34th Asian Conference on Remote Sensing 2013
20-24 October 2013, Bali, Indonesia

URBAN LAND COVER MAPPING USING OPTICAL AND SAR IMAGES

D.Amarsaikhan[1]*, S.Erdenesukh[2], V.Battsengel[2], and V.Batsaikhan[3]
[1]Institute of Informatics, Mongolian Academy of Sciences
av.Enkhtaivan-54B, Ulaanbaatar-51, Mongolia
[2]School of Geography and Geology, National University of Mongolia
Ikh Surguuliin gudamj-6, Ulaanbaatar-46, Mongolia
[3]Disaster Research Institute, National Emergency Management Agency of Mongolia
Partizany gudamj-6, Ulaanbaatar, Mongolia

*Corresponding author: amar64@arvis.ac.mn

ABSTRACT
The aim of this study is to conduct urban land cover mapping in Mongolia using optical and synthetic aperture radar (SAR) images. To extract the reliable urban land cover information from the available remote sensing (RS) data sets, a refined maximum likelihood classification algorithm that uses spectral and spatial thresholds defined from the local knowledge is constructed. Overall the research indicates that the spectral and spatial thresholds defined from the local knowledge could considerably improve the performance of the classification.

Keywords: optical, SAR, urban land cover, classification

INTRODUCTION
Generally, high resolution optical RS data sets taken from different Earth observation satellites such as Landsat and SPOT have been successfully used for a land-cover mapping since the operation of the first Landsat launched in 1972, whereas high resolution SAR images taken from space platforms have been widely used for different thematic applications since the launch of the ERS-1/2, JERS-1 and RADARSAT satellites (Amarsaikhan *et al.* 2004). It is clear that the combined application of optical and SAR data sets can provide unique information for different thematic studies, because passive sensor images will represent spectral variations of various surface features, whereas microwave data with their penetrating capabilities can provide some additional information (Amarsaikhan *et al.* 2007).

Traditionally, multispectral RS data sets have been widely used for a land cover mapping and for the generation of land cover information, different supervised and unsupervised classification methods have been applied. However, the emergence of microwave images has given new opportunities for the users and researchers dealing with processing and analysis of remotely sensed data sets. Unlike single-source data, data sets from multiple sources have proved to offer better potential for discriminating between different land cover types. Many authors have assessed the potential of multisource images for the classification of different land cover classes (Serpico *et al.* 1995, Benediktsson *et al.* 1997, Amarsaikhan and Douglas 2004, Wu *et al.* 2009).

Recently, mapping of urban areas, specifically at regional and global scales has become an important task due to the increasing pressures from rapid urbanization processes (Cao *et al.* 2009). However, in most cases urban areas are complex and diverse in nature and many features have similar spectral characteristics and it is not easy to separate them by the use of common feature combinations or by applying ordinary techniques. In order to successfully extract urban land cover classes, reliable features derived from multiple sources and an efficient classification technique should be used (Amarsaikhan *et al.* 2012).

The aim of this study is to classify the multisource features using a refined maximum likelihood classifier. For the final analysis, optical and microwave data sets of Ulaanbaatar, the capital city of Mongolia have been used. The analysis was carried out using PC-based ERDAS Imagine 10.1 and ENVI 4.8.

TEST SITE AND DATA SOURCES

As a test site, Ulaanbaatar, the capital city of Mongolia has been selected. Ulaanbaatar is situated in the central part of Mongolia, on the Tuul River, at an average height of 1350m above sea level and currently has about 1.28 million inhabitants. Although, the city is extended from the west to the east about 30km, and from the north to the south about 20km, the study area chosen for the present study covers an area of 11.3km long and 8.7km wide. In the selected image frame, it is possible to define such classes as built-up area, ger area, forest, grass, soil and water. The built-up area includes buildings of different sizes, while ger area includes mainly gers surrounded by fences. Figure 1 shows a Landsat image of the test site, and some examples of its land cover.

In the present study, for the urban land cover studies, a Landsat TM image of 31 July 2010 and an Envisat C-band image of 25 May 2010 have been used. The Landsat ETM+ data has seven multispectral bands (B1: 0.45–0.52μm, B2: 0.52–0.60μm, B3: 0.63–0.69μm, B4: 0.76–0.90μm, B5: 1.55–1.75μm, B6: 10.40–12.50μm and B7: 2.08–2.35μm). The spatial resolution is 30 m for the reflective bands, while it is 120 m for the thermal band. In the current study, channels 2,3,4,5,7 have been used. The Envisat is a European satellite carrying a cloud-piercing, all weather free polarimetric radar which is designed to monitor the Earth from a distance of about 790km. The characteristics of the Envisat data used in the current study are shown in Table 1.

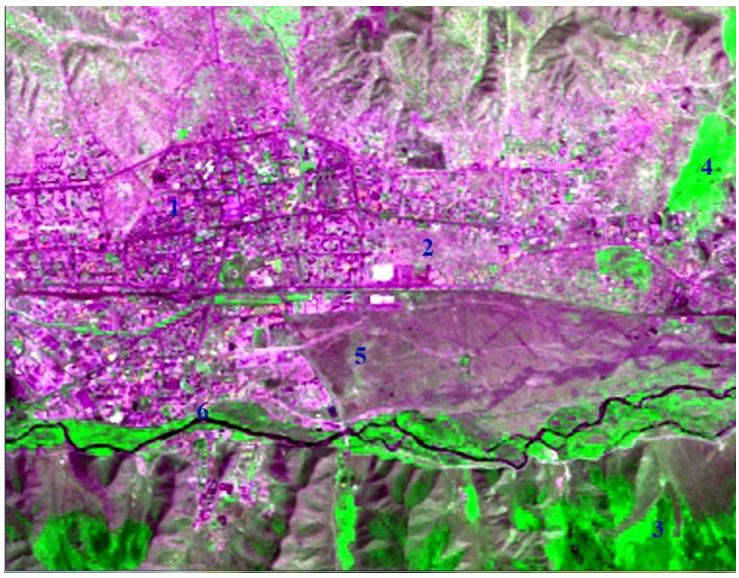

Figure 1. 2010 Landsat image of the selected part of Ulaanbaatar.
1-built-up area; 2-ger area; 3-forest; 4-grass; 5-soil; 6- water

Table 1.The characteristics of the Envisat data

Parameter	C-band
Polarization	HH
Frequency	5.36GHz
Wavelength	5.6cm
Spatial resolution	30m

COREGISTRATION OF THE IMAGES & SPECKLE SUPPRESSION OF THE SAR DATA

In order to perform accurate data fusion, good geometric correlation between the images are needed. As a first step, the Landsat image was georeferenced to a Gauss-Kruger map projection using 12 ground control points (GCPs) defined from a topographic map of the study area. The GCPs have beenselected on clearly delineated crossings of roads, streets and city building corners. For the transformation, a second-order transformation and nearest-neighbour resampling approach were applied and the related rootmean square (RMS) error was 0.83 pixel. Then, the Envisat image was geometrically corrected and its coordinates were transformed to the coordinates of the georeferenced

Landsat image. In order to correct the SAR image, 18more regularly distributed GCPs were selected from different parts of the image. For the actual transformation, a second-order transformation was used. As a resampling technique, the nearest-neighbour resampling approach was applied and the related RMS error was 1.16 pixel.

As microwave images have a granular appearance due to the speckle formed as a result of the coherent radiation used for radar systems; the reduction of the speckle is a very important step in further analysis. The analysis of the radar images must be based on the techniques that remove the speckle effects while considering the intrinsic texture of the image frame (ENVI 1999). In this study, four different speckle suppression techniques such as local region, lee-sigma, frost and gammamap filters of 3x3 and 5x5 sizes were compared in terms of delineation of urban features and texture information. After visual inspection of each image, it was found that the 3x3 gammamap filter (ERDAS 1999) created the best image in terms of delineation of different features as well as preserving content of texture information. In the output image, speckle noise was reduced with very low degradation of the textural information.

STANDARD MAXIMUM LIKELIHOOD CLASSIFICATION

Initially, to define the sites for the training signature selection from the combined optical and microwave images, several areas of interest (AOI) representing the available six classes (built-up area, ger area, forest, grass, soil and water) have been selected through accurate analysis. The separability of the training signatures was firstly checked in feature space and then evaluated using Jeffries–Matusita distance. The values of Jeffries–Matusita distance range from 0 to 2.0 and indicate how well the selected pairs are statistically separate. The values greater than 1.9 indicate that the pairs have good separability. After the investigation, the samples that demonstrated the greatest separability were chosen to form the final signatures.

The final signatures included about 80–827 pixels. For the classification, the following feature combinations were used:

1. All original spectral bands of the Landsat TM image.
2. Red, near infrared and first middle infrared (i.e., 3,4 and 5) bands of the Landsat TM image.
3. The Envisat and original spectral bands of the Landsat TM data.

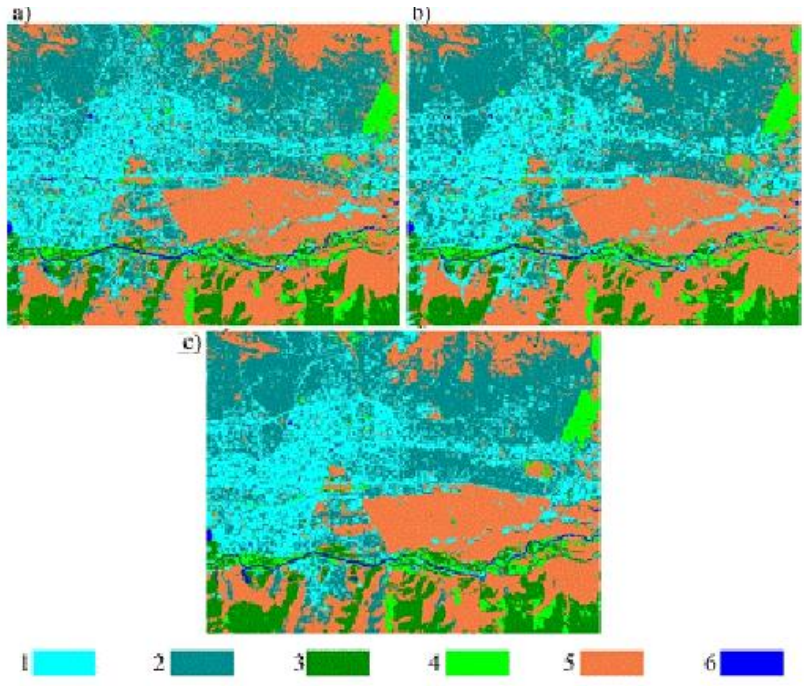

Figure 2. Comparison of the standard classification results for the selected classes (1-built-up area; 2-ger area; 3-forest; 4-grass; 5-soil; 6-water). Classified images (a) using Landsat TM bands, (b) using bands 3, 4 and 5 of the Landsat TM image, (c) using multisource bands.

For the actual classification, a maximum likelihood classifier (Richards and Xia 1999) has been used. The final classified images are shown in figure 2(a–c). As seen from figure 2(a–c), the classification result of all bands of Landsat TM gives the worst result, because there are high overlaps among two urban classes: built-up area and ger area. However, these overlaps decrease on the classified image of red and infrared bands. It can be explained by a fact that a fewer bands with statistically separable features can produce a better result than many bands with high overlaps. As could be seen, although multisource images give some improvement, it is still very difficult to obtain a reliable land cover map by the use of the standard technique, specifically on decision boundaries of the statistically overlapping classes.

For the accuracy assessment of the classification results, the overall performance has been used. This approach creates a confusion matrix in which reference pixels are compared with the classified pixels and as a result an accuracy report is generated indicating the percentages of the overall accuracy (Mather and Koh 2011). As ground truth information, different AOIs containing 1239 purest pixels have been selected. AOIs were selected on a principle that more pixels to be selected for the evaluation of the larger classes such as ger area and soil than the smaller classes such as water. The overall classification accuracies for the selected classes are shown in table 2.

Table 2. The overall classification accuracy of the classified images

No	The used band combinations	Overall accuracy (%)
1	All Landsat TM bands	84.12
2	Bands 3, 4 and 5 of the Landsat TM	86.29
4	Optical and SAR bands	89.37

THE REFINED MAXIMUM LIKELIHOOD CLASSIFICATION

For many years, single-source multispectral data sets have been efficiently used for a land cover mapping. Since the appearance of the first single polarization microwave data sets, multisource images have proved to offer better potential for discriminating between different land cover types. In general, it is very important to design an appropriate image processing procedure in order to successfully classify any digital data into a number of class labels. The effective use of different features derived from different sources and the selection of a reliable classification technique can be a key significance for the improvement of classification accuracy (Solberg *et al.* 1996, Amarsaikhan and Douglas 2004, Amarsaikhan *et al.* 2012, Laurin *et al.* 2013). In the present study, for the classification of urban land cover types, a refined algorithm has been constructed. As the features, bands 3,4,5 and 7 of Landsat TM and Envisat HH polarization images have been used. The green band (i.e., band 2) of Landsat was excluded, because it had a high correlation with the red band (i.e., band 3).

Unlike the traditional classification, the constructed classification algorithm uses spectral and spatial thresholds defined from the local knowledge. The local knowledge was defined on the basis of the spectral variations of the land surface features on the fused images. It is clear that a spectral classifier will be ineffective if applied to the statistically overlapping classes such as built-up area and ger area because they have very similar spectral characteristics. For such spectrally mixed classes, classification accuracies should be improved if the spectral and spatial thresholds could be incorporated into the classification decision-making. The idea of the spatial threshold is that it uses a polygon boundary to separate the overlapping classes and only the pixels falling within the threshold boundary are used for the classification. In that case, the likelihood of the pixels to be correctly classified will significantly increase, because the pixels belonging to the class that overlaps with the class to be classified using the threshold boundary are temporarily excluded from the decision making process. In such a way, the image can be classified several times using different threshold boundaries and the results can be merged (Amarsaikhan *et al.* 2012).

In the present study, the spatial thresholds have been applied for differentiation of the spectrally similar classes such as built-up area and ger area as well as forest and grass. The result of the classification using the refined method is shown in figure 3. For the accuracy assessment of the classification result, the overall performance has been used, taking the same number of sample points (i.e., 1239 purest pixels) as in the previous standard classifications. The confusion matrix produced for

the refined classification method showed overall accuracy of 91.83%. As could be seen from figure 3, the result of the classification using the refined classifier is much better than results of the standard method. As the overall accuracy exceeds 90%, this kind of result can be directly used for a spatial decision-making or update a thematic layer within a spatial information system.

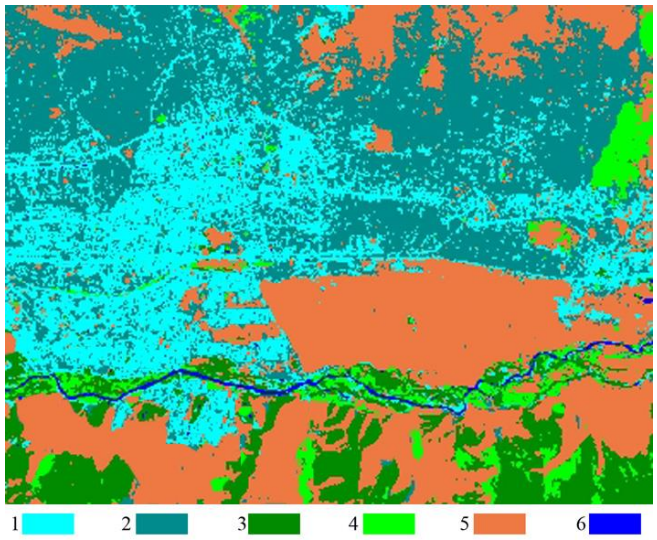

Figure 3. Classification result obtained by the refined method (1-built-up area; 2-ger area; 3-forest; 4-grass; 5-soil; 6-water).

CONCLUSIONS

The aim of the research was to classify urban land cover types and construct a refined maximum likelihood classification algorithm that uses spectral and spatial thresholds defined from the local knowledge. For this purpose, as a test area Ulaanbaatar city was selected and as data sources multispectral Landsat and Envisat SAR images were used. Overall, the study demonstrated that multisource information can considerably improve the classification of land cover types and the refined maximum likelihood classifier is a powerful tool to produce a reliable land cover map.

REFERENCES

[1] Amarsaikhan, D. and Douglas, T., 2004, Data fusion and multisource data classification, International Journal of Remote Sensing, Vol.25, No.17, pp.3529-3539.

[2] Amarsaikhan, D., Ganzorig, M., Batbayar, G., Narangerel, D. and Tumentsetseg, S.H., 2004, An integrated approach of optical and SAR images for forest change study. Asian Journal of Geoinformatics, 3, pp.27–33.

[3] Amarsaikhan, D.,Ganzorig,M., Ache, P. and Blotevogel, H., 2007, The Integrated Use of Optical and InSAR Data for Urban Land Cover Mapping, International Journal of Remote Sensing, 28, pp.1161-1171.

[4] Amarsaikhan, D., Ganzorig, M., Saandar, M., Blotevogel, H.H., Egshiglen, E., Gantuya, R., Nergui, B. and Enkhjargal, D., 2012, Comparison of multisource image fusion methods and land cover classification, International Journal of Remote Sensing, Vol.33, No.8, pp.2532-2550.

[5] Benediktsson, J. A., Sveinsson, J. R., Atkinson, P. M., and Tatnali, A., 1997, Feature extraction for multisource data classification with artificial neural networks.International Journal of Remote Sensing, 18, pp.727–740.

[6] Cao, X., Chen, J., Imura, H. and Higashi, O., 2009, A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data, Remote Sensing of Environment, 10, pp.2205-2209.

[7] ENVI, 1999, User's Guide, Research Systems.

[8] ERDAS 1999, Field Guide, 5th edn (Atlanta, Georgia: ERDAS, Inc.).

[9] pp.435–456.

[10] Laurin, G.V., Liesenberg, V., Chen, Q., Guerriero, L., Frate, F.D., Bartolini, A., Coomes, D., Wilebore, B., Lindsell, J. and Valentini, R., 2013, Optical and SAR sensor synergies for forest and

land cover mapping in a tropical site in West Africa, International Journal of Applied Earth Observation and Geoinformation, 21, pp.7-16.

[11] Mather, P.M. and Koh, M., 2011, Computer Processing of Remotely-Sensed Images: an Introduction, Fourth edition (Wiley-Blackwell).

[12] Richards, J. A. and Xia, S., 1999, Remote Sensing Digital Image Analysis—An Introduction, 3rdedn(Berlin: Springer-Verlag).

[13] Serpico, S. B. and Roli, F., 1995, Classification of multisensor remote sensing images bystructural neural networks. IEEE Transactions on Geoscience and Remote Sensing, 33, pp.562–578.

[14] Solberg, A. H. S., Taxt, T. and Jain, A. K., 1996, A Markov random field model forclassification of multisource satellite imagery. IEEE Transactions on Geoscience and Remote Sensing, 34, pp.100–112.

[15] Wu, Z., Yi, L. and Zhang, 2009, Uncertainty analysis of object location in multisource remote sensing imagery, International Journal of Remote Sensing, Vol.30, No.20, pp.5473-5487.