Assessing Traditional Classifications Ability in Identification of Flowering Trees in Urban Ecosystems

Nurul Farehah MAHZAN[1,2], Noordyana HASSAN*[1,2], Mazlan HASHIM[1,2], Shinya NUMATA[3]

[1]Geoscience and Digital Earth Centre (INSTeG), Faculty of Geoinformation and Real Estate, Universiti Teknologi Malaysia, 81310 UTM Johor, Johor, Malaysia
[2]Department of Geoinformatics, Faculty of Geoinformation and Real Estate, Universiti Teknologi Malaysia, 81310 UTM Johor, Johor, Malaysia
(email: nurulfarehah329@gmail.com, *noordyana@utm.my, mazlanhashim@utm.my)
[3] Tokyo Metropolitan University, Graduate School of Environmental Science, Department of Tourism Science, Minami-Osawa 1-1, Hachiouji, Tokyo 192-0397, Japan
(email:nmt@tmu.ac.jp)

Abstract

Spring flowering dates of cherry blossoms have been slowly advancing due to global warming and urban heat island effect. Monitoring the changes of cherry tree flowering timing is quite challenging especially when cherry tree were planted in high heterogeneous urban area making it difficult to identify the flowering cherry trees. We hypothesized that Support Vector Machine (SVM) has ability to identify pixels of flowering cherry tree (*Prunus x yedoensis*) better compared to Maximum Likelihood Classifier (MLC). Thus, to test the effectiveness of traditional classification of MLC and SVM, both classifications were assessed by employing on multispectral IKONOS image in identification of flowering cherry trees at small spatial scale and was upscale to large spatial extend. The overall classification accuracy of SVM is 84.5% (kappa coefficient: 0.79) compared with that of using MLC with 54.5% (kappa coefficient: 0.44) at large spatial scale. While 64.5% (kappa coefficient: 0.35) for SVM and 58.8% (kappa coefficient: 0.3) for MLC. The results suggest that SVM perform better classification result compared to MLC at both small and large spatial scales. However, the accuracy of the classification changes according to the scale of the selected area. It is therefore concluded that the study conducted at small spatial scale area offers prominent application for mapping of flowering tree in urban area.

Keywords: Flowering trees, Urban Ecosystems, Traditional Classification, Maximum Likelihood (MLC), Support Vector Machine (SVM)

1 Introduction

Phenological event of the flowering cherry tree in Japan is important as this national flower have a symbolic meaning, strong cultural significance and economically important (Sakurai et al. 2006). However, these cherry tree flowering date is highly depends on the urban climate conditions as some research found that spring blooming tends to bloom earlier in the city compared to the rural area (Neil and Wu 2006). Identification of flowering cherry tree by using ground measurements are time consuming, expensive and resulting to lack of data information. Thus, the potential of remote sensing observation and technique can be apply to overcome the lacking of conventional method to develop an effective monitoring solution for the flowering cherry tree. Typically high spatial resolution multispectral data is used in previous study and mostly focused on green leaves because of its dominant in vegetation and forest monitoring (Clark, Roberts, and Clark 2005; Gross and Heumann 2014). Classification on flowering tree is not often explored due to its weak spectral signal. In contrast, some studies has proved that classification accuracies are much better when using imagery collected during peak flowering time because spectral variation of flower could increase classification accuracy.

Various approaches of traditional classification has been utilized in the past regarding on plant such as extracting vegetation information (Xie, Sha, and Yu 2008), plant disease detection (Kaur and Kang 2016), and monitoring urban greenness (Gan et al. 2014). Traditional classification works by selecting that class label with the greatest likelihood of being correct and only assigned one class per pixel because the feature space decision boundaries for this method are well-defined (Schowengerdt 2007). The most popular conventional method used is MLC and it is implemented in many cases regarding on plant. In addition, SVM also gradually becomes popular because this method able to find the optimal separating hyperplane between classes effectively. Even with small training samples, SVM can works very well and this yield to high accuracy output (Horning et al. 2014; Zheng et al. 2015; Gil et al. 2013).

To address the issues of finding suitable method to be used, this study hypothesize that SVM is able to identify pixels of flowering cherry tree better than MLC. Therefore, this study was carried out by (i) exploring the ability of traditional classification approach in identification of flowering trees at small spatial scale, and (ii) to apply the classification at a larger spatial scale. Thus, this study is carried out to identify Somei Yoshino (Prunus × yedoensis) cherry tree in urban ecosystems with the utilization of MLC and SVM. All of the classification used were compared through the accuracy assessment obtained by using high-spatial resolution IKONOS imagery.

2 Materials and Methodology

2.1 Study area

The study area of this research is located at Tama and Hachioji city in the state of Tokyo. The study area was confined to Yanagisawanoike Park in Hachioji City (35.6154° N, 139.3767° E) with altitude of 128 m (Figure 1). This park is one of the popular spot of cherry blossom viewing which attract locals and visitors every year in the west parts of Tokyo, Japan. Most of the main tree species found in this park is known as Somei Yoshino (Prunus × yedoensis), a deciduous cherry tree. There are some other species like Kanzakura (Prunus sato-zakura “Sekiyama”), Mamezakura (Prunus incisa), and Shidarezakura (Prunus sapchiana), as well as other deciduous trees, such as Japanese red pine (Pinus densiflora) and hornbeam (Carpinus laxifolia), and evergreen trees, including camphor (Cinnamomum camphora), Chinese evergreen oak (Quercus mysinaefolia), and Japanese black pine (Pinus thunbergii). The mean canopy size of flowering SY tree during full bloom was 5 meters and the mean height of the tree was 3 meters

2.2 Materials

2.2.1 Remotely sensed satellite data

In this study, multispectral IKONOS imagery were used. The image has a resolution of 4-m and consists of 4 spectral bands: blue (445–516 nm), green (506–595 nm), red (632–698 nm), and near infra-red (NIR; 752–853 nm). The data chosen is dated on April 1, 2006 as the cherry tree was in full bloom at that time based on the information given by Japanese Meteorological Agency (JMA). The data obtained were already applied with radiometric and geometric correction.

2.2.2 In-situ data collection

In-situ data collection that has been collected were used in this study. The ground data of Somei Yoshino trees, soil, dry grass, and evergreen trees using a handheld GPS unit (Garmin GPSmap 60CSx) has been collected on April 1, 2014 by Hassan (2015). These data were used to assess classification accuracy.

2.3 Methodology

2.3.1 Maximum Likelihood Classifier (MLC)

The information about land cover was then extracted using traditional classification method, maximum likelihood classifier (MLC). The classification is used because of its probability based decision rule to assign each pixel based on sample data (Shafri, Suhaili, and Mansor 2007). Samples of classes (flowering SY tree, dry grass, soil, and evergreen tree) were randomly selected from known areas by using the region of interest (ROI) tools based on Google Earth, experience and knowledge about the actual distribution of land cover types of the previous researcher. Then, MLC was conducted on the combination of texture (contrast) and spectral images with eight bands to improve the classification accuracy.

2.3.2 Support Vector Machine (SVM)

SVM was chosen because of its capability to find optimal hyperplane that increases the separation among classes which can minimize misclassification, only needs minimal training area, and requires less processing time [15], [16]. Training samples of flowering SY tree, dry grass, soil, and evergreen tree created before were used to perform image classification using SVM. SVM classification were configured with default setting of polynomial kernel types to classify all classes using a combination of texture (contrast) and spectral images as well.

3 Results

The overall accuracy for MLC of IKONOS image for small spatial scale area – Yanagisawanoike Park is 58.82% with kappa coefficient is 0.3045 (Table 1). While overall accuracy for by using SVM on the same area is 64.58% with kappa coefficient is 0.3539. Kappa coefficient of both classifications at small spatial scale obtained in this study shows a poor agreement. Besides, MLC and SVM producer's accuracy of flowering SY tree were 63.89% and 70.59%. User's accuracy of MLC classification of flowering SY tree (88.46%) was higher than SVM classification (85.71%). The overall accuracy and kappa coefficient of the flowering SY tree land cover was improved by using SVM classifier compared to MLC classifiers. The overall accuracy improved about 5.76% from MLC to SVM classifiers while kappa coefficient increased about 0.05% when using SVM. This indicates that SVM classified flowering SY tree in urban ecosystem much better than MLC (Figure 1). In contrast with other traditional classifiers, SVM is capable to deal with high dimensionally data and thus high accuracy can be achieved although small training samples were used. Besides, SVM able to generate optimally separate hyperplanes each target classes can be classified better (Otukei and Blaschke 2010). However, the user's accuracy of flowering SY tree for SVM slightly lower than MLC.

Table 1 Accuracy assessment for Maximum Likelihood (ML) and Support Vector Machine (SVM) classifications of flowering SY trees in Yanagisawanoike Park. The results was compared based on the overall accuracy, kappa coefficient, and producer's accuracy.

	MLC				**SVM**			
	Overall Accuracy (100%)	**Kappa Coefficient**	**User's Accuracy (%)**	**Producer's Accuracy (%)**	**Overall Accuracy (100%)**	**Kappa Coefficient**	**User's Accuracy (%)**	**Producer's Accuracy (%)**
Somei Yoshino (SY)			88.46	63.89			85.71	70.59
Evergreen Tree	58.82	0.3045	25.00	40.00	64.58	0.3539	27.27	75.00
Dry Grass			22.22	40.00			100.00	40.00
Soil			37.50	60.00			28.57	40.00

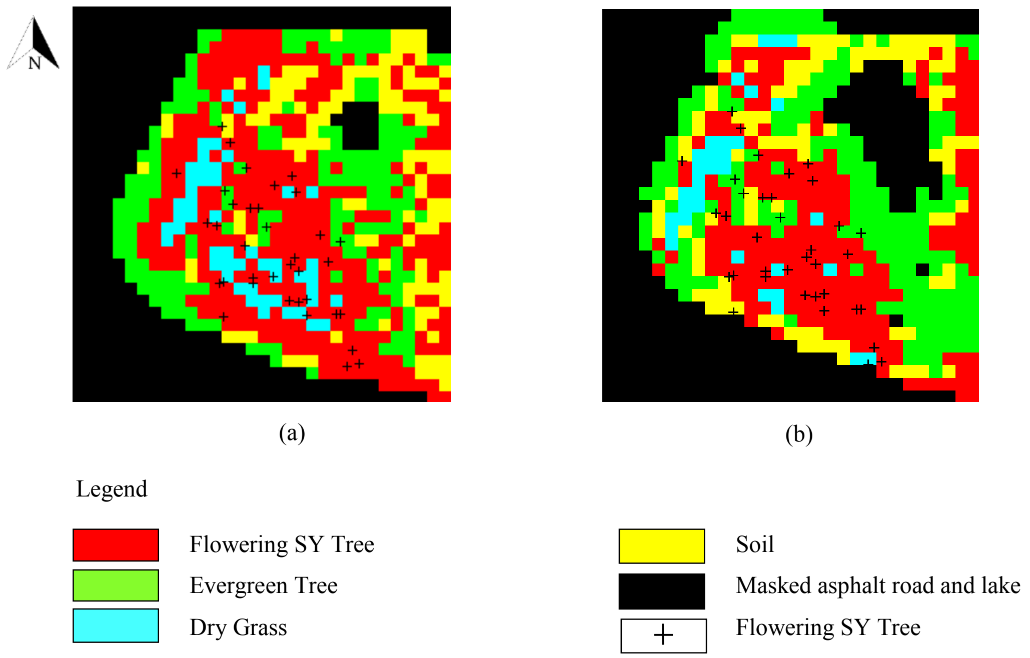

Figure 1 Comparison of classification results of land cover for small spatial scale - Yanagisawanoike Park by using different classification approaches: (a) ML classification of the IKONOS image, (b) SVM classification of the IKONOS image.

Meanwhile, the MLC and SVM classifications accuracies at large spatial area (Tama and Hachioji City) where MLC achieved an overall accuracy of 54.55% with kappa coefficient of 0.4355. Producer's accuracy of flowering SY tree were 86.05%, while user's accuracy were 60.66%. In comparison, SVM produced overall accuracy of 84.50% with kappa coefficient of 0.7923. Producer's accuracies and user's accuracies of flowering SY tree were 73.33% and 95.65% respectively (refer Table 2). Thus, in this case the SVM outperformed MLC in term of overall accuracy, kappa coefficient, and user's accuracy very well. However, the value of producer's accuracies of MLC outperformed when compare with SVM. Thus, this result shows that SVM classifier is much better than MLC at large spatial scale area. The overall accuracy and kappa coefficient of flowering SY tree is greatly increased by 29.95% from MLC to SVM. But, the producer's accuracy of flowering SY tree using SVM decreased to 12.72% than using MLC. This is may be because of misclassification of flowering SY tree pixels as evergreen tree and dry grass which can be visibly observed in Figure 2(b).

Table 2 Accuracy assessment for Maximum Likelihood (ML) and Support Vector Machine (SVM) classifications of flowering SY trees in Yanagisawanoike Park. The results was compared based on the overall accuracy, kappa coefficient, and producer's accuracy.

	Large Spatial Scale - Tama and Hachioji City							
	MLC				**SVM**			
	Overall Accuracy (100%)	**Kappa Coefficient**	**User's Accuracy (%)**	**Producer's Accuracy (%)**	**Overall Accuracy (100%)**	**Kappa Coefficient**	**User's Accuracy (%)**	**Producer's Accuracy (%)**
Somei Yoshino (SY)			60.66	86.05			95.65	73.33
Evergreen Tree	54.55	0.4355	100.00	22.73	84.50	0.7923	90.00	100.00
Dry Grass			70.45	67.39			68.63	93.10
Soil			40.00	85.71			100.00	71.43

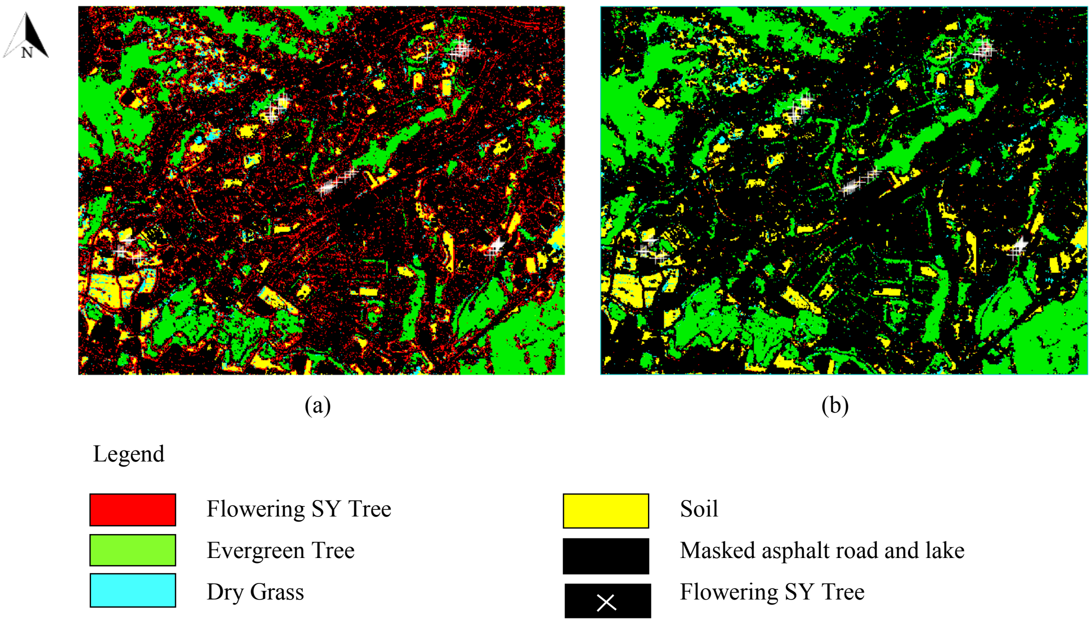

Figure 2 Comparison of classification results of land cover for large spatial scale – Tama and Hachioji City by using different classification approaches: (a) ML classification of the IKONOS image, (b) SVM classification of the IKONOS image.

4 Discussion

Results suggest that classified flowering SY using MLC can be improved by using SVM at small spatial scale area (Table 1 and 2). However, the user's accuracy of flowering SY tree for SVM is lower than MLC. This is may be due to the confusion of flowering SY tree with evergreen tree and soil which lower the values of user's accuracy of SVM. The confusion due to weak spectral response from flower compare to evergreen tree and dry grass. Thus, to overcome this advanced technique of classification such as Mixture Tuned Matched Filtering (MTMF) and Linear Spectral Unmixing (LSU) can be consider especially study area is located in high heterogeneous urban area (Abdel-Rahman et al. 2015; Andrew and Ustin 2008).

The occurrences of misclassification at large spatial scale is high especially in classifying flowering SY trees in high heterogeneity urban park as it was classified as evergreen trees and dry grass. This misclassification is due to limited number of bands (four bands) of IKONOS (Figure 2). Therefore, there was a limitation in distinguishing flowering SY trees, evergreen trees and dry grass especially when employing MLC. MLC has its own characteristics to selects the class label with the greatest likelihood of being correct and unambiguously assigns each pixel to a single class (Foody 2002; Hassan et al. 2015). In addition, the possibility of overlapping between classes increased when the number of classes increased. Thus, additional band is required to distinguish the flowering SY tree with other features (Hassan et al. 2015).

Meanwhile, results on two different spatial scales tested in this study suggest that it is important to conduct a study at small and large spatial scale even the performance of classification applied in small spatial scale is lower than large spatial scale (Table 1). At large spatial scale, the result obviously will describes the abstract features on earth. Besides, better understanding can be obtained on the changes and general trend in large spatial scale study area (Wu and Li 2009). While study area at small spatial scale can portray the details of that particular area. Through this, researcher can focused on the target features and give rational explanations through small spatial scale studies (Wu and Li 2009). It also helps in predicting or monitoring of target feature when it comes to small spatial scale.

5 Conclusion

Results of this study suggest that SVM is more advantageous than MLC in classifying flowering SY trees in urban park at both small and large spatial scales. However, the performance of SVM was limited due to limited number of IKONOS number of bands. Thus, hyperspectral imagery utilization for example AVIRIS would be helpful in improving SVM classification on flowering trees in urban park.

Acknowledgement

We gratefully acknowledge a Fundamental Research Grant Scheme (FRGS)- Spectral Characterisation of Triggering Biophysical Properties of General Flowering using Satellite Remote Sensing Data -R.J130000.7827.4F902, Ministry of Higher Education Malaysia, Universiti Teknologi Malaysia and Tokyo Metropolitan University, the sponsor of this study, and for the facilities provided to conduct this study.

References

Abdel-Rahman, Elfatih M., David M. Makori, Tobias Landmann, Rami Piiroinen, Seif Gasim, Petri Pellikka, and Suresh K. Raina. 2015. "The Utility of AISA Eagle Hyperspectral Data and Random Forest Classifier for Flower Mapping." *Remote Sensing* 7 (10): 13298–318. doi:10.3390/rs71013298.

Andrew, Margaret E., and Susan L. Ustin. 2008. "The Role of Environmental Context in Mapping Invasive Plants with Hyperspectral Image Data." *Remote Sensing of Environment* 112 (12): 4301–17. doi:10.1016/j.rse.2008.07.016.

Clark, M, D Roberts, and D Clark. 2005. "Hyperspectral Discrimination of Tropical Rain Forest Tree Species at Leaf to Crown Scales." *Remote Sensing of Environment* 96 (3–4): 375–98. doi:10.1016/j.rse.2005.03.009.

Foody, G M. 2002. "Hard and Soft Classifications by a Neural Network with a Non-Exhaustively Defined Set of Classes." *International Journal of Remote Sensing* 23 (18). Taylor & Francis: 3853–64. doi:10.1080/01431160110109570.

Gan, Muye, Jinsong Deng, Xinyu Zheng, Yang Hong, and Ke Wang. 2014. "Monitoring Urban Greenness Dynamics Using Multiple Endmember Spectral Mixture Analysis." *PLoS ONE* 9 (11). doi:10.1371/journal.pone.0112202.

Gil, Artur, Agustin Lobo, Mohamed Abadi, Luis Silva, and Helena Calado. 2013. “Mapping Invasive Woody Plants in Azores Protected Areas by Using Very High-Resolution Multispectral Imagery.” *European Journal of Remote Sensing* 46 (1): 289–304. doi:10.5721/EuJRS20134616.

Gross, John W, and Benjamin W Heumann. 2014. “Can Flowers Provide Better Spectral Discrimination between Herbaceous Wetland Species than Leaves ?,” no. February 2015: 37–41. doi:10.1080/2150704X.2014.973077.

Hassan, Noordyana, Shinya Numata, Tetsuro Hosaka, and Mazlan Hashim. 2015. “Remote Detection of Flowering Somei Yoshino (Prunus × Yedoensis) in an Urban Park Using IKONOS Imagery: Comparison of Hard and Soft Classifiers.” *Journal of Applied Remote Sensing* 9 (1): 96046. doi:10.1117/1.JRS.9.096046.

Horning, N., Andy Liaw, Matthew Wiener, Yuguo Qian, Weiqi Zhou, Jingli Yan, Weifeng Li, et al. 2014. “Support Vector Machine Classification of Object-Based Data for Crop Mapping , Using Multi-Temporal Landsat Imagery.” *R News* XXXIX (1): 54.1-54.10. doi:10.5244/C.22.54.

Kaur, Rajleen, and Sandeep Singh Kang. 2016. “An Enhancement in Classifier Support Vector Machine to Improve Plant Disease Detection.” In *Proceedings of the 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education, MITE 2015*, 135–40. doi:10.1109/MITE.2015.7375303.

Neil, Kaesha, and Jianguo Wu. 2006. “Effects of Urbanization on Plant Flowering Phenology: A Review.” *Urban Ecosystems* 9 (3): 243–57. doi:10.1007/s11252-006-9354-2.

Otukei, J. R., and T. Blaschke. 2010. “Land Cover Change Assessment Using Decision Trees, Support Vector Machines and Maximum Likelihood Classification Algorithms.” *International Journal of Applied Earth Observation and Geoinformation* 12 (SUPPL. 1). doi:10.1016/j.jag.2009.11.002.

Sakurai R., Jacobson S. K., Hiromi Kobori Richard Primack Kohei Oka Naoya Komatsu Ryo Machida. 2011. “Culture and Climate Change: Japanese Cherry Blossom Festivals and Stakeholders’ Knowledge and Attitudes about Global Climate Change.” *Biological Conservation* 144 (1). Elsevier: 654–58. doi:10.1016/J.BIOCON.2010.09.028.

Schowengerdt, Robert A. 2007. “Remote Sensing: Models and Methods for Image Processing.” *Remote Sensing*, 560. doi:10.1016/B978-012369407-2/50014-0.

Shafri, Helmi Zulhaidi Mohd, Affendi Suhaili, and Shattri Mansor. 2007. “The Performance of Maximum Likelihood, Spectral Angle Mapper, Neural Network and Decision Tree Classifiers in Hyperspectral Image Analysis.” *Journal of Computer Science* 3 (6): 419–23. doi:10.3844/jcssp.2007.419.423.

Wu, Hua, and Zhao Liang Li. 2009. “Scale Issues in Remote Sensing: A Review on Analysis, Processing and Modeling.” *Sensors*. doi:10.3390/s90301768.

Xie, Y., Z. Sha, and M. Yu. 2008. “Remote Sensing Imagery in Vegetation Mapping: A Review.” *Journal of Plant Ecology* 1 (1): 9–23. doi:10.1093/jpe/rtm005.

Zheng, Baojuan, Soe W. Myint, Prasad S. Thenkabail, and Rimjhim M. Aggarwal. 2015. “A Support Vector Machine to Identify Irrigated Crop Types Using Time-Series Landsat NDVI Data.” *International Journal of Applied Earth Observation and Geoinformation* 34 (1): 103–12. doi:10.1016/j.jag.2014.07.002.