.org/10.1080/01431160412331270858>
- Zhu, Z., & Waller, E. (2003). Global Forest cover mapping for the United Nations Food and Agriculture Organization Forest Resources Assessment 2000 Program. *Forest science*, 49(3), 369-380.

## Web references

w1 GLCC website

<http://edc2.usgs.gov/glcc/glcc.php>

w2 GLC2000 website

<http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php>

w3 GlobCover

<http://due.esrin.esa.int/globcover/>

w4 ISCGM website

<http://www.iscgm.org/>

w5 DMSP website

<http://ngdc.noaa.gov/eog/dmsp.html>

w6 UMDLC website

<http://glcf.umd.edu/data/landcover/data.shtml>

w7 GLCNMO website

<http://www.iscgm.org/gmd/download/glcnm.html>

w8 CEReS website

[http://www.cr.chiba-u.jp/~database-jp/wiki/wiki.cgi?page=GEOinfoDB\\_global](http://www.cr.chiba-u.jp/~database-jp/wiki/wiki.cgi?page=GEOinfoDB_global)  
[GG-10], [GG-11], [GG-12]

w9 CEReS Gaia website (with the keyword of “GLCNMO2008”)

<http://gaia.cr.chiba-u.jp/portal/>

w10 CORINE website

<http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster-2>

w11 NALCMS website

<http://www.cec.org/Page.asp?PageID=924&ContentID=25501>

w12 LC Canada website

<http://www.geobase.ca/geobase/en/data/landcover/index.html>

w13 NLCD website

<http://www.mrlc.gov/nlcd2006.php>

w14 LCCS website

[http://www.glcg.org/sof\\_1\\_en.jsp](http://www.glcg.org/sof_1_en.jsp)

(All websites above were accessed on June 2, 2014)

#### **Appendix A: Additional data used in this study**

The following data were used in this study in addition to the ones described in the section “2. Data used.”

**1. Ramsar Sites Database: (used for wetland mapping)**

The Ramsar Sites Database provides information of all wetlands of international importance. It is a searchable database, fully accessible through the internet with a password protected data entry system, and a reporting system for public use.

(<http://ramsar.wetlands.org/Database/AbouttheRamsarSitesDatabase/tabid/812/Default.aspx>).

**2. Köppen-Geiger climate classification map: (used for wetland mapping)**

It is a frequently used climate classification map of Wladimir Köppen, presented by Rudolf Geiger (Kottek et al., 2006) (<http://koeppen-geiger.vu-wien.ac.at/>).

**3. GTOPO30: (used for wetland mapping, mangrove mapping)**

Global DEM with a horizontal grid spacing of 30 arc seconds (approximately 1 km). The data is available from U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center.

**4. U.S. National Wetlands Inventory data: (used for wetland mapping)**

The National Wetlands Inventory database, produced by the U.S. Fish and Wildlife Service, provides digital wetland data for the United States (approximately 82% of the conterminous states) (Tiner, 1997). The vector

data is available from the Fish and Wildlife Service National Wetlands Inventory website (<http://www.fws.gov/wetlands/>).

**5. Canadian Wetland Inventory data: (used for wetland mapping)**

The Environment Canada–Canadian Wildlife Service (CWS) produced the Canadian Wetland Inventory to provide digital wetland data for parts of Canada via the website of Duck Unlimited Canada (<http://maps.ducks.ca/cwi/>).

**6. Land Cover, circa 2000-Vector (LCC2000-V) data: (used for wetland mapping)**

It is the vectorized land cover data originating from classified Landsat 5 and Landsat 7 ortho-images for Canada, were used (Centre for Topographic Information 2009). The LCC2000-V data were downloaded through the GeoBase website overseen by the Canadian Council on Geomatics (CCOG). (<http://www.geobase.ca/>).

**7. SRTM: (used for mangrove mapping in Japan and Cambodia and for global water mapping)**

1 arc second (30m) DEM and 30 arc second (1km) DEM

available from Global Land Cover Facility (GLCF) (<http://glcf.umd.edu>).

**8. Aerial photographs at a nominal scale of 1:20,000 obtained in 1987 (used for mangrove mapping in Japan and Cambodia)**

**9. Vegetation maps published by Natural Environment Biodiversity Center, Ministry of the Environment at scales 1:50,000 and 1:25,000 acquired in 2001–2006, respectively (used for mangrove mapping in Japan)**

**10. Cambodia Atlas published in 2007 (used for mangrove mapping in Cambodia)**

(<http://www.cambodiaatlas.com/map>).

**11. Ground survey data in Cambodia (used for mangrove mapping in Cambodia)**

**12. LandScan 2008™ population data: (used for urban mapping)**

Global population distribution at approximately 1 km resolution (30" X 30").

LandScan 2008™ High Resolution global Population Data Set copyrighted by UT-Battelle, LLC, operator of Oak Ridge National Laboratory under Contract No.DE-AC05-00OR 22725 with the United States Department of Energy. (<http://www.ornl.gov/landscan/>).

**13. Global Distribution and Density of Constructed Impervious Surfaces Area (ISA): (used for urban mapping)**

This data present the global inventory of the spatial distribution and density of constructed impervious surface area (ISA) on 1 km<sup>2</sup> grid. Examples of ISA include roads, parking lots, buildings, driveways, sidewalks and other manmade surface.

([http://ngdc.noaa.gov/eog/dmsp/download\\_global\\_isa.html](http://ngdc.noaa.gov/eog/dmsp/download_global_isa.html)).

**14. GDP per capita data of 2008: (used for urban mapping)**

Gross domestic product based on purchasing power parity per capita data of 2008 from The International Monetary Fund (IMF).

(<http://www.imf.org/external/pubs/ft/weo/2008/02/weodata/index.aspx>).

**15. MODIS water mask data (MOD44W): (used for water mapping)**

This water mask was produced by the University of Maryland using a combination of the SRTM Water Body Dataset (SWBD) and MODIS 250 m data from 2003 to 2007 (Carroll et al., 2009). The MOD44W data is not suitable for monitoring the frequent changes in global water coverage because it was developed using data observed over a long period (from 2000 to 2007).

**16. Global volcano data: (used for water mapping)**

This is Smithsonian Global Volcanism Program data <http://www.volcano.si.edu/products.cfm?p=9>

The KML format data of the location of volcanos (Filename: "GVPWorldVolcanoes.kml") can be downloaded from: [www.volcano.si.edu/ge/GVPWorldVolcanoes.kml](http://www.volcano.si.edu/ge/GVPWorldVolcanoes.kml).

## **Appendix B: Collaborating national geospatial information authorities (or national mapping organizations)**

The following 14 NGIAs have collaborated with the authors in the evaluation process of intermediate classification step.

Botswana: Surveys and Mapping, Ministry Lands and Housing,

Chile: Instituto Geografico Militar

Cote d'Ivoire: Centre de Cartographie et de Télédétection (CCT), Bureau National D'Etudes Techniques et de Développement (BNETD)

El Salvador: Instituto Geografico y del Catastro Nacional

Japan: Geospatial Information Authority of Japan

Macao, S.A.R., China: Direcção dos Serviços de Cartografia e Cadastro, Governo da Região Administrativa Especial de Macau

Macedonia: Agency of Real Estate Cadastre (AREC)

Mexico: Instituto Nacional de Estadística Geografía e Informática

Oman: National Survey Authority

Pakistan: Survey of Pakistan

Serbia: Republic Geodetic Authority

Sri Lanka: Survey Department of Sri Lanka

Thailand: Royal Thai Survey Department

United States: U.S. Geological Survey

## **Appendix C: Published data products by this study**

1. GLCNMO2008, land cover of 30 degree by 30 degree areas and global area:  
from Web references, w4
2. GLCNMO2008, land cover of global and continental areas:  
from Web references, w8 [GG-11]
3. MODIS2008 data used for the production of GLCNMO2008:  
from Web references, w8 [GG-10]
4. Land cover training data used for the production of GLCNMO2008:  
from Web references, w8 [GG-12]

The above data '2' and '4' are also available from CEReS Gaia [w9] with the keyword of "GLCNMO2008", and the above data '3' with the keyword of "MODIS2008".

## **Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).