

# Trend of global land cover mapping and global land cover ground truth database

Ryutaro TATEISHI(\*), Hiroshi P. SATO(\*\*) and ZHU Lin(\*)

(\*)Center for Environmental Remote Sensing (CEReS), Chiba University  
1-33 Yayoi-cho Inage-ku Chiba 263-8522 Japan  
E-mail: [tateishi@ceres.cr.chiba-u.ac.jp](mailto:tateishi@ceres.cr.chiba-u.ac.jp)

(\*\*)Geography and Crustal Dynamics Research Centre, Geographical Survey Institute,  
Tsukuba 305-0811, Japan  
E-mail: [hsato@gsi.go.jp](mailto:hsato@gsi.go.jp)

## Abstract

There are many global/continental or large area land cover mapping projects because land cover is one of key parameters in environmental studies. Though ground truth collection is an important and difficult task in land cover mapping, it is usually performed independently in each project without any cooperation between them. This is the background of the development of Global Land Cover Ground Truth (GLCGT) database by the cooperation of many projects and researchers. The developed GLCGT database will be used freely by any researcher. This cooperative and common development of GLCGT database will realize reliable and continuously improved land cover ground truth data. It also eliminates duplicated efforts of ground truth collection among projects.

**Key words:** land cover, ground truth

## 1. Trend of global land cover mapping

The necessity of large-area land cover mapping and history of large-area land cover mapping including the global one before 2000 are well described by Loveland, et al. (2000). The most representative global land cover mapping was attempted using AVHRR data of 1992 under the auspices of the IGBP.

The present on-going global land cover mapping projects are the following three.

- MODIS land cover project by NASA using TERRA/MODIS data
- Global Land Cover 2000(GLC2000) project by Joint Research Center (JRC) using SPOT-4/VEGETATION data
- GLI land cover project by NASDA using ADEOS-II/GLI data
- Global Mapping project by Geographical Survey Institute, Japan

The recent trend of global land cover mapping is to establish experts network. Food and Agriculture Organization (FAO) of the United Nations and United Nations Environmental Programme(UNEP) endorsed Artimino Declaration on Global Land Cover Network on 8 May 2002. The declaration calls for building a coalition among governmental and inter-governmental agencies, academic institutions, the private sector and other interested stakeholders in the process of institutional networking, strengthening national capacity, harmonizing land cover classification and mapping, and developing global land cover databases. Further information about this should be directed to the Chief, SDRN, FAO <[GLCN@fao.org](mailto:GLCN@fao.org)>.

The other networking initiative was started by Joint Research Center as the expansion of the partnership of GLC2000 project. "The objective of this proposed Network of Excellence is to build on the established GLC2000 partnership to sustain and strengthen a European Scientific and Technical role in the field of monitoring Global Land Cover Dynamics. A multidisciplinary, international network will be set up to share experience concerning new approaches to describe land cover and land cover change, assess the accuracy of such observations and to generate products targeted at the long term requirements of users as described in the rationale. Arising from

community consensus the network will ultimately provide internationally accepted reference databases documenting global land cover dynamics.”

## 2. Global land cover ground truth (GLCGT) database

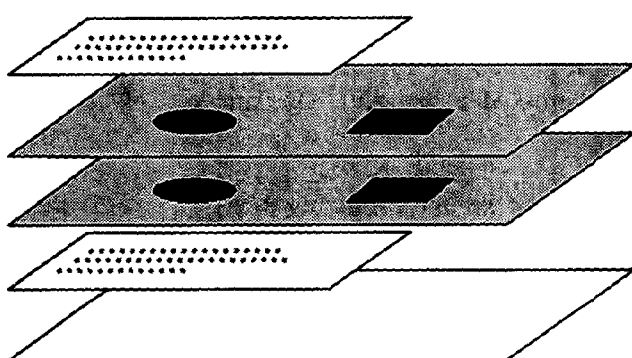
In any global land cover mapping projects, ground truth data are necessary as training sample data of classification and validation of the classified result. The collection of ground truth data is one of the most difficult and time consuming work. The stress of this paper is put on the proposal of the development of ground truth database for global land cover mapping by widely cooperative input and for open use.

The proposed global land cover ground truth (GLCGT) database consists of regional land cover ground truth (RLCGT) data. The geographical size of a RLCGT data is flexible, from a city size to a continental size.

### 2.1 Regional land cover ground truth (RLCGT) data

The regional land cover ground truth (RLCGT) data consists of

- metadata of RLCGT data (text) dataset name: “meta\*\*\*\*”
- ground truth (GT) land cover code data (raster data) dataset name: “lrgt\*\*\*\*”
- ground truth (GT) site code data (raster data) dataset name: “site\*\*\*\*”
- description of GT sites (text) dataset name: “source\*\*\*\*”
- optional data (raster data or text)



Metadata of RLCGT data
Text
Ground truth (GT) land cover code data
Raster
Ground truth (GT) site code data
Raster
Description of GT sites
Text
Optional data
Text, raster or any

Figure 1 Architecture of global land cover ground truth (GLCGT) database

### 2.2 Metadata of RLCGT data

The metadata of RLCGT data includes the following items.

- 1 RLCGT code (this code is assigned by the coordinator of the GLCGT database project)
- 2 the organization or project which produced a specific RLCGT data
- 3 representative person which produced this RLCGT data (name, affiliation)
- 4 contact person for this RLCGT data (name, affiliation, postal address or fax or email)
- 5 date of the production of this RLCGT data
- 6 latitude of the north edge of this region  
latitude of the south edge of this region  
longitude of the west edge of the region  
longitude of the east edge of the region
- 7 pixel size (by arc second)
- 8 number of pixels (east-west)

- number of pixels (north-south)
- 9 definition of land cover code
- 10 definition of the additional land cover code if any (recommended codes are 201-254)
- 11 description of optional data if any
- 12 any other description about this RLCGT data

### 2.3 Ground truth (GT) land cover code data

The GT land cover code data is a main part of ground truth data. The GT land cover code data is a raster data which cover a rectangular geographical region. The geographical region is defined in the metadata of RLCGT data. The north **edge** (not a center) of the north end pixel is recommended to be a multiple of 30 second (example: 60 degree north latitude, 60 degree 35 minute north latitude, or 61 degree 20 minute 30 second north latitude). This is the same in the case of south, west, and east edges. This makes overlay of different RLCGT data with different pixel size easier.

The geodetic coordinate system must be based on ITRF94 + GRS80 which is practically the same as **WGS84** for this application. In other words, the used geodetic coordinate system is the same as the one for Global Positioning System (GPS).

The pixel size (or interval) is **30 arc second or less**. It is arbitrarily determined but is recommended to be a factor of 30 such as 15, 10, 6, 5, 3, 2, 1 arc second. It makes the overlay of different RLCGT data with different pixel size easier.

Land cover class must be based on the Land Cover Classification System (**LCSS**) by Food and Agriculture Organization (FAO) of the United Nations (UN). The GT land cover code (1-254) must be clearly defined and described in the metadata of RLCGT data. Though contributors to input GT data can define any GT land cover code, the use of land cover code of the **ST** (Sato-Tateishi) **land cover guideline legend** is recommended. It is described in the Appendix at the end of this paper. The **additional** GT land cover code (201-254), if any, can be defined by contributors for each RLCGT data under the condition that the additional code must be clearly defined in the metadata of RLCGT data.

The background value of GT land cover code data must be 0. A pixel value is recorded in **one byte**. The ground truth (GT) is recorded as a pixel value of land cover code. A geographical unit of one GT can be any number of pixels and have any shape, but a unit of GT more than 2 x 2 pixels is recommended. This is because one pixel GT may not be reliable due to the misregistration of the reference image used in the work of GT collection. A unit of GT is called a "**GT site**" here.

### 2.4 Ground truth (GT) site code data

The GT site code data has exactly the same size and format data as the GT land cover code data. A GT site code data cover the same rectangular geographical region as the corresponding GT land cover data. A unit of GT (or a GT site) has a unique site code (1-254) in the RLCGT data. This site code must be recorded in the same geographic region as a unit of GT. The background value of GT site code data must be 0. A pixel value is recorded in one byte. When a number of GT site exceeds 254, another RLCGT data must be defined.

### 2.5 Description of GT sites

The description of GT sites is a text description of each unit of GT. The recorded information is as follows.

- site code
- approximate latitude and longitude of the GT site
- land cover code of the GT site
- any description about the land cover of this GT site
- **information source** by which a land cover type of the GT site was recognized

The information source is the most important part of this description because, by this information, the reliability of the ground truth is estimated.

(example)

- site code: 3
- lat/long: N42degree 56minute, E130degree 50minute
- land cover code: 12
- land cover: broadleaf deciduous forest
- information source: field survey, Jul-Aug 2001 (by Tateishi)  
FS -Point No. 5  
+ Land use map of China 1:1,000,000 published by the Science Press, 1990
- or  
Landsat TM (24 Feb 2000 and 16 Jan 2000)  
+ knowledge (Ekaterina Rachkovskaya)
- or  
vegetation map of Kazakhstan and Middle Asia, 1:2,500,000, 1995  
+ field survey in early 1990 (by Ekaterina Rachkovskaya)
- or  
land use map 1:200,000, Geographical Survey Institute, Japan 1982  
+ knowledge (Tateishi)

### 3. Conclusion

This paper describes the concept of global land cover ground truth(GLCGT) database which is proposed to Global Land Cover 2000(GLC2000) project, GLI land cover project, and Global Mapping project. The recommended land cover legend for GLCGT database is Sato-Tateishi (ST) land cover guideline legend which is based on the Land Cover Classification System (LCCS) by Food and Agriculture Organization (FAO) of the United Nations (UN).

### References

Di Gregorio,A. and Jansen,L.J.M., 2000, Land cover classification system(LCCS) Classification concepts and user manual version 1.0, FAO, 179p.

Loveland,T.R., et al., 2000, Large-area land cover characterization, Chapter 5 of "Global Environmental Databases", Eds. Tateishi and Hastings, 233p., Geocarto International Centre

Sato,P.H. and R. Tateishi, 2001, Global land use, land cover and vegetation classification systems: a review, Report of Geographical Survey Institute, Japan, No.96, pp.69-99

**Appendix Sato-Tateishi (ST) land cover guideline legend and its LCC label by FAO**

ST-LCG	LCC label
1. Broadleaf Evergreen Forest 101 ~ 111 [Crop list]	20089 Tree crops in [Crop list]
2. Broadleaf Deciduous Forest 112 ~ 117 [Crop list]	20090 Tree crops in [Crop list]
3. Needleleaf Evergreen Forest	20092 / 20134-1
4. Needleleaf Deciduous Forest	20093 / 20135-1
5. Mixed Forest	Mixture of 20089, 20090, 20092 / 20134-1, 20093 / 20135-1, Tree crop in [Crop list]
6. Tree Open	20013
31. Broadleaf Evergreen Woodland	20131
32. Broadleaf Deciduous Woodland	20132
33. Needleleaf Evergreen Woodland	20134-3012
34. Needleleaf Deciduous Woodland	20135-3012
7. Shrub 118, 119 [Crop list]	20017 / 20021 Shrub crops in [Crop list]
8. Herbaceous, single layer	20409 / 20454
9. Herbaceous with Sparse Tree/Shrub	20410 / 20455 / 20413 / 20458
10. Sparse Herbaceous / Shrub	20058 / 20055
11. Cropland 120 Rice, paddy 121 ~ 166 [Crop list]	0003 / 0006 3001-S0308 Herbaceous crops in [Crop list]
12. Cropland / Natural Vegetation Mosaic	Mixture of the item #11 and items of the natural vegetation in the left column.
13. Tree-Water (Brackish to Saline)	40003-R2 / 40003-R3 / 40007-R2 / 40007-R3
14. Wetland	40003-R1 / 40007-R1 / 40009 / 40011 / 40013 / 40019 / 40973 / 40985
15. Lichens / Mosses	21435 / 21438
16. Bare Areas	6001 / 6004
35. Consolidated Bare Areas	6001
36. Bare Rock	6002-1
37. Gravels, Stones and Boulders	6002-2
38. Hardpan	6003
39. Unconsolidated Bare Areas	6004
40. Bare Soil / Other Unconsolidated Materials	6005
41. Loose and Shifting Sands	6006
17. Urban	0010
18. Snow / Ice	8005 / 8008
19. Water	0013 / 8001

Note: LCC label is based on FAO's LCCS.

[Crop list]

ST-LCG code	Crop name	LCC label	FAOSTAT
112	<i>Almonds</i>	10001-1891-S0601	0221
113	<i>Apples</i>	10001-1891-S0602	0515
121	<i>Bananas</i>	10025-S0604	0486
122	<i>Barley</i>	10025-S0302	0044
123	<i>Beans</i>	10025-S0502 / 10025-S0503	0176+0414
124	<i>Broad Beans</i>	10025-S0599Zs001*	0181+0420
125	<i>Buckwheat</i>	10025-S0304	0089
126	<i>Cabbages</i>	10025-S0504	0358
127	<i>Cantaloupes &amp; other Melons</i>	10025-S0510	0568
101	<i>Cashew Nuts</i>	10001-1-S0605	0591
128	<i>Cassava</i>	10025-S0401	0125
129	<i>Castor Beans</i>	10025-S0902	0265
130	<i>Chick Peas</i>	10025-S0506	0191
102	<i>Chillies &amp; Peppers, Green</i>	10001-1-S13Zs002*	0401
103	<i>Cocoa Beans</i>	10001-1-S0801	0661
104	<i>Coconuts</i>	10001-1-S0607	0249
105	<i>Coffee, Green</i>	10001-1-S0802	0656
131	<i>Cow Peas, Dry</i>	10025-S13Zs003*	0195
132	<i>Cucumbers &amp; Gherkins</i>	10025-S0607	0397
133	<i>Eggplants</i>	10025-S13Zs004*	0399
114	<i>Grapes</i>	10001-1891-S0610	0560
134	<i>Groundnuts in Shell</i>	10025-S0611	0242
118	<i>Jute</i>	10013-S0908	0780
135	<i>Lentils</i>	10025-S0508	0201
136	<i>Linseed</i>	10025-S0999Zs005*	0333
137	<i>Lupins</i>	10025-S13Zs006*	0210
138	<i>Maize</i>	10025-S0305	0056
106	<i>Mangoes</i>	10001-1-S0615	0571
139	<i>Millet</i>	10025-S0306	0079
107	<i>Natural Rubber</i>	10001-1-S0999Zs007*	0836
140	<i>Oats</i>	10025-S0307	0075
108	<i>Oil Palm Fruit</i>	10001-1-S0909	0254
109	<i>Olives</i>	10001-1-S0910	0260
141	<i>Onions</i>	10025-S0511	0402+0403
110	<i>Oranges</i>	10001-1-S0606	0490
115	<i>Peaches &amp; Nectarines</i>	10001-1891-S0617	0534
116	<i>Pears</i>	10001-1891-S0618	0521
142	<i>Peas</i>	10025-S0512	0187+0417
143	<i>Pigeon Peas</i>	10025-S0599Zs008*	0197
111	<i>Pimento, Allspice</i>	10001-1-S13Zs009*	0689
144	<i>Plantains</i>	10025-S13Zs010*	0489
117	<i>Plums</i>	10001-1891-S0621	0536
145	<i>Potatoes</i>	10025-S0402	0116
146	<i>Pumpkins, Squash, Gourds</i>	10025-S0513	0394
147	<i>Rapeseed</i>	10025-S0999Zs011*	0270
120	<i>Rice, paddy (Oryza, spp.)</i>	3001-S0308	0027
148	<i>Rye</i>	10025-S0309	0071
149	<i>Safflower Seed</i>	10025-S0911	0280
150	<i>Seed Cotton</i>	10025-S0903	0328
151	<i>Sesame Seed</i>	10025-S0912	0289
152	<i>Sorghum</i>	10025-S0310	0083
153	<i>Soybeans</i>	10025-S0914	0236
154	<i>Sugar Beets</i>	10025-S13Zs012*	0157
155	<i>Sugar Cane</i>	10025-S0999Zs013*	0156
156	<i>Sunflower Seed</i>	10025-S0999Zs014*	0267
157	<i>Sweet Potatoes</i>	10025-S0403	0122
158	<i>Tang.Mand.Clement.Satsma</i>	10025-S0499Zs015*	0495
159	<i>Taro (Coco Yam)</i>	10025-S0499Zs016*	0136
119	<i>Tea</i>	10013-S0804	0667
160	<i>Tobacco Leaves</i>	10025-S0805	0826
161	<i>Tomatoes</i>	10025-S0514	0388
162	<i>Triticale</i>	10025-S13Zs017*	0097
163	<i>Vetches</i>	10025-S13Zs018*	0205
164	<i>Watermelons</i>	10025-S13Zs019*	0567
165	<i>Wheat</i>	10025-S311	0015
166	<i>Yams</i>	10025-S0405	0137

Note: LCC label is based on FAO's LCCS. LCC label with asterisk is defined in this paper by using LCCS.