Chin. Geogra. Sci. 2017 Vol. 27 No. 5 pp. 827–835
doi: 10.1007/s11769-017-0894-6

Springer Science Press
www.springerlink.com/content/1002-0063

Effects of RapidEye Imagery's Red-edge Band and Vegetation Indices on Land Cover Classification in an Arid Region

LI Xianju[1], CHEN Gang[2], LIU Jingyi[3], CHEN Weitao[1], CHENG Xinwen[4], LIAO Yiwei[5]

(1. *College of Computer Science and Hubei Key Laboratory of Intelligent Geo-Information Processing*, *China University of Geosciences*, *Wuhan* 430074, *China*; 2. *College of Marine Science and Technology*, *China University of Geosciences*, *Wuhan* 430074, *China*; 3. *Laboratory of Geographic Information and Spatial Analysis*, *Department of Geography and Planning*, *Queen's University*, *Kingston ON K7L3N6*, *Canada*; 4. *Faculty of Information Engineering*, *China University of Geosciences*, *Wuhan* 430074, *China*; 5. *School of Electrical and Electronic Engineering*, *Huazhong University of Science and Technology*, *Wuhan* 430074, *China*)

Abstract: Land cover classification (LCC) in arid regions is of great significance to the assessment, prediction, and management of land desertification. Some studies have shown that the red-edge band of RapidEye images was effective for vegetation identification and could improve LCC accuracy. However, there has been no investigation of the effects of RapidEye images' red-edge band and vegetation indices on LCC in arid regions where there are spectrally similar land covers mixed with very high or low vegetation coverage information and bare land. This study focused on a typical inland arid desert region located in Dunhuang Basin of northwestern China. First, five feature sets including or excluding the red-edge band and vegetation indices were constructed. Then, a land cover classification system involving plant communities was developed. Finally, random forest algorithm-based models with different feature sets were utilized for LCC. The conclusions drawn were as follows: 1) the red-edge band showed slight contribution to LCC accuracy; 2) vegetation indices had a significant positive effect on LCC; 3) simultaneous addition of the red-edge band and vegetation indices achieved a significant overall accuracy improvement (3.46% from 86.67%). In general, vegetation indices had larger effect than the red-edge band, and simultaneous addition of them significantly increased the accuracy of LCC in arid regions.

Keywords: arid region; land cover classification; RapidEye; red-edge band; vegetation indices; random forest; Dunhuang Basin

Citation: Li Xianju, Chen Gang, Liu Jingyi, Chen Weitao, Cheng Xinwen, Liao Yiwei, 2017. Effects of RapidEye imagery's red-edge band and vegetation indices on land cover classification in an arid region. *Chinese Geographical Science*, 27(5): 827–835. doi: 10.1007/s11769-017-0894-6

1 Introduction

Land desertification in arid regions is one of the most serious environmental and socio-economic problems in the world, and this issue has been receiving continuous attention from governments and scientists (Batterbury and Warren, 2001; Han *et al*., 2015; Liu *et al*., 2015). Climate variability and human activities, such as inappropriate and unsustainable land use, are the main driving factors of desertification (Wang *et al*., 2006; Zhang *et al*., 2010; Wu *et al*., 2011; Ge *et al*., 2013). Therefore, land cover classification (LCC) in arid regions is of great significance to the assessment, prediction, and management of land desertification.

Satellite remote sensing technology can be used to obtain accurate up-to-date surface information over a large area, and this technology has been widely used in studies of LCC, vegetation classification and mapping at the community or species levels (Langley *et al*., 2001; Nordberg and Everson, 2003; Cerna and Chytry, 2005; Xie *et al*., 2008; Li and Shao, 2014). Moreover, some studies have focused on LCC in arid regions (e.g., Ste-

Received date: 2016-05-13; accepted date: 2016-09-08
Foundation item: Under the auspices of Fundamental Research Funds for Central Universities, China University of Geosciences (Wuhan) (No. CUGL150417), National Natural Science Foundation of China (No. 41274036, 41301026)
Corresponding author: CHEN Gang. E-mail: ddwhcg@cug.edu.cn

fanov *et al*., 2001; Alrababah and Alhamad, 2006; Rozenstein and Karnieli, 2011; Galletti and Myint, 2014; Namdar *et al*., 2014). However, there are three problems in arid regions, which increase the difficulty of LCC: 1) the mixed-pixel problem; 2) the confusion of spectral signatures among bare soils, sparse density shrub lands, and impervious surface areas (Namdar *et al*., 2014); and 3) the vegetation coverage is either extremely high or very low (Asner and Heidebrecht, 2003). As a result, more effective land cover classification systems, which may weaken the effect of the mixed-pixel problem, must be developed based on the actual conditions in specific study areas to improve the accuracy of classification (Stefanov *et al*., 2001; Alrababah and Alhamad, 2006; Galletti and Myint, 2014; Namdar *et al*., 2014). Additionally, for other two problems more use should be made of available features, especially for vegetation identification. RapidEye satellite imagery's red-edge band can be used to identify the type and growth state of vegetation (e.g., Schuster *et al*., 2012; Tigges *et al*., 2013; Adelabu *et al*., 2014; Kim and Yeow, 2014), which may be positive to the LCC in arid regions. Especially, some of these studies have assessed the effect of the red-edge band on some special classifications. For example, Schuster *et al*. (2012) first examined the red-edge band from RapidEye imagery for improving land-use classification of a nature conservation heathland area. Adelabu *et al*. (2014) evaluated the impact of the red-edge band for classifying insect defoliation levels. Kim and Yeom (2014) evaluated the effects of the red-edge band and texture features on the classification of paddy rice crops at different growth stages based on an object-oriented method. All the three studies revealed that the red-edge band contributed to the classification. However, it has not been examined whether the RapidEye imagery's red-edge band could improve the accuracy of LCC in arid regions. Moreover, the effect of vegetation indices and the joint effect of the red-edge band and vegetation indices have not yet been investigated.

The present study focused on a typical inland arid region located in the Dunhuang Basin of northwestern China, which has suffered from increasingly serious desertification in recent years (Han *et al*., 2015). First, RapidEye satellite imagery was processed, and the normalized difference vegetation index (NDVI) and its red-edge adaptation (NDVI-RE) were extracted. Then a land cover classification system involving plant communities was developed. Random forest (RF) algorithm based models with different feature sets including or excluding the red-edge band and vegetation indices were utilized for LCC and to evaluate the effects of the red-edge band and vegetation indices on the classification accuracy.

2 Materials and Methods

2.1 Study area and data sources

The study area is located in the discharge zone of the Xihu National Nature Reserve of Dunhuang City, Gansu Province, China (Chen *et al*., 2014b). It has an area of 4.83 km^2, and the latitudes and longitudes are in the range of 40°07′32″N–40°08′34″N and 93°13′18″E–93°15′04″E, respectively (Fig. 1) (Li *et al*., 2017). The Xihu National Nature Reserve experiences a typical continental climate and is in the warm desert climate zone. The perennial average annual precipitation here is only 39.90 mm, and there is high seasonal variability in rainfall. The study area is a typical inland arid region, characterized by groundwater dependent ecosystems (Hatton and Evans, 1998; Chen *et al*., 2016). The groundwater here mainly comes from the runoff area of the alluvial fan of the Dang River and the alimentation area of the fan-roof area of the Dang River alluvial fan (Chen *et al*., 2014b). The vegetation coverage is very high in the spring-overflow area where there are mainly two kinds of natural vegetation, i.e., *Tamarix* and *Phragmites*. The vegetation is sparse with extremely low coverage in the areas with high groundwater depth (e.g., bare rock land, sandy desertified land, and wind-eroded land). In areas with low groundwater depth, the soil salinization is severe and thus vegetation coverage is low.

In this study, RapidEye 1B level satellite imagery taken on August 18, 2010, was first obtained. Then, a geometric correction was made by using topographic maps on a scale of 1∶100 000 and a second order polynomial transformation with a root-mean-square error of less than 1 pixel. The imagery was then resampled to a spatial resolution of 5-m using a cubic convolution method (Table 1). The reflectance was not used and thus the atmosphere correction was not conducted (Li *et al*., 2016). Finally, the study area was extracted from the processed imagery.

2.2 Methods

2.2.1 Analyses of feature sets

NDVI and its red-edge adaptation NDVI-RE (Schuster

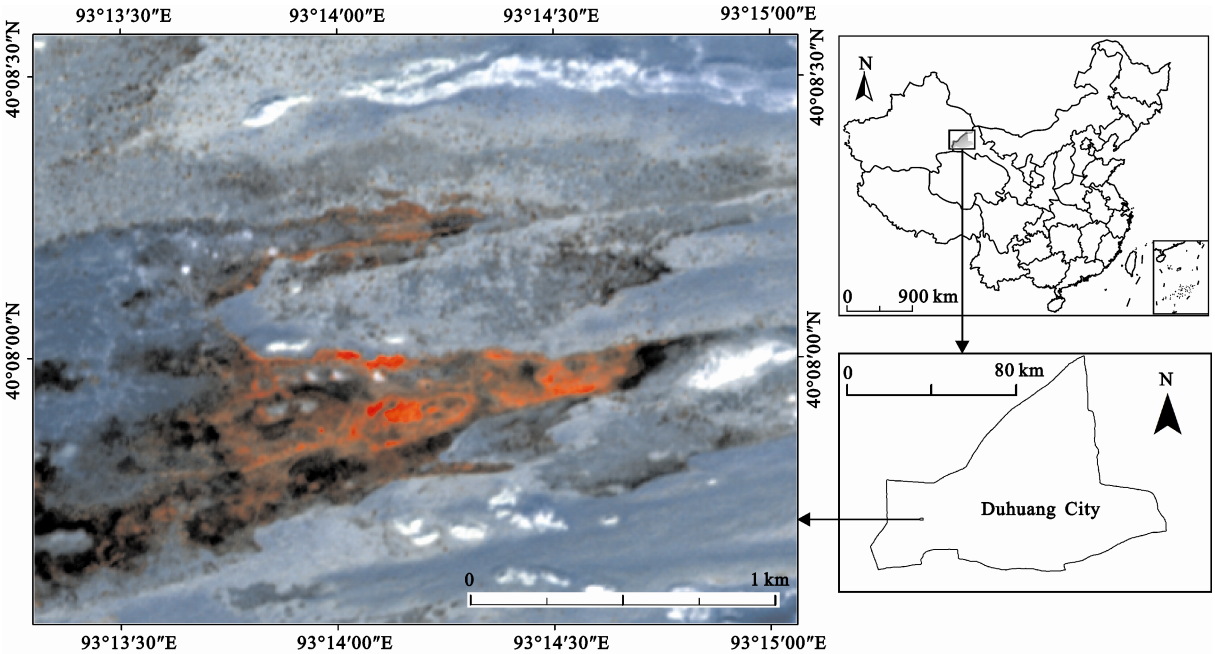

Fig. 1 Geographical location of study area and RapidEye image used in this study (RGB bands are near-infrared, red-edge, and green, respectively) (Li *et al*., 2017)

Table 1 Characteristics of RapidEye imagery

Sensor and data characteristics	RapidEye imagery
Collection date	August 18, 2010
Spectral bands	Blue (440–510 nm)
	Green (520–590 nm)
	Red (630–685 nm)
	Red-edge (690–730 nm)
	NIR (760–850 nm)
Ground sampling distance	6.5 m at nadir
Spatial resolution	5 m
Radiometric resolution	12-bit
Revisit time	5.5 days at nadir

Note: NIR: near-infrared

et al., 2012) were extracted based on the red, red-edge, and near-infrared (NIR) bands of the RapidEye imagery. NDVI-RE was calculated by replacing the NIR band with the red-edge band, i.e., NDVI-RE = (red-edge – red) / (red-edge + red). Based on RapidEye's spectral bands and the above-mentioned two vegetation indices, five feature sets (FSs A, B, C, D, and E) with different combinations of seven features were developed, and the following analyses were performed to evaluate the effects of the red-edge band and vegetation indices on LCC:

(1) Analysis 1: four bands (A) (excluding the red-edge) vs. all bands (B). This is used to evaluate the effect of the red-edge band on LCC.

(2) Analysis 2: all bands (B) vs. all bands + vegetation indices (E). This examines the effect of vegetation indices.

(3) Analysis 3: four bands (A) (excluding the red-edge) vs. all bands + vegetation indices (E). This checks the joint effect of the red-edge band and vegetation indices.

(4) Analysis 4: all bands + NDVI (C) vs. all bands + NDVI-RE (D). This enables a comparison of the relative importance of NDVI and NDVI-RE.

2.2.2 Determination of land cover types

The study area is mainly covered by bare land, sandy desertified land, wind-eroded land, saline and alkaline land, and vegetation (Qin *et al*., 2014). The bare land is cyan in the imagery (Fig. 1) with some gray dead shrubs; the area includes the Gobi Desert, which contains large amounts of exposed bare rock, sand and stone particles, and which is a rain shadow desert because the Himalaya range blocks the arrival of rain-carrying clouds. The sandy desertified land is grayish-white in the imagery (Fig. 1) and is comprised of sand dunes. The wind-eroded land is bright white in the imagery (Fig. 1) with little vegetation, and is created by forces exerted by the wind, which detaches soil particles from the land surface. The saline and alkaline land is comprised of clay soils with poor soil structure and low

infiltration capacity; it is blue-green in the imagery (Fig. 1) with some white crystallization spots of salts and alkalis. Based on a field sampling survey of vegetation in May and June of 2011, it was discovered that there are mainly four kinds of natural vegetation in the study area, i.e., *Tamarix*, *Populus*, *Phragmites*, and *Alhagi*. The natural vegetation is mixed and forms different plant community types that are named according to the dominant species (Chen *et al*., 2014a). Therefore, the land cover classification system in this study was defined as follows: bare land, sandy desertified land, wind-eroded land, saline and alkaline land, *Tamarix-Phragmites* community, and *Populus-Tamarix-Alhagi* community. Although there were only two types of vegetation land cover, each of the other four contained weak vegetation information, e.g. the sandy desertified land and saline and alkaline land showed some isolated and dotted thickets. As a result, the red-edge band and vegetation indices that have been shown to be useful for vegetation identification and mapping could potentially improve the LCC accuracy in this study.

2.2.3 Land cover classification and accuracy evaluation

The RF algorithm was used in this study for LCC, which is based on a large number of randomly grown decision trees and can avoid over-fitting with high classification accuracy (Breiman, 2001). Belgiu and Drăguț (2016) offer a clear explanation of the principles of the RF algorithm. It was realized in the R software platform (R Development Core Team, 2015) using the randomForest package (Liaw and Wiener, 2002). The RF algorithm employs the following two main parameters: 1) *ntree*, which refers to the number of trees used in the model, set to the default value of 500; and 2) *mtry*, which refers to the number of randomly selected variables at each split in the model-building process. The *mtry* parameter requires optimization. The parameter optimization of RF was based on the e1071 package (Meyer *et al*., 2015).

The training data polygons were obtained based on visual interpretation of RapidEye imagery and a field survey in May and June of 2011. A stratified random sampling method was used to choose 400 pixels from each class as a training set. The accuracy evaluation was based on a test set that was independent of the training data. It was obtained using two steps: 1) the training data were first erased from the classification result derived from the RF-based model with all the seven features; 2) a stratified random sampling method was then used to select 100 pixels in each class as test samples. Then, visual identification was used to determine their land cover types. The accuracy evaluations for RF-based models with all the five feature sets were all based on this test set. The classification results were evaluated using the F1-measure (Daskalaki *et al*. 2006), overall accuracy (OA) (Liu *et al*., 2007), and percentage deviation (Schuster *et al*., 2012). The McNemar test employed in this study is a statistical test that can be used to evaluate the performance of the classifiers (Manandhar *et al*., 2009; Petropoulos *et al*., 2012). The test is based on the error matrices of two classifications, and the chi-square statistic value (χ^2) is calculated as follows:

$$\chi^2 = \frac{(f_{12} - f_{21})^2}{f_{12} + f_{21}} \tag{1}$$

where f_{ij} represents the number of samples that classifier i misclassifies but classifier j correctly classifies (i = 1, 2; j = 1, 2). The difference between two classifications is statistically significant at the 95% confidence level (P = 0.05) when the χ^2 value is larger than or equal to 3.84. Specifically, the McNemar test was used to evaluate: 1) whether the red-edge band and vegetation indices had significant effects on LCC in arid regions; and 2) whether there were significant differences in the classification accuracy between two RF-based models with feature sets including all bands and one vegetation index.

3 Results

3.1 Correlation analyses of spectral bands and vegetation indices

Adding the red-edge band and vegetation indices meant that more information were obtained; however, more information do not always result in higher classification accuracies, which is partly owing to the correlation between features. For example, Adelabu *et al.* (2014) reported that, when using all the bands of RapidEye imagery, adding NDVI decreased the classification accuracy. In this study, the result of Pearson correlation analyses performed on the training data showed the following (Table 2): 1) the blue band, green band, red band, and red-edge band showed strong inter-correlations with

correlation coefficients larger than 0.60; 2) modest correlations existed between the NIR band and the other four spectral bands with correlation coefficients ranging from 0.25 to 0.60; 3) strong correlations dominated between vegetation indices and spectral bands; 4) the two vegetation indices were strongly correlated. In general, the features used in this study had different degrees of correlation. As a result, it was worthwhile to investigate whether adding the correlated red-edge band and vegetation indices could always improve the classification accuracies. Moreover, different land cover types showed diverging correlations, which suggested that to some degree the features were able to classify the different types of land cover.

3.2 Parameter optimization of RF algorithms and land cover classification map

For RF-based classifications, the *mtry* parameter ranged from 1 to the numbers of features in each feature set (A: 1–4, B: 1–5, C: 1–6, D: 1–6, and E: 1–7). The values of 1, 3, 1, 4, and 3 were selected for the five sets based on the RF models with the lowest errors (2.62%, 2.67%, 2.29%, 2.96%, and 2.54%, respectively).

The results of classification based on feature set E and the RF algorithm with an OA of 89.67% is shown in Fig. 2. Overall, the use of the RapidEye spectral bands and vegetation indices achieved a high degree of visual accuracy of LCC in this arid region. However, in the northeast corner, there were some commission and omission errors among sandy desertified land, saline and alkaline land, and *Populus-Tamarix-Alhagi* community. In the center of the study area, some *Tamarix-Phragmites* community were misclassified as *Populus-Tamarix-Alhagi* community.

3.3 Classification based on four bands (excluding red-edge band) vs. all bands

For analysis 1, the feature sets excluding (FS A) and

Table 2 Pearson correlation analysis for the training data

	F1	F2	F3	F4	F5	F6	F7
F1	1	0.99	0.99	0.96			
F2		1	0.99	0.98			
F3			1	0.97			
F4				1			
F5	0.25	0.32	0.27	0.48	1		
F6	–0.81	–0.78	–0.82	–0.67	0.33	1	
F7	–0.84	–0.81	–0.85	–0.71	0.27	0.99	1

Notes: F1, F2, …, F7 represent the following seven features: blue, green, red, red-edge, and NIR bands, NDVI, and NDVI-RE

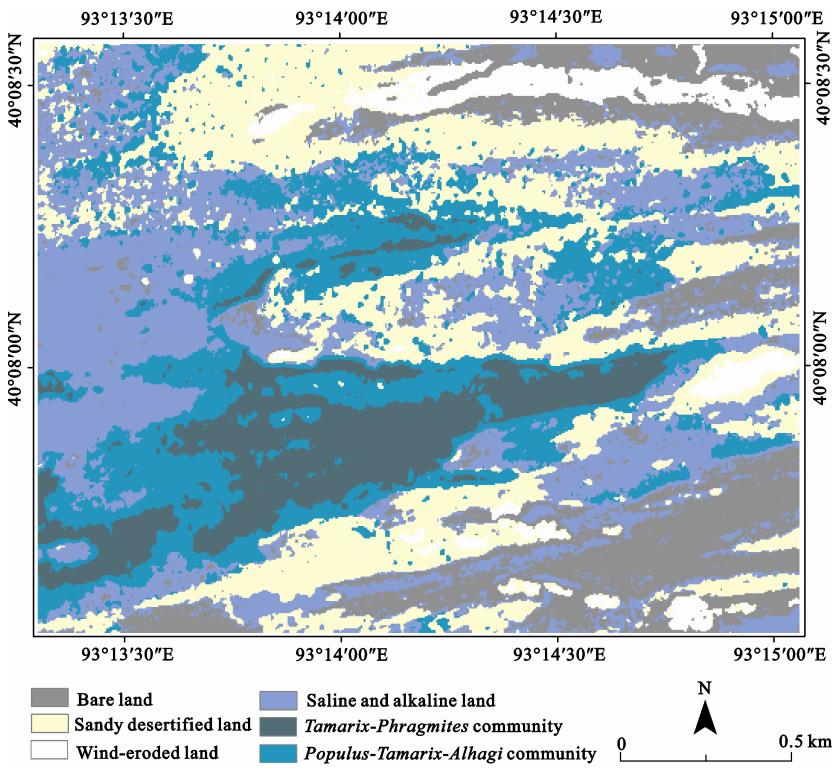

Fig. 2 Classification map of study area based on random forest algorithm and all features

including (FS B) the red-edge band were used to train the RF algorithm, and the accuracy evaluation results are shown in Table 3. The OA values based on A and B were 86.67% and 87.00%, respectively. After adding the red-edge band, OA increased by 0.38% (percentage deviation). Each class exhibited different sensitivities to adding the red-edge band. For example, the accuracies of bare land and saline and alkaline land decreased, and those of others improved. The sandy desertified land (2.24%) and saline and alkaline land (–2.40%) showed higher deviations, however the deviations for other classes were almost all less than 1.00%. The McNemar test was performed on the classification results of the test set based on feature sets A and B, and the result is shown in Table 4. It shows that there was no significant difference ($\chi^2 = 0.15$ and $P = 0.70$) between the classification results derived from RF-based models with features A and B. In general, the red-edge band could only slightly improve the LCC accuracy in the arid region studied.

3.4 Classification based on all bands vs. all bands plus vegetation indices

The F1-measure, OA, and percentage deviation for analysis 2 based on feature sets B and E (using vegetation indices as additional features) are shown in Table 3. The corresponding statistical test result is shown in Table 4. The results in Table 3 show that adding vegetation indices increased the OA by 3.07% (from 87.00% using all bands to 89.67% using all bands plus the vegetation indices). Each class also exhibited different sensitivities to adding in the vegetation indices. All classes showed positive deviations. However, half of the classes achieved large improvements, ranging from 4.46% (*Populus-Tamarix-Alhagi* community) to 8.33% (saline and alkaline land), while the other half only increased by 0.47% to 1.29%. The results in Table 4 show that there was a significant difference ($\chi^2 = 7.11$ and $P = 0.01$) between the classification results derived from RF-based models with features B and E. In general, it seems that the vegetation indices could significantly improve the accuracy of LCC in arid regions.

3.5 Classification based on four bands (excluding red-edge band) vs. all bands and vegetation indices

Results from analysis 3 indicate that the OA value increased by 3.46% when all the features (E) were used compared to when just four bands (A) (excluding red-edge) were used (Table 3). All the classes achieved accuracy improvements larger than 1.00%, though different sensitivities were observed. The saline and alkaline land showed the largest deviation (5.72%), followed by *Populus-Tamarix-Alhagi* community (5.13%), bare land (4.00%), and sandy desertified land (3.56%); only minor deviations were seen for wind-eroded land (1.36%) and *Tamarix-Phragmites* community (1.48%). The test result in Table 4 shows that there was a significant difference ($\chi^2 = 7.36$ and $P = 0.01$) between the classification results derived from RF-based models with features A and E. In general, adding the red-edge band and vegetation indices could also significantly improve the LCC accuracy in arid regions.

3.6 Classification based on all bands adding NDVI vs. NDVI-RE

Analysis 4 revealed that the addition of NDVI-RE to all bands (D) outperformed classification based on the addition of NDVI (C), with an improvement of 0.37%

Table 3 Results of accuracy evaluation for analyses 1–4

Land cover	F1-measure					Percentage deviation (%)			
	A	B	C	D	E	Analysis 1	Analysis 2	Analysis 3	Analysis 4
Bare land	84.82	84.38	87.10	85.71	88.21	–0.52	4.54	4.00	–1.60
Sandy desertified land	87.04	88.99	89.40	88.99	90.14	2.24	1.29	3.56	–0.46
Wind-eroded land	93.26	93.33	93.94	93.81	94.53	0.08	1.29	1.36	–0.14
Saline and alkaline land	77.39	75.53	79.81	80.41	81.82	–2.40	8.33	5.72	0.75
Tamarix-Phragmites community	94.29	95.24	94.23	94.74	95.69	1.01	0.47	1.48	0.54
Populus-Tamarix-Alhagi community	82.72	83.25	84.15	86.77	86.96	0.64	4.46	5.13	3.11
Overall accuracy (%)	86.67	87.00	88.17	88.50	89.67	0.38	3.07	3.46	0.37

Notes: Analysis 1: feature sets A vs. B. Analysis 2: B vs. E. Analysis 3: A vs. E. Analysis 4: C vs. D. Feature set A: four bands (excluding red-edge). B: all bands. C: all bands + NDVI. D: all bands + NDVI-RE. E: all bands + NDVI + NDVI-RE. NDVI: normalized difference vegetation index. NDVI-RE: the red-edge adaptation of NDVI

(Table 3). The OA values derived using feature sets C and D were 88.17% and 88.50%, respectively. Bare land, sandy desertified land, and wind-eroded land showed negative deviations and the others positive. Only two deviations were larger than 1.00%, i.e., bare land (–1.60%) and *Populus-Tamarix-Alhagi* community (3.11%). The χ^2 and P values for the comparison of classifications using feature sets C and D were 0.08 and 0.77, respectively (Table 4). This means that there was no significant difference between feature sets C- and D-based classifications. In general, NDVI-RE only slightly outperformed NDVI.

4 Discussion

In this study, the red-edge band could only slightly improve the LCC accuracy in the arid region studied. Similarly, Kim and Yeom (2014) reported that the red-edge band could slightly improve the classification accuracy of paddy rice crops with no statistically significant difference, as examined by the Z-test when using a single season image. Conversely, Adelabu *et al*. (2014) showed that the red-edge band could significantly increase the classification accuracy of insect defoliation levels according to the McNemar test. Moreover, Schuster *et al*. (2012) showed that adding the red-edge band increased the average OA by 2.70% (from 72.80%), but statistical test was not performed in this study. In general, the red-edge band could always improve the accuracies of various classifications, and whether there was statistically significant depended on the specific applications. Compared to Adelabu *et al*. (2014), the other two studies (Schuster *et al*., 2012; Kim and Yeom, 2014) and this study were more easily classified. Especially, in this study the land cover classification system involving plant communities, which may weaken the effect of the mixed-pixel problem, contributed to the LCC. As a result, a similar conclusion with Schuster *et al*. (2012) and Li *et al*. (2016) could be drawn that the applications that were hard to classify were sensitive to the addition of the red-edge band, i.e., the red-edge band was prone to significantly increase the accuracies of the applications that were hard to classify. The three related studies previously cited (Schuster *et al*., 2012; Adelabu *et al*., 2014; Kim and Yeom, 2014) have not analyzed the effect of vegetation indices and the joint effect of the red-edge band and vegetation indices. In this study, it seems that the vegetation indices could significantly improve the accuracy of LCC in arid regions. Simultaneously adding the red-edge band and vegetation indices could also achieve significant accuracy improvement of LCC in arid regions. Conversely, although Adelabu *et al*. (2014) did not analyze the effect of vegetation indices, they mentioned that adding NDVI and NDVI-RE decreased the classification accuracies. It may be subject to the correlation between vegetation indices and spectral bands.In this study, NDVI-RE only slightly outperformed NDVI. The same conclusion has been drawn in Schuster *et al*. (2012). On the contrary, Adelabu *et al*. (2014) noted that NDVI-RE significantly outperformed NDVI. Obviously, the relative importance of NDVI and NDVI-RE depended specific applications.

Table 4 Results of the McNemar test for analyses 1–4

Analysis	Feature sets	f_{12}	f_{21}	χ^2	P
1	A vs. B	14	12	0.15	0.70
2	B vs. E	26	10	7.11	0.01
3	A vs. E	31	13	7.36	0.01
4	C vs. D	25	23	0.08	0.77

Notes: Feature set A: four bands (excluding red-edge). B: all bands. C: all bands + NDVI. D: all bands + NDVI-RE. E: all bands + NDVI + NDVI-RE. NDVI: normalized difference vegetation index. NDVI-RE: the red-edge adaptation of NDVI. f_{12}: the number of samples that the first feature set-based classification model misclassified but the second feature set-based model correctly classified in each analysis. f_{21}: the number of samples that the second feature set-based classification model misclassified but the first feature set-based model correctly classified in each analysis. χ^2: the chi-square statistic value. P: the probability value

5 Conclusions

The present study focused on part of the Xihu National Nature Reserve in the Gansu Province of China, which is a typical inland arid region that has suffered from severe desertification in recent years. The study first evaluated the effects of RapidEye imagery's red-edge band and vegetation indices on LCC in this arid region. The RapidEye spectral bands, NDVI, and NDVI-RE were used to form five feature sets including or excluding the red-edge band and vegetation indices. Based on a new land cover classification system involving plant communities, the RF algorithm-based models with five feature sets were used for LCC and their accuracies were assessed. The following conclusions were drawn: 1) the red-edge band only slightly increased the accuracy of LCC; 2) vegetation indices significantly improved the classification accuracy; 3) the red-edge band

and vegetation indices also achieved a significant OA improvement (3.46% increases from 86.67%); 4) NDVI-RE slightly outperformed NDVI. In general, separately adding the vegetation indices and simultaneously adding the red-edge band and vegetation indices significantly contributed to improve LCC in arid regions. Future work will focus on the effects of the object-oriented classification method and other red-edge vegetation indices, and some other indirect evaluation methods such as feature ranking and class similarity measures.

Acknowledgements

We thank Kong Lingfeng, Zhao Duohui, Liu Bo, and Yan Zezhou for helping to conduct the field investigation.

References

Adelabu S, Mutanga O, Adam E, 2014. Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels. *ISPRS Journal of Photogrammetry and Remote Sensing*, 95: 34–41. doi: 10.1016/j.isprsjprs.2014.05.013

Alrababah M A, Alhamad M N, 2006. Land use/cover classification of arid and semi-arid Mediterranean landscapes using Landsat ETM. *International Journal of Remote Sensing*, 27(13): 2703–2718. doi: 10.1080/01431160500522700

Asner G P, Heidebrecht K B, 2003. Imaging spectroscopy for desertification studies: comparing AVIRIS and EO-1 Hyperion in Argentina drylands. *IEEE Transactions on Geoscience and Remote Sensing*, 41(6): 1283–1296. doi: 10.1109/TGRS.2003.812903

Batterbury S, Warren A, 2001. The African Sahel 25 years after the great drought: assessing progress and moving towards new agendas and approaches. *Global Environmental Change*, 11(1): 1–8. doi: 10.1016/S0959-3780(00)00040-6

Belgiu M, Drăguț M, 2016. Random forest in remote sensing: a review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114: 24–31. doi: 10.1016/j.isprsjprs.2016.01.011

Breiman L, 2001. Random forests. *Machine Learning*, 45(1): 5–32. doi: 10.1023/A:1010933404324

Cerna L, Chytry M, 2005. Supervised classification of plant communities with artificial neural networks. *Journal of Vegetation Science*, 16(4): 407–414. doi: 10.1111/j.1654-1103.2005.tb02380.x

Chen W, Wang Y, Li X *et al*., 2016. Land use/land cover change and driving effects of water environment system in Dunhuang Basin, northwestern China. *Environmental Earth Science*, 75:1027. doi: 10.1007/s12665-016-5809-9

Chen Weitao, Sun Ziyong, Li Xianju *et al*., 2014a. Natural plant communities mapping in inland arid regions: a case in Dunhuang Basin, northwestern China. *Arid Land Geography*, 37(6): 1257–1263. (in Chinese)

Chen Weitao, Wang Yanxin, Sun Ziyong *et al*., 2014b. Groundwater-dependent ecosystems in arid inland zones: A case study at the Dunhuang Basin, northwestern China. *Quaternary Sciences*, 34(5): 950–958. (in Chinese)

Daskalaki S, Kopanas I, Avouris N, 2006. Evaluation of classifiers for an uneven class distribution problem. *Applied Artificial Intelligence*, 20(5): 381–417. doi: 10.1080/08839510500313653

Galletti C S, Myint S W, 2014. Land-use mapping in a mixed urban-agricultural arid landscape using object-based image analysis: a case study from Maricopa, Arizona. *Remote Sensing*, 6(7): 6089–6110. doi: 10.3390/rs6076089

Ge X, Ni J, Li Z *et al*., 2013. Quantifying the synergistic effect of the precipitation and land use on sandy desertification at county level: a case study in Naiman Banner, northern China. *Journal of Environmental Management*, 123: 34–41. doi: 10.1016/j.jenvman.2013.02.033

Han L, Zhang Z, Zhang Q *et al*., 2015. Desertification assessments in the Hexi corridor of northern China's Gansu Province by remote sensing. *Natural Hazards*, 75(3): 2715–2731. doi: 10.1007/s11069-014-1457-0

Hatton T, Evans R, 1998. Dependence of Ecosystems on Groundwater and its Significance to Australia. In: *Land and Water Resources Research and Development Corporation*. CSIRO, Clayton, Australia.

Kim H O, Yeom J M, 2014. Effect of red-edge and texture features for object-based paddy rice crop classification using RapidEye multi-spectral satellite image data. *International Journal of Remote Sensing*, 35(19): 7046–7068. doi: 10.1080/01431161.2014.965285

Langley S K, Cheshire H M, Humes K S, 2001. A comparison of single date and multitemporal satellite image classifications in a semi-arid grassland. *Journal of Arid Environments*, 49(2): 401–411. doi: 10.1006/jare.2000.0771

Li X, Chen W, Cheng X *et al*., 2016. A comparison of machine learning algorithms for mapping of complex surface-mined and agricultural landscapes using ZiYuan-3 stereo satellite imagery. *Remote Sensing*, 8(6): 514. doi: 10.3390/rs8060514

Li X, Chen W, Cheng X *et al*., 2017. Comparison and integration of feature reduction methods for land cover classification with RapidEye imagery. *Multimedia Tools and Applications*. doi: 10.1007/s11042-016-4311-4

Li X, Shao G, 2014. Object-based land-cover mapping with high resolution aerial photography at a county scale in Midwestern USA. *Remote Sensing*, 6(11): 11372–11390. doi: 10.3390/rs61111372

Liaw A, Wiener M, 2002. Classification and regression by randomforest. *R News*, 2(3): 18–22.

Liu C, Frazier P, Kumar K, 2007. Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, 107(4): 606–616. doi: 10.1016/j.rse.2006.10.010

Liu S, Wang T, Kang W *et al*., 2015. Several challenges in monitoring and assessing desertification. *Environmental Earth Sci-*

ences, 73(11): 7561–7570. doi: 10.1007/s12665-014-3926-x

Manandhar R, Odeh I O A, Ancev T, 2009. Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. *Remote Sensing*, 1(3): 330–344. doi: 10.3390/rs1030330

Meyer D, Dimitriadou E, Hornik K *et al*., 2015. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. Available at: https://cran.r-project.org/web/packages/e1071/index.html

Namdar M, Adamowski J, Saadat H *et al*., 2014. Land-use and land-cover classification in semi-arid regions using independent component analysis (ICA) and expert classification. *International Journal of Remote Sensing*, 35(24): 8057–8073. doi: 10.1080/01431161.2014.978035

Nordberg M L, Evertson J, 2003. Vegetation index differencing and linear regression for change detection in a Swedish mountain range using Landsat TM and ETM+ imagery. *Land Degradation & Development*, 16(2): 139–149. doi: 10.1002/ldr.660

Petropoulos G P, Kalaitzidis C, Vadrevu K P, 2012. Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery. *Computers & Geosciences*, 41: 99–107. doi: 10.1016/j.cageo.2011.08.019

Qin Xuwen, Chen Weitao, Li Xianju *et al*., 2014. Effects of water environmental system on desertification in an inland region of northwestern China: a case study in Dunhuang Basin. *Safety and Environmental Engineering*, 21(5), 39–45. (in Chinese)

R Development Core Team, 2015. R: A language and environment for statistical computing and graphics. Available at: https://cran.r-project.org/src/base/R-3/

Rozenstein O, Karnieli A, 2011. Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. *Applied Geography*, 31(2): 533–544. doi: 10.1016/j.apgeog.2010.11.006

Schuster C, Förster M, Kleinschmit B, 2012. Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. *International Journal of Remote Sensing*, 33(17): 5583–5599. doi: 10.1080/01431161.2012.666812

Stefanov W L, Ramsey M S, Christensen P R, 2001. Monitoring urban land cover change: an expert system approach to land cover classification of semiarid to arid urban centers. *Remote Sensing of Environment*, 77(2): 173–185. doi: 10.1016/S0034-4257(01)00204-8

Tigges J, Lakes T, Hostert P, 2013. Urban vegetation classification: benefits of multitemporal RapidEye satellite data. *Remote Sensing of Environment*, 136: 66–75. doi: 10.1016/j.rse.2013.05.001

Wang X, Chen F, Dong Z, 2006. The relative role of climatic and human factors in desertification in semiarid China. *Global Environmental Change*, 16(1): 48–57. doi: 10.1016/j.gloenvcha.2005.06.006

Wu Xiuqin, Liu Hongmeng, Huang Xiulan *et al*., 2011. Human driving forces: analysis of rocky desertification in karst region in Guanling County, Guizhou Province. *Chinese Geographical Science*, 21(5): 600–608. doi: 10.1007/s11769-011-0496-7

Xie Y, Sha Z, Yu M, 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology*, 1(1): 9–23. doi: 10.1093/jpe/rtm005

Zhang Panpan, Hu Yuanman, Xiao Duning *et al*., 2010. Rocky desertification risk zone delineation in karst plateau area: a case study in Puding County, Guizhou Province. *Chinese Geographical Science*, 20(1): 84–90. doi: 10.1007/s11769-010-0084-2