

Journal of Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Accuracy assessment model for classification result of remote sensing image based on spatial sampling

Dongmei Huang
Shoujue Xu
Jingqi Sun
Suling Liang
Wei Song
Zhenhua Wang

SPIE.

Dongmei Huang, Shoujue Xu, Jingqi Sun, Suling Liang, Wei Song, Zhenhua Wang, "Accuracy assessment model for classification result of remote sensing image based on spatial sampling," *J. Appl. Remote Sens.* **11**(4), 046023 (2017), doi: 10.1117/1.JRS.11.046023.

Accuracy assessment model for classification result of remote sensing image based on spatial sampling

Dongmei Huang, Shoujue Xu, Jingqi Sun, Suling Liang,
Wei Song, and Zhenhua Wang*

Shanghai Ocean University, College of Information, Shanghai, China

Abstract. The classification accuracy of a remote sensing image should be assessed before the classification result is used for scientific investigation and policy decision. We proposed an accuracy assessment model based on spatial sampling to reflect region sensitivity of a remote sensing image. The proposed model aims to solve the following problems: (1) what sampling size should be selected for accuracy assessment; (2) where sample points should be distributed in a region; and (3) how to analyze the result of accuracy assessment. This assessment model was proposed based on gray-level co-occurrence matrix (GLCM) and considered both sampling size calculation and sample points distribution during the assessment. The overall accuracy and kappa coefficient derived from this model were very close to the true value derived from the total assessment, suggesting that the assessment accuracy of the model is close to that of total assessment. Compared with the percent sampling model, the model could quantify the relationship between GLCM-correlation parameter and sample size, thereby allowing producer and user to determine sample size according to spatial uniformity and heterogeneity. Compared with the random sampling model, the model could ensure that the sample points are uniformly distributed in the spatial region and proportionally distributed in different types of land cover. Taken together, the proposed model is suitable for the accuracy assessment of the classification result of a remote sensing image. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JRS.11.046023](https://doi.org/10.1117/1.JRS.11.046023)]

Keywords: sampling size; sample points distribution; gray-level co-occurrence matrix.

Paper 170422 received Jun. 4, 2017; accepted for publication Nov. 30, 2017; published online Dec. 15, 2017.

1 Introduction

Remote sensing images could provide the representation of object surface at different spatial and temporal scales. They are widely used in a great number of fields, including predicting epidemiology and burned area,^{1,2} detecting forest and cultivated land changes, monitoring soil erosion and environmental change,^{3,4} and mapping land cover and species distribution.⁵⁻⁸ In particular, the majority of a remote sensing image should be conducted image classification before their applications, which can be achieved by either visual or computer-aided analysis. A key concern during image classification is whether the classification result derived from the remote sensing image has sufficient quality for operational application. Thus, it is required to propose accuracy assessment model to judge whether the accuracy of classification result meets the requirement of user's applications.

Currently, several methods have been used for accuracy assessment of remote sensing classification result, including population-based statistical framework,⁹ multiple-objective accuracy assessments,¹⁰ geographically weighted accuracy measures,¹¹ and stratified random sample for the National Land Cover Database.¹² Some studies take sampling size calculation as the major concern, whereas other studies take sample points distribution as the major concern. However, the classification result of remote sensing is a special product. Both sampling size calculation and sample points distribution are crucial for the classification accuracy. During image classification,

*Address all correspondence to: Zhenhua Wang, E-mail: zh-wang@shou.edu.cn

it is required to determine sample size based on spatial autocorrelation, select sample points based on spatial heterogeneity, and qualify classification accuracy by comparing sample points and reference data.

In this paper, we proposed an accuracy assessment model for a classification result of a remote sensing image based on spatial sampling. This model considered both sampling size calculation and sample points distribution. It would allow producer and user to determine sampling rate according to spatial uniformity and heterogeneity. Moreover, it could ensure that sample points are uniformly distributed in the spatial region and proportionally distributed in different types of land cover.

2 Materials and Methods

2.1 Remote Sensing Data and Study Region

The study region is located in Sichuan Province, Western China. The data set is a fusion image of multispectral and panchromatic images based on the Landsat-8/OLI image obtained on August 24, 2015, with 15-m spatial resolution [Fig. 1(a)]. The image has 256 different gray levels. The original image is available at <http://www.gscloud.cn>. The reference data are aero high spatial resolution images obtained on August 10, 2015, with 0.6-m spatial resolution [Fig. 1(b)]. The two data follow the same coordinate system WGS_1984_UTM_zone_48N.

2.2 Accuracy Assessment Model

The accuracy is typically used to express the degree of “correctness” of a classification result. We proposed an accuracy assessment model to reduce data redundancy and ensure assessment precision based on two parameters, sampling size (n) and optimal distance (d). In the model, each pixel was defined as an assessed item. Supposed that the remote sensing image was rectangular, which had N_x columns and N_y rows, the lot size (N) of accuracy assessed items was $N = N_x \times N_y$.

According to the first law of geography,¹³ each pixel had the spatial autocorrelation with each other. The closer autocorrelation was more strongly related than that of more distant ones. In this paper, the spatial autocorrelation was calculated by gray-level co-occurrence matrix (GLCM). The sampling size n and optimal distance d were then deduced based on the model of accuracy assessment.

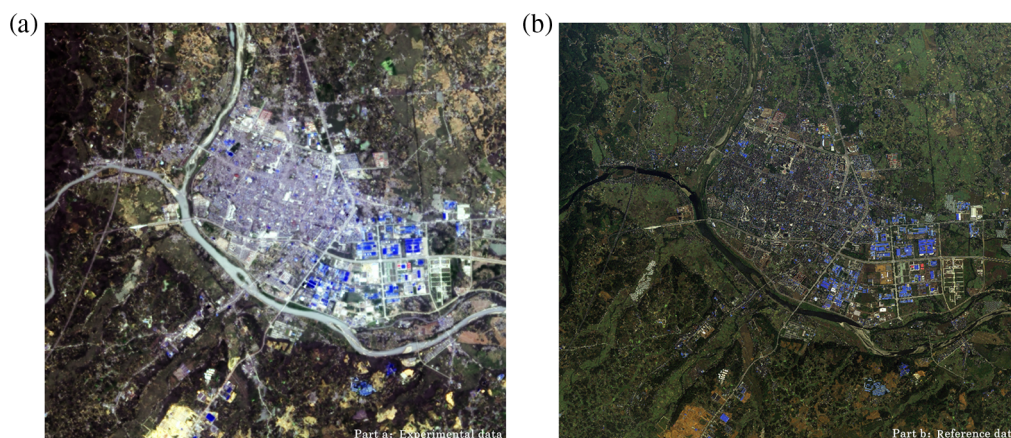


Fig. 1 Two remote sensing images in the study: (a) remote sensing image of study region with 15-m spatial resolution based on the Landsat-8/OLI image obtained on August 24, 2015, and (b) remote sensing image of reference data with 0.6-m spatial resolution obtained on August 10, 2015.

2.2.1 Gray-level co-occurrence matrix

Supposed that the gray level at each pixel was quantified as N_g levels. $G_x = \{0, 1, \dots, N_g - 1\}$ was the set of N_g quantified gray levels. The remote sensing image, H , indicated a function that assigned some gray level in G to each pixel or pair of coordinates in $N = N_x \times N_y$.

The texture-context information was specified by the matrix of relative frequencies (P_{ij}) with two neighboring pixels separated by distance d in the remote sensing image, where one pixel with gray level i and the other pixel with gray level j ($i, j \in G_x$).

The matrices of gray-level co-occurrence frequency (P_{ij}) were represented as a function of the angular relationship (θ) and distance (d) among the neighboring pixels as

$$p(i, j, d, \theta) = \#\{(k, l), (m, n) \in (N_y \times N_x)(N_y \times N_x)\} \\ \{(k - m = 0, |1 - t| = d) \times (|k - m| = d, 1 - t = 0), H(k, l) = i, H(m, t) = j\}, \quad (1)$$

where # was the item number, (k, l) and (m, t) were the rows and columns information of the pixel with i and j gray, respectively, and d was the number of interval pixels between (k, l) and (m, t) on angular (θ) in the practical calculation.

GLCM-correlation parameter (r) of each pixel was calculated by the following equation:

$$r = \frac{\sum_i \sum_j (ij) p(i, j, d, \theta) - \mu_x \mu_y}{\sigma_x \sigma_y}, \quad (2)$$

where $P_{(i,j,d,\theta)}$ was the entry in a normalized GLCM. The mean (μ) and standard deviations (σ) for the rows and columns of the matrix were calculated as follows:

$$\mu_x = \sum_i \sum_j i \cdot p(i, j, d, \theta), \quad (3)$$

$$\mu_y = \sum_i \sum_j j \cdot P(i, j, d, \theta), \quad (4)$$

$$\sigma_x = \sum_i \sum_j (i - \mu_x)^2 \cdot p(i, j, d, \theta), \quad (5)$$

$$\sigma_y = \sum_i \sum_j (j - \mu_y)^2 \cdot p(i, j, d, \theta). \quad (6)$$

GLCM-correlation parameter (r) ranged from -1 to 1 . When r was close to 1 , the pixels had strong spatial correlation, which were located at (k, l) and (m, t) . Otherwise, the pixels had weak spatial correlation.

2.2.2 Accuracy assessment model

Based on the GLCM-correlation parameter (r), the sampling size (n), and optimal distance (d) were deduced as shown below:

$$\begin{cases} \min_n \varepsilon^2 \\ \text{s.t.} \left[\frac{\sum_i \sum_j (ij) p(i, j, d, \theta) - \mu_x \mu_y}{\sigma_x \sigma_y} - r_0 \right] = \varepsilon, \\ n = \left\lfloor \frac{N_x \cdot N_y}{n_0 \text{ deg} \cdot n_{90} \text{ deg}} \right\rfloor \end{cases} \quad (7)$$

where ε was an arbitrarily small value, r_0 was the critical value of GLCM-correlation parameter (r) provided by the users and producers to balance data redundancy and accuracy, θ was defined as the value with four different orientation information, including 0 deg and 90 deg. Here, for

simplified calculation, only two different orientations were considered. n was the optimal sample size. $n_{0 \text{ deg}}$ and $n_{90 \text{ deg}}$ were the number of the interval pixels at 0 deg and 90 deg, respectively.

2.3 Accuracy Analysis and Comparison

The feasibility and advantage of our proposed accuracy assessment model were assessed by comparing with total assessment, percent sampling model, and random sampling model. The overall accuracy, producer accuracy, user accuracy, commission, omission, and kappa coefficient were used as the assessment parameters during these comparisons.¹⁴⁻¹⁹

3 Results

3.1 Classification Result of Remote Sensing Images

Five different types of land cover were classified from the two above-mentioned images, including building, agriculture, bare, water, and forest based on the support vector machine (SVM) in ENVI 5.1 software. Two classification results in vector form were shown in Fig. 2. (a) was the classification result of experiment data, and (b) was the classification result of the reference data.

SVM method consisted of finding a separation hyperplane among the training samples with the larger margins. The separating hyperplane was the geometric place where the following linear function was zero

$$f(x) = \langle x, w \rangle + b, \tag{8}$$

where w represented the orthogonal vector to the hyperplane, $f(x) = 0; b/\|w\|$ was the distance from the hyperplane to the origin, and $\langle x, w \rangle$ denoted that x inner products w . The parameters of Eq. (8) were obtained from the following quadratic optimization problem:

$$\sum_{i=1}^m \lambda_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j y_i y_j \langle \varphi(x_i), \varphi(x_j) \rangle \quad \text{subject to: } \begin{cases} 0 \leq \lambda_i \leq C; i = 1, \dots, m \\ \sum_{i=1}^m \lambda_i y_i = 0 \end{cases}, \tag{9}$$

where λ_i was the Lagrange multipliers, $y_i = \{-1, +1\}$ defined the class of x_i , since SVM was a binary classifier, C acted as an upper bound of λ values, and $\varphi(x)$ was a function adopted to remap the input vectors into a higher dimensionality space. The inner product $\langle \varphi(x_i), \varphi(x_j) \rangle$ was known as the kernel function. A popular example of kernel was the radial basis function, expressed by $\langle \varphi(x_i), \varphi(x_j) \rangle = \exp(-\|x_i - x_j\|^2 / 2\sigma^2), \sigma \in \mathbb{R}_+$, which was adopted in this study.^{20,21} The parameters C and σ were performed as $C = 100$ and $\sigma = 0.25$, respectively.

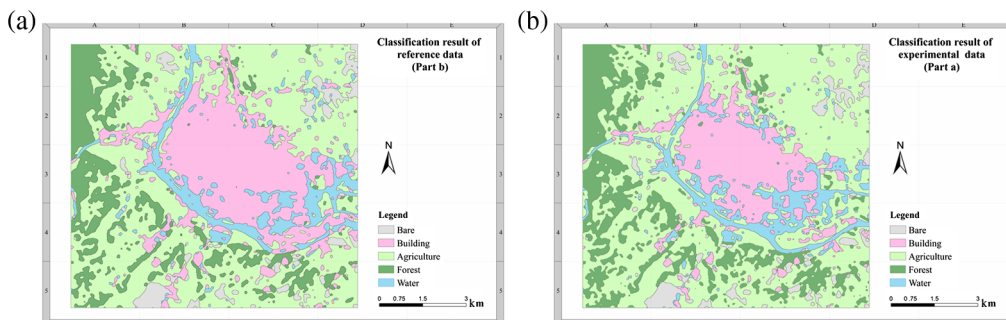


Fig. 2 Classification result of two images: (a) classification result of experimental region and (b) classification result of the reference data.

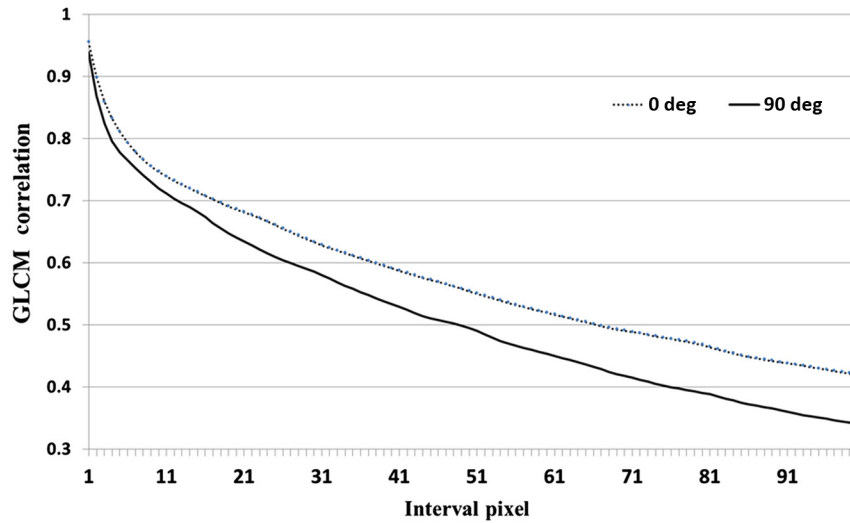


Fig. 3 GLCM-correlation parameters with different interval pixels.

3.2 Accuracy Assessment Model

3.2.1 Sample size calculation

The pixels of the studied remote sensing image (N) were 401,888 totally. The sampling rate was the proportion covered by the sample size (n) in the total size of this image data (N). Calculated by Eq. (2), the quantitative relationship of study region between distances (interval pixel) and GLCM correlation was shown in Fig. 3.

Taken the GLCM-correlation parameter $r = 0.9, 0.85, 0.8, 0.75, 0.7, 0.65, 0.6, 0.55,$ and 0.5 as example, the values of the optimal number of interval pixels and the optimal distance in both 90- and 0-deg orientations were shown in Table 1.

Based on Table 1 and Fig. 3, we knew that GLCM-correlation parameters were negatively related with the number of interval pixels. If the number of interval pixels became large enough, the GLCM-correlation parameter would be close to 0. GLCM-correlation parameters had a different gradient in different orientations. In this study, the gradient was sharper at 90-deg

Table 1 Sample size with different correlations.

GLCM correlation	Optimal number of interval pixels		Optimal distance D (m)		Sample size (pixel)	Sample rate (%)
	0 deg	90 deg	0 deg	90 deg		
0.9	2	1	30	15	203,072	50
0.85	3	2	45	30	67,792	16.7
0.8	6	4	90	60	16,872	4.15
0.75	10	7	150	105	5829	1.44
0.7	17	12	255	180	2379	0.59
0.65	27	19	405	285	800	0.20
0.6	38	27	570	405	414	0.10
0.55	51	37	765	555	208	0.05
0.5	66	49	990	735	120	0.03

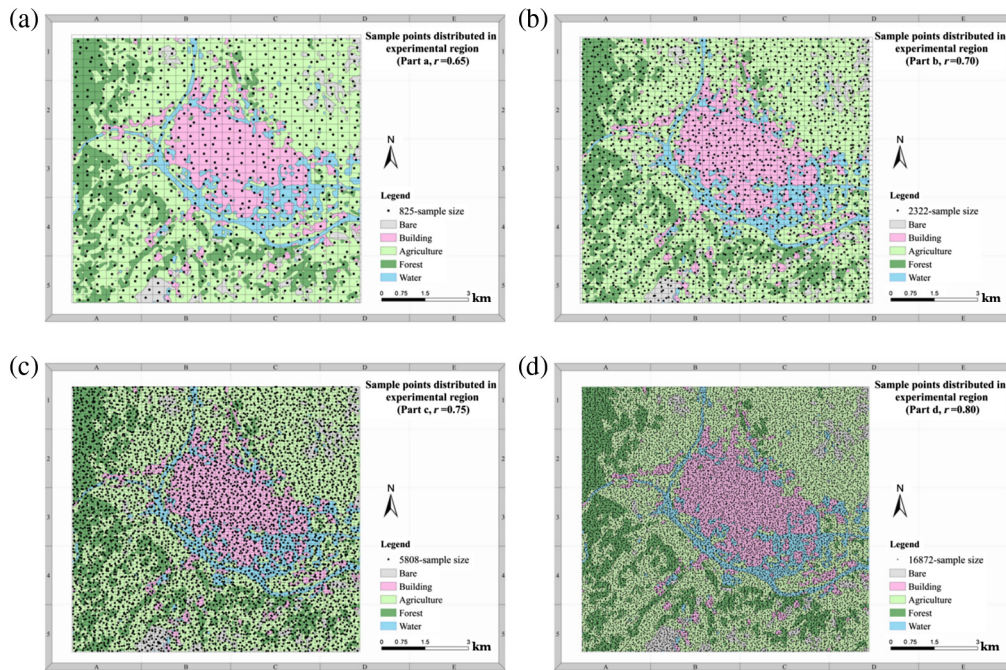


Fig. 4 Sample points located in study region with different correlation parameters: (a) correlation parameter $r = 0.6$ and the sample size $n = 825$, (b) correlation parameter $r = 0.70$ and the sample size $n = 2322$, (c) correlation parameter $r = 0.75$ and the sample size $n = 5808$, and (d) correlation parameter $r = 0.80$ and the sample size $n = 16872$.

orientation than that at 0-deg orientation. If the GLCM-correlation parameters had a large value, larger sample size should be selected for the accuracy assessment of land cover.

3.2.2 Sample points distribution

The distribution of sample points affected the assessment precision. In this study, the principle of sample points selection was uniformity and heterogeneity. Based on the optimal distance (D) in Table 1, the experimental region was divided into n rectangles and one's area was $D \times D$. One sample point was then selected in each rectangle region. Thus, n sample points were selected. Taken GLCM-correlation parameters $r = 0.85, 0.8, 0.75, 0.7$, and 0.65 as example, sample points located in the region were shown in Fig. 4.

Based on Fig. 4 and Table 2, we concluded that: (1) the sample points are uniformly distributed in the studied region, which were not associated with sample size (Fig. 3) and (2) the sample points are uniformly distributed in different types of land cover, which were consistent with the area of different types of land cover (Table 2). Thus, the result showed that the proposed model could ensure that the sample points are uniformly distributed in the spatial region and different types of land cover, which were unrelated with the definition of GLCM-correlation parameter and the size of land-cover area.

3.2.3 Accuracy analysis of classification result of remote sensing image

In this study, we took the land-cover classified from high-resolution image as reference data. We then selected the points located at the same positions from the high-resolution image and studied image, respectively. If the type of land cover from the two different images was consistent, the variable was assigned as 1. Otherwise, the variable was assigned as 0. The confusion matrix of accuracy assessment was shown in Table 3 (GLCM-correlation parameters $r = 0.85$). Overall accuracy, kappa coefficient, and other assessment parameters could be obtained from the above-mentioned confusion matrix.

Table 2 Sample points distributed in different types of land cover.

	Land cover classification													
	Water			Agriculture			Forest			Bare			Building	
	Sample size	Sample rate		Sample size	Sample rate		Sample size	Sample rate		Sample size	Sample rate		Sample size	Sample rate
GLCM correlation	$R = 0.65$	$n = 76$	9.21	$n = 431$	52.24	$n = 145$	17.58	$n = 45$	5.45	$n = 128$	15.52			
	$R = 0.70$	$n = 204$	8.37	$n = 1215$	52.33	$n = 399$	17.18	$n = 106$	4.57	$n = 398$	17.14			
	$R = 0.75$	$n = 500$	8.61	$n = 3114$	53.26	$n = 991$	17.06	$n = 269$	4.63	$n = 934$	16.08			
	$R = 0.80$	$n = 1517$	8.99	$n = 8935$	52.96	$n = 2770$	16.42	$n = 835$	4.95	$n = 2815$	16.68			
Area rate (%)		8.91		52.52		17.01		4.91						16.65

Table 3 The confusion matrix for accuracy assessment.

Classified	Reference data					Total
	Water	Forest	Agriculture	Bare	Building	
Water	1002	1	5	1	508	1517
Forest	2	2256	478	0	34	2770
Agriculture	433	21	7547	43	891	8935
Bare	9	0	48	743	35	835
Building	1	0	0	0	2814	2815
Total	1447	2278	8078	787	4282	16,872

Table 4 Comparison of the overall accuracy and kappa coefficient.

Accuracy value	Different GLCM-correlation parameters				True value
	$R = 0.65$	$R = 0.7$	$R = 0.75$	$R = 0.8$	
Overall accuracy	0.8497	0.8570	0.8569	0.8524	0.8540
Kappa coefficient	0.7799	0.7879	0.7807	0.7859	0.7837

The accuracy parameters obtained from total assessment (401,888 pixels) were taken as the true value. The accuracy parameters obtained from our model were taken as the assessment values (Tables 4 and 5).

The rate of deviation (r) was calculated by Eq. (8). Figure 5 showed the rate of deviation comparison of each GLCM-correlation parameters

$$r = \left| \frac{\tilde{P}}{P} - 1 \right| * 100\%, \quad (10)$$

where r was the rate of deviation, \tilde{P} denoted the accuracy value of each GLCM-correlation parameters, which was overall parameter or kappa coefficient, and P was the overall parameter or kappa coefficient of true value.

Based on Tables 4 and 5 and Fig. 5, we knew that the overall accuracy and kappa coefficient derived from our model were very close to the true value. The greatest rate of deviation was only 0.54%. As the GLCM-correlation parameter increased, the rate of deviation of overall parameter and kappa coefficient decreased. Thus, the assessment accuracy of our proposed model was close to the accuracy of total assessment.

3.3 Comparison Results of Different Assessment Models

In this section, we used three different assessment models to conduct accuracy assessment for the classification result of the above-mentioned remote sensing image, including percent sampling model, random sampling model, and our proposed model.

3.3.1 Compared with percent sampling model

Taking 2% as the sampling rate, the percent sampling was used to assess the accuracy of land cover. As shown in Fig. 6, we knew that the percent sampling model had a fixed sampling rate. The autocorrelation among different pixels was ignored in the remote sensing image. Thus, it was different to define the sampling rate for percent sampling model. However, our model could

Table 5 Comparison of different accuracy parameters.

Classification result	R = 0.65			R = 0.70			R = 0.75			R = 0.8			Total assessment								
	P	U	O	P	U	O	P	U	O	P	U	O	P	U	O						
Water	72.60	69.74	30.26	27.40	70.37	70.37	34.80	29.63	69.09	66.60	33.40	30.91	69.25	66.05	33.95	30.75	70.63	66.69	33.31	29.37	
Agriculture	94.39	85.85	14.15	5.61	92.68	92.68	84.44	15.56	7.32	94.21	84.71	15.29	5.79	93.43	84.47	15.53	6.57	93.60	84.47	15.53	6.40
Forest	98.41	85.52	14.48	1.59	99.09	99.09	81.70	18.30	0.91	98.92	83.15	16.85	1.08	99.03	81.44	18.56	0.97	98.93	82.76	17.24	1.07
Bare	90.70	86.67	13.33	9.30	92.08	92.08	87.74	12.26	7.92	94.88	89.59	10.41	5.12	94.41	88.98	11.02	5.59	95.53	89.04	10.96	6.47
Building	67.02	100.00	0	32.98	66.78	66.78	100.00	0.00	33.22	64.84	99.89	0.11	35.16	65.72	99.96	0.04	34.28	65.82	99.95	0.05	34.18

Note: P, producer accuracy; U, user accuracy; C, commission; and O, omission.

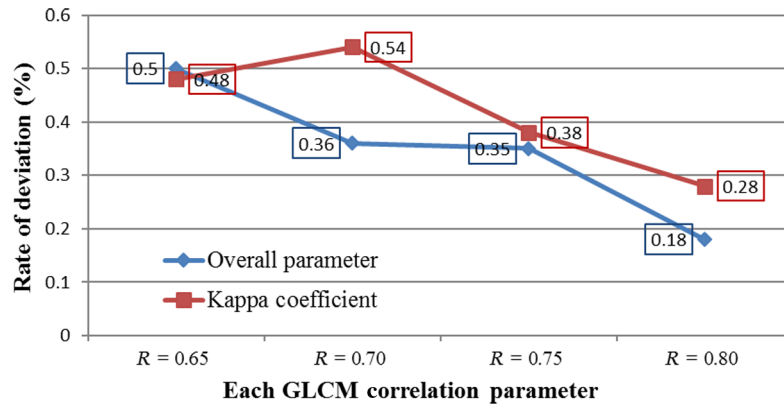


Fig. 5 Rate of deviation comparison of each GLCM-correlation parameters.

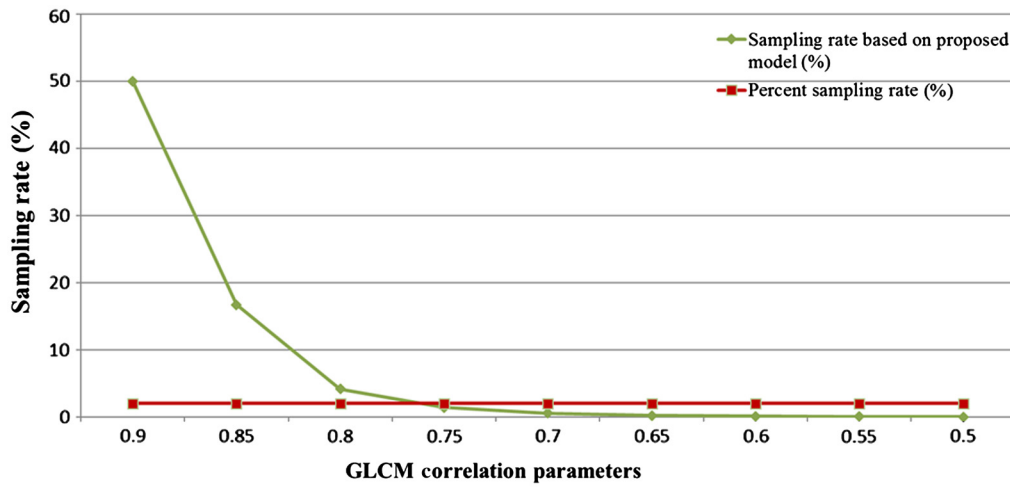


Fig. 6 Comparison sampling rate with each GLCM-correlation parameter and percent sampling.

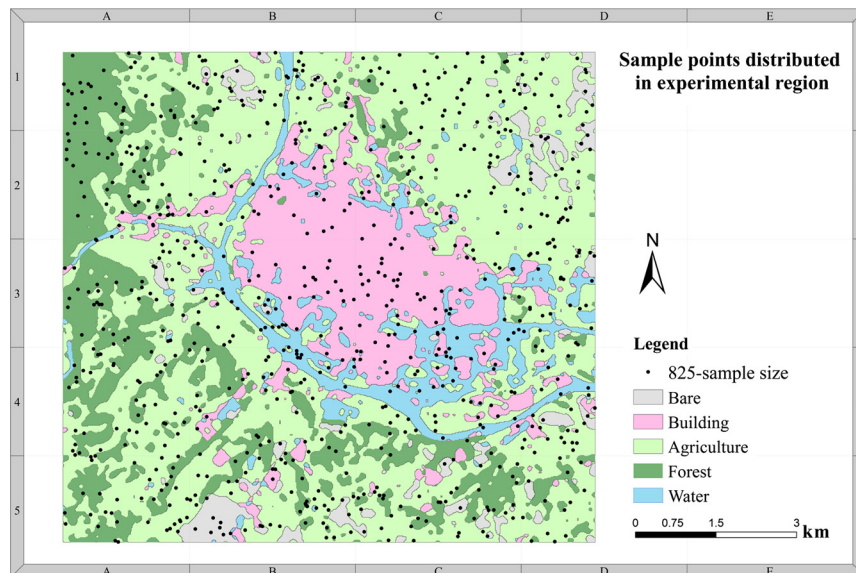


Fig. 7 Sampled pixels divided in the region by random sampling ($n = 825$).

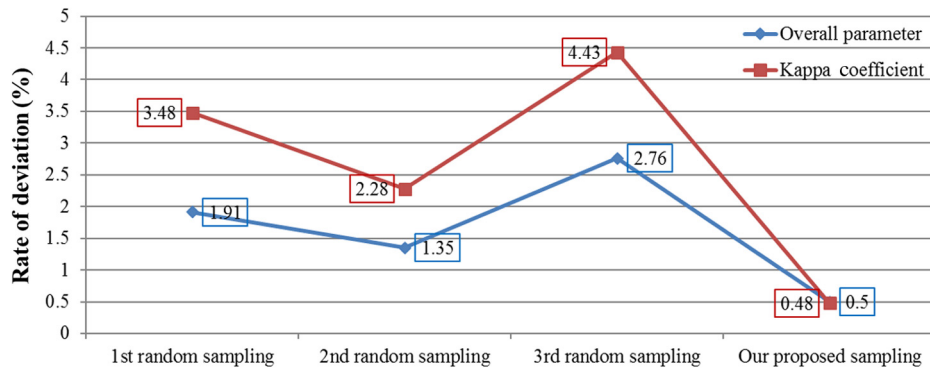


Fig. 8 Rate of deviation comparison of each sampling methods.

quantify the relationship between GLCM-correlation parameter and sampling rate. Thus, the producers and users could easily determine the sampling rate according to the spatial autocorrelation and heterogeneity.

3.3.2 Compared with random sampling model

Given the sampling size of 825, the sample points were randomly selected in the region three times. Figure 7 showed the result of the random sample sampling at one time. Figure 8 showed the rate of deviations for random sampling model and our proposed model. We knew that the result of accuracy assessment for random sampling model was not consistent. Sample distribution is an important determinant in accuracy assessment. If sample distribution was considered, it would lead to sample choice preference and could not provide an objective result. As shown Table 4, the sample rate was consistent across different experiments in our model. Moreover, the sample rate deviation of our model was less than that of random sampling model.

4 Discussions

The classification accuracy of the remote sensing image is very necessary before the application for scientific investigation and policy decision. In this study, we proposed an accuracy assessment model based on spatial sampling. This model considered both sample size calculation and sample points distribution during the accuracy assessment. Compared with percent sampling model, the proposed model could quantify the relationship between GLCM-correlation parameter and sample size. Compared with random sampling model, the proposed model ensured that the sample points are uniformly distributed in the spatial region and proportionally distributed in different types of land cover. Overall, our model is suitable for the accuracy assessment of the classification result of the remote sensing image.

During the classification accuracy assessment of the remote sensing image, our model could not only consider sample size calculation but also consider sample points distribution. As for sample size calculation, we used the GLCM to quantify the relationship between spatial autocorrelation and sample size. This matrix could provide useful information about the spatial relationships of pixels in an image. Compared with percent sampling, which has a fixed sampling rate, our model could allow the producers and users to determine the sampling rate according to the spatial autocorrelation and heterogeneity. As for sample point distribution, our method considered both the uniformity and heterogeneity of sample points distribution. It ensures that the sample points are uniformly distributed in the spatial region and proportionally distributed in different types of land cover. Compared with random sampling model, our model has great advantage on accuracy consistence and sample rate deviation.

However, there are some limitations for our proposed model. We only calculated the GLCM-correlation parameter (r) at two different orientations, including 0 deg and 90 deg. More directions should be considered in future study.

5 Conclusions

In this study, we proposed an accuracy assessment model for remote sensing classification result based on spatial sampling. This model calculates the sample size required for accuracy assessment, determines the sample points distributed in a region, and analyzes the result of accuracy assessment. This model considers both sampling size calculation and sample points distribution during the classification accuracy assessment. Our model could allow producer and user to easily determine sample size. Moreover, our model ensures that the sample points are uniformly distributed in the spatial region and proportionally distributed in different types of land cover. Thus, our proposed model is a suitable model for the accuracy assessment of the classification result of the remote sensing image.

Acknowledgments

This work was generously supported by the grants from the National Natural Science Foundation of China (Grant No. 41671431 to H.D.M. and Grant No. 41501419 to W.Z.H.) and the grants from the Capacity Development for Local College Project (Grant No. 15590501900 to H.D.M. and Grant No. 17050501900 to S.W.). The authors declare no conflicts of interest.

References

1. A. L. Buczak et al., "A data-driven epidemiological prediction method for dengue outbreaks using local and remote sensing data," *BMC Med. Inf. Decis. Making* **12**, 124 (2012).
2. M. Padilla et al., "Comparing the accuracies of remote sensing global burned area products using stratified random sampling and estimation," *Remote Sens. Environ.* **160**, 114–121 (2015).
3. H. Nagendra et al., "Remote sensing for conservation monitoring: assessing protected areas, habitat extent, habitat condition, species diversity, and threats," *Ecol. Indic.* **33**, 45–59 (2013).
4. D. D. Alexakis, D. G. Hadjimitsis, and A. Agapiou, "Integrated use of remote sensing, GIS and precipitation data for the assessment of soil erosion rate in the catchment area of 'Yialias' in Cyprus," *Atmos. Res.* **131**, 108–124 (2013).
5. M. Hussain et al., "Change detection from remotely sensed images: from pixel-based to object-based approaches," *ISPRS J. Photogramm. Remote Sens.* **80**, 91–106 (2013).
6. S. V. Stehman et al., "A global land cover validation data set, II: augmenting a stratified sampling design to estimate accuracy by region and land cover class," *Int. J. Remote Sens.* **33**, 6975–6993 (2012).
7. W. T. Bean, R. Stafford, and J. S. Brashares, "The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models," *Ecography* **35**, 250–258 (2012).
8. J. Wang et al., "Spatial sampling design for monitoring the area of cultivated land," *Int. J. Remote Sens.* **23**, 263–284 (2002).
9. S. V. Stehman and J. D. Wickham, "Pixels, blocks of pixels, and polygons: choosing a spatial unit for thematic accuracy assessment," *Remote Sens. Environ.* **115**, 3044–3055 (2011).
10. S. V. Stehman and R. L. Czaplewski, "Design and analysis for thematic map accuracy assessment: fundamental principles," *Remote Sens. Environ.* **64**, 331–344 (1998).
11. A. Comber et al., "Spatial analysis of remote sensing image classification accuracy," *Remote Sens. Environ.* **127**, 237–246 (2012).
12. J. D. Wickham et al., "Accuracy assessment of NLCD 2006 land cover and impervious surface," *Remote Sens. Environ.* **130**, 294–304 (2013).
13. W. R. Tobler, "A computer movie simulating urban growth in the Detroit region," *Econ. Geogr.* **46**, 234–240 (1970).
14. S. S. Rwanga and J. M. Ndambuki, "Accuracy assessment of land use/land cover classification using remote sensing and GIS," *Int. J. Geosci.* **08**(4), 611–622 (2017).

15. D. Marmanis et al. "Deep learning earth observation classification using imagenet pretrained networks," *IEEE Geosci. Remote Sens. Lett.* **13**(1), 105–109 (2016).
16. A. Nawaz, Z. Iqbal, and S. Ullah, "Performance analysis of supervised image classification techniques for the classification of multispectral satellite imagery," in *Fourth Int. Conf. on Aerospace Science and Engineering*, IEEE, pp. 1–5 (2016).
17. A. Bastarrika, E. Chuvieco, and M. P. Martín, "Mapping burned areas from Landsat TM/ETM+ data with a two-phase algorithm: balancing omission and commission errors," *Remote Sens. Environ.* **115**(4), 1003–1012 (2011).
18. N. Koutsias et al. "A rule-based semi-automatic method to map burned areas: exploring the USGS historical Landsat archives to reconstruct recent fire history," *Int. J. Remote Sens.* **34**(20), 7049–7068 (2013).
19. B. N. Soomro et al. "Bilayer elastic net regression model for supervised spectral–spatial hyperspectral image classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **9**(9), 4102–4116 (2016).
20. E. Zhu, W. Song, and J. Dai, "Road extraction of high-resolution remote sensing images based on improved SVM," *Sci. Surv. Mapp.* **12**, 044 (2016).
21. P. Liu et al., "SVM or deep learning? A comparative study on remote sensing image classification," *Soft Comput.* **21**(23), 7053–7065 (2017).

Dongmei Huang is a professor, PhD supervisor, and the leading person of Digital Ocean of Information College, Shanghai Ocean University. She obtained the 2016 Annual Second Class Prizes of the Shanghai Scientific and Technological Progress Award. Her current research interests include ocean GIS, RS, and big data analysis. She is a member of the China Computer Federation.

Zhenhua Wang is an associate professor, master supervisor, and the deputy dean of Digital Ocean of Information College, Shanghai Ocean University. Her research mainly focuses on big data processing and analysis.

Biographies for the other authors are not available.