UNCERTAINTY ESTIMATION OF IMAGE CLASSIFICATION IN HAZE-CONTAMINATED AREA THROUGH WEIGHTS OF EVIDENCE MODEL

Yi-Shiang Shiu[1], Meng-Lung Lin[2], Yu-Chi Kao[1], Yung-Chung Chuang[1]

[1] Department of Urban Planning and Spatial Information, Feng-Chia University, No. 100, Wenhwa Rd., Seatwen Dist., Taichung, 40724 Taiwan
Email: ysshiu@fcu.edu.tw
[2] Department of Tourism, Aletheia University, No.32, Chen-Li St., Danshui District, New Taipei City, 25103 Taiwan
Email: mllin1976@mail.au.edu.tw

ABSTRACT: Haze is an inevitable interference when mapping land use/cover with optical satellite imagery. In this study, we applied weights of evidence model to locate possible area with omission and commission errors and estimate the degree of uncertainty for image classification in haze-contaminated area. Agricultural area in Kansas, U. S. is chosen as the study area. We used Landsat-5 TM images and hybrid classification to classify the land cover into four dominant classes in the study area: urban, cropland, grassland and water. Haze thickness and distances to clusters are assumed as the key factors for uncertainty estimation. For haze thickness, we applied Fourier analysis and high-pass filter to filter the haze-contaminated image and then spectral change vector analysis was used to acquire the magnitude of the difference between pre-filtered and post-filtered images. The magnitude of the difference can be considered as the haze thickness information. For distance to clusters, distances in multi-spectral space between a pixel's signature and all training clusters can be computed during the procedure of maximum likelihood classification. We chose distance to the first, second and third closest clusters as the representation of classification confidence. Finally, the weights of evidence model was used to combine the four key factors to map the uncertainty estimation results. Model assessment with the receiver operating characteristics (ROC) shows the area under curve (AUC) is 0.576. The model also indicates the two most significant factors are distance to first cluster and haze thickness, which is different from the findings of the previous study. Future applications include providing the possible misclassification areas for human-computer interactive interpretation. It can be helpful to improve the accuracy by using the human interpretation method to re-interpret the possible misclassification areas.

1. INTROCUTION

Image classification is considered to be the important process to map land use / land cover. However, misclassification is inevitable and can be a significant source of error when integrating remote sensing data with other geographic information system (GIS) data. Potential error sources come from the inability of classification systems to categorize mixed classes, transition zones, or dynamic systems; poorly defined or ambiguous class definitions; human subjectivity; and the lack of compatibility among different classification systems used with both remote sensing and traditional data types(Issues et al., 1991). Performances of classifiers and their associated uncertainties are commonly assessed with the confusion matrix and its derived measures including overall accuracy (OA), user's accuracy (UA), producer's accuracy (PA) (Congalton, 1991) and the kappa statistic (κ) (Cohen, 1960). However, this evaluation is based on the existence of ground truth data. Accuracy assessment with high degree of statistical confidence may need plenty of ground truth data, which also means taking plenty of time and manpower is inevitable.

From 1990's, several studies have proposed the concepts of "uncertainty" in image classification (Fisher, 1994; Van der Wel et al., 1998). Uncertainty comes from the discreteness nature of remotely sensed data which neglects the fuzzy character of the environment. Cloud or haze also decreases the image classification quality significantly because it influences the spectral characteristics of the ground features and makes them easily confused with other feature type (Cristóbal and Gabarda, 2007; Meng et al., 2009; Zhang et al., 2002). Haze-off methodologies mainly include histogram matching (Richter, 1996), regression tree (Helmer and Ruefenacht, 2005), image fusion (Gabarda and Cristóbal, 2007), dark channel prior method (DCPM)(Kaiming et al., 2011) and wavelet analysis of time-series imagery (Yong et al., 2002). These haze-off methodologies can help derive the degree of haze influence for image uncertainty estimation. Generally, uncertainty estimation can benefit the image classification process in two ways. First, it allows measuring the classification reliability before accuracy assessment with ground truth data. Second, it enhances accuracy assessments and improves classification confidence. Entropy, originating from information theory, was commonly used to define the uncertainty as the information content of a piece of information that would reveal this value with perfect accuracy (Van der Wel et al., 1998). In addition, derived information from classification

models can also help estimate uncertainty. For example, several classifiers, such as maximum likelihood, compute distances between the unique spectral signature of a given pixel and all possible clusters within an *n*-dimensional feature space that represents discrete land cover categories to decide the possible category of that given pixel. The distance to the second closest cluster can reflect the classification confidence (Mitchell et al., 2008). On the other hand, soft classification models with adaptive and flexible manner can also deal with the uncertainty and improve the classification accuracy (Binaghi et al., 2003; Comber et al., 2012; Wang, 1990). Similarly, decision tree classifiers, such as random forests (RF) and extremely randomised trees (ERT), are also able to model the uncertainty with Monte Carlo simulation (Barrett et al., 2014; Loosvelt et al., 2014).

For risk management, the most popular models used to carry out uncertainty estimation are certainty factor (CF) and weights of evidence (WoE). CF and Woe models with GIS spatial factors can be beneficial for landslide (Aboye, 2009; Binaghi et al., 1998; Devkota et al., 2013; Lan et al., 2004; Pourghasemi et al., 2012; Regmi et al., 2014) or flooding susceptibility mapping (Regmi et al., 2014; Tehrany et al., 2014). The above instances demonstrate that risk assessment can help to delineate environmentally sensitive areas and has been widely used to investigate the probability of risk occurrence and evaluate risk levels for hazard prevention (Zhang et al., 2009). In this study, we applied WoE model to map the uncertainty of image classification. Haze thickness and distances to clusters are assumed as the key factors for uncertainty estimation.

2. DATA AND METHODS

2.1 Study Area and Data

We chose part of Kansas State, United States as the study area (Fig. 1). The study area is about 101,524 ha. Cropland, pasture, orchards and groves dominate the land cover and show up as wide and homogeneous areas. We used Landsat-5 imagery with 6 bands besides thermal band and 30 m spatial resolution. In order to study the uncertainty of image classification in haze-contaminated area, we chose the image with haze significantly influencing the image quality and hindering the land use/cover analysis. The acquisition date of the image was April 24th, 2005. The image was converted from the digital numbers (DNs) to top-of-atmosphere (TOA) reflectance (Chander and Markham, 2003; Chander et al., 2007). The Kansas Gap Analysis Program (GAP) database and the attributed U.S. Department of Agriculture (USDA) Common Land Unit (CLU) dataset were used for training and validation (Fig.1). In order to avoid the geometric problem between the image and the CLU dataset, the image was orthorectified and registered with road network and CLU dataset to yield a root mean square error (RMSE) of 0.5 pixels.
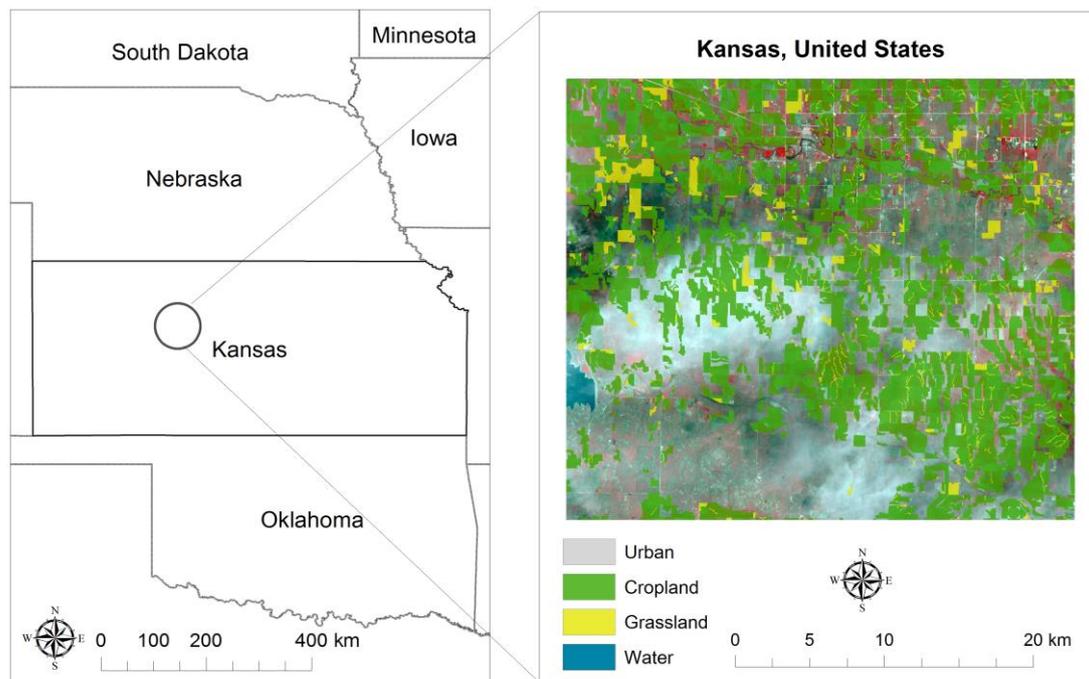
Fig. 1 Study area in Kansas, U. S. and the USDA CLU dataset.

2.2 Haze Detection

We revised the approach of hazy area recovery from Shiu et al. (2011) to model the haze distribution and thickness. Haze could be assumed as the low spatial frequency component in a hazy image because the spatial variation of its distribution is slower (on the scale of km) than for land cover which generally changes at higher frequencies (Feng et al., 2004; Yong et al., 2002). Therefore, Fourier analysis and filter can be used to reduce the haze and recover ground information. If we assumed the haze is homogeneous in each 1 km square area in the 30-meter resolution image, noise caused by haze would be periodic. Based on the assumption above, the image was diced into $M \times N$ pixels grids. We chose $M = 25$ and $N = 25$ for our image materials. Each grid was transformed to frequency domain with Fourier analysis and calculated as Fourier spectrum. The haze-off filter was derived from a series of image pairs in the training group in Shiu et al. (2011). Let the Fourier spectrum of one grid in a hazy image be denoted by A_{fi}, spectrum of one grid in a clear image be denoted by $A_{fi'}$, $i \in \{1,\ldots,m\}$ and m is the number of the grid pairs, $m \in N$. The haze filter, H_f, is given by the equation:

$$H_f = \frac{\sum_{i=0}^{m} \dfrac{A_{fi}}{A_{fi'}}}{m} \tag{1}$$

The haze filter H_f was used to be a moving window filter to remove the haze and recovery ground information in the hazy image. The filtered result was then transformed back to spatial domain with inverse Fourier transform to get the haze-off image. Spectral change vector analysis was then used to acquire the magnitude of the difference between hazy and haze-off images, which represents the degree of haze influence. The magnitude was utilized as the factor for the uncertainty estimation of image classification.

2.3 Image Classification

Based on the hybrid approach, with both unsupervised and supervised classification, a pixel-based process was carried out (Turner and Congalton, 1998). The hybrid approach used ISODATA (iterative self-organizing data analysis) and maximum likelihood classification techniques to overcome the high spectral heterogeneity and overlap caused by difference of cultivating time in the study area. Four classes including urban, cropland, grassland and water were classified. Mahalanobis distances between the measurement vector of one given pixel and the mean vector of all possible clusters were calculated. Pixels with higher Mahalanobis distance values are more likely to be misclassified. Contrarily, the pixels with lower values are spectrally nearer, and more likely to be classified correctly. We chose the distances to the first, second and third closest clusters for each pixel as the factors for the uncertainty estimation of image classification.

2.3 Uncertainty Estimation of Image Classification

With the selected factors from haze detection and Mahalanobis distances, the WoE model was then used to estimate the uncertainty of image classification. The uncertainty map can help find the potential areas where misclassification happened. A detailed description of WoE is available in Bonham-Carter et al. (2013). The method calculates the weight for each factors (F) based on the presence or absence of the misclassification (M) within the area as follows (Bonham-Carter et al., 2013; Dahal et al., 2008):

$$W_i^+ = \ln \frac{P\{F \mid M\}}{P\{F \mid \overline{M}\}}, \tag{1}$$

$$W_i^- = \ln \frac{P\{\overline{F} \mid M\}}{P\{\overline{F} \mid \overline{M}\}}; \tag{2}$$

where P is the probability and ln is the natural log. Similarly, F is the presence of a given factor, $\overline{F}$ is the absence of a given factor, M is the presence of the misclassified area and $\overline{T}$ is the absence of the misclassified area. A positive weight (W_i^+) indicates that the predictable variable is present at the misclassified locations and the magnitude of this weight is an indication of the positive correlation between the presence of the factors and the misclassified areas. A negative weight (W_i^-) indicates the absence of the factors and shows the level of negative correlation. The difference between the two weights is known as the weight contrast, W_f ($W_f = W_i^+ - W_i^-$); the magnitude of the contrast reflects the overall spatial association between the factors and the misclassified areas. To evaluate the contribution of each factor towards the misclassified areas, the layer of misclassified areas was compared to various factor layers separately. For this purpose, Eqs. (1) and (2) were written in a number of pixel format as follows:

$$W_i^+ = \ln \frac{\dfrac{Npix_1}{Npix_1 + Npix_2}}{\dfrac{Npix_3}{Npix_3 + Npix_4}} , \quad (3)$$

$$W_i^- = \ln \frac{\dfrac{Npix_2}{Npix_1 + Npix_2}}{\dfrac{Npix_4}{Npix_3 + Npix_4}} ; \quad (4)$$

where $Npix_1$ is the number of pixels representing the presence of both a given factor and the misclassified areas, $Npix_2$ is the number of pixels representing the presence of the misclassified areas and absence of a given factor $Npix_3$ is the number of pixels representing the presence of a given factor and absence of the misclassified areas, $Npix_4$ is the number of pixels representing the absence of both a given factor and the misclassified areas.

3. RESULTS AND DISCUSSION

Table 1 shows the accuracy assessment of hybrid classification. Generally, cropland and water areas show better classification results while urban areas exhibit lower accuracy. Although the overall accuracy can achieve over 80%, the kappa statistic is only 0.348. Low kappa statistic may reflect the influence of haze. The misclassified and correct areas were separated into training and validation groups with 1:1 proportion respectively in order to generate the uncertainty map and validate the WoE model. The results of haze detection and the Mahalanobis distances to the first, second and third closest clusters for each pixel are shown in Fig. 2. Like most uncertainty estimation approaches, the generalization of the factors with reclassification is inevitable because of the basic assumption of the models (Aboye, 2009; Lan et al., 2004; Pourghasemi et al., 2012). To be more objective, we reclassified each factors into 5 classes with natural-breaks (Jenk) method, which can minimize within-class variance and maximized between-class variance in an iterative series of calculations (Jenks and Coulson, 1963). Class 1 stands for the area with the lowest degree of haze influence or the lowest value of Mahalanobis distance while class 5 represents the highest. According to the Eqs. (3) and (4), W_i^+, W_i^- and W_f of each factors were calculated. The summary of the evidence classes identified as indicators for the potential misclassified areas is shown in Table 2. High W_f value indicates the given class of the given factor has high positive correlation with the presence of the misclassified areas, and vice versa. Different from the concept proposed by Mitchell et al. (2008), Table 2 shows that the first closest distance has positive relationship with the W_f; however, the second closest distance does not. This result implies that the first closest distance may be the better measure for uncertainty estimation of image classification.

The uncertainty estimation map for image classification can be generated by summing up all W_f value layers (Fig. 3). The uncertainty estimation map was classified into five levels: very low, low, uncertain, high and very high potential based on the natural-breaks method. Compared the original image in Fig. 1 with the uncertainty estimation map in Fig. 3, hazy areas generally show higher potential level of misclassification, which matches our assumption; the urban areas also show higher potential level than the other land features, which corresponds to the low accuracy of urban areas in Table 1. However, model assessment with the receiver operating characteristics (ROC) shows the area under curve (AUC) is 0.576. We speculate the low AUC may result from two reasons. First, the areas with ground truth data concentrate in the clear part rather than the hazy areas; meanwhile, the area proportion of the cropland class is much larger than the other classes. Second, we have not considered other factors dominating the classification uncertainty such as shadow caused by cloud and high land features.

4. CONCLUSIONS

This study provides an alternative solution for uncertainty estimation of image classification. Future applications include semiautomatic classification with human-computer interaction. In other words, the most possible misclassified areas can be defined for the artificial classification after the automatic approach. Future studies can focus on three issues. First, the proposed WoE model should apply to different study areas with more factors included. Second, the reclassification of each factor and the uncertainty estimation result would be a tricky step in the WoE model. Different reclassification method can lead different result. More objective substitute step has to be proposed to improve this uncertainty. Finally, methodologies of uncertainty estimation suitable for other classifiers are also necessary to develop for practical use.

Table 1 The error matrix of the hybrid classification result.

		Reference Data				
		Urban	Cropland	Grassland	Water	User's Acc.
Classified	Urban	255	3480	831	2	5.58%
	Cropland	115	358102	9727	42	97.31%
	Grassland	224	78967	32864	27	29.32%
	Water	0	197	26	1977	89.86%
	Producer's Acc.	42.93%	81.25%	75.64%	96.53%	80.77%

Table 2 Summary of the evidence classes identified as indicators for the misclassified areas (only two of the factors are shown here).

Factor	Class	$Npix1$	$Npix2$	$Npix3$	$Npix4$	W^+	W	W_f
First Closest Distance	1	1	46837	2173	1079038	-4.545	0.002	-4.547
	2	1048	45789	57090	1024121	-0.859	0.032	-0.890
	3	29387	17450	698861	382350	-0.030	0.052	-0.082
	4	13625	33212	272478	808733	0.144	-0.053	0.197
	5	2777	44060	50609	1030602	0.236	-0.013	0.250
Second Closest Distance	1	1	46837	466	1080745	-3.005	0.000	-3.005
	2	49	46788	6925	1074286	-1.812	0.005	-1.817
	3	20042	26795	440834	640377	0.048	-0.035	0.083
	4	21540	25297	493483	587728	0.008	-0.006	0.014
	5	5206	41631	139502	941709	-0.149	0.020	-0.169
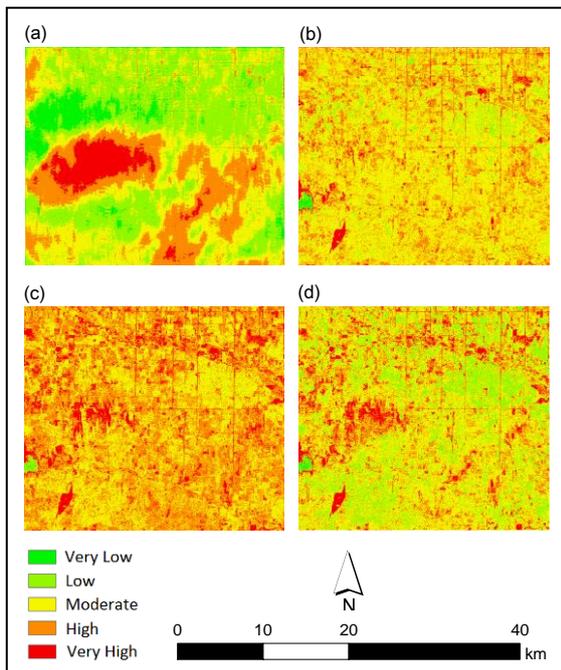
Fig. 2 Factors for the WoE model: (a) shows the degree of haze influence; and (b) to (d) represent the Mahalanobis distances to the first, second and third closest clusters.
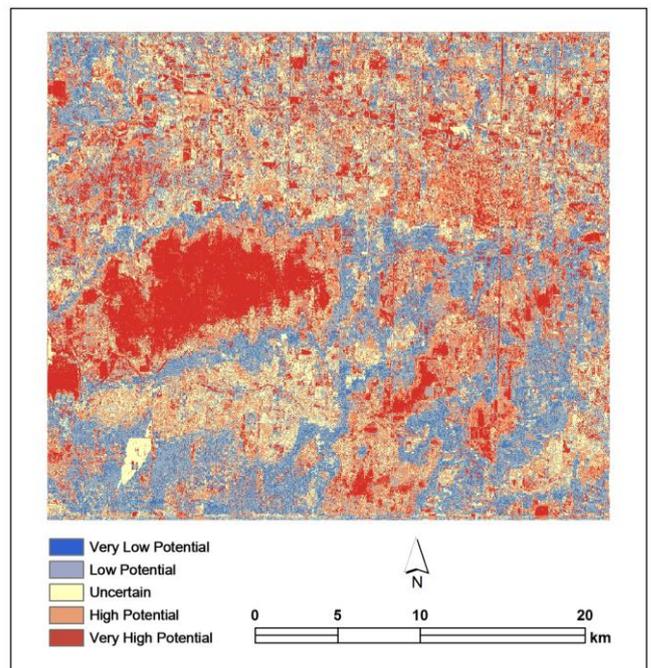
Fig. 3 Uncertainty estimation of image classification generated from the WoE model.

5. REFERENCES

[1] Aboye, S.A., 2009. Slope stability analysis using GIS and numerical modeling techniques. In: Physical Land Resources (p. 145). Brussels: University Ghent.

[2] Barrett, B., Nitze, I., Green, S., and Cawkwell, F., 2014. Assessment of multi-temporal, multi-sensor radar and ancillary spatial data for grasslands monitoring in Ireland using machine learning approaches. *Remote Sensing of Environment, 152*(0), pp. 109-124.

[3] Binaghi, E., Gallo, I., and Pepe, M., 2003. A cognitive pyramid for contextual classification of remote sensing images. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 41*(12), pp. 2906-2922.

[4] Binaghi, E., Luzi, L., Madella, P., Pergalani, F., and Rampini, A., 1998. Slope instability zonation: a Comparison between certainty factor and fuzzy Dempster–Shafer approaches. *Natural Hazards, 17*(1), pp. 77-97.

[5] Bonham-Carter, G.F., Agterberg, F.P., and Wright, D.F., 2013. Integration of Geological Datasets for Gold Exploration in Nova Scotia. *Digital Geologic and Geographic Information Systems* (pp. 15-23): American Geophysical Union.

[6] Chander, G., and Markham, B., 2003. Revised Landsat-5 TM radiometric calibration procedures and postcalibration dynamic ranges. *Geoscience and Remote Sensing, IEEE Transactions on, 41*(11), pp. 2674-2677.

[7] Chander, G., Markham, B.L., and Barsi, J.A., 2007. Revised Landsat-5 Thematic Mapper Radiometric Calibration. *Geoscience and Remote Sensing Letters, IEEE, 4*(3), pp. 490-494.

[8] Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), pp. 37-46.

[9] Comber, A., Fisher, P., Brunsdon, C., and Khmag, A., 2012. Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment, 127*, pp. 237-246.

[10] Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*(1), pp. 35-46.

[11] Cristóbal, G., and Gabarda, S., 2007. Cloud covering denoising through image fusion. *Image and Vision Computing, 530*, pp. 523-530.

[12] Dahal, R., Hasegawa, S., Nonomura, A., Yamanaka, M., Masuda, T., and Nishino, K., 2008. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environmental Geology, 54*(2), pp. 311-324.

[13] Devkota, K., Regmi, A., Pourghasemi, H., Yoshida, K., Pradhan, B., Ryu, I., Dhital, M., and Althuwaynee, O., 2013. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. *Natural Hazards, 65*(1), pp. 135-165.

[14] Feng, C., Ma, J.-w., Dai, Q., and Chen, X., 2004. An improved method for cloud removal in ASTER data change detection. In: Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International (pp. 3387-3389 vol.3385)

[15] Fisher, P.F., 1994. Hearing the Reliability In Classified Remotely Sensed Images. *Cartography and Geographic Information Systems, 21*(1), pp. 31-36.

[16] Gabarda, S., and Cristóbal, G., 2007. Cloud covering denoising through image fusion. *Image and Vision Computing, 25*(5), pp. 523-530.

[17] Helmer, E.H., and Ruefenacht, B., 2005. Cloud-Free Satellite Image Mosaics with Regression Trees and Histogram Matching. *Photogrammetric Engineering and Remote Sensing, 71*, pp. 1079-1089.

[18] Issues, R., Fenstermuker, L.K., McGwire, K.C., and Tinney, L.R., 1991. Remote sensing and geographic information system data integration: error sources and research issues. *PE & RS, 57*, pp. 677-687.

[19] Jenks, G.F., and Coulson, R.C., 1963. Class intervals for statistical maps. *International Yearbook of Cartography, 3*, pp. 119-134.

[20] Kaiming, H., Jian, S., and Xiaoou, T., 2011. Single Image Haze Removal Using Dark Channel Prior. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 33*(12), pp. 2341-2353.

[21] Lan, H.X., Zhou, C.H., Wang, L.J., Zhang, H.Y., and Li, R.H., 2004. Landslide hazard spatial analysis and prediction using GIS in the Xiaojiang watershed, Yunnan, China. *Engineering Geology, 76*(1-2), pp. 109-128.

[22] Loosvelt, L., De Baets, B., Pauwels, V.R.N., and Verhoest, N.E.C., 2014. Assessing hydrologic prediction uncertainty resulting from soft land cover classification. *Journal of Hydrology, 517*(0), pp. 411-424.

[23] Meng, Q., Borders, B.E., Cieszewski, C.J., and Madden, M., 2009. Closest Spectral Fit for Removing Clouds and Cloud Shadows. *Photogrammetric Engineering and Remote Sensing, 75*(5), pp. 569-576.

[24] Mitchell, S.W., Remmel, T.K., Csillag, F., and Wulder, M.A., 2008. Distance to second cluster as a measure of classification confidence. *Remote Sensing of Environment, 112*(5), pp. 2615-2626.

[25] Pourghasemi, H., Pradhan, B., Gokceoglu, C., Mohammadi, M., and Moradi, H., 2012. Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran. *Arabian Journal of Geosciences*, pp. 1-15.

[26] Regmi, A., Devkota, K., Yoshida, K., Pradhan, B., Pourghasemi, H., Kumamoto, T., and Akgun, A., 2014. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arabian Journal of Geosciences, 7*(2), pp. 725-742.

[27] Richter, R., 1996. Atmospheric correction of satellite data with haze removal including a haze/clear transition region. *Computers & Geosciences, 22*(6), pp. 675-681.

[28] Shiu, Y.-S., Lin, M.-L., and Chu, T.-H., 2011. Mapping and recovering cloud-contaminated area in multispectral satellite imagery with visible and near-infrared bands. In: Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International (pp. 543-546)

[29] Tehrany, M.S., Pradhan, B., and Jebur, M.N., 2014. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *Journal of Hydrology, 512*, pp. 332-343.

[30] Turner, M.D., and Congalton, R.G., 1998. Classification of multi-temporal SPOT-XS satellite data for mapping rice fields on a West African floodplain. *International Journal of Remote Sensing, 19*(1), pp. 21-41.

[31] Van der Wel, F.J.M., Van der Gaag, L.C., and Gorte, B.G.H., 1998. Visual exploration of uncertainty in remote-sensing classification. *Computers & Geosciences, 24*(4), pp. 335-343.

[32] Wang, F., 1990. Fuzzy Supervised Classification of Remote Sensing Images. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 28*, pp. 194-210.

[33] Yong, D., Guindon, B., and Cihlar, J., 2002. Haze detection and removal in high resolution satellite image with wavelet analysis. *Geoscience and Remote Sensing, IEEE Transactions on, 40*(1), pp. 210-217.

[34] Zhang, Y., Guindon, B., and Cihlar, J., 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. *Remote Sensing of Environment, 82*(2–3), pp. 173-187.

[35] Zhang, Z.J., Yang, Z.X., Zhang, P., Mao, Z.J., and Hao, J.Y., 2009. Hierarchical network-based safety assessment decision support system for thermal power plants. In: Networking, Sensing and Control, 2009. ICNSC '09. International Conference on (pp. 592-596). Okayama City, Japan