*Journal of Environmental Sciences* Vol.15, No.4, pp.520—524, 2003 ISSN 1001—0742 CN11-2629/X

**Article ID**: 1001-0742(2003)04-0520-05 **CLC number**: X171.1 **Document code**: A

# Certainties and uncertainties of land cover statistics in China

ZHAO Jing-zhu[1], SUN Yi[1], BAI Guang-xin[1], WU Gang[1], SHAO Guo-fan[2]*

(1. Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China; 2. Department of Forestry and Natural Resources, 1159 Forestry Building, Purdue University, West Lafayette, IN 47907, USA. E-mail: gshao@fnr.purdue.edu)

**Abstract**: Vegetation or land cover maps have been made directly or indirectly available for the entire territory of China. The certainties and uncertainties of land cover statistics were analyzed by comparing three data sources: China's Vegetation Map, IGPB DISCover, and University of Maryland Product. Great similarities in the statistics of 7 aggregated land cover types were found among the three data sources, particularly between the two global land cover maps. The per-pixel agreement between any 2 of 3 maps was between 38.0%—51.4%; the per-pixel agreement among all three maps was only 27.1%. Certainties were found in regions where vegetation types are typical and human land use practice is relatively homogenous; the uncertainties occurred to either vegetation transition zones or regions where land cover types and land use practice are relatively diversified. Systematic and multidisciplinary efforts are necessary to promote accurate mapping of nationwide land cover types in China. Intensive efforts should be made in regions where uncertainties of land cover information are found.

**Keywords**: uncertainties of land cover; remote sensing; spatial pattern; landscape characterization

## Introduction

Vegetation or land use and land cover maps are important data sources for providing regional land cover information useful for many purposes, such as planning and evaluating long-term land cover change, analyzing and modeling interactions between land cover change and environments, and designing and detecting changes in human land use practice. China is a huge country. Mapping land cover types for the entire country is a time- and resources-consuming effort. Various nationwide maps, including vegetation map and forestry division map, have been produced during the past decades. Tremendous efforts have been made in mapping land cover types with remote sensing data. To demonstrate effectiveness and potential problems associated with remotely sensed mapping technology, it is useful to examine the existing land cover maps derived from remote sensing data.

Mapping land cover types at global or regional scales are becoming a relatively easy task as remote sensing technology advances. There are currently two global 1 km resolution land cover data available, both derived from remotely sensed data of the Advanced Very High Resolution Radiometer(AVHRR) acquired between 1992 and 1993. The first was created by the U.S. Geological Survey for the International Geosphere-Biosphere Programme(IGBP) and referred as IGBP DISCover; the second was created by the Laboratory for Global Remote Sensing Studies at the University of Maryland(UMD) and referred as UMD product. Because of the differences in input data and classification processes, global area totals of aggregated vegetation types have a per-pixel agreement of 74%(Hansen, 2000).

A spatial resolution of 1 km is by far the highest for global land cover maps. Such a fine resolution is good enough for many applications at regional scales. Land cover statistics can be easily extracted from the global land cover data. The questions are: how are the statistics derived from the global land cover data agreeable with each other? How are these statistics agreeable with other existing data sources? One of the existing data is the digital vegetation map digitized in 1996 by the State Key Laboratory of Resources and Environmental Information System(LREIS), Chinese Academy of Sciences. The vegetation map contains enough vegetation types that can be re-grouped to match the aggregated classes from the global land cover data.

It is important to quantify the accuracy or certainty of any land cover data for planning, evaluation,

Foundation item: The National Natural Science Foundation of China(Nos. 79930800; 40173033 and 39970591); * Corresponding author and designing purposes with the data. Due to the difficulties in assessing the accuracy of the nationwide land cover data, an alternative approach is to compare land cover data from various sources. Such comparisons will help to provide information at least on the certainties and uncertainties of land cover types. This paper intends to quantify the similarities and differences in land cover statistics among three sources of land cover data for the mainland of China as well as Hainan and Taiwan islands.

# 1 Method

The 1 km resolution global land cover data were downloaded from http://edcdaac.usgs.gov/glcc/glcc.html for the IGBP Product and from http://www.geog.umd.edu/landcover/global-cover.html for the UMD Product. The China's Vegetation Map was provided by the State Key Laboratory of Resources and Environmental Information System(LREIS), Chinese Academy of Sciences. Detailed descriptions on the processes of developing these two global data products are provided by Loveland *et al*. (Loveland, 1999) and Hansen and Reed (Hansen, 2000). Both land cover maps were produced from the AVHRR data acquired between April 1992 and March 1993. The IGBP product was made by using unsupervised classification method with 12 monthly NDVI (Normalized Difference Vegetation Index) composite whereas the UMD product was made by using supervised classification tree method with 41 metrics derived from NDVI and bands 1—5.

All the data products were imported into ArcInfo GRID format and re-projected into Albers Equal-Area Conic system with parameters of 25°N for the first parallel, 47°N for the second parallel, and 110°E for the central meridian. The global land cover data were clipped with the provincial boundary data published by CPHC(CPHC, 1988). All the land cover or vegetation types were re-grouped into 7 classes(Table 1).

**Table 1 Re-grouping land cover or vegetation types into 7 aggregated categories**

| Aggregated class | IGBP product | UMD product | China's vegetation map |
|---|---|---|---|
| Forest | Evergreen needleleaf forest | Evergreen needleleaf forest | Needleleaf forest |
| | Evergreen broadleaf forest | Evergreen broadleaf forest | Broadleaf forest and woodland |
| | Deciduous needleleaf forest | Deciduous needleleaf forest | |
| | Deciduous broadleaf forest | Deciduous broadleaf forest | |
| | Mixed forest | Mixed forest | |
| | Woody savannas | Woodland | |
| | Savannas | | |
| Shrubland | Closed shrublands | Closed shrublands | Scrub and coppicewood |
| | Open shrublands | Open shrublands | |
| Grassland | Grasslands | Wooded grassland, grassland | Steppe and savanna, meadow and swamp |
| Cropland | Croplands, cropland/natural vegetation mosaic | Cropland | One crop annually, two crops annually or three crops in two years |
| Bare ground | Snow and ice, barren or sparsely vegetated | Bare ground | Desert, land without vegetation |
| Urban & build-up | Urban and built-up | Urban and built-up | None |
| Water | Water bodies, permanent wetlands | Water | Lake |

Overlay operations were performed to the three data products to quantify the similarities and differences in the spatial locations of various lands cover types between each pair of them and among all three of them. The agreement levels were measured, based on which, all the provinces and autonomous regions were grouped into sub-classes.

# 2 Results

There are great similarities among the three data products in characterizing the coverage of three land cover types, including forest(15.5%—16.1%), shrubland(19.7%—22.3%), cropland(16.3%—18.7%) but clear differences in the coverage of grassland(23.7%—29.2%) and bare ground(14.7%—

20.5%) (Fig. 1). The differences between the two global data products are < 15% whereas the differences between other two pairs can exceed 30% for grassland and bare ground classes. In other words, there are greater differences between the global data products and China's Vegetation Map than those between the two global data products. Water covers 1.2%—1.4% for the global land cover products and 0.7% for China's Vegetation Map.

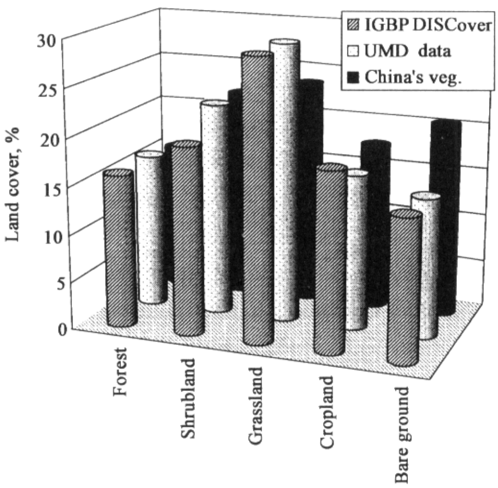


Fig.1 A comparison of major land cover statistics in China

The overlay operations showed that the two global land cover data have a per-pixel agreement of 51.4%, the IGBPDIScover and China's Vegetation Map 46.9%, and the UMD Product and China's Vegetation Map 38.0%. By overlaying all three maps, the per-pixel agreement of the three maps is 27.1% and the per-pixel disagreement of the three maps is 17.7%.

Assigning value 1 for the agreement of all three data sources, 3 for the disagreement among three data sources, and 2 for the rest, all the provinces were divided into three "equal-area" regions by the means of the assigned values within each province (Fig.2). The best agreement occurs to regions where typical land cover types are located and human land use practice is relatively homogenous. These regions include desert or grassland dominated provinces (e.g., Inner Mongolia and Xinjiang) and agriculture provinces (e.g., Henan, Shandong and Jiangsu). The worst agreement occurs to either vegetation transition zones or regions (e.g., Gansu, Ningxia and Shaanxi) where land cover types and human land use practice are relatively diversified (e.g., southwest and southeast provinces).

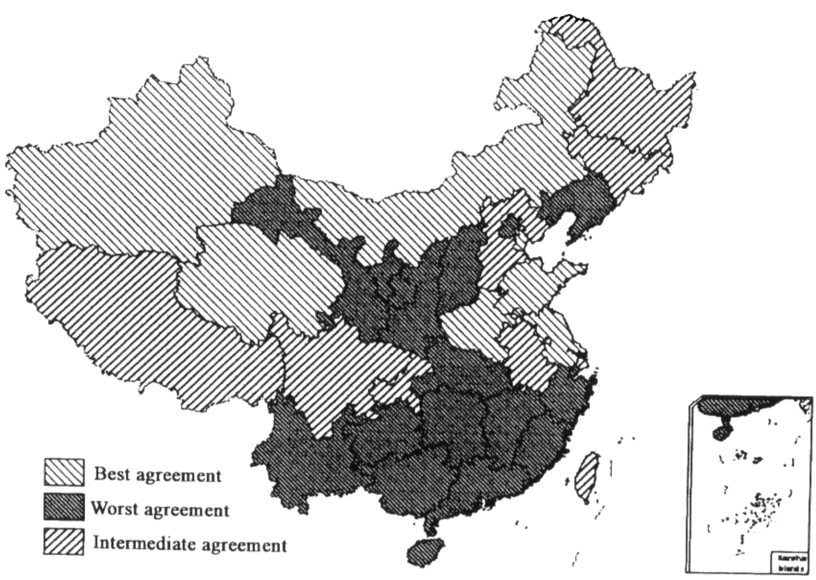


Fig.2 Divisions of the certainties and uncertainties of land cover statistics in mainland of China, Hainan and Taiwan Islands

## 3 Discussion

The overall accuracy of the IGBP DISCover data with original 17 classes was 73.5% at the global level(Scepan, 1999). By aggregating the 17 into 7 classes, the overall accuracy is supposed to be higher than the original overall accuracy. The UMD product was not assessed and its exact accuracy is unknown. Because the per-pixel agreement is only 51.4%, the overall accuracy of the UMD product cannot be higher than 75% for China's territory. Despite of the relatively low accuracy of both global 1 km land cover data, the coverage areas for various land cover types are very close to each other and close to some national statistics results. For example, the forest coverage in China was reported to be 16.7% when using 10%-canopy-coverage standard. The total forest areas from the two global land cover data are close to this reported value.

The accuracy of coverage areas is not the same as the accuracy of classification. Although similar coverage areas for several land cover types are derived both global land cover data, the map has nearly 50% per-pixel disagreement. This means that the similarities in statistic values derived from the global land cover map do not support the similarities in the locations of land cover types. Therefore, the map may lead to different outcomes for analyzing, planning or designing land use practice in China. All these differences are caused mainly by the differences in processing remote sensing data rather than in the data themselves. Different processing method results in different errors, including omission and commission errors(Congalton, 1999). When these two types of errors compensate, the information on areas of various land cover types is still accurate(Czaplewski, 1992). However, if both omission and commission errors are high, land cover types will have unacceptable location errors. The two global lands cover products represent the state-of-the-art of global scale remote sensing applications in the end of 20th century (DeFries, 2000). Even advanced remote sensing technology can be easily misused and result in misleading interpretations.

It is not theoretically sound practice to directly compare China's Vegetation Map with remote sensing land cover maps, such as the global land cover data. This is because that any vegetation map is made based on information on vertical vegetation structure whereas satellite sensors only detect radiations from the top layers vegetation. The differences between China's Vegetation Map and global landscape data reveals the differences in the way the data were prepared, definitions of vegetation or land cover classes, and temporal changes of land. When aggregated classes are used, the two types of data can be quantitatively compared. An agreement of 38%—49% indicates a lot of similarities between the two types of data sources.

The agreement of the three vegetation/land cover maps reveals that the relatively reliable information on land cover types is available for only 27.1% of China's total land area. There are at least 17.7% of lands where vegetation/land cover types are the most uncertain. Research priorities should be given to clarify vegetation/land cover types in the confused regions, including the Loess Plateau, southwest, and southeast China. Higher resolution and multi-scale data sources should be used for mapping reasonably accurate vegetation/land cover types in China.

# 4 Conclusions

Great similarities in the statistics of 7 aggregated land cover types were found among the three data sources, particularly the global land cover map. The per-pixel agreement between any 2 of 3 maps was between 38.0%—51.4%; the per-pixel agreement among all maps was only 27.1%. Certainties were found in regions where vegetation types are typical or human land use practice is relatively homogenous; the uncertainties occurred to vegetation transition zones or regions where land cover types and human land use practice are relatively diversified.

Information about the areas of various land cover types and their distributions in China is available from various resources but only a small portion of information is spatially accurate. The careless uses of the existing information may lead to misleading interpretations. The accuracy in estimating the areas of land cover types is not the same as that in quantifying the patterns of land cover types. In other words, the map with similar area statistics of various land cover types may look differently due to the differences in location errors. The use of map with different location errors would propagate differences in the downstream interpretations. Remote sensing technology is a useful tool as long as it is used properly. We can not simply rely on efforts from outside China to obtain useful land cover maps in China; we can not simply rely on efforts from individual institutions to accomplish the comprehensive mapping tasks for the entire country; we can not simply rely on only new remote sensing technology and ignore the existing resources from various agencies. Systematic and multidisciplinary efforts are necessary to promote accurate mapping of nationwide land cover types in China. Intensive efforts should be made in regions where uncertainties of land cover information are found.

## References:

Cartographic Publishing House of China(CPHC), 1988. Administrative division map of China at 1:4000000[M]. Beijing: Cartographic Publishing House of China.

Congalton R G, Green K, 1999. Assessing the accuracy of remotely sensed data: principles and practices[M]. New York: Lewis Publishers.

Czaplewski R L, 1992. Missclassification bias in a real estimates[J]. Photogrammetric Engineering & Remote Sensing, 58(2): 189—192.

DeFries R S, Belward A S, 2000. Global and regional land cover characterization from satellite data: an introduction to the special issue[J]. International Journal of Remote Sensing, 21(6 &7): 1083—1092.

Hansen M C, Reed B, 2000. A comparison of the IGBP DISCover and University of Maryland 1 km global land cover products[J]. J Remote Sensing, 21(6&7): 1365—1373.

Loveland T R, Zhu Z, Ohlen D O *et al*., 1999. An analysis of the IGBP global land-cover characterization process[J]. Photogrammetric Engineering & Remote Sensing, 65(9): 1021—1032.

Scepan J, 1999. Thematic validation of high-resolution global land-cover data sets[J]. Photogrammetric Engineering & Remote Sensing, 65(9): 1051—1060.

(Received for review April 11, 2002. Accepted June 21, 2002)