Towards a comprehensive evaluation of V-I-S sub-pixel fractions and land surface temperature for urban land-use classification in the USA

Quan Tang , Lei Wang , Bin Li & Jaehyung Yu

Taylor & Francis
Taylor & Francis Group

Towards a comprehensive evaluation of V-I-S sub-pixel fractions and land surface temperature for urban land-use classification in the USA

QUAN TANG*†, LEI WANG†, BIN LI‡ and JAEHYUNG YU§¶

†Department of Geography and Anthropology, Louisiana State University, Baton Rouge, LA 70803, USA
‡Department of Experimental Statistics, Louisiana State University, Baton Rouge, LA 70803, USA
§Department of Physics and Geosciences, Texas A&M University Kingsville, Kingsville, TX 78363, USA

Remote-sensing image classification based on the vegetation–impervious surface–soil (V-I-S) model and land-surface temperature (LST) has proved to be more efficient in characterizing the urban landscape than conventional spectral-based classification. However, current literature emphasizes discussion of the classifier's accuracy improvement achieved by the input of V-I-S fractions and LST over conventional spectral-based classification while ignoring the stability evaluation. Hence, this study proposes an evaluation framework for exploring the superiority of the input features and the stability of classifiers by integrating statistical randomization techniques and a kappa-error diagram. The evaluation framework was applied to case studies for demonstrating the superiority of V-I-S fractions and LST in the context of urban land-use classification with five different types of classifiers, including the maximum likelihood classifier (MLC), the tree classifier, the Bagging classifier, the random forest (RF) and the support vector machine (SVM). It followed that the use of V-I-S fractions and LST (1) could alleviate the 'salt and pepper' effect; (2) is preferred by tree and tree-based ensembles for branch splitting; (3) could produce classification trees with less complexity; (4) could benefit the stability of classifiers in addition to the accuracy improvement; and (5) could allow histograms following nearly normal distribution in its feature space, which boosts the performance of MLC. It is shown that MLC becomes comparable with modern classifiers when trained with V-I-S fractions and LST combination. Because of its adequacy and simplicity, MLC is recommended for urban land-use classification when V-I-S fractions and LST are used as the only input features. However, replacing them with, or including, the band reflectance might degrade MLC. A direct use of spectral band reflectance is not recommended for any of the classification approaches being considered in this study, except for SVM, which is the most robust classifier as it has a consistently high performance for all the input feature combinations. We recommend using tree-based ensemble classifiers or SVM when V-I-S fractions and LST as well as the band reflectance

---

*Corresponding author. Email: tang_quan_1025@hotmail.com
¶Present address: Department of Geology and Environmental Sciences, Chungnam National University, Daejon, South Korea

are all used in the classification. The proposed evaluation framework can also be applied to the assessment of input features and classifiers in other remote-sensing classification endeavours.

1. Introduction

The urban landscape is composed of dynamic and complex land-use (LU) features. Remote-sensing image classification has been widely used to obtain historical and present urban land-use and land-cover (LULC) conditions. Previous spectral-based classification mainly produced land-cover (LC; land material) classes and has received criticism due to its lack of consideration of the spatial configuration and arrangement of pixels, especially when high-resolution remote-sensing images were used. The 'salt and pepper' effect in classified images was also produced in this manner. Despite being referred to together and sometimes interchangeably, the concept of 'LU' and the concept of 'LC' are intrinsically distinct with respect to the natural and anthropogenic urban landscape. LC refers to the physical properties of the Earth's surface, which can be directly identified from the spectral characteristics; while LU, mostly being associated with large-scale studies, represents a higher-level understanding of the Earth's surface and should be inferred from the LC composition and configuration (Barnsley *et al.* 2001, pp. 103–104, Mesev 2010, p. 144). In urban remote sensing, the LU-level classification is generally more desirable as LU is directly related to social–economic processes and is of greater interest (Barnsley *et al*. 2001, p. 103, Lu and Weng 2006). Three research tracks have been documented in literature to make the use of spatial information for LU-level classification, namely sub-pixel analysis and subsequent classification with medium-resolution images (Ridd 1995, Ward *et al*. 2000, Small 2001, Phinn *et al*. 2002, Lu and Weng 2004, Lee and Lathrop 2006, Lu and Weng 2006, Weng *et al*. 2006), texture and spatial metrics aided classification (Haralick *et al*. 1973, Gong *et al*. 1992, Herold *et al*. 2003, Wu *et al*. 2006) and object-oriented classification (Baatz and Schäpe 2000, Shackelford and Davis 2003, Myint *et al*. 2011).

As a representative of sub-pixel analysis, the vegetation–impervious surface–soil (V-I-S) model (Ridd 1995) opened a new avenue for urban LULC studies and was extended and applied by many others (Small 2001, Phinn *et al*. 2002, Lu and Weng 2004, 2006, Small and Lu 2006). Different V-I-S configurations directly translate to different LU patterns, and the subsequent LULC classification based on V-I-S fractions (abundance) has proved to be more accurate and efficient for urban land characterization (Lu and Weng 2006, Weng *et al*. 2007, Weng and Quattrochi 2007). However, there is a scarcity of studies focusing on the stability of classifiers built with V-I-S fractions. Although the benefit of using V-I-S fractions over the conventional way of classifying urban LU from multi-spectral band reflectance can be found in the literature, discussion mainly emphasizes the accuracy improvement; seldom is the superiority evaluated in terms of the classifiers' stability or other aspects. Moreover, classification methods applied to V-I-S fractions are mostly conventional approaches, such as the maximum likelihood classifier (MLC) (Tarek *et al*. 2001, Lu *et al*. 2003, Lu and Weng 2004, 2005, Lu and Weng 2006) and the decision tree (Tarek *et al*. 2001, Lu and Weng 2004), failing to embrace other promising classifiers that have attracted the interest of the remote-sensing community due to their lesser demand for assumptions and their robustness to outliers, such as support vector machine (SVM) classifiers (Brown *et al*. 1999, Huang *et al*. 2002, Foody and Mathur 2004, Mathur and Foody 2008) and ensemble classifiers (DeFries and Chan 2000, Chan *et al*. 2001, Gislason

et al. 2006, Foody *et al.* 2007, Chan and Paelinckx 2008). Many studies have compared SVM with MLC (Huang *et al.* 2002), decision trees (Huang *et al.* 2002, Foody and Mathur 2004) and neural networks (Huang *et al.* 2002, Foody and Mathur 2004), and many of them have demonstrated the superiority of SVM. Other than the SVM method, ensemble classifiers are generally superior to the conventional classifiers as the variation of the base classifiers is reduced during the ensemble process. In addition, the land-surface temperature (LST) derived from thermal bands has also been found to be closely related to urban biophysical characteristics and has been applied to LULC mapping and urban sprawl studies (Lo *et al.* 1997, Weng 2001, Weng *et al.* 2007, Weng 2009). How much improvement of the classifiers that a synergic performance of the V-I-S fractions and LST could bring to the urban LU classification has also remained an unanswered though interesting question.

2. Evaluation framework

Previously, comparisons of input features and the classifier selection were limited to the overall classification accuracy and a single degree-of-agreement kappa statistic while ignoring the stability assessment. This research introduced an evaluation framework in which a classification tree and a randomization technique were used (1) to evaluate the importance/superiority of input features and (2) to evaluate and compare classifiers' performance for both accuracy and stability. Figure 1(*a*) displays three criteria for evaluating input features. Specifically, the tree structure and the sequence of variable being selected for branch splitting of trees are used to indicate the input feature's superiority. A random forest of trees is considered by using these two criteria to draw conclusions with statistical significance. The increase of the badness-of-fit on the out-of-bag samples in the random forest is used to reveal the superiority of input features. The evaluation of classifiers is depicted in figure 1(*b*) from two aspects: the kappa-error diagram and the analysis of variance (ANOVA) test. The following sub-sections elaborate the evaluation framework in detail.

2.1 *Variable selection sequence and the complexity of the classification tree*

The first two criteria for evaluating the superiority of input features originated in the classification and regression tree (CART). The CART produces classification rules by recursively seeking the most significant variable and the associated cut-off value for splitting. The fitted tree tends to select variables with great variability and the ability to separate classes in the early splits. The algorithm proceeds until no further splitting can be found to reduce the node impurity or classification error or gain information by a pre-defined threshold (controlled by the complexity parameter (cp)). Hence, the sequence of variables chosen for splitting is regarded as an indicator of the variables' importance and superiority. A tree's complexity highlighted by its depth and the number of nodes also indicates the capability of input variables in separating designated classes. In this sense, the simpler a fitted tree is, the more superior and relevant the input features are when used for growing such a tree in the given classification scheme.

2.2 *Randomization technique*

Randomization of training samples (the so-called bootstrap) creates multiple training data sets through drawing samples with replacement from the original sample pool,

(*a*)

(*b*)
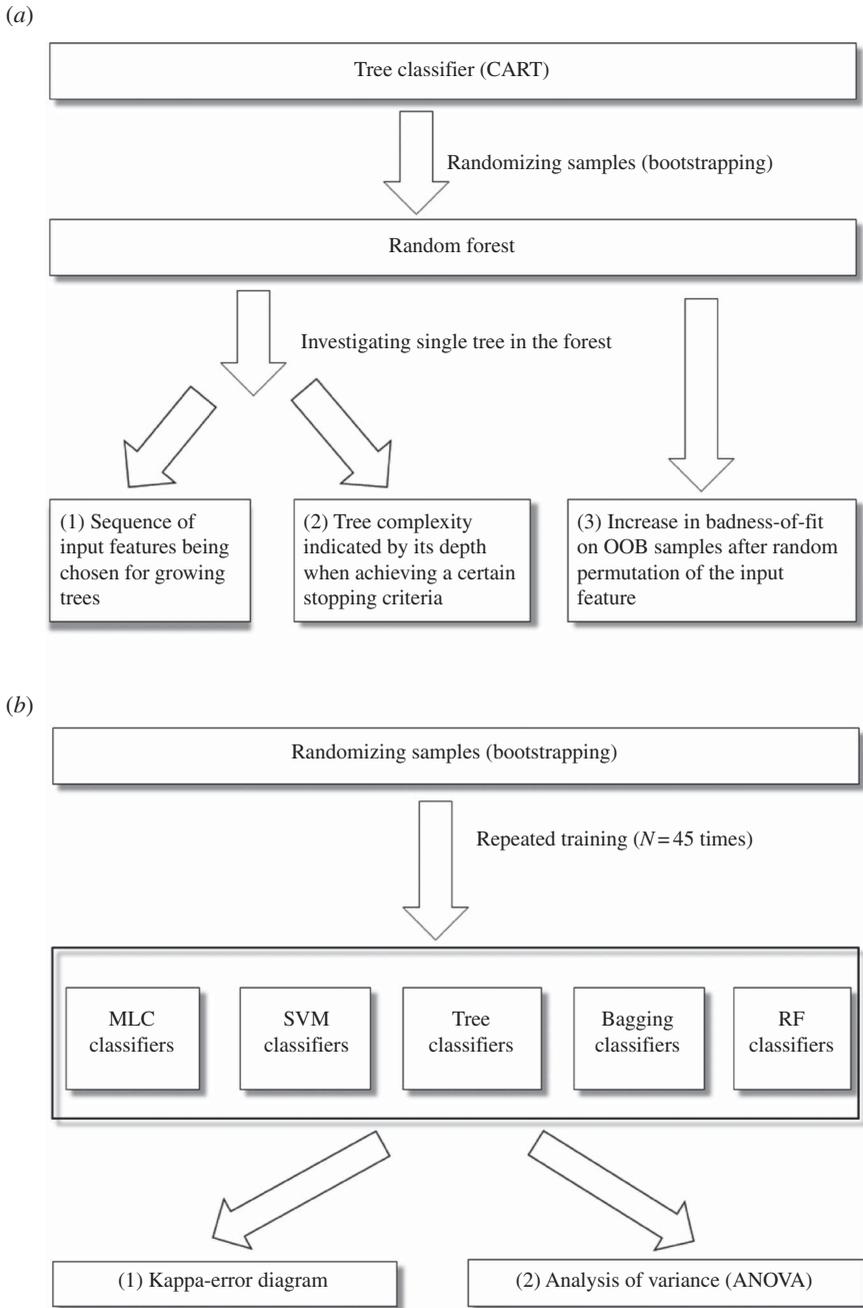
Figure 1.   Evaluation framework. (*a*) The evaluation of input features; (*b*) the evaluation of classifiers.

and the bootstrap samples are used for producing an ensemble of classifiers (Hastie *et al*. 2009, p. 249). The contribution of the randomization technique to the evaluation framework is threefold. First, these produced classifiers cannot only be considered

together to obtain a final classification but also, and more importantly, be investigated individually for the stability assessment of base classifiers that constitute the ensemble. Second, the conclusion regarding the superiority of input features by criteria introduced in §2.1 can be drawn repeatedly from an ensemble of tree classifiers (random forest); hence, the statistical significance of the conclusion is obtained. Third, randomization provides an innovative out-of-bag (OOB) method (Hastie *et al*. 2009, p. 593) to look at the variable's importance in the random forest. In the random forest, each tree is built with bootstrap samples. The prediction of a given observation is constructed by averaging only those tree classifiers corresponding to bootstrap samples in which this observation did not appear. Hence, the OOB idea is similar to the N-fold cross-validation (Hastie *et al*. 2009, p. 593). The variable's importance is measured by the increase in the badness-of-fit on OOB samples after a random permutation of the variable's value. In other words, if a given input feature is important, a random permutation would lead to a significant degradation of the model. Conversely, if a given input feature is not important in the first place, the random permutation will not make a big difference to the model fit.

2.3 *Kappa-error diagram*

Previously, the variability of the overall accuracy was adopted as a measure of stability of classifiers (Huang *et al*. 2002). However, this indicator still causes problems because classifiers with a similar overall accuracy might still have discrepant or even distinct classification results; hence, it is not reasonable to infer stability just from consistent accuracy reports. A kappa-error diagram (Margineantu and Dietterich 1997) is superior to the variability of the overall accuracy as a measure of the stability of classifiers. The kappa-error diagram visualizes the accuracy and stability of a classifier in a single scatter plot. First, n sets of bootstrap samples are produced for fitting a classifier n times, producing $C_n^2 = n(n-1)/2$ pairs of classifiers. Then, the kappa-error diagram is constructed by taking the mean error rate of pairwise classifiers on the validation data set as y and the corresponding degree of agreement (indicated by the kappa statistic as x). The kappa-error pattern for a stable and accurate classifier will display a compact point cloud located at the lower-right corner in the diagram indicating a low error rate and a high kappa statistic (meaning consistent and stable). The kappa-error diagram has not been generally adopted in the remote-sensing community albeit a few exceptions, such as DeFries and Chan (2000) and Chan *et al*. (2001). In this research, we promote the use of the kappa-error diagram by demonstrating its value in classifier evaluation.

2.4 *Classifiers*

We have tested five types of classifiers (figure 1(*b*)), including MLC, classification tree, tree-based ensemble classifiers and SVM. The tree-based ensembles include Bagging (bootstrapped aggregation) and random forest (RF).

As a parametric classifier, MLC relies on multi-variate normality assumptions and is mathematically equivalent to the Bayesian quadratic discriminant analysis (QDA; Richards and Jia 2006, pp. 194–199). MLC is the most popular classifier used by the remote-sensing community and is implemented in most remote-sensing packages, yet selections of training samples and classes that satisfy the normality assumptions remain a challenge (Myint and Lam 2005). Owing to its popularity, MLC is used in this research as a benchmark classifier.

Table 1. R packages with the implementation of different classification algorithms.

Classification method	R package
Maximum likelihood classification	MASS
Classification tree	rpart
Random forest	randomForest
Bagging	randomForest
Support vector machine	e1071

The ensemble classifier is a committee formed by fitting a collection of base classifiers that makes classification decisions by a (weighted) majority vote (Hastie *et al*. 2009, p. 605). The base classifiers are produced as independent classifiers uncorrelated to each other, which is possible through randomization to reduce the estimation variation, which generally leads to improved classification performance. Bagging generates an ensemble via randomization, in which the training samples are bootstrapped for fitting the base classifier repeatedly (Breiman 1996). RF is an extension of the Bagging method but differs from it in two aspects. First, it uses only trees as the base classifiers; second, only a random subset of all input features is considered for each branch split when growing tree classifiers. The selected input feature subsets vary over the entire forest, which further reduces the correlation among trees (Breiman 2001).

The SVM classifier belongs to the 'kernel-trick' family, which projects the data set into a higher dimensional feature space to achieve the ability to separate classes linearly by using kernel functions. Recent literature has witnessed a growing interest in applying the SVM classifier in remote-sensing classification. A recent thorough review of SVM applied in the remote-sensing area was furnished by Mountrakis *et al.* (2011). A Gaussian radial basis function (RBF) was used in this study since it has been proved effective in many classifications problems (Bruzzone and Carlin 2006).

The statistical programming language R available in the public domain was used to build the evaluation framework since it has implemented all the aforementioned classification algorithms (table 1).

3. Case study

3.1 *Study area*

The case study area is the suburban and core city areas ($\sim$360 km^2 in area) of New Orleans, LA (see figure 2), a port city to the north of the Gulf of Mexico and home to 1.24 million people in 2010 (according to US Census Bureau). The flow of the Mississippi River across the city resembles an upside-down Ω symbol. New Orleans is a metropolitan area with modern industrial and ethnic development despite its geographical location being around a hurricane-threatened community. New Orleans is in a sub-tropical environment and undergoes an evident urban heat island (UHI) effect, especially in warm weather, due to the large amounts of paved and dark-coloured surfaces and greenhouse gas release in the urban communities (Baseline Greenhouse Gas Emission Profile 2001). Several LU classes have been identified in the study area. The southwest contains major vegetated areas composed of forest, agricultural land and wetland; both the City Park to the south shore of Lake Pontchartrain

(*a*)

(*b*)
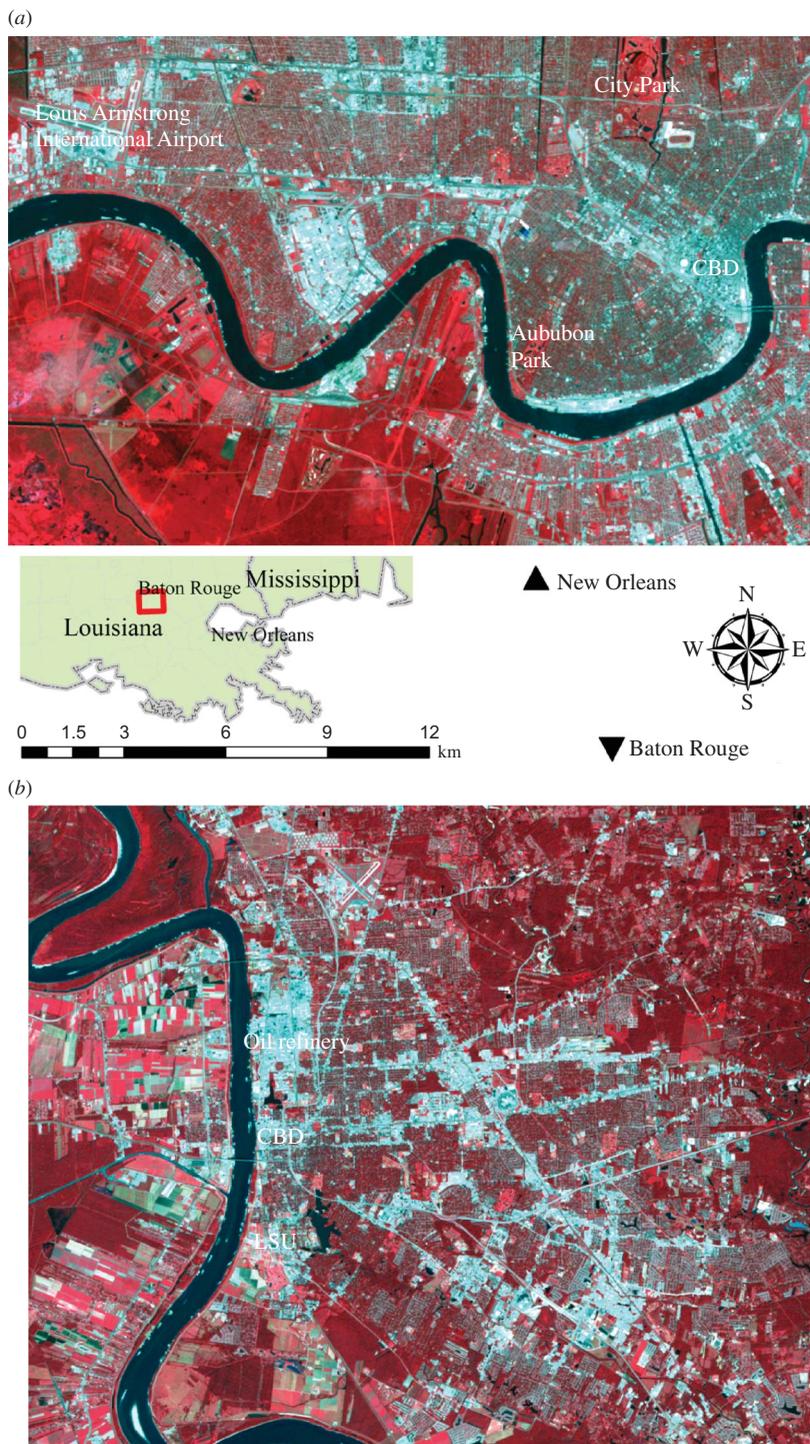
Figure 2.   Study area of New Orleans and Baton Rouge (Landsat 22/39 and 23/39).

and the Audubon Park in the uptown neighbourhood also have a large coverage of vegetation; the residential area is mainly located to the north and southeast of the Mississippi River; the commercial and industrial areas are predominantly located along the Mississippi River and in the central city (the northeastern part of the upside-down Ω symbol, such as the Superdome Stadium). The waterbody areas are evident as the Mississippi River across the city and canals as well as bayous in the southwest. The diversity of LU patterns and the evident UHI effect in New Orleans are the main reasons that we selected it for the case study.

Baton Rouge (see figure 2), the capital and the second largest city of Louisiana State, is also included as an additional test site to verify major conclusions and to demonstrate the generalizability of this study. The economic development of Baton Rouge depends on the petrochemical industry. The central business district (CBD) area is to the north of Baton Rouge and a large area in the north of the CBD is occupied by oil refineries. The residential areas have been established mainly in the east and south of the city. Baton Rouge is a dynamic urban area with a diversity of LU patterns that is similar to New Orleans. These two urban areas are representative of the majority of the US urban environment.

3.2 *Data source and classification scheme*

The Landsat Thematic Mapper (TM) image used in this analysis (Row/Path: 22/39 for New Orleans and 23/39 for Baton Rouge) was obtained in August 2005 since a better UHI effect is expected in this season than in other seasons. Bands 1–5 and 7 were calibrated with image metadata to compute the spectral reflectance. The 1 m resolution digital orthophoto quarter quadrangles (DOQQs) obtained in that time of the year were used for validation. A rigorous image co-registration between DOQQs and the TM images was performed in the UTM Zone 15, North American Datum 1983 projection system, to ensure the locational accuracy of our analysis. A classification scheme with four classes of 'Commercial', 'Residential', 'Vegetated' and 'Waterbody' was adopted (table 2). These LU classes are commonly encountered in US urban areas.

Training samples were specified by identifying polygons of homogeneous LU areas in the Landsat TM scene in New Orleans, which comprised 628 pixels of 'Commercial' LU, 541 pixels of 'Residential' LU, 471 pixels of 'Vegetated' LU and 173 pixels as 'Waterbody'. For the Baton Rouge site, the training samples covered 195 pixels of 'Commercial', 368 pixels of 'Residential', 481 pixels of 'Vegetation' and 133 pixels of 'Waterbody'. Random points were generated and visually interpreted from the DOQQ to identify the reference data of LU classes.

Table 2. Description of land-use classes adopted in this study.

Category	Description
Commercial	Areas predominantly constructed for human activities associated with commercial and industrial events, including, buildings, parking lots, shopping centres, transportation roads, etc.
Residential	Areas predominantly constructed for human dwelling and residential purposes
Vegetated	Large homogeneous areas of vegetation cover, including forests, forested wetlands, grassland, shrubs, etc.
Waterbody	Mississippi River, lakes, ponds, canals and bayous

3.3   *Feature extraction*

The fraction images were derived from the Landsat TM image by using the linear spectral mixture analysis (LSMA) method (Wu and Murray 2003). A normalization procedure (Wu 2004) was performed prior to the LSMA in order to reduce the spectra variance of the end-members, followed by a minimum noise fraction (MNF) transformation to reduce the data dimensionality and to eliminate noise in the TM image. End-members were extracted based on the scatter plots of MNF components. The impervious surfaces were mainly man-made features with a high albedo, such as buildings, transportation roads, parking lots and river deck. The spectrum of soil was mixed with that of low-albedo features. Vegetation was predominately found as forests and grassland. LSMA produced three fraction images: impervious surface + high albedo fraction image (figure 3(*a*)), low albedo + soil fraction image (figure 3(*b*)) and vegetation fraction image (figure 3(*c*)). The overall root mean square error (RMSE) of the LSMA was 0.045.

Various algorithms were developed to recover the absolute value of LST from thermal infrared (TIR) bands, including the radiative transfer equation (RTE)-based algorithm (Berk *et al*. 1989, Schmugge *et al*. 1998, Sobrino *et al*. 2004), the mono-window algorithm (Qin *et al*. 2001) and the single-channel algorithm (Jiménez-Muñoz and Sobrino 2003). These models usually require additional data input (water vapour content, etc.) and prior knowledge of the underlying surface for atmospheric correction and emissivity adjustment. However, in this study, LST has been used for classification. Therefore, the relative measurement was sufficient for mapping the LST spatial variation and relating it to the LULC patterns (Weng 2009); hence, LST was computed by using the standard calibration procedure (Landsat Project Science Office 2002). Figure 3(*d*) shows the LST image obtained from TM band 6. The UHI effect
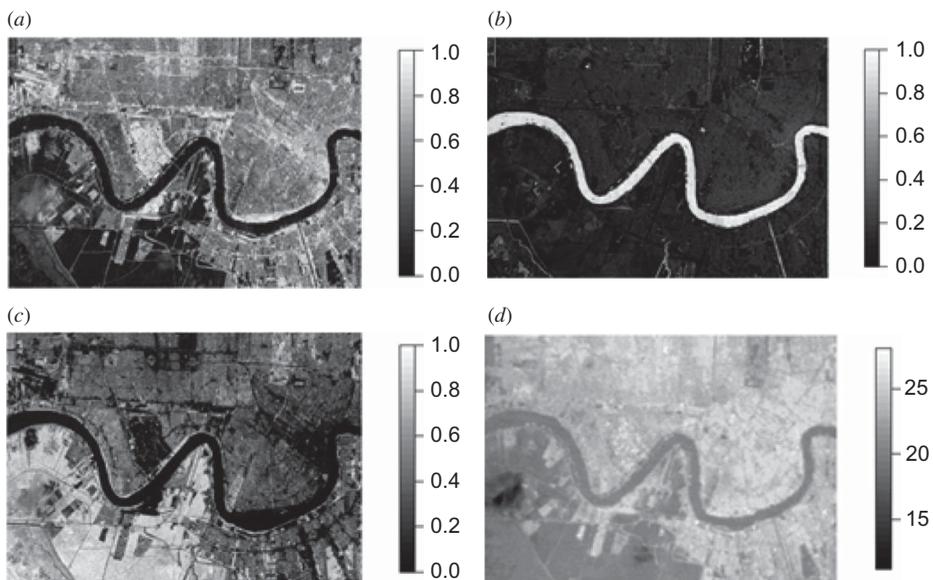
Figure 3.   (*a*) Impervious surface + high-albedo fraction image; (*b*) Low-albedo fraction image; (*c*) Vegetation fraction image; (*d*) LST image.

was evident in New Orleans during the period of the study. The average LST of the southern forest area was much lower than the commercial area.

A ten-dimensional composite input data set for classification was made by stacking three fraction images, six multi-spectral reflectance images from TM bands 1–5 and 7 and the LST image. Various combinations of these input bands are possible for classification. Specifically, we were interested in testing the following three commonly used input feature configurations:

1. six original spectral bands of TM only;
2. V-I-S fractions and LST only; and
3. a combination of six original spectral bands of TM and V-I-S fractions and LST (the full data dimension, ten bands).

4. Results and discussion

4.1 *Classification results*

The classification results using the six-band multi-spectral reflectance data (TM bands 1–5 and 7) are shown in figure 4. Regardless of the classifier being used, multi-spectral reflectance input produced an evident 'salt and pepper' effect as marked in the maps, which is a common problem associated with pixel-based classification. Two salient misclassification regions were noticed. One was the bayou area with vegetation coverage in the southwest of the scene; and the other was the vegetated area in the north of the southwest canals. Both the regions were misclassified as 'residential' LU. Note that the sliver areas in the southern forest area were also misclassified as 'residential' and those pixels appeared isolated. These misclassifications revealed the incapability of spectral reflectance in distinguishing the urban LU patterns. In contrast, as seen in figure 5, classifiers built with V-I-S fractions and LST were notably improved by alleviating the 'salt and pepper' problem, producing a reasonable and homogeneous vegetated area in the southwest. Furthermore, the accuracy of classification also improved when using V-I-S fractions and LST instead of the six-band multi-spectral data (table 3).

4.2 *Evaluation framework-guided selection of input features*

The following sections evaluate the input features by using the randomization framework. As mentioned before, we are interested in testing three specific combinations of input data: (1) multi-spectral reflectance only, (2) V-I-S fractions and LST combination and (3) a composite of (1) and (2).

4.2.1 Multi-spectral reflectance only. All classifiers were trained repeatedly (by $n = 45$ times) with bootstrapping samples for producing the kappa-error diagram. Figure 6(*a*) shows the accuracy of the performances of all the classifiers trained with different input feature combinations. When used with the multi-spectral reflectance as the only input feature, all the classifiers gave a low accuracy (indicated by the low-elevated green boxes in figure 6(*a*)). The tree classifier appeared to be the most spread-out box, indicating that it is the most unstable classifier when using only the spectral reflectance data. Owing to the instability and weakness of the single-tree classifier, tree-based ensemble classifiers can improve the stability of individual tree classifiers. However, the improvement was statistically significant only in stability, not
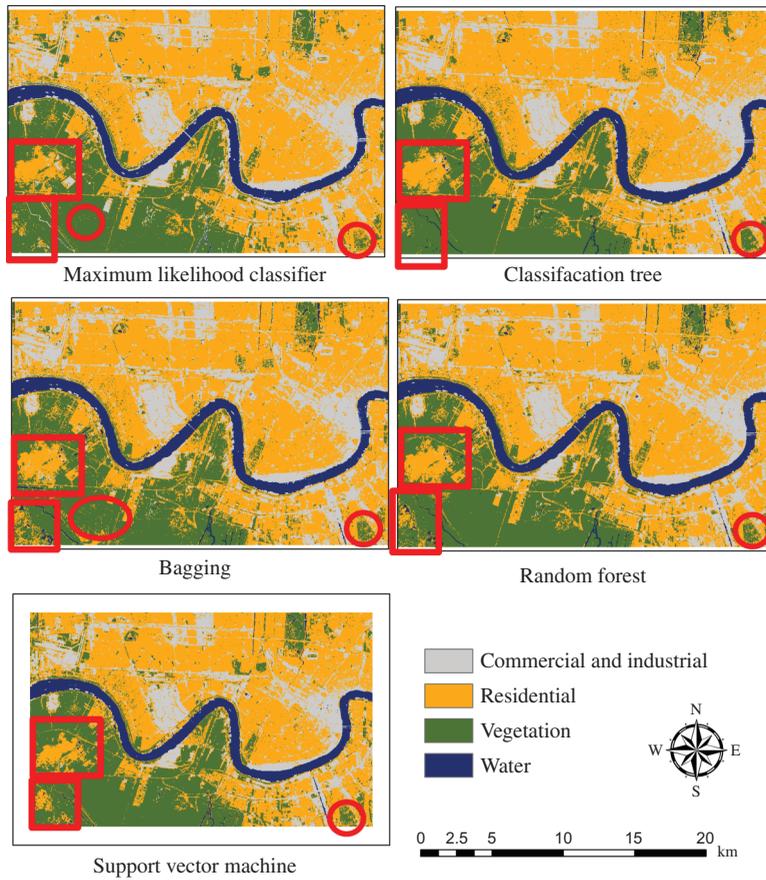
Figure 4. Classification results with only the multi-spectral reflectance variable used as the input. Red circles mark the 'salt and pepper' effect resulting from the image classification. Red rectangles mark the salient misclassification regions.

much in accuracy, as could be seen from the ANOVA tests using multiple comparisons with Tukey–Kramar adjustment (Freund and Wilson 2002, pp. 256–257; table 4). The ANOVA test revealed that the SVM classifier stood out for both accuracy and stability. In the kappa-error diagram (figure 7(a)), classifiers trained with band spectral reflectance depicted a more compact point cloud for those tree-based ensemble methods when compared with the single-tree classifier. Again, the SVM classifier demonstrated its power by arriving at the lower-right corner of the kappa-error diagram.

4.2.2 Synergy of V-I-S fractions and LST. The performance of MLC benefits significantly from the replacement of the band reflectance input with the V-I-S fraction and LST combination input (see figure 6(a) for a clearly elevated blue box for the MLC and table 5 for significant p-values when testing the accuracy of the performance of MLC against other classifiers). The performance was even better than a synergy of
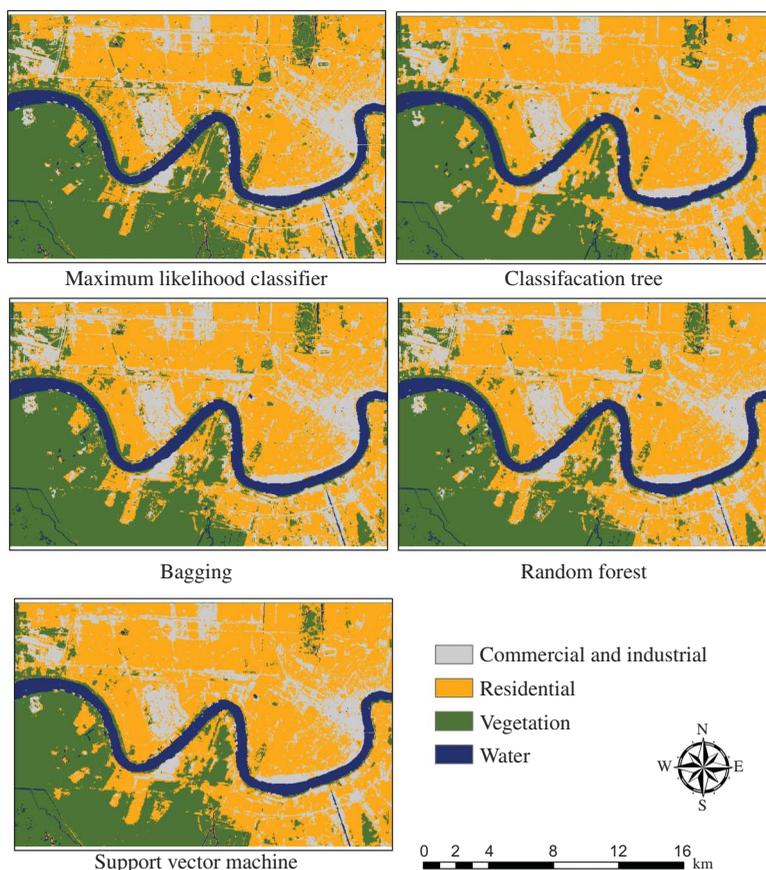
Figure 5. Classification results with V-I-S fractions and land-surface temperature. When compared with figure 4, the 'salt and pepper' problem is alleviated. A more homogeneous area of vegetated area to the southwest of the study area is produced. The salient misclassification regions in figure 4 are also eliminated in figure 5.

Table 3. The comparison of accuracy performance between two input feature configurations: (a) multi-spectral reflectance only; (b) the synergy of V-I-S fractions and LST.

Multi-spectral reflectance only			Synergy of V-I-S fractions and LST		
Classifier	Accuracy	Kappa	Classifier	Accuracy	Kappa
MLC	0.69	0.55	MLC	0.75	0.64
CART	0.69	0.55	CART	0.70	0.57
Bagging	0.69	0.55	Bagging	0.72	0.60
RF	0.69	0.55	RF	0.72	0.60
SVM	0.72	0.60	SVM	0.70	0.55

V-I-S fractions and LST and all spectral reflectance variables taken together. A possible explanation of this is as follows: according to Ridd's V-I-S model (Ridd 1995) and Lu–Weng model (Lu and Weng 2006), certain urban LU types can be identified as cluster of points at certain locations in the V-I-S feature space (see figure 8 for our
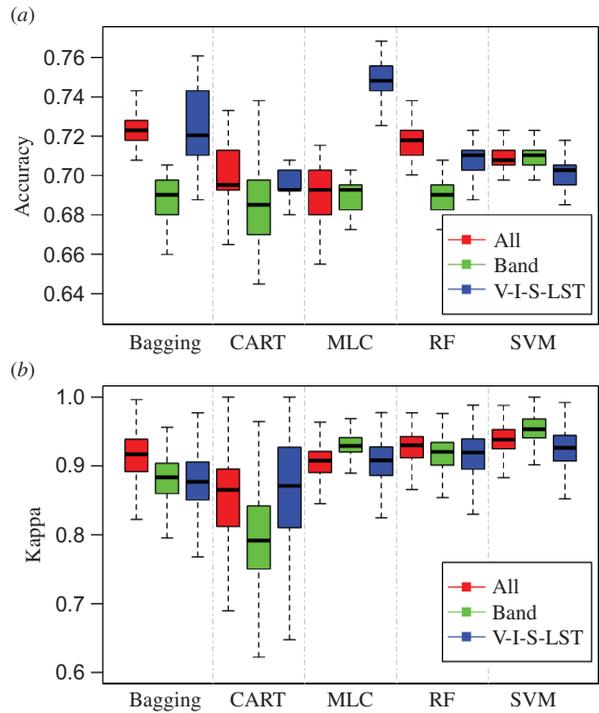
Figure 6. Accuracy and stability comparisons of five classifiers for three different input feature configurations (New Orleans, LA).

Table 4. Multiple comparisons in terms of accuracy and stability of different classifiers trained by band spectral reflectance variables only.

Classifier	Accuracy difference (%)	*p*-Value	Implication
CART–Bagging	0.8	0.9994	Randomization achieves no
CART–RF	0.2	0.9513	significant accuracy improvement on the single-tree classifier
SVM–Bagging	2.0	<0.0001	SVM significantly outperforms
SVM–MLC	2.0	<0.0001	all other classifiers
SVM–RF	1.9	<0.0001	
SVM–CART	2.0	<0.0001	

Classifier	Kappa difference	*p*-Value	Implication
CART–Bagging	−0.08	<0.0001	Randomization improves the
CART–RF	−0.12	<0.0001	stability of the single-tree classifier significantly
SVM–Bagging	0.07	<0.0001	SVM is significantly more stable
SVM–CART	0.15	<0.0001	than any other classifiers,
SVM–RF	0.04	<0.0001	especially as opposed to the
SVM–MLC	0.02	<0.0001	single-tree classifier
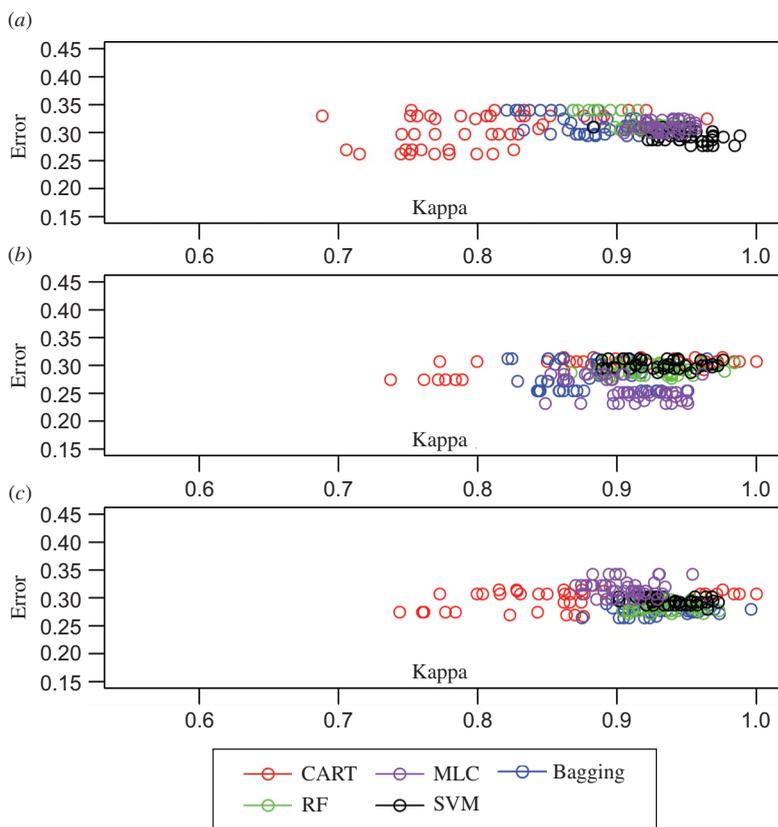
Figure 7. Kappa-error diagrams produced by three different input feature configurations: (a) Multi-spectral reflectance only; (b) V-I-S fractions and LST combination; (c) Composite of (a) and (b), e.g. all features being included.

case). Hence, the distribution of V-I-S fractions and LST in the feature space displays several bell-shaped (Gaussian-like) surfaces with each individual peak corresponding to a certain urban LU type, which matches with the MLC's normality assumption and thus leads to the optimal Bayesian decision boundary of the classes. However, a similar pattern cannot be found for the case in which the multi-spectral reflectance is used as the data input. A similar finding was also stated in Pal and Mather (2003): ' . . . the ML algorithm is preferred unless there is particular reason believing that the data do not follow a Gaussian distribution'. Hence, despite being criticized in the literature, MLC is a good classifier as long as its basic assumption is met by a careful selection of the input features. In this study, we found that the use of V-I-S fractions and LST data satisfied the normality assumption more easily.

The tree classifiers built with V-I-S fractions and LST had more stability than their counterparts built with the multi-spectral reflectance input, as highlighted by the fact that the 'Accuracy' box for CART was less spread-out (figure 6(a)) and the 'kappa' box for CART was more elevated (figure 6(b)) and a significant p value ($p < 0.001$) in table 6. The point cloud for the tree classifier was more compact in figure 7(b) than in figure 7(a). The commonly reported improvement caused by using tree-based ensembles over the single-tree classifier was, however, not observed (table 4). This result is different

Table 5. Multiple comparisons in terms of accuracy and stability of different classifiers built with V-I-S fractions and LST combination.

Classifier	Accuracy difference (%)	p-Value	Implication
CART–Bagging	−2.2	<0.0001	Ensembles slightly improve the accuracy of CART. The base tree classifier was sufficiently stable, making less room for potential improvement
CART–RF	−1.2	0.0026	
MLC–Bagging	2.7	<0.0001	MLC is the most accurate classifier when built with V-I-S fractions and LST
MLC–CART	5.0	<0.0001	
MLC–RF	3.7	<0.0001	
MLC–SVM	4.3	<0.0001	

Classifier	Kappa difference	p-Value	Implication
CART–Bagging	−0.02	<0.0001	Improvements in stability achieved by ensembles are less than as in table 3 or even none. The base tree classifier is sufficiently stable, making less or no room for potential improvement
CART–RF	−0.06	<0.0001	
MLC–Bagging	0.03	0.0224	MLC is also the top stable classifier when built with V-I-S fractions and LST
MLC–CART	0.04	0.5475	
MLC–RF	−0.01	<0.0001	
MLC–SVM	0.02	<0.0001	
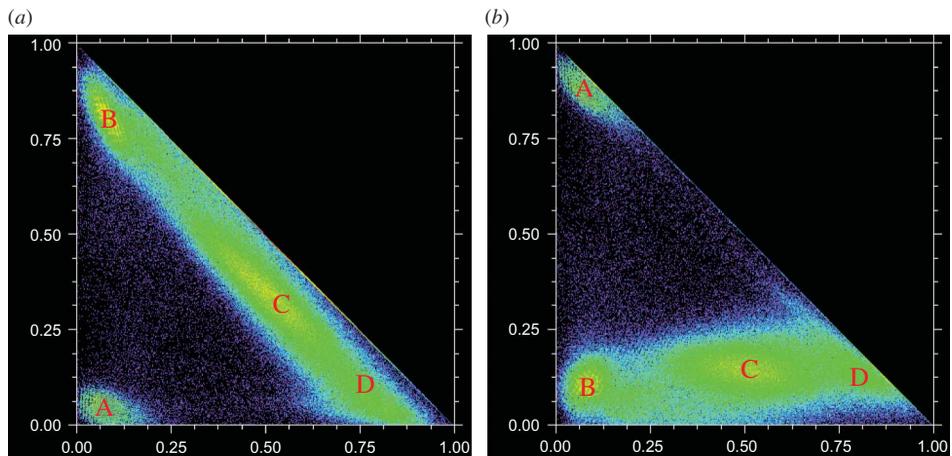
Figure 8. Slice density maps of distribution of V-I-S fractions and LST in the study area. The distribution is Gaussian-like and each bump represents a certain land-use pattern. (*a*) 'Impervious surface' fraction as *x* and 'Vegetation' fraction as *y*; (*b*) 'Impervious surface' fraction as *x* and 'Low albedo + soil' fraction as *y*. Land-use classes are labelled with (A) Waterbody; (B) Vegetated; (C) Residential; (D) Commercial.

Table 6. Performance comparison between (a) V-I-S fractions and LST (b) band spectral reflectance only.

Classifier	Accuracy difference (%)	*p*-Value	Implication
Bagging	3.0	<0.0001	The Bagging, RF and MLC
CART	0.8	0.1767	classifiers improve when band
MLC	6.0	<0.0001	spectral reflectance input is
RF	1.9	<0.0001	replaced by V-I-S fractions
SVM	−0.6	0.6843	and LST

Classifier	Kappa difference	*p*-Value	Implication
CART	0.06	<0.0001	More stable trees are produced when trained with V-I-S fractions and LST

Note: The differences were calculated by subtracting (b) from (a).

from the case when the multi-spectral reflectance was used. Hence, the improvement of tree-based ensembles over the single-tree classifier was suppressed because the tree classifier as the base classifier was adequately accurate and stable when trained with V-I-S fractions and LST combinations. This conforms with the conclusion drawn in the previous theoretical (Breiman 1996) and empirical studies (Chan *et al*. 2001). RF and Bagging had 1.9 and 3.0% accuracy increases, respectively.

Commercial and industrial areas have an evident UHI effect due to the thermal property of underlying surfaces (buildings, transportation network, parking lots, etc.) and massive human activities generating extra heat. These areas usually have higher surface temperatures as opposed to more vegetated areas in the suburban area (forest, grassland, etc.). Conversely, the 'vegetated' areas are less disturbed by human activities and undergo transpiration that lowers the surface temperature. Hence, LST has the ability to differentiate the vegetated areas from the built-up areas. However, 'Waterbody' with high thermal inertia also appears cool, making it not separable from the 'vegetated' areas merely by using LST. The 'residential' area has intermediate surface temperature in between, so it might be confused with 'commercial' LU or 'vegetated' LU depending on the relative amount of vegetation coverage (the 'V' component) and man-made materials (the 'I' component). The lack of ability to separate in terms of LST was clearly seen in figure 3(*d*). Although these LU classes ('Vegetated' and 'Waterbody', 'Residential' and 'Commercial', etc.) have similar LSTs, their V-I-S configuration differs considerably. This is where the V-I-S fractions input comes into play to increase the ability to separate. As such, V-I-S fractions and an LST are complementary to each other and make an ideal input data configuration for urban LU classification.

4.2.3 Full data dimension. Table 7 demonstrates the extent of gain that could be obtained by the addition of V-I-S fractions and LST to the conventional spectral-based classification. MLC has little benefit from the use of the spectral reflectance variables, as they violate the assumptions of MLC and the impact of this was great enough to choose to use the V-I-S fractions and LST. On the contrary, the tree and

Table 7. Performance comparison between (a) all input features (b) band spectral reflectance only.

Classifier	Accuracy difference (%)	*p*-Value	Implication
Bagging	3.0	<0.0001	Adding V-I-S fractions and an LST variable would not enhance MLC as the assumption still remains violated. Tree classifier gained accuracy; tree-based ensemble classifiers achieved the highest accuracy
CART	1.7	<0.0001	
MLC	0.3	0.9983	
RF	2.4	<0.0001	
SVM	0.0	1.0000000	

Classifier	Kappa difference	*p*-Value	Implication
Bagging	0.03	<0.0001	A slight but statistically significant stability improvement is achieved by adding V-I-S fractions and LST
CART	0.06	<0.0001	
MLC	0.03	<0.0001	
RF	0.01	<0.0001	
SVM	−0.02	<0.0001	

Note: The differences were calculated by subtracting (b) from (a).

tree-based ensembles were enhanced in accuracy due to the addition of V-I-S fractions and LST. The improvement in accuracy for CART, Bagging and RF was 1.7, 3.0 and 2.4%, respectively. The best classifier to be used for this input feature configuration was the Bagging method. SVM retained its good performance regardless of the addition of input features.

4.3 *Discussion on the superiority of V-I-S fractions and LST*

In addition to the stability and accuracy improvement revealed, the superiority of V-I-S fractions and LST was also explored from two other aspects: (1) the tree structures (e.g. the sequence of variable being selected and the tree size) of a single classification tree and trees in the forest; (2) the variable importance determined by OOB samples. The latter would provide an alternative way for band selection, such as the application in hyper-spectral remote sensing (Chan and Paelinckx 2008).

In our case study, the tree classifier built with the ten-dimensional data set using the 'Gini index' criteria picked up all the V-I-S fractions and LST (figure 9(*a*)) in the earlier splits, which, however, only adopted one spectral band reflectance (B5), indicating that not much decrease in the overall lack-of-fit could be achieved by using spectral reflectance data. The 'information gain' criterion was also specified and the fitted tree picked up the V-I-S fractions and LST for their early splits and ignored most of the spectral reflectance variables (figure 9(*b*)). Tree was pruned to cp = 0.001 because beyond this cp value little accuracy improvement would result.

Delving into the grown forest to view each single fitted tree further confirmed the relevance of V-I-S fractions and LST with the urban LU classes. All ten input features were used for growing the RF. The forest size (number of random trees) was set to 300 as with this number the RF can achieve the least OOB error rate. The increase in the badness-of-fit can be represented by either a decrease in accuracy (equivalent to the
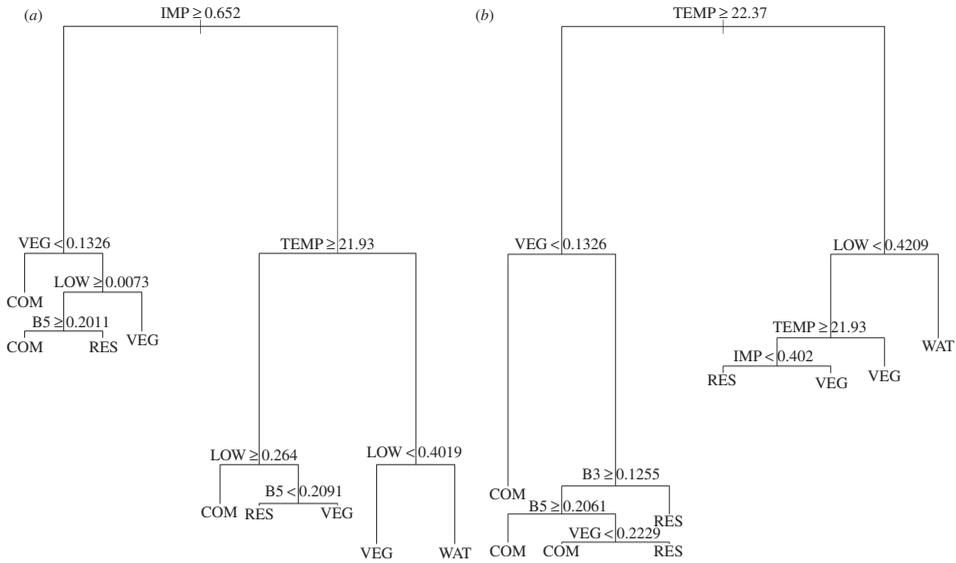
Figure 9.   (*a*) Tree fitted with cp = 0.001 and 'Gini index' criteria. (*b*) Tree fitted with cp = 0.001 and 'information gain' criteria.
Note: IMP, impervious surface fraction; VEG, vegetation fraction; LOW, low albedo + soil fraction; TEMP, LST; COM, Commercial; RES, Residential; VEG, Vegetated; WATER, Waterbody.

Table 8. Variable importance comparisons based on the increase in the badness-of-fit for OOB samples in a random forest.

Criteria	Sequence
Mean increase in node impurity	TEMP > IMP > VEG > LOW > B5 > B7 > B4 > B3 > B1 > B2
Mean decrease in accuracy	TEMP > VEG > IMP > B5 > B3 > LOW > B7 > B4 > B1 > B2

Notes: LST is the most important input feature. Band spectral reflectances are less important than the LST and V-I-S fractions in urban land-use classification.
IMP, impervious surface fraction; VEG, Vegetation fraction; LOW, Low albedo/soil fraction; TEMP, LST.

increase in OOB error rate) or an increase in 'node impurity'. The greater importance of V-I-S fraction and LST can be observed in table 8.

To evaluate the input feature's superiority with statistical significance, a forest consisting of 500 over-fitted trees was grown for individual inspections. The average size of the 500 trees when only spectral reflectance was used was about 39 nodes; and this number was reduced to 25 when V-I-S fractions and LST were used instead. The ANOVA test returned a significant *p*-value ($p < 0.001$). The frequencies of the ten input variables being used as the first five splitting variables are summarized in table 9. The first five splits were generally based on the use of V-I-S fractions and LST, suggesting their superiority over band reflectance for urban LU studies. Another finding in the analysis is that the infrared bands (B4, B5 and B7) of Landsat TM carry more information for separating urban LU classes than the visible bands.

Table 9. Deep investigation of individual trees by the Bagging method.

Input feature	Root node	Second node	Third node	Fourth node	Fifth node
B1	6	13	30	2	6
B2	2	10	9	11	2
B3	6	10	5	45	1
B4	22	15	0	90	1
B5	101	11	3	76	11
B7	188	46	23	67	43
VEG	107	134	72	50	103
IMP	34	180	132	46	151
LOW	28	81	161	29	119
TEMP	6	0	65	13	43

Notes: The random forest method was replaced by the Bagging method to guarantee fairness. In a random forest, the probability of V-I-S fractions and LST being selected for growing a single tree compared to that of band spectral reflectance is 4:6, which is not fair. A preference for using V-I-S fractions and LST for splitting a tree should be noted.

4.4 *Generalizability and limitation*

To test the generalizability of the results, the same analysis was also conducted on the test site of Baton Rouge, LA. The SVM classifier was still the best classifier among all classifiers being considered for conventional spectral-based classification. Except for SVM, all the classifiers were boosted when we replaced the band reflectance with V-I-S fractions and LST combination as input features. This was indicated by the prevalently more elevated and less spread blue boxes when compared with their green counterparts (see figure 10). Especially, the improvement was most significant for MLC. The difference in accuracy for MLC was 0.10 with a significant p value. This could be attributed again to the superiority of the V-I-S fractions and LST combined input as it has nearly multi-variate normal distribution and meets the MLC assumption. Given that the V-I-S fractions and LST combination is the input, although MLC was not the best classifier of the five tested classifiers for the Baton Rouge test site, it still managed to achieve a comparable level of accuracy to those modern machine-learning classifiers.

The tree classifier was again made more stable (but not more accurate) by either using the ensemble approach or replacing the band reflectance input with the V-I-S fractions and LST, which can be clearly seen in figure 10(*a*), where the 'Accuracy' boxes, B (band reflectance + Bagging), C (band reflectance + RF) and D (V-I-S fractions and LST combination + CART), are less spread-out than A (band reflectance + CART), as well as in figure 10(*b*), where the 'kappa' boxes, B (band reflectance + Bagging), C (band reflectance + RF) and D (V-I-S fractions and LST combination + CART), are more elevated than A (band reflectance + CART). The addition of V-I-S fractions and LST to the multi-spectral reflectance also improved the tree-based ensemble classifiers significantly, which was also the same as the case for New Orleans.

The V-I-S model is the most suitable input when the classes of interest have distinguishable differences in their V-I-S structures. This is true for our two study areas and most US urban areas where residential parcels are predominantly composed of individual houses with adequate vegetation coverage. Therefore, the V and I fraction quantities can distinguish different LU classes, e.g. the 'residential class' is in

(*a*)

(*b*)
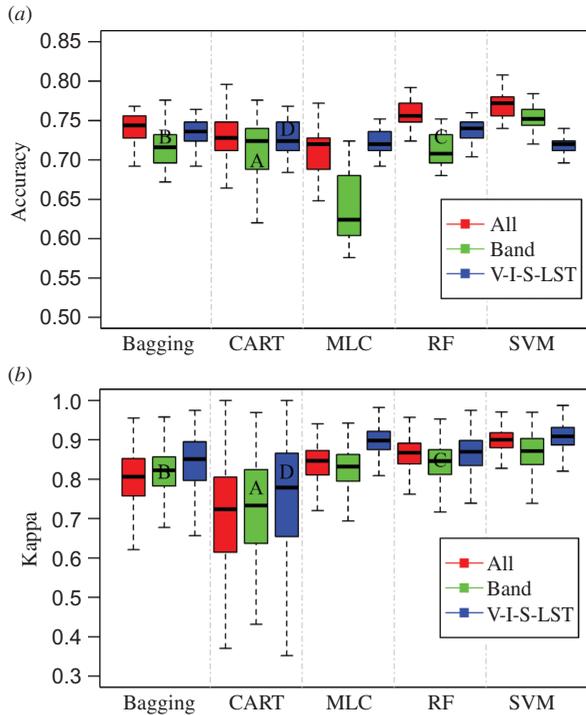
Figure 10. Accuracy and stability comparisons of five classifiers for three different input feature configurations (Baton Rouge, LA).

between the 'commercial class' and 'vegetation class'. However, for some developing countries, such as Haiti and China, the majority of the population lives in multi-storey apartments and the average vegetation coverage is limited. Hence, the boundary between 'residential class' and 'commercial class' becomes fuzzy. Hence, the superiority of V-I-S-derived LU classification might be applicable only to the US urban environment and the like. The conclusion may also not be applicable if a different or more detailed classification schema is to be adopted, such as a breakdown of vegetation class to 'Forest' and 'Pasture' class. However, the current classification schema is very common in the literature of US urban studies. These four LU classes belong to Anderson Level-I classes (Anderson *et al*. 1976). Given the medium-resolution image, this is a reasonable and applicable LU classification schema. The classification results of this study are useful for a Level-II classification with higher spectral/spatial resolution images. For example, the LU class boundary could be used as constraints for the segmentation in a subsequent object-oriented classification with a more detailed classification schema.

5. Conclusions

The selection of classifiers is closely related to the selection of input features for urban LU classification. This article proposes an evaluation framework based on the classification tree and statistical randomization methods to offer a comprehensive evaluation of the superiority of input features and the performance of classifiers. A case study using the framework showed the superiority of the V-I-S fractions and

LST over the direct use of multi-spectral band reflectance in urban LU classification. The investigations and discussions were made based on (1) the sequence of variable being selected in growing trees, (2) tree complexity, (3) kappa-error diagram, (4) increase in the badness-of-fit for the OOB samples and (5) ANOVA tests. Three major conclusions could be drawn.

First, the results of the analysis in this research promote the use of V-I-S fraction images and LST in urban LU studies, because in addition to the improvement in the overall accuracy of classification, which has already been documented in the previous literature and also confirmed in this study by using the evaluation framework, the use of V-I-S fractions and LST combination as the input for urban LU classification can also: (1) alleviate the 'salt and pepper' problem; (2) be preferred by tree and tree-based ensembles for branch splitting; (3) lead to less complex trees when achieving the stopping criteria; (4) improve the stability of tree classifiers; (5) appear to be nearly a normal distribution for urban LU classes and make the parametric MLC a reliable classifier. These advantages of V-I-S fractions and LST revealed from the evaluation framework complete a comprehensive assessment of their superiority in the context of urban LU classification.

Second, MLC is comparable with the modern statistical learning classifiers or even outperforms them when the V-I-S fractions and LST are used instead of the direct use of spectral bands. This is consistent with previous reports that MLC could be superior to non-parametric methods if the normality assumption is properly met. In our case, the histogram of V-I-S fractions and LST displays adequate normality, and the use of other complex classifiers becomes unnecessary. It is recommended that MLC be used in conjunction with V-I-S fractions and LST as the input in urban LU classification due to its adequacy and simplicity. However, the MLC is only suitable when V-I-S fractions and LST are used as the only input features. Replacing them with, or including, the band reflectance, regardless of the enrichment of the data input, might degrade the multi-variate normal distribution and subsequently degrade MLC.

Third, the direct use of spectral bands is not recommended for any classification approach being considered, except for the SVM, which maintains a consistently high classification accuracy and is the most robust classifier. The tree classifier trained with multi-spectral band reflectance lacks stability, leaving room for potential improvements through randomization achieved by tree-based ensembles. Specifically, tree-based ensembles increase the classification stability, but no statistically significant accuracy improvement is found. The SVM achieves the highest overall accuracy and stability when multi-spectral band reflectance is the only input. Therefore, SVM is recommended for those urban LU studies that do not use the derivation of V-I-S fractions and LST. The addition of V-I-S fractions and LST to the multi-spectral reflectance contributes to the improvement of ensemble classifiers in the urban LU classification significantly. The tree-based ensemble classifiers and SVM are suitable for cases in which the full data dimension is used.

Despite the fact that the conclusions of this study have some minor limitations on locations and classification schema, this work is among the few that have investigated the stability of V-I-S fractions and LST in urban LU classification. The evaluation framework developed in this article could also be applied to other urban environments and considered for use in the assessment of input features and classifiers in other remote-sensing classification endeavours. Future studies may aim to apply the evaluation framework for the assessment of texture variables and object-oriented

classification to urban LU investigations. A different classification schema may also be considered.

References

ANDERSON, J.R., HARDY, E.E., ROACH, J.T. and WITMER, R.E., 1976, *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. Geological Survey Professional Paper No. 964 (Washington, DC: US Geological Survey).

BAATZ, M. and SCHÄPE, A., 2000, Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informationsverarbeitung*, vol. XII, J. Strobl, T. Blaschke and G. Griesebner (Eds.) (Heidelberg: Wichmann-Verlag).

BARNSLEY, M.J., JENSEN, L.M.L. and BARR, S., 2001, Inferring urban land use by spatial and structural pattern recognition. In *Remote Sensing and Urban Analysis*, J.-P. Donnay, M.J. Barnsley and P.A. Longley (Eds.) (New York: CRC Press).

BERK, A., BERNSTEIN, L.S. and ROBERTSON, D.C., 1989, *MODTRAN: A Moderate Resolution Model for LOWTRAN 7.* Final report GLTR-89-0122. 12 May 1986–April 1989.

BREIMAN, L., 1996, Bagging predictors. *Machine Learning*, **24**, pp. 123–140.

BREIMAN, L., 2001, Random Forests. *Machine Learning*, **45**, pp. 5–32.

BROWN, M., GUNN, S.R. and LEWIS, H.G., 1999, Support vector machines for optimal classification and spectral unmixing. *Ecological Modelling*, **120**, pp. 167–179.

BRUZZONE, L. and CARLIN, L., 2006, A Multilevel context-based system for classification of very high spatial resolution images. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, pp. 2587–2600.

CHAN, J. and PAELINCKX, D., 2008, Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment*, **112**, pp. 2999–3011.

CHAN, J.C.-W., HUANG, C. and DEFRIES, R., 2001, Enhanced algorithm performance for land cover classification from remotely sensed data using bagging and boosting. *IEEE Transactions on Geoscience and Remote Sensing*, **39**, pp. 693–695.

DEFRIES, R.S. and CHAN, J.C.-W., 2000, Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. *Remote Sensing of Environment*, **74**, pp. 503–515.

FOODY, G.M., BOYD, D.S. and SANCHEZ-HERNANDEZ, C., 2007, Mapping a specific class with an ensemble of classifiers. *International Journal of Remote Sensing*, **28**, pp. 1733–1746.

FOODY, G.M. and MATHUR, A., 2004, A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, **42**, pp. 1335–1343.

FREUND, R.J. and WILSON, W.J., 2002, *Statistical Methods*, 2nd ed. (Amsterdam: Elsevier Science).

GISLASON, P.O., BENEDIKTSSON, J.A. and SVEINSSON, J.R., 2006, Random Forests for land cover classification. *Pattern Recognition Letters*, **27**, pp. 294–300.

GONG, P., MARCEAU, D.J. and HOWARTH, P.J., 1992, A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data. *Remote Sensing of Environment*, **40**, pp. 137–151.

HARALICK, R.M., SHANMUGAM, K. and DINSTEIN, I.H., 1973, Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, **6**, pp. 610–621.

HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J., 2009, *The Elements of Statistical Learning, Data Mining, Inference, and Prediction*, 2nd ed. (New York: Springer).

HEROLD, M., LIU, X. and CLARKE, K.C., 2003, Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering & Remote Sensing*, **69**, pp. 991–1001.

HUANG, C., DAVIS, L.S. and TOWNSHEND, J.R.G., 2002, An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, **23**, pp. 725–749.

JIMÉNEZ-MUÑOZ, J.C. and SOBRINO, J.A., 2003, A generalized single-channel method for retrieving land surface temperature from remote sensing data. *Journal of Geophysical Research*, **108**, p. 4688.

LANDSAT PROJECT SCIENCE OFFICE, 2002, *Landsat 7 Science Data User's Handbook* (Washington, DC: Goddard Space Flight Center, NASA).

LEE, S. and LATHROP, R.G., 2006, Subpixel analysis of Landsat ETM+ using self-organizing map (SOM) neural networks for urban land cover characterization. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, pp. 1642–1654.

LO, C.P., QUATTROCHI, D.A. and LUVALL, J.C., 1997, Application of high-resolution thermal infrared remote sensing and GIS to assess the urban heat island effect. *International Journal of Remote Sensing*, **18**, pp. 287–304.

LU, D., MORAN, E. and BATISTELLA, M., 2003, Linear mixture model applied to Amazonian vegetation classification. *Remote Sensing of Environment*, **87**, pp. 456–469.

LU, D. and WENG, Q., 2004, Spectral mixture analysis of the urban landscape in Indianapolis with Landsat ETM+ imagery. *Photogrammetric Engineering & Remote Sensing*, **70**, pp. 1053–1062.

LU, D. and WENG, Q., 2005, Urban classification using full spectral information of Landsat ETM+ imagery in Marion County, Indiana. *Photogrammetric Engineering and Remote Sensing*, **71**, pp. 1275–1284.

LU, D. and WENG, Q., 2006, Use of impervious surface in urban land-use classification. *Remote Sensing of Environment*, **102**, pp. 146–160.

MARGINEANTU, D.D. and DIETTERICH, T.G., 1997, Pruning adaptive boosting. In *Fourteenth International Conference on Machine Learning*, pp. 211–218 (San Francisco, CA: Morgan Kaufmann).

MATHUR, A. and FOODY, G.M., 2008, Crop classification by support vector machine with intelligently selected training data for an operational application. *International Journal of Remote Sensing*, **29**, pp. 2227–2240.

MAYOR'S OFFICE OF ENVIRONMENT AFFAIRS, 2001, *City of New Orleans Baseline Greenhouse Gas Emissions Profile*, pp. 1–11 (New Orleans, LA: City of New Orleans).

MESEV, V., 2010, Classification of urban areas: inferring land use from the interpretation of land cover. In *Remote Sensing of Urban and Suburban Areas, Remote Sensing and Digital Image Processing*, vol. 10, T. Rashed and C. Jürgens, p. 144 (New York: Springer).

MOUNTRAKIS, G., IM, J. and OGOLE, C., 2011, Support vector machines in remote sensing: a review. *ISPRS Journal of Photogrammetry and Remote Sensing*, **66**, pp. 247–259.

MYINT, S.W., GOBER, P., BRAZEL, A., GROSSMAN-CLARKE, S. and WENG, Q., 2011, Per-pixel versus object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, **115**, pp. 1145–1161.

MYINT, S.W. and LAM, N.S.-N., 2005, Examining lacunarity approaches in comparison with fractal and spatial autocorrelation techniques for urban mapping. *Photogrammetric Engineering & Remote Sensing*, **71**, pp. 927–937.

PAL, M. and MATHER, P.M., 2003, An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, **86**, pp. 554–565.

PHINN, S., STANFORD, M., SCARTH, P., MURRAY, A.T. and SHYY, P.T., 2002, Monitoring the composition of urban environments based on the vegetation-impervious surface-soil (VIS) model by subpixel analysis techniques. *International Journal of Remote Sensing*, **23**, pp. 4131–4153.

QIN, Z., KARNIELI, A. and BERLINER, P., 2001, A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel–Egypt border region. *International Journal of Remote Sensing*, **22**, pp. 3719–3746.

RICHARDS, J.A., and JIA, X., 2006, *Remote Sensing Digital Image Analysis: an Introduction*, 4th ed. (Berlin: Springer).

RIDD, M.K., 1995, Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. *International Journal of Remote Sensing*, **16**, pp. 2165–2185.

SCHMUGGE, T., HOOK, S.J. and COLL, C., 1998, Recovering surface temperature and emissivity from thermal infrared multispectral data. *Remote Sensing of Environment*, **65**, pp. 121–131.

SHACKELFORD, A.K. and DAVIS, C.H., 2003, A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 2354–2363.

SMALL, C., 2001, Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing*, **22**, pp. 1305–1334.

SMALL, C. and LU, J.W.T., 2006, Estimation and vicarious validation of urban vegetation abundance by spectral mixture analysis. *Remote Sensing of Environment*, **100**, pp. 441–456.

SOBRINO, J.A., JIMÉNEZ-MUÑOZ, J.C. and PAOLINI, L., 2004, Land surface temperature retrieval from LANDSAT TM 5. *Remote Sensing of Environment*, **90**, pp. 434–440.

TAREK, R., WEEKS, J.R., GADALLA, M.S. and HILL, A.G., 2001, Revealing the anatomy of cities through spectral mixture analysis of multispectral satellite imagery: a case study of the Greater Cairo Region, Egypt. *Geocarto International*, **16**, pp. 4–15.

WARD, D., PHINN, S.R. and MURRAY, A.T., 2000, Monitoring growth in rapidly urbanizing areas using remotely sensed data. *The Professional Geographer*, **52**, pp. 371–386.

WENG, Q., 2001, A remote sensing-GIS evaluation of urban expansion and its impact on surface temperature in the Zhujiang Delta, China. *International Journal of Remote Sensing*, **22**, pp. 1999–2014.

WENG, Q., 2009, Thermal infrared remote sensing for urban climate and environmental studies: methods, applications, and trends. *ISPRS Journal of Photogrammetry and Remote Sensing*, **64**, pp. 335–344.

WENG, Q., LIU, H. and LU, D., 2007, Assessing the effects of land use and land cover patterns on thermal conditions using landscape metrics in city of Indianapolis, United States. *Urban Ecosystems*, **10**, pp. 203–219.

WENG, Q., LU, D. and LIANG, B., 2006, Urban surface biophysical descriptors and land surface temperature variations. *Photogrammetric Engineering and Remote Sensing*, **72**, pp. 1275–1286.

WENG, Q. and QUATTROCHI, D.A., 2007, *Urban Remote Sensing*, 1st ed. (Boca Raton, FL: Taylor & Francis Group, LLC).

WU, C., 2004, Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery. *Remote Sensing of Environment*, **93**, pp. 480–492.

WU, C. and MURRAY, A.T., 2003, Estimating impervious surface distribution by spectral mixture analysis. *Remote Sensing of Environment*, **84**, pp. 493–505.

WU, S.-S., XU, B. and WANG, L., 2006, Urban land-use classification using variogram-based analysis with an aerial photograph. *Photogrammetric Engineering & Remote Sensing*, **72**, pp. 813–822.