

How Does Land Use/Land Cover Map's Accuracy Depend on Number of Classification Classes?

Truong Van Thinh¹, Phan Cao Duong^{1,2}, Kenlo Nishida Nasahara³, and Takeo Tadono⁴

¹Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

²Hydraulic Construction Institute - Vietnam Academy for Water Resources, Hanoi, Vietnam

³Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

⁴Earth Observation Research Center, Japan Aerospace Exploration Agency (JAXA), Tsukuba, Japan

Abstract

A land use/land cover map is an important input for different applications. However, the accuracy of land cover maps remains a great uncertainty and mapping accuracy assessment is not well-documented. The objective of this paper is to examine the relationship between overall accuracy and the number of classification classes by conducting a literature review of land cover/land use studies. The results revealed a weak negative correlation between the map's accuracy and the number of classes. The paper suggests a decrease of 0.77% map's overall accuracy with respect to the increase of 1 land cover class. The average overall accuracy produced by 05 sensor types does not show the big difference. In addition, high spatial resolution sensor such as Airborne might not be always advantageous for producing high overall accuracy map since its accuracy depends on several factors including the number of land cover classes.

(Citation: Thinh, T. V., P. C. Duong, K. N. Nasahara, and T. Tadono, 2019: How does land use/land cover map's accuracy depend on number of classification classes? *SOLA*, **15**, 28–31, doi:10.2151/sola.2019-006)

1. Introduction

The land cover/land use is an important contributing factor to the climate system. For example, agriculture, forestry, and other land use contributed 24.5% of total greenhouse gas emission globally (IPCC 2014). Another example, land cover impacts on precipitation (Sugimoto et al. 2015), generation of dust in the atmosphere (Kimura 2012), water resource (Sawaya et al. 2003), distribution of wild-fire (Keramitsoglou et al. 2008), ecosystems, and biodiversity (Delalieux et al. 2012). Therefore, producing accurate land cover/land use map are essential to the different fields of study and management processes.

However, the accuracy of land cover products is inconsistent from one to another. For example, the accuracy of some global scale land cover products such as MODIS 1 km, GLC2000 1 km and IGBP DISCover 1.1 km are quite different, ranging from 67% to 78% (Herold et al. 2008). Another example, a wide range of overall accuracy from 42% to 98% is seen in studies using images from different sensors such as Landsat, Hyperion, IKONOS, Quick Bird (Laba et al. 2002; Sawaya et al. 2003; Arroyo et al. 2010; El-Zeiny and Effat 2017).

The inconsistency of the map accuracy might depend on some factors such as spatial resolution, the homogeneity or heterogeneity of land surface (Ma et al. 2017) or classification algorithms (Gómez et al. 2016). Among the factors, the number of land cover classes is crucial. Generally speaking, the map accuracy can decrease when the number of classes (e.g., water, urban, forest) increases, because of more chances of misclassifications among classes. In fact, few studies have shown the negative correlation

between overall accuracy and the number of land cover classes (Dronova 2015; Ma et al. 2017) for some specific methods or targets (such as object-based classification or wetland study). This paper aims to build a benchmark of the relation between overall accuracy and the number of land cover classification classes in a wide range of studies without restriction about classification method, targets, sensor types, or spatial scale. The paper firstly describes how the data is collected and analyzed, then the result about the relationship of map accuracy and number of classes will be showed and discussed.

2. Method

2.1 Publications collection

The selected papers in this study were collected from the Science Direct website of Elsevier publishing company. The website provides the accessibility to a large database of scientific research including many papers on well-known remote sensing journals such as "ISPRS Journal of Photogrammetry and Remote Sensing", "Remote Sensing of Environment" and "International Journal of Applied Earth Observation and Geo-information". The relevant literature was obtained by using the advanced search function in the website. The search keywords are: "land cover mapping", "land use mapping" and "accuracy assessment". By searching based on these keywords, the website returned 163 relevant publications. From these publications, a quick scanning for titles and abstracts was implemented to eliminate irrelevant papers. After the quick scanning process, 99 papers were selected for further reading and data collection. In addition, several rules were applied to manually screen out papers which provided necessary data for overall accuracy assessment. The rules are described as following:

- Exclusion of studies focusing on an individual land cover class which affects significantly to the result of analyzing the correlation between land cover accuracy and the number of land cover classes, for example, Zhang et al. (2015).
- Removing articles failed to provide useful data such as overall accuracy, number of land cover classes which will be used for quantitative analysis in this paper.
- Exclusion of review papers relates to land cover, and land use accuracy assessment. For example, Costa et al. (2018).

After the careful reading based on the above rules, 64 related papers were selected for further analysis.

2.2 Data analysis

A data file containing 10 fields (Table 1) was established to collect information from 64 selected papers. Specifically, the file includes some general information such as title, author, year of publication, and journal. In addition, it contains other fields including sensor types, number of classes and overall accuracy. Within 10 fields of data, this paper concentrated on synthesizing and analyzing the relationship between overall accuracy and number of land cover classes. The sensor types information then was added as a complement factor to overall accuracy assessment. The other factors such as maps' resolution and classification algorithm will not be described in detail. With the hypothesis that

Corresponding author: Truong Van Thinh, Graduate School of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8572, Japan. E-mail: vanhinh1711@gmail.com.

Table 1. Fields of data used for accuracy assessment.

ID	Fields	Definition	Type
1	Title	Title of the article	Free text
2	Authors	Authors	Free text
3	Year	Year of publication	Numeric
4	Journal	Name of Journal	Free text
5	Volume and page	Publication volume and page number	Numeric
6	Sensor	Type of sensor	Free text
7	Class number	Number of classification classes	Numeric
8	Map's resolution	Resolution of land cover map	Numeric
9	Classification algorithm	Algorithm used for land cover classification	Free text
10	Overall accuracy	Best overall accuracy value	Numeric

the number of class was among the most important factors to land cover maps overall accuracy, the study conducted the analysis of variance (ANOVA) separately for four factors including number of classes, map's resolution, classification algorithm, and sensor types. The p -value produced by ANOVA process showed that the number of classes is the most affected factor (p -value = 0.005), following by classification algorithm (p -value = 0.029), map's resolution (p -value = 0.296) and sensor types (p -value = 0.723). Therefore, the "Result and discussion" section will mainly present the result of analyzing the dependence of overall accuracy on the number of classes. Additionally, the analysis of sensor types and overall accuracy will aim to emphasize the importance of the number of classes to map's accuracy.

3. Result and discussion

3.1 Land cover map's accuracy in individual studies

To a single study, if mappers increase the number of classification classes, it can be a high possibility of decreasing the overall accuracy. For example, Gessner et al. (2015) showed that the accuracy of a multi-sensor land cover map was 80% at 9 classes while it decreased to 73% in 14 classes. The similar results were shown in other studies such as Colditz et al. (2011), Van Lier et al. (2011), and Parent et al. (2015).

To demonstrate the relationship between overall accuracy and the number of classes, two studies were selected including Herold et al. (2008) (using MODIS images) and McCombs et al. (2016) (using Landsat image). From the error matrices in these studies, the number of land cover classes were reduced at every 1 class interval and the overall accuracy was recalculated accordingly. The results in Fig. 1 and Fig. 2 show a clear strong negative correlation between overall accuracy and the number of the defined classes with R^2 equal to 0.93 and 0.95 for the MODIS study and the Landsat study respectively. Although the number of classification classes is 23 and 13 for two studies respectively, the higher 95% of overall accuracy can only be achieved when the number of classes is less than 6. Interestingly, the combination of some land cover classes did not significantly improve the overall accuracy of both studies. For example, it is recognized that the number of classes from 10 to 13 in Fig. 1 show approximately equal overall accuracy and a similar situation for the classes from 8 to 12 in Fig. 2. Particularly, based on the error matrix of the study of Herold et al. (2008), the map overall accuracy remains at 91.5% when combining the class "Palustrine Emergent Wetland" and the class "Estuarine Forested Wetland". This is because of none misclassification between these two classes. In the study of McCombs et al. (2016), we obtained a 0.3% increase of overall accuracy when the "Deciduous Forest" classes and the "Evergreen Forest" classes are combined. This might be explained by the 1.4% misclassifying "Evergreen Forest" into "Deciduous Forest" and 7.5% misclassifying "Deciduous Forest" into "Evergreen Forest". Overall, the combination of classes that have low a chance of misclassification from each other could not considerably increase overall accuracy."

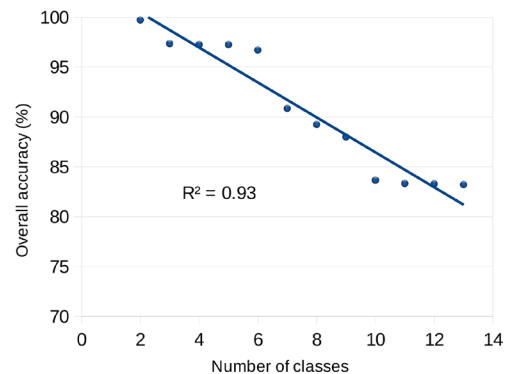


Fig. 1. The correlation between overall accuracy and the number of classes derived from the error matrix in the study of Herold et al. (2008).

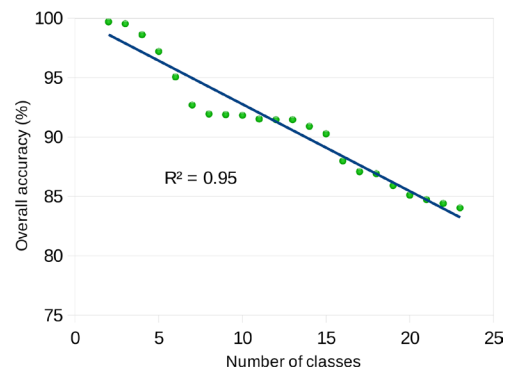


Fig. 2. The correlation between overall accuracy and the number of classes derived from the error matrix in the study of McCombs et al. (2016).

3.2 Land cover map's accuracy synthesized from multiple studies a. The dependence of overall accuracy on the number of classification classes

This paper used data derived from a variety of land cover studies to examine the dependence of overall accuracy on the number of classification classes. The selected publications are different in method, classification algorithms, sensor types and study areas. Overall, the mean accuracy of all 64 studies is 83.7% with 10 classes on average. The highest accuracy achieves 98.7% with 4 classified classes (Wardlow and Egbert 2008) while the lowest accuracy is 42% with 29 classes (Laba et al. 2002).

Figure 3 presents the correlation between overall accuracy and the number of classes for three cases: "all studies", "MODIS studies", and "Landsat studies". Obviously, the "MODIS studies" and the "Landsat studies" are also included in the "all studies" case. The result shows a negative correlation between overall accuracy and the number of classes for all three cases. Furthermore, a weak

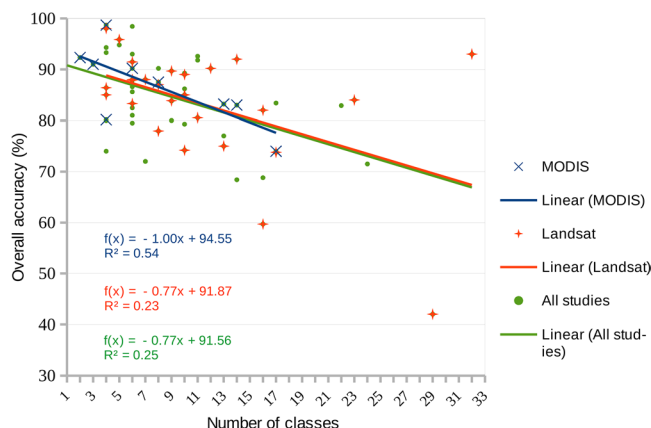


Fig. 3. The correlation between overall accuracy and the number of land cover classes (for 64 studies including MODIS and Landsat).

correlation is seen for both “all studies” and “Landsat studies” with R^2 equal to 0.25 and 0.23 respectively. This weak negative correlation is consistent with the conclusion of Ma et al. (2017) and Dronova (2015). In contrast, the MODIS case studies showed a negative moderate correlation with R^2 equal to 0.54. However, this moderate correlation might be affected by the small number of samples (9 studies).

Statistically, the overall accuracy is 100% when the number of class is 1. Therefore, initially, this paper considered using the restricted regression line with intercept point is $x_0 = 1, y_0 = 100$. However, this regression line showed a not good fit with the given set of data (with $R^2 = -0.047$). Alternatively, using the regression line without the restricted interception point provided a better fit to sample data ($R^2 = 0.25, p < 0.001$). The slope value 0.77 of the equation in Fig. 3 means the decrease of 0.77% overall accuracy with respect to the increase of 1 land cover class.

Two outlier values in Fig. 3 are of Landsat studies. One is Laba et al. (2002), in which the authors mapped 29 land cover classes for 12 million hectares area with 42% of overall accuracy. The other is Sesnie et al. (2008), in which 93.3% of overall accuracy was achieved for mapping 32 land cover classes of 8 hundred thousand hectares area. There is a big gap in overall accuracy between these two studies despite the approximately equal number of classes. By carefully considering the method provided in the studies, the reason of this gap might be explained by the intensive use of reference data (62,154 pixels) to map land cover for a small area of 8,000 km² in the study of Sesnie et al. (2008). In this study, the inflation of overall accuracy is likely to happen due to the spatial correlation between training pixels and validation pixels.

b. The average overall accuracy performing by sensor types

Figure 4 shows the average overall accuracy performed by 5 different sensor categories. The studies using images from more than 2 sensors were categorized in “Multiple sensors”. In addition, “Other” category consists of the studies using a single sensor apart from “Landsat”, “MODIS” and “Airborne”. As can be seen in Fig. 4, Landsat studies accounted for 40.6% (26 studies), followed by “Multiple sensors”, “Others”, “MODIS” and “Airborne” with the proportion are 20.3%, 17.2%, 14.1%, and 7.8% respectively. The majority of Landsat images might because of the free-downloading and its 30 m spatial resolution is popular for many studies at the national and local scale.

Most of the sensor types performed the higher overall accuracy than 80% except Airborne sensors (79.3%). Surprisingly, the mean classification accuracy of studies using Airborne images is the lowest; however, their spatial resolution is much higher than Landsat and MODIS images. Specifically, the resolution of images in 05 studies using Airborne images are from 0.4 m to 4 m (Chan and Paelinckx 2008; Lyons et al. 2018) compared to 30 m of Landsat and 250 m of MODIS. The reason might be the higher

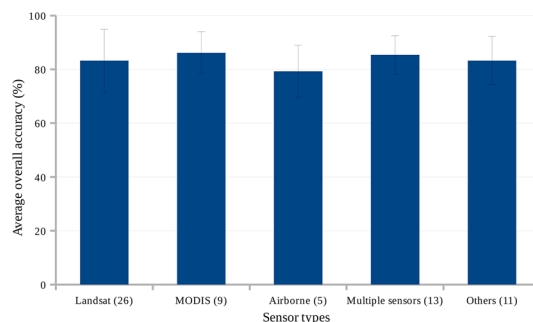


Fig. 4. The average accuracy with respect to sensor types (the error bars are the standard deviation).

average number of land cover classes in Airborne studies (14.4 classes) than Landsat and MODIS studies (11.3 classes and 7.9 classes accordingly). This reason might also be explained for the highest average accuracy of MODIS-studies (86.1%) despite its 250 m coarse spatial resolution.

There is a high percentage of multiple sensor studies (20.3%) which account for 13 studies out of 64 in total. The emerging of multiple sensors approach in land cover classification might be encouraged by the development of many satellites mission such as ALOS, Quickbird, IKONOS and so forth. To some extent, the overall accuracy of studies using images from different sensors has improved accordingly. For example, by using the combination of optical and SAR images, Zhang et al. (2015) reached the overall accuracy of the land cover map to 98.4%.

4. Conclusion

Through assessing the relationship between land cover overall accuracy and the number of classification classes, some conclusions are drawn as follows:

(1) A negative correlation exists between overall accuracy and the number of classification classes. It is easy to recognize the strong negative correlation between overall accuracy and the number of class in a single study while it shows a weak negative correlation when it comes to multiple studies. It is the trade-off between accuracy and the detail of land cover map and this remains as inner uncertainty of land cover mapping. However, as the suggestion in this paper, the land cover mappers should consider about 0.77% decrease of overall accuracy when they increase 1 land cover classification class.

(2) There is a small gap in the average overall accuracy produced by different sensor types. In addition, high spatial resolution sensor such as airborne might not be always advantageous for producing high overall accuracy map since its accuracy depends on several factors in which the number of land cover classes played an important role.

Due to the incomplete understanding about land cover map accuracy, we highly recommend more practical and review studies on accuracy assessment of land cover and land use mapping.

Acknowledgments

We thank the Project for Human Resource Development Scholarship (JDS) by Japanese Grant Aid for providing the scholarship to authors. This paper is supported under the ecosystem group of Earth Observation Research center (EORC) and ALOS-3 project of JAXA.

Edited by: M. Huang

References

- Arroyo, L. A., K. Johansen, J. Armston, and S. Phinn, 2010: Integration of LiDAR and QuickBird imagery for mapping riparian biophysical parameters and land cover types in Australian tropical savannas. *Forest Ecol. Manage.*, **259**, 598–606, doi:10.1016/j.foreco.2009.11.018.
- Chan, J. C. W., and D. Paelinckx, 2008: Evaluation of random forest and adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sens. Environ.*, **112**, 2999–3011, doi:10.1016/j.rse.2008.02.011.
- Colditz, R. R., M. Schmidt, C. Conrad, M. C. Hansen, and S. Dech, 2011: Land cover classification with coarse spatial resolution data to derive continuous and discrete maps for complex regions. *Remote Sens. Environ.*, **115**, 3264–3275, doi:10.1016/j.rse.2011.07.010.
- Costa, H., G. M. Foody, and D. S. Boyd, 2018: Supervised methods of image segmentation accuracy assessment in land cover mapping. *Remote Sens. Environ.*, **205**, 338–351, doi:10.1016/j.rse.2017.11.024.
- Delalieux, S., B. Somers, B. Haest, T. Spanhove, J. Vanden Borre, and C. A. Múcher, 2012: Heathland conservation status mapping through integration of hyperspectral mixture analysis and decision tree classifiers. *Remote Sens. Environ.*, **126**, 222–231, doi:10.1016/j.rse.2012.08.029.
- Dronova, I., 2015: Object-based image analysis in wetland research: A review. *Remote Sens.*, **7**, 6380–6413, doi:10.3390/rs70506380.
- El-Zeiny, A. M., and H. A. Effat, 2017: Environmental monitoring of spatiotemporal change in land use/land cover and its impact on land surface temperature in El-Fayoum governorate, Egypt. *Remote Sens. Appl. Soc. Environ.*, **8**, 266–277, doi:10.1016/j.rsase.2017.10.003.
- Gessner, U., M. Machwitz, T. Esch, A. Tillack, V. Naeimi, C. Kuenzer, and S. Dech, 2015: Multi-sensor mapping of West African land cover using MODIS, ASAR and TanDEM-X/TerraSAR-X data. *Remote Sens. Environ.*, **164**, 282–297, doi:10.1016/j.rse.2015.03.029.
- Gómez, C., J. C. White, and M. A. Wulder, 2016: Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.*, **116**, 55–72, doi:10.1016/j.isprsjprs.2016.03.008.
- Herold, M., P. Mayaux, C. E. Woodcock, A. Baccini, and C. Schmullius, 2008: Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sens. Environ.*, **112**, 2538–2556, doi:10.1016/j.rse.2007.11.013.
- Keramitsoglou, I., C. Kontoes, O. Sykioti, N. Sifakis, and P. Xofis, 2008: Reliable, accurate and timely forest mapping for wildfire management using ASTER and Hyperion satellite imagery. *For. Ecol. Manage.*, **255**, 3556–3562, doi:10.1016/j.foreco.2008.01.077.
- Kimura, R., 2012: Effect of the strong wind and land cover in dust source regions on the Asian dust event over Japan from 2000 to 2011. *SOLA*, **8**, 77–80, doi:10.2151/sola.2012-020.
- Laba, M., S. Gregory, J. Braden, D. Ogurcak, E. Hill, E. Fegraus, J. Fiore, and S. DeGloria, 2002: Conventional and fuzzy accuracy assessment of the New York Gap Analysis Project land cover map. *Remote Sens. Environ.*, **81**, 443–455, doi:10.1016/S0034-4257(02)00020-2.
- Lyons, M. B., D. A. Keith, S. R. Phinn, T. J. Mason, and J. Elith, 2018: A comparison of resampling methods for remote sensing classification and accuracy assessment. *Remote Sens. Environ.*, **208**, 145–153, doi:10.1016/j.rse.2018.02.026.
- Ma, L., M. Li, X. Ma, L. Cheng, P. Du, and Y. Liu, 2017: A review of supervised object-based land-cover image classification. *ISPRS J. Photogramm. Remote Sens.*, **130**, 277–293, doi:10.1016/j.isprsjprs.2017.06.001.
- McCombs, J. W., N. D. Herold, S. G. Burkhalter, and C. J. Robinson, 2016: Accuracy assessment of NOAA coastal change analysis program 2006–2010 land cover and land cover change data. *Photogramm. Eng. Remote Sens.*, **82**, 711–718, doi:10.14358/PERS.82.9.711.
- Parent, J. R., J. C. Volin, and D. L. Civco, 2015: A fully-automated approach to land cover mapping with airborne LiDAR and high resolution multispectral imagery in a forested suburban landscape. *ISPRS J. Photogramm. Remote Sens.*, **104**, 18–29, doi:10.1016/j.isprsjprs.2015.02.012.
- Sawaya, K. E., L. G. Olmanson, N. J. Heinert, P. L. Brezonik, and M. E. Bauer, 2003: Extending satellite remote sensing to local scales: Land and water resource monitoring using high-resolution imagery. *Remote Sens. Environ.*, **88**, 144–156, doi:10.1016/j.rse.2003.04.006.
- Sesnie, S. E., P. E. Gessler, B. Finegan, and S. Thessler, 2008: Integrating Landsat TM and SRTM-DEM derived variables with decision trees for habitat classification and change detection in complex neotropical environments. *Remote Sens. Environ.*, **112**, 2145–2159, doi:10.1016/j.rse.2007.08.025.
- Sugimoto, S., T. Sato, and T. Sasaki, 2015: Impact of land-use change on winter precipitation in Hokkaido, Japan. *SOLA*, **11**, 95–99, doi:10.2151/sola.2015-023.
- Van Lier, O. R., J. E. Luther, D. G. Leckie, and W. W. Bowers, 2011: Development of large-area land cover and forest change indicators using multi-sensor Landsat imagery: Application to the Humber River Basin, Canada. *Int. J. Appl. Earth Obs. Geoinf.*, **13**, 819–829, doi:10.1016/j.jag.2011.05.019.
- Wardlow, B. D., and S. L. Egbert, 2008: Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sens. Environ.*, **112**, 1096–1116, doi:10.1016/j.rse.2007.07.019.
- Zhang, G., X. Xiao, J. Dong, W. Kou, C. Jin, Y. Qin, Y. Zhou, J. Wang, M. A. Menarguez, and C. Biradar, 2015: Mapping paddy rice planting areas through time series analysis of MODIS land surface temperature and vegetation index data. *ISPRS J. Photogramm. Remote Sens.*, **106**, 157–171, doi:10.1016/j.isprsjprs.2015.05.011.

Manuscript received 24 August 2018, accepted 4 January 2019
 SOLA: <https://www.jstage.jst.go.jp/browse/sola/>