Chin. Geogra. Sci. 2011 21(3) 312–321
doi: 10.1007/s11769-011-0465-1
www.springerlink.com/content/1002-0063

Land Cover Classification with Multi-source Data Using Evidential Reasoning Approach

LI Huapeng[1, 2], ZHANG Shuqing[1], SUN Yan[1, 2], GAO Jing[3]

(*1. Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130012, China; 2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China; 3. School of Computer and Information Science, University of South Australia, Adelaide 5095, Australia*)

Abstract: Land cover classification is the core of converting satellite imagery to available geographic data. However, spectral signatures do not always provide enough information in classification decisions. Thus, the application of multi-source data becomes necessary. This paper presents an evidential reasoning (ER) approach to incorporate Landsat TM imagery, altitude and slope data. Results show that multi-source data contribute to the classification accuracy achieved by the ER method, whereas play a negative role to that derived by maximum likelihood classifier (MLC). In comparison to the results derived based on TM imagery alone, the overall accuracy rate of the ER method increases by 7.66% and that of the MLC method decreases by 8.35% when all data sources (TM plus altitude and slope) are accessible. The ER method is regarded as a better approach for multi-source image classification. In addition, the method produces not only an accurate classification result, but also the uncertainty which presents the inherent difficulty in classification decisions. The uncertainty associated to the ER classification image is evaluated and proved to be useful for improved classification accuracy.

Keywords: evidential reasoning; Dempster-Shafer theory of evidence; multi-source data; geographic ancillary data; land cover classification; classification uncertainty

Citation: Li Huapeng, Zhang Shuqing, Sun Yan, Gao Jing, 2011. Land cover classification with multi-source data using evidential reasoning approach. *Chinese Geographical Science*, 21(3): 312–321. doi: 10.1007/s11769-011-0465-1

1 Introduction

Remote sensing research focusing on image classification is always an ongoing topic, because classification results are the basis for converting remote sensing images to available geographic data (Wilkinson, 2005). The land use/cover information derived from satellite image classification is very important in the Earth Surface Science (ESS). For example, many landscape metrics are built on the land cover classification (Zhang *et al.*, 2006; Zhang *et al.*, 2009). It has been confirmed that multi-source data have the potential for improving image classification accuracy, since they can provide much more information compared to single satellite imagery for classification decisions (Kim and Swain, 1995; Chiuderi, 1997; Le Hegarat-Mascle *et al.*, 1997; Datcu *et al.*, 2002; Franklin *et al.*, 2002; Ozesmi and Bauer, 2002; Chitroub, 2003; Tzeng *et al.*, 2007; Camps-Valls *et al.*, 2008; Na *et al.*, 2009).

Geospatial data used for image classification include four types: interval data, ratio data, nominal data and ordinal data (Franklin *et al*., 2002). Traditional statistical classifiers (e.g., maximum likelihood classifier) can only use the data meeting some normal distribution such as satellite image data (Mertikas and Zervakis, 2001). Interval data (e.g., DEM), ordinal data (e.g., land degradation map) and nominal data (e.g., soil map) are difficult to be incorporated into these classifiers, in fact, these data are crucial to some classification tasks. For example, hydrological and soil maps contribute to wet-

Received date: 2010-05-13; accepted date: 2010-11-10
Foundation item: Under the auspices of National Natural Science Foundation of China (No. 40871188), Knowledge Innovation Programs of Chinese Academy of Sciences (No. INFO-115-C01-SDB4-05)
Corresponding author: ZHANG Shuqing. E-mail: shqzhang@263.net

land mapping (Li and Chen, 2005; Vaiphasa *et al.*, 2006; Castaneda and Ducrot, 2009; Davranche *et al.*, 2010), and vegetation inventory map improves the mapping of grizzly bear habitat (Franklin *et al.*, 2002). Therefore, it is necessary to explore a new classification method which can combine data regardless of the number of sources and the scale of measurement.

Currently, there are mainly two effective methods, including the decision tree and evidential reasoning (ER), to handle multi-source data in a classification task (Franklin *et al.*, 2002). The ER method, built on Dempster-Shafer theory of evidence, is a method of inexact reasoning (Shafer, 1967; Giarratano and Riley, 1998). The method has the ability to incorporate all four types of geospatial data, so it has been widely used in a variety of classification tasks (Le Hegarat-Mascle *et al.*, 2000; Foucher *et al.*, 2002; Cohen and Shoshany, 2005; Al Momani *et al.*, 2006; Sun *et al.*, 2008). Additionally, the method can produce interpretive measures (such as degree of support, plausibility and uncertainty) which are useful to classification (Lein, 2003; Cohen and Shoshany, 2005). However, few researches have been devoted to the relationship between uncertainty and classification accuracy produced by the ER method.

In this paper, we selected the Dazhanhe National Nature Reserve in Heilongjiang Province, China as the study area, combined TM images and geographic ancillary data (altitude and slope), to assess the potential of the ER method for multi-source image classification and analyze the relationship between uncertainty and classification accuracy. In addition, the maximum likelihood classifier (MLC) method was also achieved based on multi-source data for a comparison with the ER method. The accurate information derived from image classification may be useful for land planning in the study area.

2 Background of Evidential Reasoning Approach

The evidential reasoning (ER) was firstly proposed by Dempster and further developed by Shafer, and it is developed based on the fact that the knowledge and information is always uncertain, incomplete, or insufficient in making decisions such as image classification (Sun *et al.*, 2008). The ER method can integrate various data sources through formal probabilistic reasoning (Cayuela *et al.*, 2006), and the classification confidence can be improved as the evidence increases (Al Momani *et al.*, 2006).

In the image classification context, the ER method has several advantages compared with traditional classification procedures. Firstly, it is a non-parametric classifier and thus can handle data which may violate the normal distribution (Srinivasan and Richards, 1990). Secondly, it can incorporate data regardless of the number of sources and the scale of measurement. Thirdly, it is a 'soft' classification method and the results can be adjusted by human interpretation. Fourthly, it can provide several interpretive measures such as support, plausibility and uncertainty which can be used to evaluate the classification results (Lein, 2003; Peddle, 1995a; Sun *et al*., 2008). The fundamental aspects of the ER method are described as follows.

2.1 Frame of discernment (FoD)

Concepts and functions of ER are based on frame of discernment (FoD). In Dempster-Shafer theory, a set of mutually exclusive and exhaustive elements constitutes the frame of discernment, denoted by Θ (Peddle, 1995a). Any element of Θ is called singleton, a set with none element is called empty set (ϕ). A set of size n has 2^n subsets. In the image classification context, all the land covers to be classified correspond to the elements of FoD. Suppose that there are three kinds of land covers (forest, grass and water) needed to be discerned in a satellite imagery, then the FoD can be defined as the following formula: $\Theta = \{ft, gs, wt\}$. If the research only focuses on singleton (i.e., each land cover class), the FoD can be simplified as: $\Theta = \{ft\}, \{gs\}, \{wt\}$. Then, the number of sets is reduced significantly (from 8 to 3) which can simplify the computational complexity and the processing time exponentially of the ER method (Gordon and Shortliffe, 1985; Wilkinson and Megier, 1990).

2.2 Basic probability assignment (BPA)

A piece of evidence can support one or more subsets. Suppose that there is a piece of evidence in support of the non-empty subset X on Θ, and a set function $m(X)$ named the Basic Probability Assignment (BPA) represents the degree of support to subset X, then the degree of support is defined in the interval [0, 1], and must sum to 1 over all possible sets:

$$\begin{cases} m(\phi) = 0 \\ \sum\limits_{A \in 2^{\Theta}} m(X) = 1 \end{cases} \tag{1}$$

2.3 Belief function and plausible function

In FoD Θ, suppose that X, Y are two sets on Θ, if the set function m meets the following equation:

$$Bel(X) = \sum_{Y \subseteq X} m(Y) \tag{2}$$

$$Pl(X) = (1 - Bel(X^-)) = \sum_{X \cap Y \neq \phi} m(Y) \tag{3}$$

where $Bel(X)$ is the belief function of set X on Θ (i.e., the lower probability); $Pl(X)$ is the plausible function set X (i.e., the upper probability), and $Bel(X^-)$ is the belief function of the complementation set of X. The interval $[Bel(X), Pl(X)]$ reflects the unknown degree of set X.

2.4 Combination of belief functions

The theory of evidence allows us to combine belief functions from different evidences on the same FoD with combination rule and generate new belief functions. Let Bel_1 and Bel_2 denote belief functions over the same FoD Θ, then the new belief function, termed as the orthogonal sum of Bel_1 and Bel_2, denoted as $Bel_1 \oplus Bel_2$, is calculated as follows:

$$Bel_1 \oplus Bel_2 = \frac{\sum\limits_{X \cap Y = Z} m_1(X) m_2(Y)}{1 - \sum\limits_{X \cap Y = \phi} m_1(X) m_2(Y)} \tag{4}$$

where m_1, m_2 are the BPA functions of Bel_1, Bel_2, composed of the possible hypotheses sets X_1, …, X_K and Y_1, …, Y_K, respectively. The sum extends over all class labels whose intersection $X \cap Y = Z$. If there are more than two pieces of evidence, the combination rule can be extended as: $Bel_1 \oplus Bel_2 \oplus Bel_3 \oplus Bel_4 \oplus Bel_5, \dots$

2.5 Interval of uncertainty

In Dempster-Shafer theory of evidence, the uncertainty is defined as: $Un(X) = Pl(X) - Bel(X)$ (Fig. 1). The uncertainty is a probability interval which neither supports nor negates subset X obviously. We suppose X is true if $Un(X) = [1, 1]$, false if $Un(X) = [0, 0]$, and unknown if $Un(X) = [0, 1]$. $Un(X)$ indicates the unknown degree of set X, and fusing evidence from multi-source data may narrow this unknown interval (Shafer, 1976).

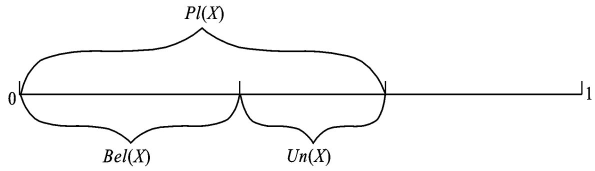


Fig. 1 Illustration of interval of uncertainty

3 Materials and Methods

3.1 Study area

In this paper, we selected the Dazhanhe National Nature Reserve (47°59′55″–48°50′12″N, 127°44′57″–128°30′14″E), located in the northern Heilongjiang Province of China, as the study area, with a total area of 430 000 ha (Fig. 2). The mean annual temperature is –2℃, and the annual precipitation is 500–700 mm which mainly occurs in July and August. The altitude ranges from 320 m to 620 m, and the slope is from 0° to 26°. The representative soil is dark brown, and non-zonal soils are meadow soil, bog soil, peat soil and black soil.

According to field investigation, there are four types of land covers including forest, wetland, water and building land (Table 1). Forest and wetland are predominant in the study area, and there are a variety of forest types and wetland types. Forest is mainly distributed in low mountains, gentle hills and highland of low watershed, while the other classes mainly in low-lying land and gently sloping areas.

3.2 Data collection and processing

Landsat 5 TM image (Row/Path: 118/26) dated on September 30, 2007 was acquired from the USGS Earth Resource Observation Systems Data Center (http://glovis.usgs.gov/). The image was rectified to the Gauss Kruger projection system (datum Xi′an 1980, zone 22) based on topographic maps using approximately 70 ground control points, primarily the intersection of ways and rivers, evenly distributed across the image. A third order polynomial model was used for the rectification with the nearest neighbor algorithm with a pixel size of 30 m × 30 m for all bands. The root mean square errors were less than 0.5 pixels (15 m).

The topographic map covered the study area was acquired from the Heilongjiang Mapping and Surveying Bureau. Altitude and slope maps were generated based on the vector contour lines and altitude points derived from the topographic map using ARCGIS 9.1 software. The geographical ancillary data (altitude and slope)

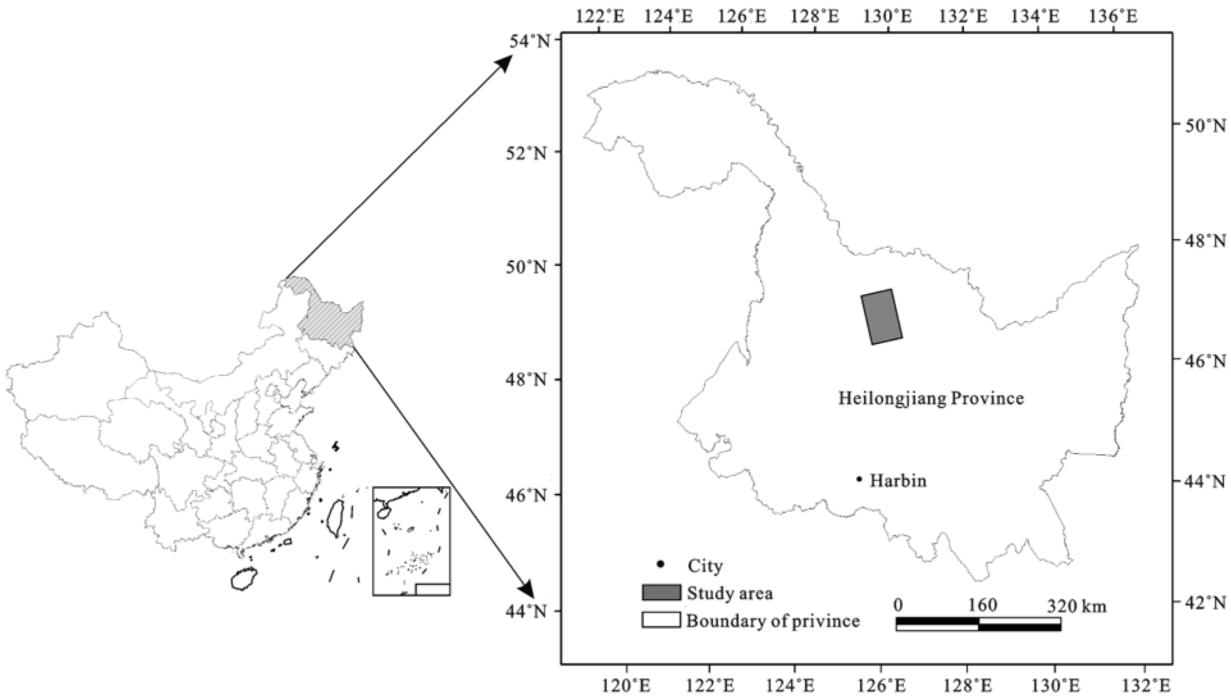


Fig. 2 Sketch map of study area

Table 1 Land cover types of study area

Land cover	Description
Forest	Broad-leaf forest, coniferous forest and mixed forest
Wetland	Sphagnum swamp, herbaceous swamp, shrub swamp and forested wetland
Water	Dazhan River, pond and comic lake
Building land	Small town, village, road, waste land, *etc*.

were then re-sampled to pixel size 30 m × 30 m, so as to match the resolution of TM image.

For supervised classification and accuracy assessment, 1000 pixels for four land cover types were identified from: 1) field locations in late August 2007 and late June 2008 referenced on the ground with a global positioning system (GPS); 2) random locations derived from imagery using stratified random sampling method. To acquire representative samples for each class, random samples were collected strictly within the polygons which are larger than 1 ha (around 11 TM pixels) (Wright and Gallant, 2007). Although the choice of pixels is arbitrary, it can reduce mixed-pixel sampling and mislabeling due to the positional errors. The number of sampling sites in the four classes was proportional to their relative areas to avoid the over-prediction of rare classes. A total of 520 sites were chosen as the training samples, including 190 forest sites, 180 wetland sites, 80 building land sites and 70 water sites.

3.3 Basic probability assignment

Due to its generality, the Dempster-Shafer theory of evidence does not discuss the generation method of evidence measures, i.e., basic probability assignment (BPA) (Sun *et al*., 2008). However, the ER method is conducted based on the BPA, so how to calculate BPA is important for satisfied classification results. Maximum likelihood classifier (MLC), the most widely used traditional classifier, can produce an accurate classification result if the dataset used is distributed normally, thus the posterior probability derived by MLC based on TM imagery was used as the BPA from satellite imagery (Cayuela *et al*., 2006). The MLC was performed by using the following equation (Richards and Jia, 1999):

$$f_i(x) = -\ln\left|\Sigma_i\right| - (x - m_i)^t \, \Sigma_i^{-1}(x - m_i) \qquad (5)$$

where f_i is the discriminant function for class i; x is the pixel of imagery; Σ_i is the variance-covariance matrix estimated from training pixels in class i, and m_i is the mean vector of the class i.

According to the knowledge derived from field investigation and related reports in the study area, forest is mainly distributed in high altitude and high slope areas, and water, wetland and building land are distributed in low altitude and low slope areas. We found that the areas with an altitude larger than 520 m and slope larger than 6° are almost covered with forest, while there is no forest in the areas where altitude is lower than 410 m and slope is lower than 2°, and the occurrence probability of forest distribution gradually increases with the altitude between 410 m and 520 m. Thus, the altitude and slope referred to the hypothesis of the forest class directly. Therefore, the knowledge was incorporated by using a linear function (Fig. 3). The given BPA for the forest class was 0.8 in the area where altitude is above 520 m and slope above 6°, and 0.05 where altitude is below 410 m and slope below 2°.

3.4 Classified models

Four classified models were constructed by using four different combinations of data sources: 1) TM imagery alone; 2) TM imagery plus slope (TM + SLO model); 3) TM imagery plus altitude (TM + DEM model); 4) TM imagery plus altitude and slope (TM + SLO + DEM model). The MLC and ER methods, using ERDAS and IDRISI software respectively, were conducted based on the four models using the same training locations as mentioned above. The difference between the two procedures was that the MLC method combined the multi-source data directly, while the ER method did it by using the evidence derived from multi-source data.

3.5 Accuracy assessment

A total of 480 reference sites not previously used for training purpose were used for accuracy assessment, including 82 building land pixels, 78 water pixels, 142 wetland pixels, and 178 forest pixels. A confusion matrix was generated, which can reflect the quality of classification directly (Foody, 2002), and three kinds of accuracy were calculated: 1) user′s accuracy for each class, which indicates the percentage of classified samples that were identified in reality; 2) producer′s accuracy for each class, indicating the percentage of reference samples that were correctly classified; and 3) overall accuracy, indicating the general level of classification results. The standard error of the mean (SEM) was calculated to assess the variation level of classification accuracy over the four classified models of the two classifiers, it is designated as S, and the equation is as follows:

$$S = \frac{\sigma}{\sqrt{N-1}} \tag{6}$$

where σ is the standard deviation of sample and N is the size of sample.

4 Results

4.1 MLC classification

The accuracy of the MLC classification decreases significantly when more data sources are added (Fig. 4c). When all data sources are accessible, the overall accuracy rate declined by 8.35% compared with that derived based on TM imagery alone. In the predominant forest class, producer′s accuracy rate (PAR) is relative constant across the four models (Fig. 4a). Similarly, in the build-

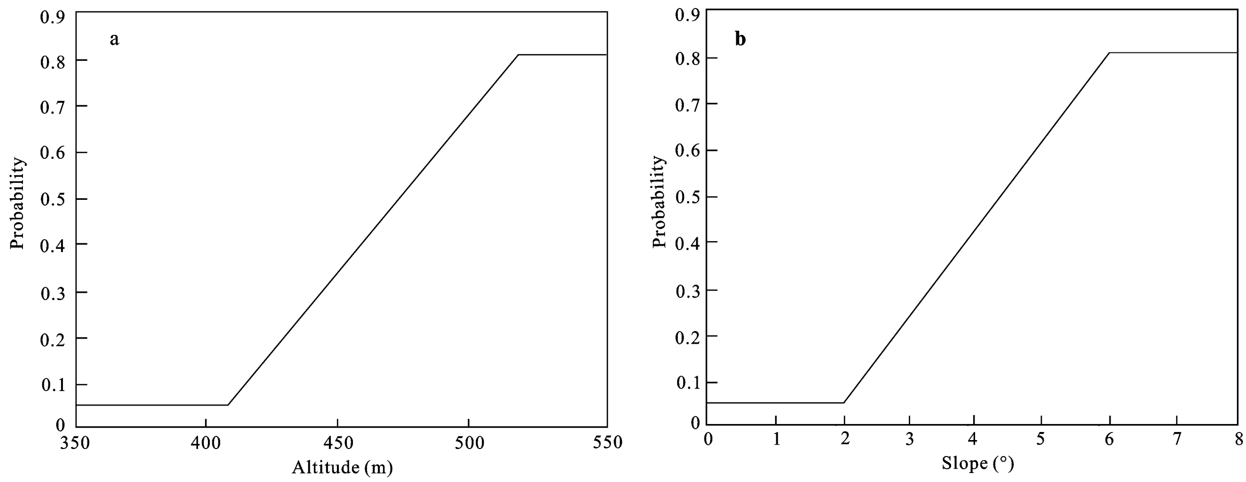


Fig. 3 Altitude-based and slope-based probabilities for forest class

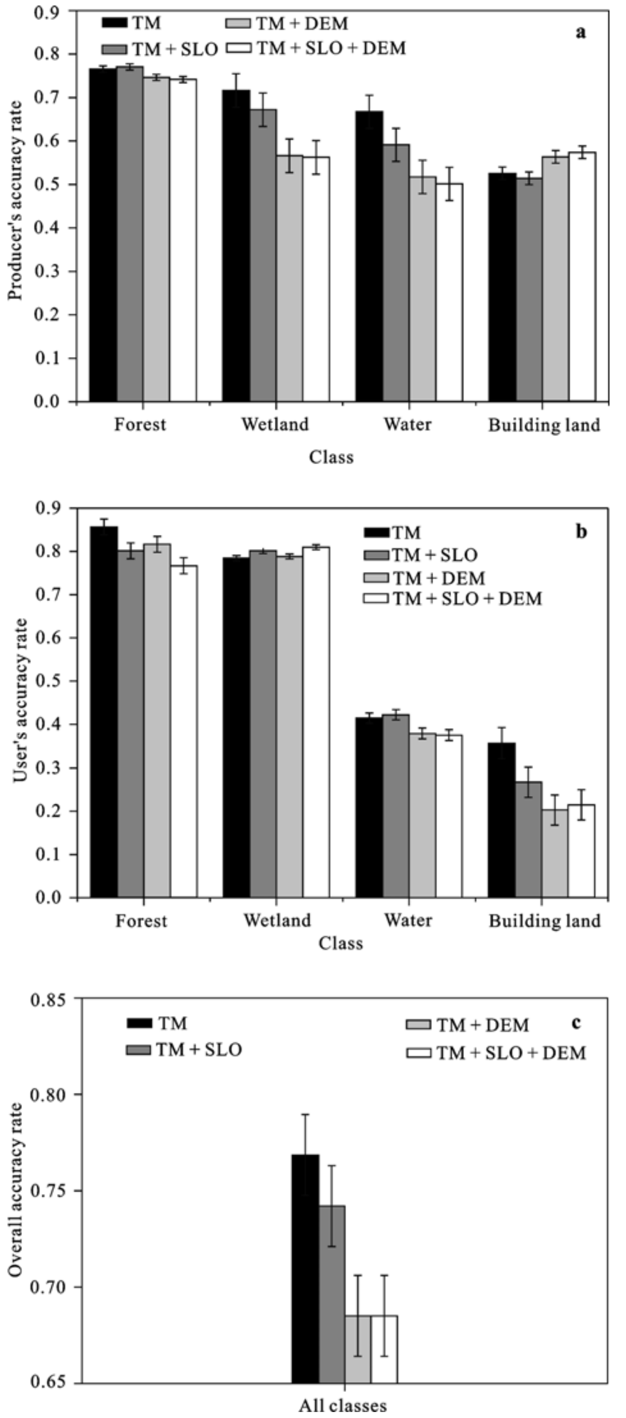


Fig. 4 Producer's accuracy rate, user's accuracy rate and overall accuracy rate observed by MLC method across four models

ing land class, the variation of PAR is slight. However, in the water and wetland classes, the PARs decline significantly when more data are added. The results for user's accuracy rate (UAR) are much poor (Fig. 4b). UARs of the building land and water classes are below 45% across the four models. The addition of the altitude and slope data does not improve the UARs of the two types of land cover. As can be seen, high UARs of the wetland and forest classes (about 80%) are consistent across the four models, and there is a decreasing trend for the forest class. The overall accuracy rate decreases sharply with the addition of the slope and altitude data (Fig. 4c). The accumulation effect of classification accuracy by the MLC method is not apparent when adding new data sources. For example, the overall accuracy rate of TM + SLO + DEM model is almost equal to that of TM + DEM model, and for the PAR of the wetland class, the result is similar. This indicates that the addition of slope data does not reduce the classification accuracy further.

4.2 ER classification

Overall accuracy rate of the ER classification increases when more data sources are added (Fig. 5c). Compared with the slope, the altitude does well in improving the overall accuracy rate. In the predominant forest class, the PARs are high, greater than 80% in the last three models. In the building land class, the poor PARs are almost constant across models, but there is an upward trend with respect to the UARs. Similarly, in the wetland class, the PARs are almost constant across the models, and the difference is that the average PAR of the wetland class is higher than that of the building land class by about 20%. In the water class, the PAR is improved by almost 20% in the last two models compared to the former two models. The UAR of the forest class decreases slightly when more data sources are added. Generally, accuracy rates for each class in the TM + SLO + DEM model perform better than those in the former three models. In this study, the overall accuracy rate of TM + SLO + DEM model is 84.52% which increases by 7.66% compared with that of TM model. The results demonstrate that the multi-source data can improve the classification accuracy significantly by using the evidential reasoning method.

4.3 Uncertainty of ER classification

The ER method can provide a richer description of the inherent uncertainty associated with the classification results. The mean uncertainty for each class of four classified models is shown in Table 2. The mean uncertainty for each class decreases gradually when more data sources are added. It indicates that multi-source data

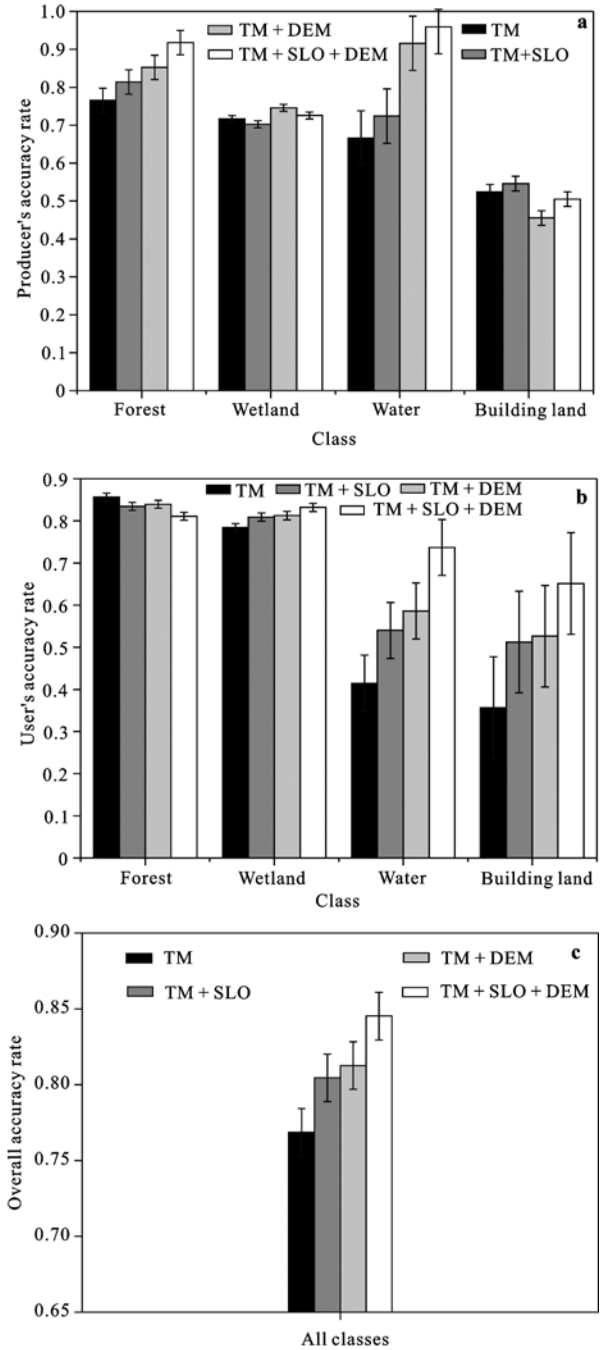


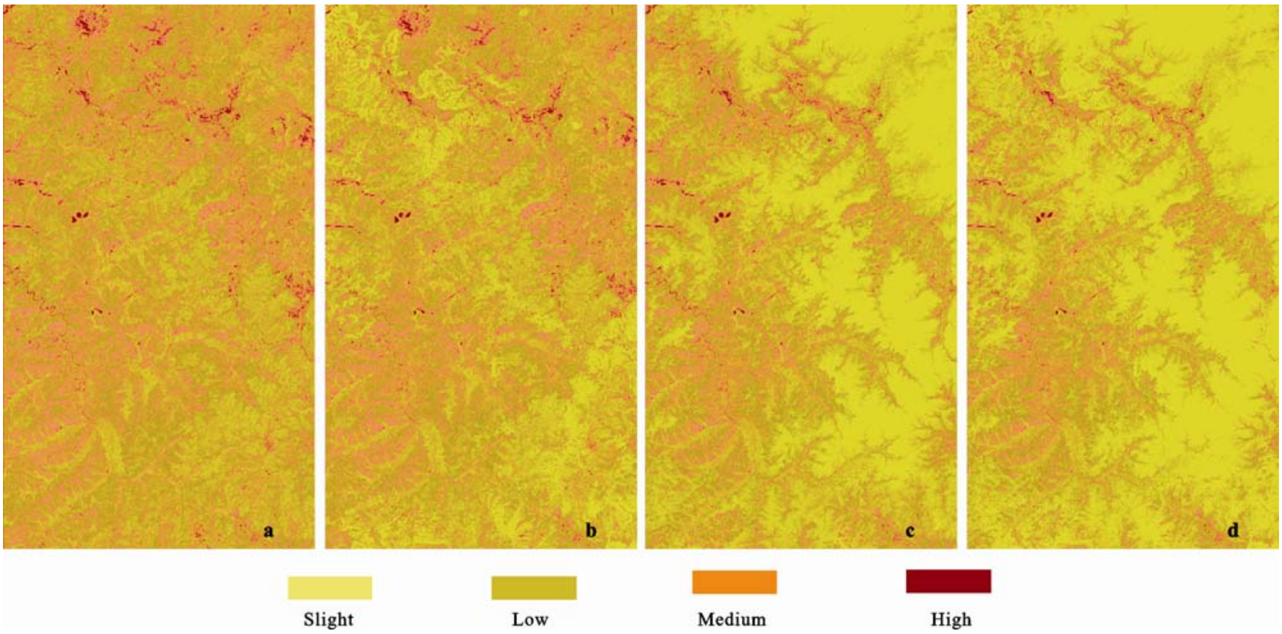
Fig. 5 Producer's accuracy rate, user's accuracy rate and overall accuracy rate observed by ER method across four models

reduce the area of image where multi-spectral data are insufficient to label pixels. The uncertainty of TM + DEM model declines than that of TM + SLO model for almost all the classes. As mentioned above, the contribution of the altitude to the accuracy rates is higher than that of the slope. Thus, low accuracy rate always, appears in high uncertainty area, although we still can not demonstrate this result in theory at present. Compared with traditional methods, this richer description may be useful to classification results. For example, we can investigate the low uncertainty area for more information which may have the potential to obtain better results. Additionally, the uncertainty is an indirectly variable for assessing the contribution to classification accuracy from a variety of data sources. We divided the uncertainty into four levels: slight level (uncertainty in the range of 0–0.075); low level (0.075–0.100); medium level (0.100–0.200); high level (0.200–1.000) (Fig. 6).

As can be seen from Fig. 6, the area of slight and low level uncertainty expands and that of medium and high level shrinks when more data sources are added. Obviously, the altitude data do better than the slope in reducing the uncertainty (Fig. 6b, Fig. 6c). When all data sources are accessible, the uncertainty is further decreased. In addition, the accumulation effect of the uncertainty is also apparent. In TM model, the high uncertainty occurs mainly in the rare building land and water which have relative poor classification accuracies. When combining the slope and altitude with TM imagery, the area of high and medium uncertainty reduces sharply, which indicates that sufficient information has supplied for most of this area in classification decisions. Although the area of high and medium uncertainty decreases, it is found that the area of medium uncertainty is also large (Fig. 6d). Therefore, in the future, much more data sources such as soil and landform maps should be added in the ER classification procedure to decrease the uncertainty of classification further.

Table 2 Mean uncertainty of ER classified models

Model	Forest (S.D.)	Wetland (S.D.)	Water (S.D.)	Building land (S.D.)	Overall (S.D.)
TM model	0.0976 (0.0378)	0.0963 (0.0463)	0.1076 (0.0477)	0.1374 (0.0507)	0.1012 (0.0441)
TM + SLO model	0.0866 (0.0442)	0.0919 (0.0322)	0.1047 (0.0460)	0.1345 (0.0472)	0.0923 (0.0430)
TM + DEM model	0.0687 (0.0405)	0.0867 (0.0316)	0.1053 (0.0513)	0.1335 (0.0528)	0.0778 (0.0424)
TM + SLO + DEM model	0.0613 (0.0376)	0.0839 (0.0308)	0.1027 (0.0507)	0.1311 (0.0505)	0.0703 (0.0402)



(a) TM model; (b) TM + SLO model; (c) TM + DEM model; (d) TM + SLO + DEM model

Fig. 6 Distribution of uncertainty level

5 Discussion

The results in this study show that the slope and altitude can not improve the classification results implemented by the MLC method. To the contrary, they play a negative role, and the altitude makes a greater impact than the slope does. It is mainly because that the classifier is designed on the statistical theory and thus can not handle data such as slope and altitude which may violate normal distribution. The results in this study are consistent with those in Cayuela *et al*. (2006) and Franklin *et al*. (2002).

The contribution of the altitude to the accuracy rates of the ER approach is higher than that of the slope, and it may be attributed to: 1) the altitude is much more useful and important in this area; 2) the knowledge formulation we designed for the generation of evidence measure from the altitude is better and more reasonable than that from the slope. Because of the well documented manner of evidence generation in the ER method (Peddle, 1995b; Mertikas and Zervakis, 2001), we can explore a much better method for converting input data into evidence, e.g., exponential functions, which may lead to better results.

As can be seen from Fig. 5a and Fig. 5b, the PAR and UAR are poor for the building land and water classes in TM model, which is mainly due to the overwhelming preponderance of the forest and wetland classes. Some studies have demonstrated that the classification accuracies of rare classes were usually poor (Wright and Gallant, 2007). It is obvious that the SEMs of PAR and UAR of the building land and water classes are higher than those of the wetland and forest classes, which indicates that the multi-source data play a positive role in increasing the accuracy rates for the rare classes, such as building land and water classes in this study.

As mentioned above, the slope and altitude data contribute to accuracy rates respectively when using the ER method for image classification. However, once combining them with TM imagery together, the accuracy rate is further improved. The accumulation effect of classification accuracy by the ER method is apparent and very important in multi-source data classification. The method can accumulate slight improvements from diverse data sources to achieve a substantial improvement of overall accuracy (Franklin *et al.*, 2002; Sun *et al.*, 2008). However, if the evidence derived from single data source plays a negative role, the overall result will be disturbed due to the accumulation effect characterization the of ER method, for example the PAR of the building land class and UAR of the forest class in this study.

Some researches determine that the accuracy rates of the hypothesis which multi-source data referred to were improved greatly (Franklin *et al*., 2002; Cayuela *et al*., 2006). However, in this study, the accuracy rates of the hypothesis set (i.e., forest class in this study) supported by multi-source data are not improved significantly, and the UAR of the class even decreases slightly. It may be the reason that the incorporation of ancillary data improves the identification of pixels which are mislabeled as forest class in low slope and low altitude areas, thus the UARs of water and building land classes improve significantly, while the classified area of forest is over-predicted in high altitude and slope areas, which reduces the UAR of forest slightly.

6 Conclusions

This paper presents a multi-source evidential reasoning method for satellite image classification. The preliminary results suggest that multi-source data fusion is a promising method in improving classification results. The ER method does better in combining multi-source data compared with traditional maximum likelihood classifier (MLC). The ER method offers some advantages over MLC: 1) significant improvement in overall accuracy rate; 2) improvements in producer′s accuracy rate and user′s accuracy rate for almost all classes; and 3) a richer description to classification results such as uncertainty.

Although better results have been obtained by the ER method, it is important to stress that the improvement is traded with much time, as it is difficult for us to derive reasonable evidence from multi-source data used in the ER classification at present, especially for ancillary data, which is an issue that requires further research. The study area in this paper is relatively small and represents only one type of landscape. It is necessary to test the application of the ER approach in other larger regions characterized by various landscapes. Furthermore, what ancillary data should be used and how to integrate these data into the ER approach is another issue deserving further research.

Acknowledgements

The authors would like to thank the staff of the Dazhanhe National Nature Reserve in Heilongjiang Province, China for access to valuable materials used in this paper.

References

Al Momani B, McClean S, Morrow P, 2006. Using Dempster-Shafer to incorporate knowledge into satellite image classification. *Artificial Intelligence Review*, 25(1–2): 161–178. doi: 10.1007/s10462-007-9027-4

Camps-Valls G, Gomez-Chova L, Munoz-Mari J *et al*., 2008. Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 46(6): 1822–1835. doi: 10.1109/TGRS.2008.916201

Castaneda C, Ducrot D, 2009. Land cover mapping of wetland areas in an agricultural landscape using SAR and Landsat imagery. *Journal of Environmental Management*, 90(7): 2270–2277. doi: 10.1016/j.jenvman.2007.06.030

Cayuela L, Golicher J D, Rey J S *et al*., 2006. Classification of a complex landscape using Dempster-Shafer theory of evidence. *International Journal of Remote Sensing*, 27(9–10): 1951–1971. doi: 10.1080/01431160500181788

Chitroub S, 2003. Optimal fusion-classification of multisource remote sensing imagery using global optimization and fuzzy logic. *Geoscience and Remote Sensing Symposium*, *2003. IGARSS '03. Proceedings. 2003 IEEE International*, 3: 1793–1795. doi: 10.1109/IGARSS.2003.1294252

Chiuderi A, 1997. Multisource and multitemporal data in land cover classification tasks: The advantage offered by neural networks. *Geoscience and Remote Sensing*, *1997. IGARSS '97. Remote Sensing—A Scientific Vision for Sustainable Development. 1997 IEEE International*, 4: 1663–1665. doi: 10.1109/IGARSS.1997.609014

Cohen Y, Shoshany M, 2005. Analysis of convergent evidence in an evidential reasoning knowledge-based classification. *Remote Sensing of Environment*, 96(3–4): 518–528. doi: 10.1016/j.rse.2005.04.009

Datcu M, Melgani F, Piardi A *et al*., 2002. Multisource data classification with dependence trees. *IEEE Transactions on Geoscience and Remote Sensing*, 40(3): 609–617. doi: 10.1109/TGRS.2002.1000321

Davranche A, Lefebvre G, Poulin B, 2010. Wetland monitoring using classification trees and SPOT-5 seasonal time series. *Remote Sensing of Environment*, 114(3): 552–562. doi: 10.1016/j.rse.2009.10.009

Foody G M, 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1): 185–201. doi: 10.1016/S0034-4257(01)00295-4

Foucher S, Germain M, Boucher J M *et al*., 2002. Multisource classification using ICM and Dempster-Shafer theory. *IEEE Transactions on Instrumentation and Measurement*, 51(2): 277–281. doi: 10.1109/19.997824

Franklin S E, Peddle D R, Dechka J A *et al*., 2002. Evidential reasoning with Landsat TM, DEM and GIS data for landcover classification in support of grizzly bear habitat mapping. *International Journal of Remote Sensing*, 23(21): 4633–4652. doi: 10.1080/01431160110113971

Giarratano J C, Riley G D, 1998. *Expert Systems*: *Principles and*

Programming. Boston, MA: PWS-KENT Publishing.

Gordon J, Shortliffe E H, 1985. A method for managing evidential reasoning in hierarchial hypothesis space. *Artificial Intelligence*, 26(3): 323–357. doi: 10.1016/0004-3702(85)90064-5

Kim H, Swain P H, 1995. Evidential reasoning approach to multisource-data classification in remote-sensing. *Ieee Transactions on Systems Man and Cybernetics*, 25(8): 1257–1265. doi: 10.1109/21.398687

Le Hegarat-Mascle S, Bloch I, Vidal-Madjar D, 1997. Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4): 1018–1031. doi: 10.1080/014311600210678

Le Hegarat-Mascle S, Quesney A, Vidal-Madjar D *et al.*, 2000. Land cover discrimination from multitemporal ERS images and multispectral Landsat images: A study case in an agricultural area in France. *International Journal of Remote Sensing*, 21(3): 435–456. doi: 10.1080/014311600210678

Lein J K, 2003. Applying evidential reasoning methods to agricultural land cover classification. *International Journal of Remote Sensing*, 24(21): 4161–4180. doi: 10.1080/01431160 31000095916

Li J H, Chen W J, 2005. A rule-based method for mapping Canada's wetlands using optical, radar and DEM data. *International Journal of Remote Sensing*, 26(22): 5051–5069. doi: 10.1080/01431160500166516

Mertikas P, Zervakis M E, 2001. Exemplifying the theory of evidence in remote sensing image classification. *International Journal of Remote Sensing*, 22(6): 1081–1095. doi: 10.1080/ 014311601300074621

Na X D, Zhang S Q, Zhang H Q *et al*., 2009. Integrating TM and ancillary geographical data with classification trees for land cover classification of marsh area. *Chinese Geographical Science*, 19(2): 177–185. doi: 10.1007/s11769-009-0177-y

Ozesmi S L, Bauer M E, 2002. Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, 10: 381–402. doi: 10.1023/A:1020908432489

Peddle D R, 1995a. Mercury circle plus: An evidential reasoning image classifier. *Computers & Geosciences*, 21(10): 1163–1173. doi: 10.1016/0098-3004(95)00047-X

Peddle D R, 1995b. Knowledge formulation for supervised evidential classification. *Photogrammetric Engineering and Remote Sensing*, 61(4): 409–417.

Richards J A, Jia X, 1999. *Remote Sensing Digital Image Analysis.* Berlin: Springer-Verlag.

Shafer G, 1967. *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.

Srinivasan A, Richards J A, 1990. Knowledge-based techniques for multisource classification. *International Journal of Remote Sensing*, 11(3): 505–525. doi: 10.1080/01431169008955036

Sun W X, Liang S L, Xu G *et al.*, 2008. Mapping plant functional types from MODIS data using multisource evidential reasoning. *Remote Sensing of Environment*, 112(3): 1010–1024. doi: 10.1016/j.rse.2007.07.022

Tzeng Y C, Chiu S H, Chen D *et al.*, 2007. Multisource remote sensing images classification data fusion using a multiple classifiers system weighted by a neural decision maker. *Igarss: 2007 IEEE International Geoscience and Remote Sensing Symposium*, Vols 1–12: 3069–3073. doi: 10.1109/IGARSS. 2007.4423493

Vaiphasa C, Skidmore A K, de Boer W F, 2006. A post-classifier for mangrove mapping using ecological data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(1): 1–10. doi: 10.1016/j.isprsjprs.2006.05.005

Wilkinson G G, 2005. Results and implications of a study of fifteen years of satellite image classification experiments. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3): 433–440. doi: 10.1109/TGRS.2004.837325

Wilkinson G G, Megier J, 1990. Evidential reasoning in a pixel classification hierarchy—A potential method for integrating image classifiers and expert system rules based on geographic context. *International Journal of Remote Sensing*, 11(10): 1963–1968. doi: 10.1080/01431169008955152

Wright C, Gallant A, 2007. Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. *Remote Sensing of Environment*, 107(4): 582–605. doi:10.1016/j.rse. 2006.10.019

Zhang S Q, Na X D, Kong B *et al*., 2009. Identifying wetland change in China′s Sanjiang Plain using remote sensing. *Wetlands*, 29(1): 302−313. doi: 10.1672/08-04.1

Zhang S Q, Zhang J Y, Li F *et al*., 2006. Vector analysis theory on landscape pattern (VATLP). *Ecological Modelling*, 193(3–4): 492–502. doi: 10.1016/j.ecolmodel.2005.08.022