

GENETIC ALGORITHM OPTIMIZED NEURAL NETWORK CLASSIFICATION AND OVERLAPPING PIXEL CHANGE DETECTION BASED ON REMOTE SENSING FOR URBAN SPRAWL: A CASE STUDY IN JIADING DISTRICT OF SHANGHAI, CHINA

X. Zhang*, X.H., Tong, M.L., Liu

Department of Surveying and Geo-informatics, Tongji University, No.1239 of Siping Road, Shanghai, 200092, China
ejau@163.com

KEY WORDS: Multi-Spectral Remote Sensing, Landuse, Image Understanding, Change Detection, Urban Planning, Back Propagation algorithm, Genetic Algorithm, Jiading District

ABSTRACT:

Taking Jiading district of Shanghai as the study area, which is one of the most typical urbanization fringe regions in Shanghai during the past decades, the urban sprawl in this region from 1989 to 2006 was studied on the basis of multi-spectral remotely sensed images. Multi-source data including four epochs of representative TM images (1989, 1995, 2001 and 2006) and the vector topographic map were used in our study. A genetic algorithm optimized back propagation neural network approach was first proposed in our study to classify sorts of land use types from the four epochs of remotely sensed images. The accuracy of the classification was assessed for the four remotely sensed images. The urban land use type was thus extracted based on the land use classification result and three urban land use change images were correspondingly derived from the extracted urban land use type. Based on this, the study area was divided into four plates: the central plate, the southern plate, the northern plate and the western plate, and the detailed temporal-spatial urban land use changes in the study area were investigated in each plate by using overlapping pixel comparison based change detection method during the three time intervals: 1989-1995, 1995-2001 and 2001-2006. From the change detection analysis, the process and pattern of the urban land use change in the Jiading district was finally revealed during nearly 20 years in our study.

1. INTRODUCTION

For better planning of future urban development, urban planners and policy makers need to acquire up-to-date and systematic information on urban land and the pattern of its change. Reliable urban sprawl detection and prediction are needed for long-term planning purposes (Jat et al. 2007). Since satellite remote sensing provides multi-spectral and multi-temporal data that can be used to quantify the type, amount and location of land use change, many efforts have been made to investigate the mapping and monitoring of land use using remote sensing techniques in the past decades (Vogelmann et al. 1998, Masek et al. 2000, Epstein et al. 2002, Ninomiya et al. 2005). Two key points involved in the study of land use using remote sensing are classification and change detection. Many types of conventional classification approaches based on statistical theories have been employed to capture land use information, such as minimum distance classifier and maximum likelihood classifier. In recent years, artificial neural networks (ANNs) have been widely applied for remote sensing classification because of their strong capabilities in nonlinear mapping, self-organization, and self-learning (Omatu and Yoshida 1991, Benediktsson et al. 1990). This is especially so for the back propagation neural network (BPN) (Benediktsson et al. 1990, Heermann and Khazenie 1992, Toshniwal 2005). In spite of its popularity as a powerful tool for image classification, there are some insufficiencies of the BPN such as slow constringency speed and local extrema (Foody and Arora 1997). Although many improvements have been made by Benediktsson et al. (1990), Toshniwal (2005), etc., there are still many drawbacks

of the BPN.

With regard to the problems of slow constringency speed and local extrema in the BPN training, a complete substitution by a genetic algorithm (GA) might be advantageous due to the strong robustness and high efficiencies of GA combining with the error backward transferring mechanism of BP algorithm. Some investigations into BPN training using GAs have been reported (Rojas et al. 2002, Nicolás et al. 2006, Liu et al. 2004). Liu et al. (2004) employed a real coded GA to evolve the connection weight vectors between the nodes of the BPN while the number of hidden layer nodes is invariable. However, the GA in our study was proposed to optimize the structure of the BPN for not only the connection weight vectors but also the number of hidden layers, the number of hidden layer nodes and the node threshold vectors of all layers. Meanwhile, with a floating coded GA strategy hybrid with a BP algorithm, the unreliability of momentum addition in the BPN (Jia 2000) and the instability of the automatic adjustment for the training pace (Benediktsson et al. 1990) can be avoided. The GA-optimized BPN approach was used in land use classification of the study area.

On the basis of land use classification results, four epochs of urban land images were conformed by the extraction of urban land fragments in four epochs of classified land use images. Thus, the change detection for these four epochs of urban land images was conducted. Many methods have been presented to detect land use change. Rajendran and Chandrasekaran (2001) used remote sensing techniques to obtain urban information

* Corresponding author: Xue ZHANG, Ph.D Student, E-mail: ejau@163.com

together with geographic information systems (GIS) for analysis of change phenomena to monitor urban sprawl. El-Gamily (2007) applied the ERDAS Imagine detection tool to determine environmental damage in the southeastern region of Kuwait through transferring imagery pixel values to a matrix. Nelson et al. (2005) discussed ordinal conversion for radiometric normalization and temporal change detection while neglecting spatial change detection. In our study, however, the study area was divided into four plates: the central plate, the southern plate, the northern plate and the western plate, and three derived urban land images were conformed by overlapping ordinal epochs of urban land images. The first is conformed by overlapping the images of 1989 and 1995, the second of 1995 and 2001 and the third of 2001 and 2006. In our study, through overlaying the vector map containing borderlines of the towns with each derived urban land image, temporal-spatial changes of urban land in the derived urban land images for each town were detected based on the overlapping pixel comparison during the three time intervals (1989-1995, 1995-2001 and 2001-2006).

2. STUDY AREA AND DATA SOURCE

The study area, Jiading district, is a typical area of urbanization in Shanghai, having an administrative area of 463.9 km² and including eleven towns, two street blocks and one industrial zone (see Figure 1).

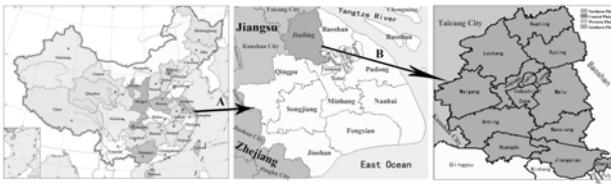


Figure 1. Location of study area.

Four epochs of cloud-free Landsat TM images with 30-m spatial resolution in the study area were used as a data source for remote sensing classification, and the images are respectively recorded on 11 August 1989, 21 March 1995, 3 July 2001 and 20 April 2006. Thus, land use changes can be detected in three time intervals: from 1989 to 1995, from 1995 to 2001, and from 2001 to 2006. Furthermore, a topographic map of 1:50,000 scale and a vector map containing borderlines of Jiading district and towns were collected for preprocessing the images and analyzing the temporal-spatial changes of land use from 1989 to 2006.

3. METHODOLOGY

In the methodology, the land use categories were first classified by the GA-optimized BPN approach after necessary preprocessing of remotely sensed data, the temporal-spatial changes of urban land were then detected, and the pattern of urban sprawl was finally analyzed. In the preprocessing process, as total 45 ground control point (GCPs) were selected from a topographic map of 1:50,000 scale in study area, the quadratic polynomial method was employed for image registration with the bilinear resampling and the root mean square error (RMSE) of the image registration is less than 0.6 pixels. Therefore, these TM images were geo-referenced to the local coordinate system

of Shanghai. After these, the images were resected by the district borderline in the vector map, and each of images has 846203 pixels. Then the Regions of Interest (ROI) were captured on screen for training.

3.1 Optimization of the BPN with Genetic Algorithms

Assuming a BPN has one input layer with n nodes, one output layer with m nodes and one hidden layer with p nodes, there are S sample signals. In the process of evolution to BPN by the GA, the optimal network is obtained while the overall output error (E) of the network is the smallest. The overall output error is calculated by

$$E = \frac{1}{S} \sum_{i=1}^S \sum_{j=1}^m (y_{ij} - d_j)^2 \quad (1)$$

where y_{ij} is the j^{th} output variable corresponding to the i^{th} sample signal inputted, and d_1, \dots, d_m are the observation vectors.

The process of using the proposed GA to train the BPN is presented as the following.

(1) Coding. In the study, each chromosome represents one kind of BPN structure. The length of the chromosome coding string is variable, and the number of hidden layers and hidden layer nodes are coded as floating numbers in each chromosome coding string. Each chromosome coding string is composed by eight units: a floating number q' corresponding to the number of hidden layers, a floating number p_k' to the number of hidden layer nodes, n threshold values to n nodes in the input layer, m threshold values to m nodes in the output layer, the threshold values to the i^{th} node in the k^{th} hidden layer ($i = 1, 2, \dots, p_k, k = 1, 2, \dots, q$), the connection weights between n nodes in the input layer and the j_1^{th} node in the first hidden layer ($j_1 = 1, 2, \dots, p_1$), the connection weights between the j_q^{th} node in the last hidden layer and m nodes in the output layer ($j_q = 1, 2, \dots, p_q$) and the connection weights between the j_k^{th} node in the k^{th} hidden layer and the j_{k+1}^{th} nodes in the $(k+1)^{\text{th}}$ hidden layer ($j_k = 1, 2, \dots, p_k, k' = 1, 2, \dots, q-1$). Where q is the number of hidden layers and p_k are the numbers of hidden layer nodes, and they can be computed according to Equation (2) and (3):

$$q = \text{Int}(q' + 0.5) \quad (2)$$

$$p_k = \text{Int}(p_k' + 0.5) \quad (3)$$

where $\text{Int}(X)$ takes the integer part of the number X .

It should be noted that the floating code (q') is rounded down to q , and the other units of the coding string are determined in accordance with q . To sum up, the length of a chromosome coding string varies and is calculated by

$$L = 1 + q + n + m + \sum_{i=1}^q p_i + np_1 + p_q m + \sum_{i=1}^{q-1} p_i p_{i+1} \quad (4)$$

where L is the length of a chromosome coding string.

(2) Selection of the fitness function $f(x)$. Since the optimal BPN is the one has a minimal output error (E) which can be calculated by Equation (1) after network training, $f(x) = 1/E$ was chosen as the fitness function of the GA. Thus, if the output error is smaller, the fitness value is larger, and the larger the fitness value is, the better a chromosome is. While the GA

catches up the best chromosome with largest fitness value, the optimal BPN which has minimal output error after training is obtained.

(3) Initialization of community. The community of N chromosomes is initialized in this step. The value of N is defined by customizing, which bases on the time when the GA-optimized BPN iterates at fixed times. For each chromosome, a floating code (q') is first randomly initialized and q is determined accordingly, and the other units of a chromosome coding string are then randomly initialized.

(4) Evolution. In the evolution process, there are three main steps:

Step 1: Selection. In this step, the roulette wheel selection method is used and the optimal chromosomes are preserved for the next evolution. 5-10% of N chromosomes with larger fitness values than those of the parent generation are preserved, and the other chromosomes are selected by roulette wheel selection from chromosomes with smaller fitness values.

Step 2: Crossover. The crossover of two parent chromosomes produces two offspring chromosomes. The crossover probability is used to determine whether a crossover will be conducted and the crossover is carried out based on:

$$\begin{aligned} I_1 &= \beta F_1 + (1 - \beta)F_2 & (5) \\ I_2 &= \beta F_2 + (1 - \beta)F_1 & (6) \end{aligned}$$

where I_1 and I_2 are the coding strings of two offspring chromosomes, F_1 and F_2 are the coding strings of two parent chromosomes. I_1 or I_2 is adaptively adjusted according to the variation in the hidden layer's node number (p) rounded from the floating code (p') that is obtained from the crossover of F_1 and F_2 , and if the length of F_1 is shorter than that of F_2 , the former should be extended as long as the later by adding random numbers. The β is the crossover degree, ranging within [0,1].

Step 3: Mutation. The mutation probability is used to determine whether a mutation will be conducted and the mutation is carried out based on

$$I_{new} = \alpha I_{old} \quad (7)$$

where I_{old} is the coding string of an old chromosome, I_{new} is the coding string of a new chromosome. I_{new} is adaptively adjusted according to the p rounded from p' that is obtained from the mutation of I_{old} . The α is a mutation degree, ranging within [0,1].

Thanks to the alterability of length of chromosome, the flexibility of the algorithm increases as a small change of q' and p_k' resulting in big change in the coding string is avoided in the evolution process.

3.2 GA-optimized BPN Classification

The classification of TM images based on the proposed GA-optimized neural network follows three steps:

Step 1: GA-optimized BPN training. There are seven bands in the TM data. In our study, the 1st, 2nd, 3rd, 4th, 5th and 7th were selected as input signals corresponding to the nodes of the input layer in the BPN. Some regions of interest to each land use categories in each image were collected manually on PC screen

for training as its sample set, and the sigmoid function (Chattopadhyay and Bandyopadhyay 2007) was chosen as the activation function in BPN training.

Step 2: BPN classification. Each image with the above mentioned 6 bands is the input data of the GA-optimized BPN, and the classification result was obtained accordingly.

Step 3: Accuracy assessment. The classification result and accuracy based on the GA-optimized BPN were assessed. The stratified systematic unaligned sampling method was first applied to assess the accuracies of the classification results. Overall accuracy and kappa coefficient were then computed from the confusion matrix (Congalton 1991).

3.3 Change Detection

Based on the land use classification results, four urban land images were first conformed by the extraction of urban land fragments in four epochs of classified land use images and three derived urban land images were then made. The idea of producing a derived urban land image based on the pixel comparison method is that: (1) the value of pixel to urban land in the urban land image is set to 1 while the pixel value to non-urban land is set to 0; (2) a derived urban land image is produced by overlapping two ordinal epochs of urban land images such as the years 1989 and 1995; and (3) the value of a pixel in the derived urban land image is then calculated by

$$P_{i,j}^{(k,k')} = P_{i,j}^{(k)} \times 2 + P_{i,j}^{(k')} \quad (k' > k) \quad (8)$$

where $P_{i,j}^{(k)}$ is the value of the (i, j) pixel in the k^{th} epoch of the urban land image, which should be 0 or 1, $P_{i,j}^{(k,k')}$ is the value of the (i, j) pixel in the derived urban land image, and there are four epochs of urban land images from 1989, 1995, 2001 and 2006, thus both k' and k are 1, 2, 3 or 4.

According to Equation (8), the value of a pixel in the derived urban land image should be 0, 1, 2 or 3. If the value is 0, then the pixels in both k^{th} and k'^{th} epoch images are non-urban land; if the value is 1, the pixel in the k^{th} epoch image is non-urban land while the pixel is urban land in the k'^{th} epoch image; if the value is 2, the pixel in the k^{th} epoch image is urban land while the pixel is non-urban land in the k'^{th} epoch image (which is an unusual case); and if the value is 3, then the pixels in both k^{th} and k'^{th} epoch images are urban land.

The number of pixels for each of the four cases was calculated accordingly. Furthermore, the vector map containing town borderlines was overlaid to the derived urban land images, and by taking each town as a statistical unit, temporal-spatial changes of urban land were detected according to the proposed pixel comparison method during the three time intervals.

4. RESULTS AND DISCUSSIONS

4.1 Classification and Accuracy Assessment

After the classifications for the four epochs of TM images in Jiading district by using the GA-optimized BPN approach, the classification results are obtained and illustrated in Figure 2. And the corresponding overall accuracies and Kappa coefficients are presented in Table 1.

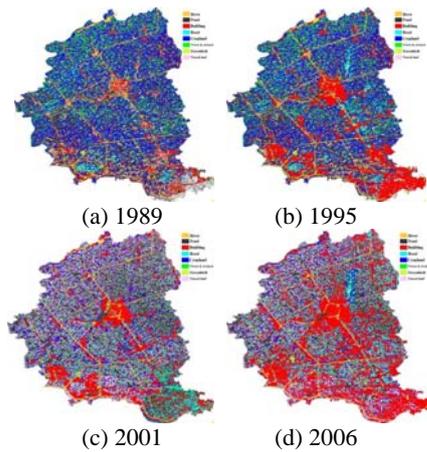


Figure 2. Results of classification in 1989, 1995, 2001 and 2006 based on the proposed approach for the study area

Statistic \ Year	Year			
	1989	1995	2001	2006
Overall accuracy	82.2%	85.6%	84.7%	84.2%
Kappa	0.7965	0.8353	0.8246	0.8194

Table 1. Overall accuracies and Kappa coefficients for the four epochs of TM images

The overall accuracy of the classification reflects the entire classification effect while the Kappa coefficient measures the consistency between classified data and sample data (Lillesand et al. 2003). From Table 1, it can be seen the overall accuracies are over 80% and Kappa coefficients are close to or over 0.8, thus it is feasible and reliable to use the above classification results to detect urban sprawl in the following discussion.

4.2 Analysis of Change Detection

4.2.1 Temporal change detection for whole urban area

Figure 3 shows the derived urban land images conformed by overlapping two epochs of extracted urban land images as discussed in Section 3.3.

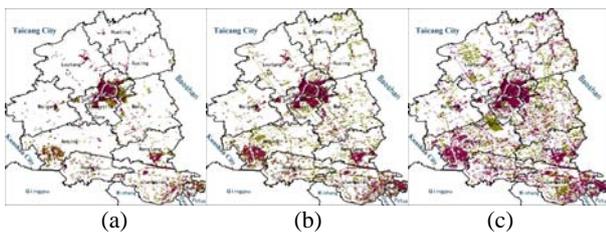


Figure 3. Three derived urban land images, in which the vector map containing town borderlines is overlaid to: (a) derived image by overlapping the urban land images from 1989 and 1995; (b) derived image by overlapping the urban land images from 1995 and 2001; and (c) derived image by overlapping the urban land images from 2001 and 2006.

Based on the three derived urban land images, urban land changes in the study area were detected through calculating the number of four types (0, 1, 2 and 3) of pixels in each derived urban land image and the results are listed in Table 2.

Interval	1989-1995		1995-2001		2001-2006	
Year	1989	1995	1995	2001	2001	2006
Urban land	67845	99157	99157	173272	173272	290394
Changed area	31312		74115		117122	
Annual growth	5900		12821		22245	
Annual rate	7.69%		12.46%		13.52%	

Table 2. Changes of urban land in the study area during the three time intervals from 1989 to 2006 (Unit: pixel)

From Figure 3 and Table 2, it can be seen that the increase in urban land in the study area is significant during the past 20 years, and the annual growths of urbanization are different during the three time intervals: (1) during the period from 1989 to 1995, the urban land sprawled in the central plate, and started growing in the western plate and southern plate (see Figure 3 (a)). Quantitatively, the urban land in the study area has increased by 31,312 pixels with an annual growth of 5,219 pixels and an annual growth rate of 7.69%; (2) during the period from 1995 to 2001, the urban land of the central plate continued sprawling into the surroundings while the urban land of the northern plate near the central plate grew slightly. The urban land of the southern plate and western plate axially expanded at high speed and became dense (see Figure 3 (b)). Quantitatively, the urban land has increased by 74,115 pixels with an annual growth of 9,019 pixels and an annual growth rate of 12.46%, which is much faster than that in the former period; and (3) during the period from 2001 to 2006, the urban lands of the central, western and southern plates experienced a series of rapid sprawls (see Figure 3 (c)). Quantitatively, the overall urban land has increased 117,122 pixels with an annual growth rate of 13.52%, which is quicker than that in 1995-2001.

4.2.2 Spatial change detection for an urban area

Through overlaying the vector map containing town borderlines with the derived urban land images, spatial changes of urban land for each town were detected. Table 3 shows the changes of urban land during the three time intervals for each town.

From Tables 3, the changes of urban land can be determined for each town during the three time intervals from 1989 to 2006:

(1) The urban land of Jiading town of the central plate in 1989 accounted for the largest proportion of town area at 79.15%, and the urban lands of Xincheng street block in the central plate and Zhenxing street block in the southern plate followed with 30.78% and 33.04% proportions, respectively. However, the proportions of urban land for Huating and Waigang towns in the northern plate were very small, being only 2.85% and 2.89%, respectively.

(2) During the period of 1989-1995, the urban land of Malu town in the central plate grew fastest with an annual growth rate of 14.66%, followed by annual growth rates over 10% for the industrial zone and Xincheng street block in the central plate and Jiangqiao and Nanxiang towns in the southern plate. However, the urban land of Jiading town of the central plate had a slower annual growth rate of only 1.13% between 1989 and 1995, but the proportion of urban land was very large and reached 84.51% by 1995.

(3) During the period of 1995-2001, the urban land of

Nanxiang town of the southern plate grew fastest with an annual growth rate of 19.09% and an urban land proportion increasing to 32.69% by 2001 from 15.24% in 1995. The annual growth rates of Malu town and Juyuan town in the central plate, Jiangqiao town in the southern plate and Anting town in the western plate followed with growth rates over 15%, while the annual growth rates of the industrial zone in the central plate and Huangdu town in the western plate were also over 10%. The urban land proportions of Xincheng street block in the central plate and Zhenxin street block in the southern plate increased from 50.51% and 48.01% in 1995 to 73.54% and 69.12% in 2001, respectively.

(4) During the period of 2001-2006, the urban land of Anting town of the western plate grew the fastest with an annual growth rate of 25.51% and the proportion increased from 19.31% in 2001 to 45.17% in 2006, followed by the

annual growth rates of over 20% of Loutang and Huating towns in the northern plate and the annual growth rate of Malu town in the central plate of 18.9%. Meanwhile, the urban land proportions of Loutang, Huating and Malu towns in 2006 have doubled compared to those in 2001, reaching 15.86%, 10% and 38.07%. From 2001 to 2006, the urban land of Jiangqiao town in the southern plate and Huangdu town in the western plate sprawled with annual growth rates of 14.60% and 14.12%. The urban land proportion of Jiangqiao town had reached 61.21% by 2006 and that of Huangdu town accounted for nearly half the total town area. Although the annual growth rates of urban sprawl were not more than 8% in Jiading town, the industrial zone, Xincheng street block of the central plate and Zhenxin street block of the southern plate, their urban land proportions increased to 96.35%, 91.67% and 93.86%, respectively. The urban land proportion in the industrial zone of the central plate was also over 50% by 2006.

Region	Year	1989		1989-1995		1995		1995-2001		2001		2001-2006		2006	
		Urban land	Growth rate												
Central plate	Whole	6784	3131	7.69%	9915	7411	12.46%	17327	11712	13.52%	29039				
	Jiading	5919	401	1.13%	6320	371	0.98%	6691	514	1.54%	7205				
	Industrial zone	3580	2505	11.66%	6085	4392	12.03%	10477	5663	5.58%	16140				
	Xinchen	2886	1849	10.68%	4735	2159	7.60%	6894	1700	4.93%	8594				
	Malu	5522	4856	14.66%	1037	1002	16.10%	20406	19284	18.90%	39690				
Southern plate	Juyuan	5771	1904	5.50%	7675	7242	15.73%	14917	6874	9.22%	21791				
	Nanxian	5563	3686	11.04%	9249	1059	19.09%	19841	9245	9.32%	29086				
	Zhenxin	3218	1458	7.55%	4676	2056	7.33%	6732	2410	7.16%	9142				
Northern plate	Jiangqiao	8357	5500	10.97%	1385	1356	16.32%	27422	20023	14.60%	47445				
	Huating	2054	1159	3.40%	2473	1009	6.80%	3482	3499	20.10%	6981				
	Xuhang	3446	911	4.41%	4357	2315	8.86%	6672	1934	5.80%	8606				
	Loutang	5383	1444	4.47%	6827	2216	5.41%	9043	9186	20.32%	18229				
Western plate	Waigang	2681	736	4.58%	3417	1757	8.57%	5174	2412	9.32%	7586				
	Anting	7513	3466	7.69%	1097	1015	15.41%	21130	26955	25.51%	48085				
	Huangdu	5952	2177	6.10%	8129	6262	12.84%	14391	10161	14.12%	24552				

Table 3. Changes of urban land for each town in the study area during the three time intervals from 1989 to 2006 (Unit: pixel)

5. CONCLUSIONS

In this paper, urban sprawl based on remote sensing classification and change detection was studied. A GA optimized BP neural network classification approach was first proposed. The proposed algorithm is not only to optimize the connection weight vectors, but also to optimize other elements of the network structure, including the number of hidden layers, the number of hidden layer nodes and the node threshold vectors of all layers. An overlapping pixel comparison based change detection method was further employed. Taking Jiading district of Shanghai as a study area, four epochs of TM images in 1989, 1995, 2001 and 2006 were used in the study. Based on the proposed classification and change detection approaches, temporal-spatial changes of urban land during the three time intervals from 1989 to 2006 were detected, and the process and pattern of urban sprawl in the study area were discovered. Conclusions relating to the process and pattern of the urban sprawl in Jiading district are drawn.

The denseness of urban land in Jiading district has enhanced at

increasing rates from 1989 to 2006. Firstly, the urban land sprawled circularly, with Jiading town of the central plate as the expansion center, and then, expanded in all directions from many centers of towns and street blocks, spreading and inter-filling along lines which formed the meshlike between town centers. Secondly, the urban land grew according to the distance from the city zone of Shanghai, i.e., the nearer to the city zone of Shanghai, the denser the distribution of urban land. Therefore, the distribution of urban land was very sparse in the northern district while it was very dense in the southern district. Moreover, industrial development oriented urban land concentration was also regarded as a form of urban sprawl in the study area.

REFERENCES

Benediktsson, J.A., Swain, P.H. and Ersoy, O.K., 1990, Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, Vol.28, 4, pp.540-552.

- Chattopadhyay, S. and Bandyopadhyay, G., 2007, Artificial neural network with backpropagation learning to predict mean monthly total ozone in Arosa, Switzerland. *International Journal of Remote Sensing*, Vol.28, 20, pp.4471-4482.
- Congalton, R.G., 1991, A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37, pp.35-46.
- El-Gamily, H.I., 2007, Utilization of multi-dates LANDSAT_TM data to detect and quantify the environmental damages in the southeastern region of Kuwait from 1990 to 1991. *International Journal of Remote Sensing*, Vol.28, 8, pp.1773-1788.
- Epstein, J., Payne, K. and Kramer, E., 2002, Techniques for mapping suburban sprawl. *Photogrammetric Engineering and Remote Sensing*, Vol.63, No.9, pp.913-918.
- Foody, G.M. and Arora, M.K., 1997, An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *International Journal of Remote Sensing*, 18, pp.799-810.
- Heermann, P.D. and Khazenie, N., 1992, Classification of multispectral remote sensing data using a back-propagation neural network. *IEEE Transactions on geoscience and remote sensing*, Vol.30, 1, pp.81-89.
- Jat, M.K., Grag, P.K. and Khare, D., 2007, Monitoring and modelling of urban sprawl using remote sensing and GIS techniques. *International Journal of Applied Earth Observation and Geoinformation*, doi:10.1016/j.jag.
- Jia, Y.H., 2000, Application of artificial neural network to classification of multi-source remote sensing imagery. *Bulletin of Surveying and Mapping*, 7, pp.7-8.
- Lillesand, T.M., Kiefer, R.W. and Chipman, J.W., 2003, *Remote sensing and image interpretation (Fourth Edition)* (John Wiley & Sons, Inc: Publishing House of Electronics Industry).
- Liu, Z.J., Liu, A.X., Wang, C.Y. and Niu, Z., 2004, Evolving neural network using real coded genetic algorithm (GA) for multispectral image classification. *Future Generation Computer Systems*, 20, pp.1119-1129.
- Masek, J.G., Lindsay, F.E. and Goward, S.N., 2000, Dynamics of urban growth in the Washington DC metropolitan area, 1973-1996, from Landsat observation. *International Journal of Remote sensing*, Vol.21, 18, pp.3473-3486.
- Nelson, T., Wilson, H.G., Boots, B. and Wulder, M.A., 2005, Use of ordinal conversion for radiometric normalization and change detection. *International Journal of Remote Sensing*, Vol.26, 3, pp.535-541.
- Nicola's, G.P., Domingo, O.B. and Ce'sar, H.M., 2006, An alternative approach for neural network evolution with a genetic algorithm: Crossover by combinatorial optimization. *Neural Networks*, 19, pp.514-528.
- Ninomiya, A., Hatamoto, K., Hoyano, A., Tanaka, S. and Sugimura, T., 2005, An analysis of Asian urban expansion and changes using Landsat data. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp.1554-1556.
- Omatu, S. and Yoshida, T., 1991, Pattern classification for remote sensing using neural network. *Neural Networks*, 1, pp.652-658.
- Rajendran, S. and Chandrasekaran, V.A., 2001, Monitoring urban expansion of Tiruchirapalli town of Tamil Nadu State, India using IRS-1C satellite data. *"New Technology for a New Century" Technical Conference during the FIG Working Week of Seoul, Korea*.
- Rojas, F., Manuel, R.Á., Carlos, G.P. and Ruben, M.C., 2002, Applying neural networks and genetic algorithms to the separation of sources. *Lecture Notes in Computer Science*, 2527, pp.420-429.
- Toshniwal, M., 2005, An optimized approach to application of neural networks to classification of multi-spectral, remote sensing data. *IEEE Networking, Sensing and Control, Proceedings*, pp.235-240.
- Vogelmann, J. E., Sohl, T. and Howard, S. M., 1998, Regional characterizations of land cover using multiple sources of data. *Photogrammetric Engineering and Remote Sensing*, 64, pp.45-57.

ACKNOWLEDGEMENTS

The work described in this paper was substantially supported by the National Natural Science Foundation of China (Project No. 40771174), the National Science and Technology Support Program (Project No. 2006BAJ09B01), the Program for New Century Excellent Talents in Universities (Project No. NCET-06-0381), and the Foundation of Shanghai ShuGuang Scholarship Program (Project No. 07SG24).