

AREA ESTIMATION OF MULTI-TEMPORAL GLOBAL IMPERVIOUS LAND COVER BASED ON STRATIFIED RANDOM SAMPLING

Yali Gong*, Huan Xie, Xiaohua Tong, Yanmin Jin, Xiong Xv and Qunming Wang

College of Surveying and Geo-Informatics, Tongji University – refi.gong@outlook.com

Commission VI, WG VI/4

KEY WORDS: Area estimation, Impervious Land Cover, Stratified sampling, Design-based Inference

ABSTRACT:

Estimating area of impervious land cover is the most useful and one of the ecological assessment indexes of urban and regional environment. Global land cover maps are inevitably misclassified, which affects the quality and application of the data. Statistical approach for assessing the accuracy is critical to understand the global change information and area estimation is usually based on sample data with a probability-based estimator. However, research on evaluation of multi-temporal global impervious land cover maps has not been implemented. In this study, spatial characteristics of the data are considered to assess the thematic map accuracy with a two-stage stratified random sampling plan. The first stage of stratification is determined by the global urban ecoregion and the second one is determined by land cover classes. Additionally, sample size of both map stage and pixel stage are calculated using a probability sampling model. A response design is constructed for a per-pixel accuracy assessment and blind interpretation is implemented using sample pixels and its surrounding area. Our method is applied to the multi-temporal global impervious land cover maps between 2000 and 2010 with a time interval of 5 years and the estimated area in different epoch are listed in detail. The main contribution of our research is illustrating the details for calculating the proportion area of impervious land cover and corresponding confidence intervals based on the reference classification. The experimental results show that the increasing area of the impervious surface according to the sample unit shows good agreement with the urbanization and descriptive accuracy assessments by user's, producer's and overall accuracy are shown respectively.

1. INTRODUCTION

Area estimation is one of the most critical indicators to assess the accuracy of the land cover data from the remotely sensed imagery, which plays an important role in the development of productive economy and establishing scientifically accounting application to policy approaches for global ecological sustainability.

Attempts to estimate the area of the single land cover map or land cover change has a far-reaching impact on societies and environments, reflecting the range of human activities on natural environment and ecosystem (Schneider et al., 2010). Area estimation should follow a good practice principle, i.e., neither over- nor underestimating global urban and global urban change area, and reduce uncertainties as far as possible (Olofsson 2014). The aim of area estimation is estimating the reference area of land cover class or that of a land cover change (Stehman, 2009).

Estimating area of global land cover products is obviously very challenging, expensive, and time consuming. Area estimation based on a sample of reference observations, and design-based inference for the entire population (Stehman, 1997) can be a good choice when the estimation refers to a large scale region or some ground surveys are difficult to reach. In addition, stratified sampling design with the consideration of data features and will provide a good result of area estimation. The stratified random sampling is defined as selecting a random sample from each stratum, and implemented with the map classes defined as a stratum.

A variety of area estimators approaches had been proposed, for example, a bias-adjusted estimator, a model-assisted area estimator, calculating area from the map and so on. For stratified random sampling, three sample-based approaches were direct estimation from the estimated sample error matrix yield the same area estimators (Stephen, 2013). Additionally, the land cover dataset produced by the classification of the remote sensing data such as Landsat and other satellites imagery has classification errors (Olofsson, 2020), indicating that area estimation directly from counting the classified pixels is totally biased. Instead of obtaining the estimated area of category i directly from the map classification, an area estimator can be constructed based on the reference classification of each sample unit. The reference classification is defined by the comparison between the classification label of land cover map with that of high resolution satellite imagery or ground truth. A confidence interval should also be provided to quantify the uncertainty of area estimation.

The objective of this research is estimating the proportion of area of global impervious land cover map and its corresponding uncertain interval at a 95% confidence level with a stratified sampling estimator (Olofsson 2014). In this paper, we present a two-stage stratified sampling for estimating the area of the global impervious land cover maps from 2000 to 2010 based on the classification error matrix of stratified sampling and validated the effectiveness of this approach. The experimental dataset is provided by the results of NUACI-based global impervious land cover classification with a resolution of 30 meters (Liu, 2018).

* Corresponding author

2. METHODOLOGY

2.1 Sampling Design

Here we present in detail the calculations of the accuracy data for estimating area and associated confidence intervals. This research follows the sampling framework proposed in the previous research (Sarndal, 1992, Stehman, 2001), and ensures the consistency of estimator and parameters. For rare categories, stratified probability sampling would be a good choice. Here, the global impervious land cover map is partitioned into two category, urban and non-urban.

In order to rigorously evaluate the accuracy of the global scale impervious land data products, a two-stage stratification is designed to obtain the samples. Generally, the main steps of the stratified sampling design are sampling design, response design and error matrix analysis (Stehman, 1997). During the sampling design, appropriate stratification is important to obtain independent samples. The pixels labelled as urbans are geographical heterogeneity and some are extremely sparse in some map grids. If independent samples are obtained by the use of the simple random sampling, then the great majority of the sample units would fall in the area where no impervious surface exists, leading to a higher variance in the estimated accuracy of the population. Here a two-stage stratified random sampling is adopted (Tong et al., 2011) to assess the accuracy of the global impervious land cover products precisely considering the spatial characteristics in different region scales.

The first-stage of the stratification is designed to select the samples geographically, which is created by urban ecoregion scheme proposed by previous research (Schneider et al, 2010), leading to 16 geographic regions within the global imperious land. They are Temperate forest in North America, Temperate forest in Europe, Temperate forest in East Asia, Temperate grassland in North-South America, Temperate grassland in Middle East Asia, Tropical broadleaf forest in south America, Tropical broad forest in Africa, Tropical and sub-tropical forest in Asia, Tropical and sub-tropical savannah in S.America, Tropical and sub-tropical forest in Africa, Tropical and sub-tropical grassland, Temperate Mediterranean, Arid semi-arid desert and shrubland, Arid semi-arid steppe in Central Asia, Boral forest and tundra and permanent ice and snow. It is worth noting that the 16th ecological zone is a perennial ice and snow cover so that there is no urban land. Thus, land cover maps within this ecological zone are discarded. This stratification is necessary since previous research has shown that geographic variations lead to inaccurate assessment when the class was homogenous.

The primary sampling units (PSU) in the first stage are map grids with a size of $10^{\circ} \times 10^{\circ}$. Then each original map grid is further divided into 100 small grids with a size of $1^{\circ} \times 1^{\circ}$ to obtain the evenly distributed map grids. Taking these new map grids as the PSU, the samples size of the first stage can be calculated the statistical sampling model proposed in (Tong et al. 2011). Given a level of significance, the sample size is obtained by minimizing the relative error between the actual classification error and the predicted classification error. Then, the total sample size of the PSU was determined and the samples in each $1^{\circ} \times 1^{\circ}$ PSU were selected randomly. Since the sample size of each ecoregion is proportional to the land area,

all PSU within the same ecoregion has the same inclusion probability.

The second stage of stratification is conducted within the two land cover classes, i.e., urban and non-urban in each PSU. The pixel is the secondary sampling unit (SSU) in this stratified sampling design and the sample size in this stage is obtained by the same method as that in the first stage. Given a fixed total sample size, the allocated sample size of each class is determined by minimizing the sum of the variances of producer's accuracy, user's accuracy and area estimation of urban for stratified random sampling. (John E.Wagner, 2015).

Note that the stratified random sampling plans are designed separately for the dataset in different epoch and samples for area estimation in different epoch are independent. Thus, the error matrix of single epoch doesn't include any information related to the area of urban change. So the area estimation of land cover change is not within the scope of this research.

2.2 Response Design

The response design is the process to evaluate the similarity between the map classification and the reference label (Stephen, 1998). Generally, the reference label is defined by the assessment of high-resolution satellite imagery or ground truth. The main elements of the response design are listed as follows.

2.2.1 Blind interpretation: Collection of reference labels was accomplished by seven interpreters who has completed training and experience in collection for land cover products. The interpreters should not know the map classification of each sample unit in advance.

2.2.2 Using Google Earth™ imagery as reference data: Generate a vector Keyhole Markup language Zipped (KMZ) file of each sample unit and overlay it on Google earth™ imagery to assist the interpreters in obtaining the reference label for the sample pixel. Regarding the potential errors of geodetic coordinates of the original dataset, each sample pixel is surrounded with a 3×3 pixel window to determine the most appropriate label for the sample pixel. The interpreters can select the corresponding date of Google Earth™ based on the Landsat imagery acquisition date for global impervious land cover to determine the reference sample labels corresponding to the three periods.

2.2.3 Judgment criteria for major decision: During the judgement of the mixed sample, the principle of area dominance is applied to determine the attribute of the sample. Feature types are used as final labels, that is, quantitative indicators to reduce the judgment errors. The consistent comparison between the time series impervious land cover map and reference data is based on the proportion of the impervious land within the reference pixels. As a result, the attribute labels of the reference data can be determined by comparing the decision from different interpreters and the majority of the results with the same judgment result are used as the final label to reduce the judgment error.

2.2.4 Credibility of the interpretation: The credibility of the interpretation results of the samples can be divided into three parts, i.e., completely correct, completely wrong or unable to judge. In the process of sample judgment, various factors, i.e., classification methods, ecological distribution, phenological conditions, original TM images, reference data, external factors, etc., lead to uncertainty of the judgment results. Credibility classification can be used to make quantitative judgments on samples, remove uncertain samples, and reduce the impact of uncertain factors.

2.3 Area estimation

Area estimation should follow the good practice principle, i.e., neither overestimating nor underestimating global urban area to reduce the uncertainties as more as possible (Olofsson, 2014). The urban area estimated from counting the pixels labelled as urban is not rigorous since the classification errors are not considered. It is better to estimate the area using a stratified sampling estimator based on the summary of the area of all stratifications with a given confident level. Once the manual inspection is completed, the error matrix can be obtained to estimate the urban area for each epoch. The probability-based estimator, as well as its variance, can be estimated using the error matrix.

The error matrix is obtained based on defined as a match between the map class and the reference label (Stephen, 1998). Comparison between the map class of the sample pixel and its corresponding reference class can be expressed as the confusion matrix. The counting sample pixel for each cell of the confusion matrix is denoted as n_{ij} to represent the number of pixels that are map class i and reference class j . Since the objective of this research is to estimate the area of global impervious land cover of three epochs, so the confusion matrix should be converted to error matrix, which is directly related to the estimated area. A more informative presentation of the error matrix is in terms of the unbiased estimator of the proportion of area in cell i, j of the error matrix, which is defined as follows (Stehman, 2013):

$$\hat{p}_{ij} = w_i \frac{n_{ij}}{n_i} \quad (1)$$

It should be noted that the weights of different stratifications are defined by their area proportions. In equation (1), w_i is the proportion of the area mapped as category i , denoted as:

$$w_i = \frac{A_i}{A_{total}} \quad (2)$$

where A_{total} denotes the total area of pixels in the map dataset, A_i is the proportion of the area mapped as category i . Supposing that the terminology symbol of n_i denotes the sum of proportion of the area mapped as category i , the direct estimator of $p_{.j}$ is the sum of the sample-based estimators p_{ij} in stratified random sampling, which is defined as follows:

$$\hat{p}_{.j} = \sum_{i=1}^m w_i \frac{n_{ij}}{n_i} \quad (3)$$

Based on previous definition, the standard error matrix can be obtained. In the error matrix, map labels refer to the rows and reference labels refer to columns, cell entries are expressed as the percent of area.

The cells of the error matrix are the area proportions with a difference between map and reference labels, where the summary of the column gives the estimated area proportions by the reference classification and the summary of the row represents that by the map classification. The diagonal elements contain the correctly classified items in each category, and the off-diagonal elements contain the confusion between the corresponding categories.

The proportion of area of category j based on the reference classification can be estimated from the total of the column. An unbiased estimator of the total area of category j is defined as:

$$\hat{A}_j = A_{total} \cdot \hat{p}_{.j} \quad (4)$$

Area is typically estimated from samples, these estimates are subject to uncertainty. The uncertainty of an estimate can be represented by calculating its standard deviation with a given confidence interval. For a stratified estimator, the estimated standard deviation of the estimated area proportion is (Cochran, 1977)

$$S(\hat{p}_{.j}) = \sqrt{\sum_{i=1}^m w_i^2 \frac{n_i}{n_i - 1} \left(\frac{n_{ij}}{n_i} \left(1 - \frac{n_{ij}}{n_i} \right) \right)} \quad (5)$$

Subsequently, the standard error of the adjusted area estimate is

$$S(\hat{A}_j) = A_{total} \cdot S(\hat{p}_{.j}) \quad (6)$$

Thus, the final estimated area with an approximate 95% confidence interval for category j is defined as follows:

$$\hat{A}_j \pm 2 \cdot S(\hat{A}_j) \quad (7)$$

3. DATA SOURCE

To address area estimation concerns of the global urban land, this research focus on a new multi-temporal global impervious surface data from 2000 to 2010 with a five-year interval, on the basis of the Landsat images. Thick black boxes denotes the original $10^\circ \times 10^\circ$ map grids. The multi-temporal global urban land cover ranges from 80 degrees north to 60 degrees south which is produced by the method of Normalized Urban Area Composite Index (NUACI) proposed by Liu et al. (2018) with the Google Earth Engine platform. This datasets utilized Landsat images with a spatial resolution of 30 meters extract the urban land cover, the classification includes urban and non-urban. According to the uniform amplitude division specification, there are 224 maps all over the world.

4. RESULTS

A two-stage stratified random sampling plan is designed for the multi-temporal global impervious surface data. The first stage of the stratified sample is selected from the 22400 map grids and the number of primary sampling units is 532. These samples are selected randomly according to the land area of each stratum, which is shown in Figure 1. These samples are selected randomly according to the land area of each stratum. Note that samples selected from the smaller map grid shows a better reflection of urban aggregation.

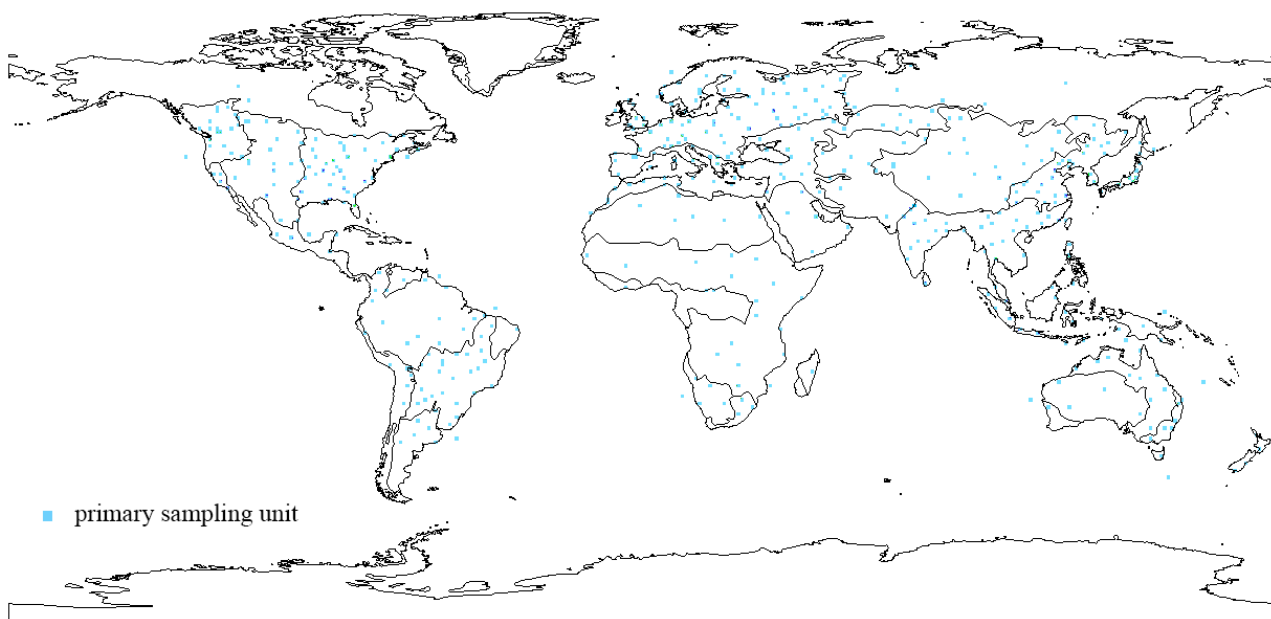


Figure 1. Simple random sample of PSUs per stratum for global impervious land cover

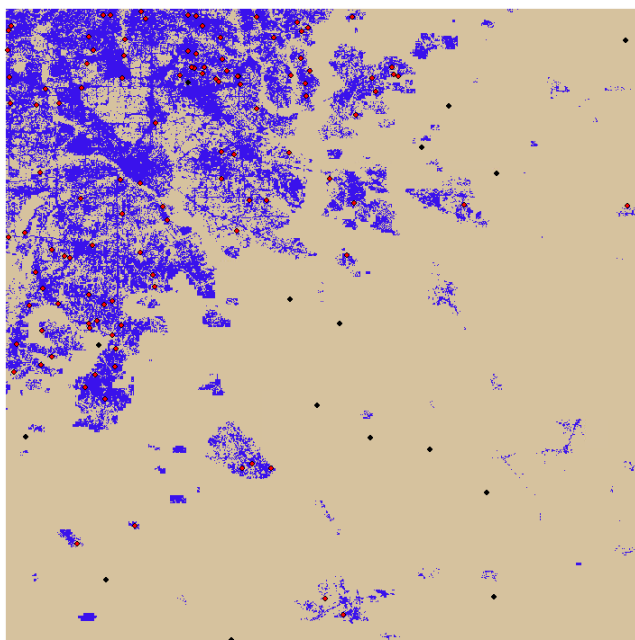


Figure 2. Magnified view of one PSU from temperate forest in North America in 2000 (red and black dots denote the urban and non-urban sample)

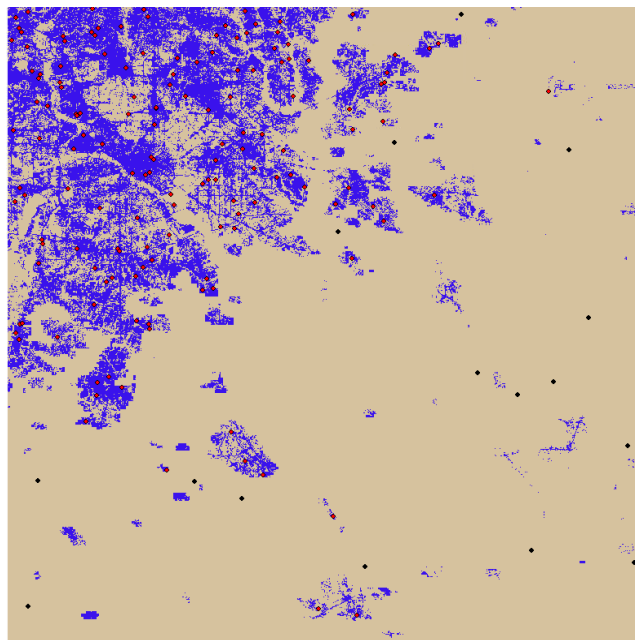


Figure 3. Magnified view of one PSU from temperate forest in North America in 2005

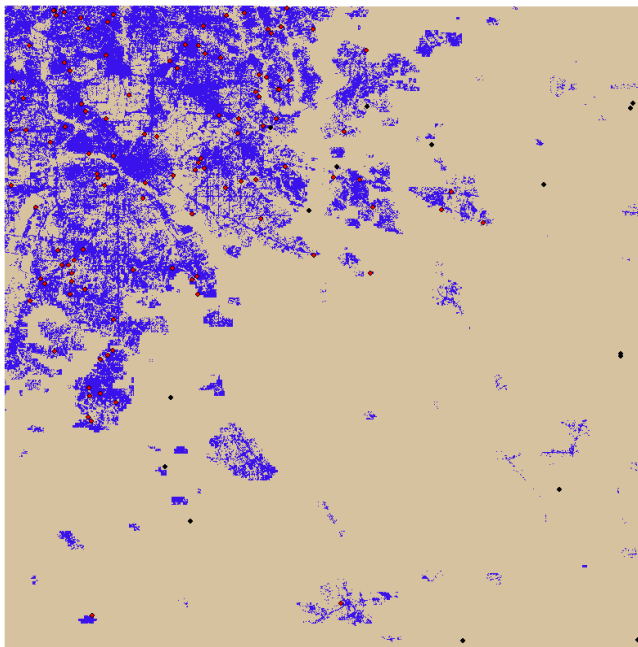


Figure 4. Magnified view of one PSU from temperate forest in North America in 2010

The second stage of the stratified sample is conducted within each selected map grid and the total number of samples for three epochs is 8708 pixels. Then the size of samples for urban and non-urban are allocated using the optimizing sample size allocation in the reference (Wagner, 2015). As a result, 4671 urban samples and 4037 non-urban samples are obtained in this case. Furthermore, the samples are also selected randomly and sample in each epoch are selected independently.

A selected PSU of three epoch from the temperate forest in North America in the second stage and its magnified views are shown in Figures 2, 3 and 4, showing that the arbitrary locations of the sample in three epochs. Following the approach in Section 2, the corresponding error matrix are shown in Tables 1, 2 and 3.

The classification accuracy increases with the improvement of the data quality, showing less variance. Estimates of area of this dataset, based on the stratified random sampling design, the result can be obtained by equations. (4)-(7). Therefore, the estimated area of global urban region and within its corresponding variance in different epochs are 813.95 ± 194.86 thousand km² in 2000, 947.6 ± 192.76 thousand km² in 2005 and 1038.93 ± 216.85 thousand km² in 2010, respectively. It is clear that the global impervious land cover increased from 2000 to 2010.

		Reference		
		no-urban	urban	Total
Map	no-urban	0.8936	0.0812	0.9748
	urban	0.0086	0.0166	0.0252
	Total	0.9022	0.0379	

Table 1 estimated error matrix with cell entries expressed as the estimated proportion of area. in 2010

		Reference		
		no-urban	urban	Total
Map	no-urban	0.8559	0.1227	0.9786
	urban	0.0081	0.0133	0.0214
	Total	0.864	0.136	

Table 2 estimated error matrix with cell entries expressed as the estimated proportion of area in 2005

		Reference		
		no-urban	urban	Total
Map	no-urban	0.8524	0.1302	0.9762
	urban	0.0059	0.0115	0.0034
	Total	0.9624	0.0057	

Table 3 estimated error matrix with cell entries expressed as the estimated proportion of area. in 2000

5. CONCLUSION

In this paper, we use a two stage stratified random sampling and estimated the area of sample-based approaches were direct estimation from sample error matrix to estimate the area of a 30-m resolution multi-temporal global impervious surface data from 2000 to 2010 with a five-year interval. The proposed two-stage stratification sampling plan is the full consideration of the geographical locations of impervious surface land cover, leading to a better-selected samples for area estimation. Analysing this dataset shows a rapid urbanization all over the world.

ACKNOWLEDGEMENTS

We sincerely thank Prof. Xiaoping Liu for offering the multi-temporal global urban land map. This work was supported by National Natural Science Foundation of China (41631178 and 41822106) and the National Key R&D Program of China (2018YFB0505400).

REFERENCES

- Cochran, W. G., 1977. Sampling techniques. New York, NY: Wiley.
- Liu X., Hu G., Chen Y., et al., 2018. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. *Remote Sens. Environ.*, 209, 227-239.
- Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.*, 148, 42-57.
- Olofsson, P., Arévalo, P., Espejo, A. B., Green, C., Lindquist, E., McRoberts, R. E., & Sanz, M. J., 2020. Mitigating the effects of omission errors on area and area change estimates. *Remote Sens. Environ.*, 236, 111492.
- Särndal, C.E., Swensson, B., Wretman, J., 1992. Model-Assisted Survey Sampling. Springer-Verlag, New York.
- Schneider, A., Friedl, M.A., Potere, D., 2010. Mapping global urban areas using MODIS 500-m data: new methods and

datasets based on 'urban ecoregions'. *Remote Sens. Environ.*, 114, 1733–1746.

Stehman, S.V., 1997. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.*, 62, 77–89.

Stephen, V, Stehman, et al., 1998. Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. *Remote Sens. Environ.*, 64(3), 331-344.

Stehman, S.V., 2001. Statistical rigor and practical utility in thematicmap accuracy assessment. *Photogramm. Eng. Remote. Sens.*, 67, 727–734.

Stehman S V., 2009. Model-assisted estimation as a unifying framework for estimating the area of land cover and land-cover change from remote sensing. *Remote Sens. Environ.*, 113, 2455-2462.

Stehman, S. V., 2013. Estimating area from an accuracy assessment error matrix. *Remote Sens. Environ.*, 132, 202-211.

Tong, X. H., Wang, Z. H., Xie, H., Liang, D., Li, J. C., 2011. Designing a two-rank acceptance sampling plan for quality inspection of geospatial data products. *Computers & Geosciences*, 37(10), 1570-1583.

Wagner J E., Stehman S V., 2015. Optimizing sample size allocation to strata for estimating area and map accuracy. *Remote Sens. Environ.*, 168, 126-133.