TRAINING SCHEME FOR LAND-COVER MAPSVALIDATION BY GROUND-TRUTH

Koki IWAO [a,], Kenlo NISHIDA[b] and Yoshiki YAMAGATA[a]

[a] National Institute for Environmental Studies, 16-2, Onogawa, Tsukuba, Ibaraki 305-8506, JAPAN –
iwao.koki@nies.go.jp
[b] Institute of Agricultural and Forest Engineering, The University of Tsukuba.
Tennnoudai 1-1-1, Tsukuba 305-8572, JAPAN -

KEY WORDS: Land Cover, Classification, Global-Environmental-Databases, Accuracy, Sampling

ABSTRACT:

Land cover maps validation data acquisition from ground-truth is proposed. Land cover maps are used in the numerical models which estimate ecosystem behaviour, water cycle, and climate in global scale. Therefore, accuracy validation of these land cover maps is important. Each of the existing land cover map has been validated with its own validation method. However the distribution is restricted by the differences in the fields made into the researchers' interests and the researchers' experiences may affect those accuracies. As a result, there is no validation method which evaluates these land cover maps with single validation dataset in global scale. This might makes less agreement if the existing land cover maps are compared. To overcome these problems, authors summarized the matter which is needed by training for land cover validation data development which makes possible for many people to participate. This gives fairly uniform ground-truth information with accurate location information worldwide.

1. INTRODUCTION

1.1 Importance of global land cover data and its accuracies

The land cover map is positioned as one of the important data with which improvement in accuracy is desired scientifically and politically most, when arguing about global environment problems. [Patennaude et al., 2005; DeFris et al., 2000] From sensitivity analysis of some terrestrial biosphere models, area total of Net Primary Production (NPP) or its spatial distribution which a model presumes may change by changing a land cover maps. So, improvements in global land cover maps are desired [Ahl et al., 2005; Kok et al., 2001].

Until now, many validation techniques about global land cover maps are advocated. For example, validation by the interpretation using an aerial photograph or a high spatial resolution satellite picture is proposed. However, subjects, such as spatial size of a sample, the number of validation points, and accuracy of the ground validation data itself, remain. [Strand, et al., 2002; Kelly et al., 1999; Rosenfield and Fitzpatrick-Lins, 1986; Hord and Brooner, 1976] The distribution is restricted by the differences in the fields made into the researchers' interests and the researchers' experiences may affect those accuracies. As a result, there is no validation method which evaluates these land cover maps with single validation dataset in global scale.

Inventory information (agricultural statistics, forestry statistics, etc.) are also used for accuracy appraisal method and it is widely adopted as accuracy validation of a land cover data. [Foody et al., 2002]. However, according to the research which carried out the relative comparison of two or more global land cover maps where accuracy validation was performed based on them, the gross area of each classification class is alike, but the spatial distribution differs greatly. [McCallum et al., 2006; Giri et al., 2005]. This shows that validation of global land cover data is not fully progressing.

1.2 Proposed validation method

To overcome these problems, authors proposed a validation method that can address this shortcoming [Iwao et al., 2006]. Our method employs information gathered by "the Degree Confluence Project (DCP)," a voluntary-based project that collects on site data from each of the degree confluence points (DCPoints) in the world [DCP, 1996]. DCPoints are located at the intersections of integer level latitude and longitude grid lines. Volunteers with the project visit the DCPoints and collect data in the form of GPS readings, pictures and descriptions of the landscape. Focusing on the IPCC LULUCF (Land Use, Land Use Change and Forestry) guidelines, we classified each DCPoints into six classes (Forest, Grassland, Crop, Wetland, Residential, and Other) manually [IPCC, 2003]. In this paper, we focus on the improvement of this DCPoints information.

2. METHODOLOGY

2.1 DCP

As DCPoints information, the present condition information and photograph of the point are exhibited. In order to evaluate whether special knowledge (knowledge, such as a types of vegetation and a technical term about cultivation) is needed on the interpretation, three persons' interpretation person was prepared. Three persons are general office worker, an ecologist and a remote sensing researcher. Among three persons, when two persons' result was the same, the method of making it into majority was taken. The visual situation of land cover may differ depending on the season (exploration time). Then, interpretation gave priority to present condition information (descriptions of the landscape). Photograph information was used in order to mainly check the accuracy of present condition information. However, when present condition information was inadequate, interpretation was tried from the photograph. On the other hand, the check of a spatial distribution was evaluated using photograph information mainly.

2.2 Visual classification with Landsat images

These DCP derived classifications were then compared to classifications derived from Landsat Thematic Mapper images by visual interpretation. High resolution ortho-rectified Landsat TM (Thematic Mapper) images which Earth Science Data Interface [University of Maryland, 2004] opens through the web site were used. The colour composite (Red: 4, Blue:3 and Green: 2) was used for visual interpretation. Interpretation was performed who has the experience which handled the remote sensing images for more than ten years. However, information related to that spot such as crop calendar was not used.

2.3 Ground Truth

We have conducted ground truth of three DCPoints already successfully completed. The reliability of the information (position information is included) indicated was checked. Then we consider the training scheme for the improvement of DCP information for validation of land cover maps.

3. RESULTS AND DISCUSSION

3.1 DCPoints derived information

We found that there are no discriminable differences in the validation results of three persons. Compared with the results from Landsat image interpretations and this accuracy were comparable as the visual interpretation by a Landsat Images.

3.2 Ground Truth

To assess the reliability of DCP information, ground truth (13N, 100E; 14N 100E; 13N 101E) was also actually performed and the reliability of the information (position information is included) indicated was also checked. Among these, in 13N, 101E points, three explorations have already been performed (May 2001, October 2004, and April 2006).Although the photograph of this point is looks same, written information differs among three dates. With the ground truth, we could confirm the land cover types (Crop). But the contents information was not correct (A coconut is indicated to be a rubber tree).14N, 100E are the points read as the typical paddy area (Crop), and have confirmed the point. 13N, 100E points are the areas (other) which were not able to be interpreted from DCP. There was a location gap (approximately 50m) between the indicated DCP information and our exploration. This can consider the setting mistake of GPS etc. However the classification class (Crop) presumed from DCP could be checked. From the above results, I found the following conclusions.

1. In some point, the names of the trees or types of crops were not correct. This was not because of the time difference. When specifying the names of trees or types of crops in DCP as text, it is good to also take a picture which can be used to confirm it.
I hope that a database (illustrated guide to flora and crops) can be built through this procedure.

2. There was a place which deviates about fifty meters from the last exploration person's position information and my exploration. Since the accuracy of the DCP must be to within 100 meters, it is still acceptable, but I am afraid that some people are not been careful enough with WGS84 settings.

3. In order to mitigate this error, I advocate using a feature of the land such as a coastal area or an intersection to carryout confirmation with satellite pictures (e.g. Google Earth, which has geographic information in the image) to confirm whether it is truly in agreement with the GPS values.
Such information may also be useful to include in a database someday to be used for validation or geometric correction of satellite data.

4. CONCLUSION

We proposed a new validation method which employs DCP information. Then we proposed a scheme to improve the information of DCP for land cover map validation. When specifying the names of trees or types of crops in DCP as text, it is recommended to take pictures which can be used to confirm it. Also, to confirm the GPS settings, using a feature of the land such as a coastal area or an intersection to carryout confirmation with satellite pictures (e.g. Google Earth, which has geographic information in the image) to confirm whether it is truly in agreement with the GPS values is also recommended.

Acknowledgements
This research has supported from S1 project "Integrated Study for Terrestrial Carbon Management of Asia in the 21st Century Based on Scientific Advancements". Moreover, gratitude is expressed to the participants in Degree Confluence Project, and a sponsor.

References from Journals:

Ahl, D. E., Gower, S. T., Mackay D. S., Burrows S. N., Norman J. M., and Diak, G. R., 2005. The effects of aggregated land cover data on estimating NPP in northern Wisconsin, *Remote Sensing of Environment*, 97, pp.1-14.

DeFris, R. S., and Belward, A. S., 2000. Global and regional land cover characterization from satellite data: an introduction to the Special Issue, *International Journal of Remote Sensing* 21, no. 6 & 7, pp.1083-1092

Foody, G. M., 2002. Status of land cover classification accuracy assessment, *Remote Sensing of Environment* 80, pp.185-201

Giri, C., Zhu, Z., and Reed, B., 2005. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets, *Remote Sensing of Environment* 94, pp.123-132

Hord, R. M., and Brooner, W., 1976. Land-use map accuracy criteria, *Photogrammetric Engineering and Remote Sensing* 42, pp.671-677

IPCC, 2003. Good Practice Guideline for Land Use, Land-Use Change, and Forestry

Iwao, K, Nishida, K and Yamagata, Y, 2006, Validation method for land cover maps using Degree Confluence Points information , Journal of the Japan Society of Photogrammetry and Remote Sensing, submitted

Kelly, M. K., Estes, J. E., and Knight, K. A., 1999. Image Interpretation Keys for Validation of Global Land-Cover Data Sets, *Photogrammetric Engineering and Remote Sensing* 65, no.9, pp.1041-1049

Kok, K. K., Farrow, A., Velkamp, A., and Verburg, P. H., 2001. A method and application of multi-scale validation in spatial land use models, *Agriculture, Ecosystems and Environment* 85, pp.223-238

McCallum, I., Obersteiner, M., Nilsson, S., and Shvidenko, A., 2006. A spatial comparison of four satellite derived 1 km global land cover datasets, *International Journal of Applied Earth Observation and Geoinformation* (in press)

Patenaude, G., Milne, R., and Dawson, T. P., 2005. Synthesis of remote sensing approaches for forest carbon estimation: reporting to the Kyoto Protocol, *Environmental Science & Policy* 8, pp.161-178

Rosenfield, G. H., and Fitzpatrick-Lins, K., 1986. A coefficient of agreement as a measure of thematic classification accuracy, *Photogrammetric Engineering and Remote Sensing* 52, pp.223-227

Strand, G.-H., Dramstad, W., and Engan G., 2002. The effect of field experience on the accuracy of identifying land cover types in aerial photographs, *International Journal of Applied Earth Observation and Geoinformatio*n 4, pp.137-146

References from websites:

Degree Confluence Project (DCP), 1996. “DCP homepage.” http://www.confluence.org/ (accessed 10 Feb. 2006)

University of Maryland, 2004. “Earth Science Data Interface.” http://glcf.umiacs.umd.edu/index.shtml (accessed 10 Feb. 2006)