APPLICATION OF GEOBIA-MULTINOMIAL LOGISTIC REGRESSION FOR LANDSLIDE VULNERABILITY ASSESSMENT IN KAYANGAN CATCHMENT INDONESIA

G. Samodra [a,b] *, G. Chen [a], J. Sartohadi [b], Z. Kouki [a], K. Kasama [a], D. S. Hadmoko [b]

[a] Civil and Structural Engineering Department, Kyushu University, Japan
[b] Environmental Geography Department, Universitas Gadjah Mada, Indonesia

KEY WORDS: GEOBIA, Multinomial logistic regression, Landslide, Vulnerability

ABSTRACT:

An effective and low-cost vulnerability assessment based on single object settlement is necessary for landslide risk assessment in medium scale mapping. GEOBIA-Multinomial logistic regression was employed to reduce the drawback of landslide vulnerability mapping based on census procedure. The aim of this research is to assess landslide vulnerability of identified settlement with application of GEOBIA-Multinomial logistic regression model in Kayangan Catchment, Yogyakarta Indonesia. Region growing algorithm was applied to segment SPOT 5 pan-sharp image into settlement and non-settlement. The household survey technique with 112 respondents was applied to generate multinomial logistic regression model. It was developed to imitate the ability of human vision to interpret the occurrence of an object related to another object. Two environmental factors i.e. slope and distances to road were employed to model the spatial prediction of vulnerability. Slope and distance to road as predictor variables were employed to predict landslide vulnerability as categorical dependent variable. The result showed that high vulnerability degree was located in the hilly area and low degree of accessibility. Meanwhile, low vulnerability degree was located in the more gentle area and high accessibility degree (near to the centre of economic, education, and government activity). The GEOBIA-Multinomial logistic regression accuracy attained 93% and 80% for image classification and spatial vulnerability prediction respectively. It performed well to assign the probability values for the settlement vulnerability in Kayangan Catchment. Expert judgment based on the local knowledge of social structure play important role to design the predictor variables in vulnerability assessment.

1. INTRODUCTION

Recently, landslide risk analysis is not only focused on landslide hazard mapping and analysis. It is necessary to recognize the element at risk, vulnerability, and social resilience (Birkmann, 2006). Vulnerability assessment is one of key point to determine risk assessment comprehensively. UN/ISDR (2004) defines vulnerability as the physical, social, economic and environmental condition of an individual or community which may increase the susceptibility when the disaster occurred. The damage of building as a consequence of landslide disaster is not enough to portray the effect of its disastrous events. Total effects might be greater than the damage of building as measured as a consequence. It would be different related to the ability of individual, community, or society to cope and to anticipate the impact of disaster. For instance, people living under poverty line will have greater effect due to the limitation access to the resources and the less capacity to recovery. Landslide vulnerability assessment is an important thing to formulate social and physical aspect of disaster as a result of interaction between landslide occurrence and potentially suffering event.

The application of vulnerability assessment is usually applied over to the large area (Arakida, 2006; Birkmann, *et al*., 2006; Bollin and Hidajat, 2006) or in site-specific location (Bell and Glade, 2004). Large area analysis is usually based on the analysis of secondary data i.e. age, gender, literacy, household size and population represented in administrative unit. However, it will face difficulties if the analysis is applied on the basis of basin or catchment scale in which its boundary does not match with the administrative boundary. In addition, high level generalization is demanded when it is applied in large scale vulnerability mapping analysis. In the other hand, site-specific vulnerability data should be collected by direct survey (primary data) through census or and community based methods. Those can portray vulnerability of a household, single building or settlement area. However, data collected using census and community based methods have limited efficiency and transferability (Ebert *et al*., 2008). It is time consuming and costly.

An approach for more effective and efficient data handling to assess vulnerability of a settlement or single building remains a challenge in medium to small scale mapping analysis. The availability of high resolution of satellite imagery shows the potential for the application of vulnerability assessment of a single building or settlement area. Enhanced image quality and new image processing methodologies have shown their high potential for the application of vulnerability assessment.

Vulnerability can be determined through specific proxies related to the specific vulnerability (Elbert and Kerle, 2009). It is much possible to interpret a single object i.e. settlement area or single building related to vulnerability by several key interpretation such as colour, shape, texture, size and

* Email: guruh.samodra@gmail.com

occurrence to other object. Image interpretation by its key interpretation has been applied widely by many disciplines concerned with geo-science since long time ago. It was applied widely by manual image interpretation of many applications such as land use mapping, geomorphology mapping, geology mapping, and soil mapping. In recent decade, the idea of interpretation based on spatial pattern rather than pixel based is transferred to the application of Object Based Image Analysis (OBIA) later called as Geographic Object based Image Analysis (GEOBIA).

GEOBIA was developed based on the concept of regionalization (Strobl, 2009). The object extraction from remotely sensed data is obtained based on the spatial information which is similar to key interpretation in manual image interpretation process. It becomes popular due to the increasing of high resolution satellite image availability and advance computer processing. Hay and Castilla (2006) proposed the definition of GEOBIA as a sub-discipline of GIScience in which the primary tasks are to classify remote sensing imagery into meaningful objects and to assess their characteristics through spatial, spectral and temporal scale. GEOBIA was developed by segments as the basic processing units, therefore called as segmentation. The basic processing of segmentation is that dividing image objects into regions based on their common properties. It involves region merging to measure homogeneity and separating objects through digital numbers (DNs) between neighboring pixels (Navulur, 2007). It was developed to imitate the ability of human vision to interpret an object of image through computer machine.

The use of spatial pattern in GEOBIA combined with GIS techniques is a promising methodology applied in vulnerability assessment. It can reduce timely and cost consuming census procedure. The understanding of vulnerability's spatial pattern with the environmental condition is a necessary to interpret the meaning of object behind the pixel in imagery. It can be modelled in GIS platform and employed in the procedure of GEOBIA. Multinomial logistic regression was applied to model the spatial pattern of vulnerability. Several sample respondents were taken to build vulnerability model. Therefore, it was used to predict and extrapolate the vulnerability in Kayangan Catchment. Multinomial logistic regression was developed to imitate the ability of human vision to interpret the occurrence of an object related to another object. Moreover, it was applied as expert knowledge to assist the assessment of landslide vulnerability. Expert knowledge can be captured as knowledge based system and is useful to assist a decision making in a particular feature (Richards and Jia, 2006). The input of "knowledge" was obtained from the spatial pattern of vulnerability assessed from household survey. It included physical, social and economic dimension of vulnerability. Thus, the aim of this research is to assess landslide vulnerability of identified settlement with application of GEOBIA-Multinomial logistic regression model in Kayangan Catchment, Yogyakarta Indonesia.

2. STUDY AREA

The research area was conducted in Kayangan Catchment Kulon Progo, Yogakarta Indonesia (Figure. 1). The area of Kayangan Catchment extends 4 sub-districts i.e. Girimulyo, Nanggulan, Samigaluh and Kaligesing. It is located in the middle Java Island and comprises 35 km^2. The average annual rainfall in Kayangan Catchment is 2478 mm. The highest rainfall intensity usually occurs from February to March with average monthly rainfall 426 mm. Landslide usually occurs in the month of November to April during wet season. The lithology is dominated by medium to heavily weathered andesitic breccias. Land use in the study area can be classified into bushes, rain fed paddy field, irrigated paddy field, people forest, settlement, and dry cultivated land.

Kulon Progo area has been traditionally dominated by agricultural sector. Around 78% household working in agricultural sector and mostly (88%) are living in rural areas. With the limitation of economic development and infrastructure, 40.31 % people are living below national poverty line (BPS, 2008). Economic and social condition of society may increase the vulnerability degree.

3. METHODOLOGY

3.1. Household Survey and Questionnaire

The sample of landslide vulnerability was conducted by household survey using observation, photograph analysis and interview. Household survey was conducted by transect walk and structured (closed question) interview. Photograph analysis of building was also used to accompany the analysis of household survey to portray the physical vulnerability of landslide (building architecture). It was directed to obtain data related to vulnerability indicators (Table 3.1). Therefore, the landslide vulnerability assessment was analyzed with ©SPSS 15 software through frequency analysis and crosstab analysis based on the method proposed by Dwyer, et al., (2004) with several modifications. Modification was intended to give alternative indicators and method that cannot be applied in Kayangan Catchment due to the uniqueness of social and local culture.

Table 3.1. Vulnerability Indicators

No.	Type of Vulnerability Indicator	Vulnerability Indicators	Data Inquiry
1.	Physical	1. Building Material 2. Construction type 3. Building age 4. Infrastructure (distance to road)	Household survey & Photo
2.	Social	1. Age 2. Gender 3. Preparedness 4. Knowledge of landslide mitigation	Household survey
3.	Economic	1. Income 2. Saving 3. Building architecture	Household survey

Physical Vulnerability was assessed through field survey using observation and photograph analysis. It was taken from 112 respondents in different settlement building and then was plotted using GPS (Global Positioning System). Social and economic vulnerability samples were also taken from 112 respondents which are also plotted using GPS. Therefore, multinomial logistic regression was employed to model the spatial distribution of the landslide vulnerability degree towards its environmental condition in Kayangan Catchment.

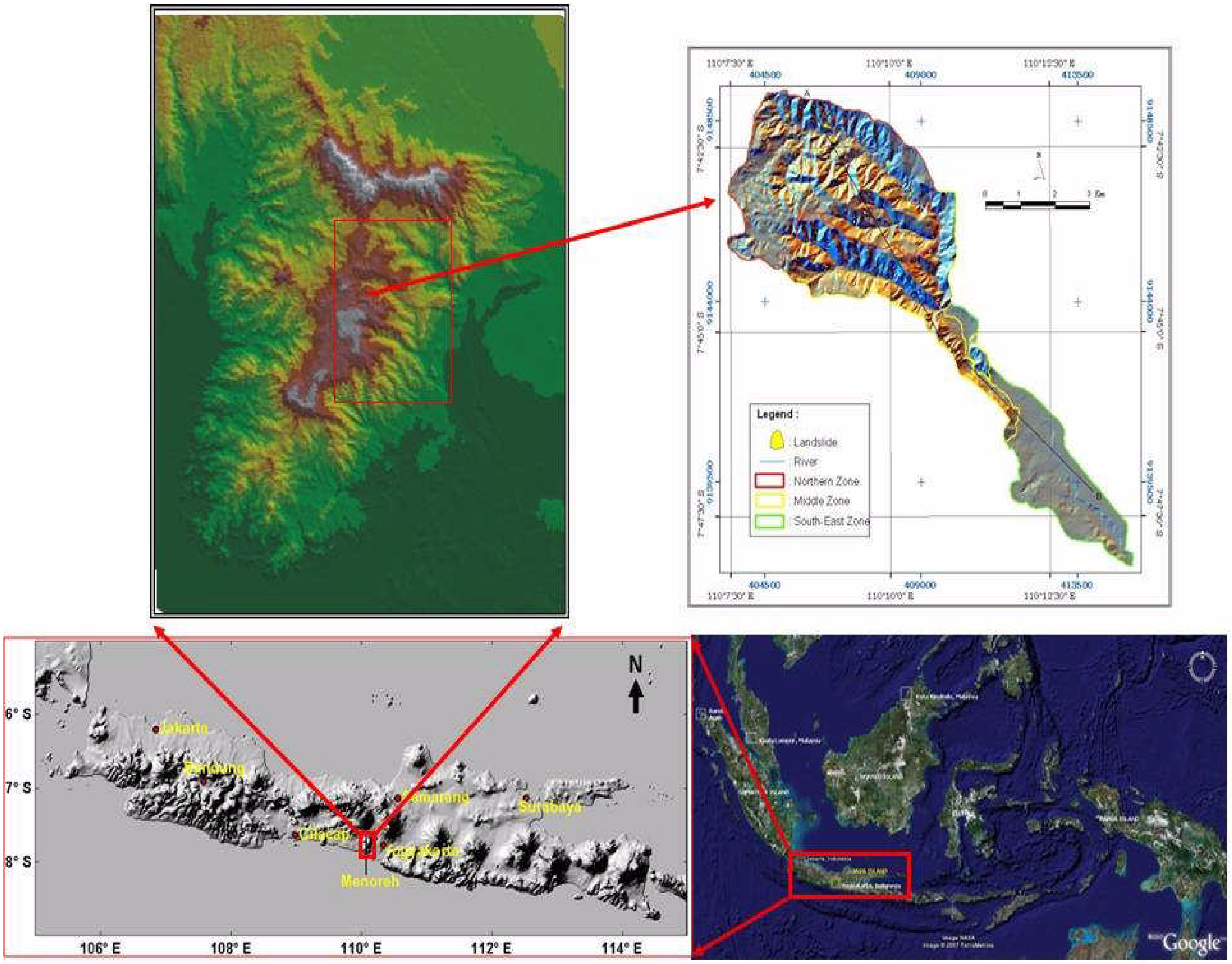

Figure 1. Location of Study Area, Kayangan Catchment Indonesia

3.2. GEOBIA-Multinomial Logistic Regression

The procedure of GEOBIA can be classified into pre processing, segmenting image, collecting sample, identifying optimum feature, calculating final threshold, and finally classification. ©ENVI 4.5 Software was used to preprocess the image, whereas segmenting to classification was done by SPRING Software (Camara et al., 2006). Image preprocessing phase of this research involved geometric correction and image fusion. Geometric correction included correcting geometric distortion caused by sensor-earth geometry and conversion of the data to the real world coordinates on the earth's surface. Imagery data used in this research was raw image-Level 1A ©SPOT 5 HRG XS and Panchromatic May 2006 obtained from LAPAN (National Institute of Aeronautics and Space Indonesia). It has been systematically radiometric corrected. Systematic errors caused by space borne sensor irregularities such as scan skew errors, mirror-scan velocity variance, panoramic distortion, platform velocity errors, and earth rotation have been corrected in this level. It means that some distortion may be removed using technique which model systematic orbital and sensors characteristic (Jensen, 1986). Unfortunately, it does not remove error produced by changes in altitude.

There is limited geometric correction performed in level 1 A SPOT 5 Image. It has minimum processing involved detector equalization which compensate the differences sensitivities between elementary detectors of the CCD (Charged Coupled Device) (Levine, 1999). Thus, orthorectification or 3D transformation was applied in this research to correct geometric errors. RPC (Rational Polynomial Coefficient) from SPOT 5 metadata/image header and DEM obtained from topographic contour were used to orthorectify the image. Orthorectification has been done by using ©ENVI 4.5 Software. RPCs were computed using a digital photogrammetry technique using a collinearity equation to construct sensor geometry (ENVI Help).

Eleven GCPs were selected to enhance the accuracy of the RPC orthorectification process. Therefore, the position accuracy of the imagery was evaluated by both quantitative and qualitative.

Quantitative accuracy was evaluated by using the value of RMSE (Root Mean Square Error), whereas qualitative accuracy was evaluated by topography map.
Furthermore, orthorectified multispectral SPOT data with 10m spatial resolution was fused with the orthorectified panchromatic SPOT data with 2.5 m resolution. It fits the suggested minimum requirement spatial resolution used for medium mapping purpose. According to Richards and Jia (2006), an image with the spatial resolution 1 m to 5 m can be used to produce a medium scale map of 1:10.000 to 1:50.000.
HIS (Hue-Saturation Index) with the combination of 432 composite was applied in this research. The basic procedure of HIS transform is that separating a standard RGB (Red, Green, Blue) image into spatial (I) and spectral (H,S) information. According to (Zhang, 2002), the steps can be classified into (1) transforming a color image composite from the RGB space into the IHS space, (2) replace the I (intensity) component by a panchromatic image with a higher resolution, (3) reversely transform the replaced components from IHS space back to the original RGB space to obtain a fused image.

Then, Region growing algorithm was applied to segment SPOT 5 pan-sharp image. Several samples are depicted to classify settlement and non-settlement. Trial and error were applied to determine final threshold. After being segmented, SPOT 5 pan-sharp image was classified into settlement and non-settlement by bhattacharya classifier algorithm and masked into building block of the element at risk. Therefore, the masked building object was validated using building object recorded by GPS plotting.

Multinomial logistic regression was employed to assess landslide vulnerability of object building. It represented the human cognition to reinterpret the spatial pattern of landslide vulnerability toward its environmental factor. Two environmental factors i.e. slope and distances to road were employed to model the spatial prediction of vulnerability. It was deducted from the analysis of vulnerability pattern (tendencies) based on expert judgment. There was an indication that high vulnerability degree was located in the hilly area and low degree of accessibility. Meanwhile, low vulnerability degree was located in the more gentle area and high accessibility degree (near to the centre of economic, education, and government facility). Therefore, slope and distance to road as predictor variables were employed to predict landslide vulnerability as categorical dependent variable in SPSS software through multinomial logistic regression. Low vulnerability class was chosen as reference category.

The Chi-square test, wald statistic, coefficient (B) and EXP(B) were used to determined the fitness of the model. Chi-square test represented the influences of the predictors to the dependent variables. Wald statistic represented the influence of a change in given predictor variables to a change in in the odds ratio of the dependent variable. Coefficient (B) represented the significance correlation between the vulnerability and the environmental factors. The interpretation of the significance correlation between the vulnerability and the environmental factors was also evaluated using EXP(B) or odds ratio. Odds ratio greater than 1 indicated positive correlation between the environmental factors and the probability of vulnerability. In the other hand, odds ratio less than 1 indicated the negative correlation between the environmental factors and the probability of vulnerability.
Furthermore, confusion matrix was employed to the evaluation of spatial prediction accuracy.

In RS-GIS platform, the probability of odds ratio represented logit or logarithmic function between (P) and (1-P)) is written as:

$$z_{ij} = \ln\left(\frac{P_{ij}}{1-P_{ij}}\right) = a_{ij} + b_{1j}X_{1i} + b_{2j}X_{2i} + \dots + b_{nj}X_{ni} \quad (1)$$

$$P_{ij} = \frac{e^{z_{ij}}}{1+\sum_{j=1}^{m-1} e^{z_{ij}}} \quad (2)$$

$$P_r = \frac{1}{1+\sum_{j=1}^{m-1} e^{z_{ij}}} \quad (3)$$

where z is logarithmic function of the ratio between the probability (P) and (1-P) in which a pixel (i) is a member of class vulnerability j, a indicates the intercept of the regression curve for vulnerability class j, $b_{1j \dots nj}$ are the coefficient of each predictor $X_{1i \dots ni}$ for the respective vulnerability class j. The n is the total number of environmental factors of vulnerability. P_{ij} is the probability of pixel i in a vulnerability class j. m stands for the total number of dependents categories, whereas the • is the summation of logits of all vulnerability for the particular pixel i except the reference vulnerability class. “r” is reference and the probability is given in Eq. 3.

Each vulnerability category (high and medium) had a value of ‘a’ and ‘b’ based on the empirical data (analysis of questionnaire data). It is used to generate probability of high and medium vulnerability based on the Eq. 2. And probability of low vulnerability was generated by Eq. 3 as reference. Therefore the probability model was calculated in ILWIS 3.7 software to produce map which showed the probability of presence each category of vulnerability in a single settlement block.

4. RESULT AND DISCUSSION

4.1. Spatial Pattern of Settlement and Spatial Vulnerability in Kayangan Catchment

The vulnerability of Kayangan Catchment more tends to have such a pattern rather randomly. Since the nature of people, social structure and culture of Kayangan Catchment were less influenced by the development of technology and modernization, the natural feature and condition would more dominantly influence the way of life of the people in Kayangan Catchment. According to Whynne-Hammond (1979) the growing of settlement (village) is based on the combination of four factors i.e. topographical, economic, historical, and cultural.

Topographical and economic factors are the most factor influence the growing of settlement or preferences of people to build a settlement in Kayangan Catchment. There were several options for a person to build a settlement, but there was limited option for the poor. Settlement type was often a reflection of prosperity. It was usually that the poor could not choose either they would build their home in more gentle topography or hilly area. Therefore, the poor tend to build settlement based on their land given from their parent or ancestor. There were usually located in more hilly, remote, and lower accessibility. For the rich, there were several options to build a settlement. They usually build a settlement in relatively gentle topography and have good accessibility. Good accessibility is the most common reason, in which place, they would build a settlement. Even, with the reason of accessibility, they would build a settlement by excavating and cutting the slope into more gentle topography.

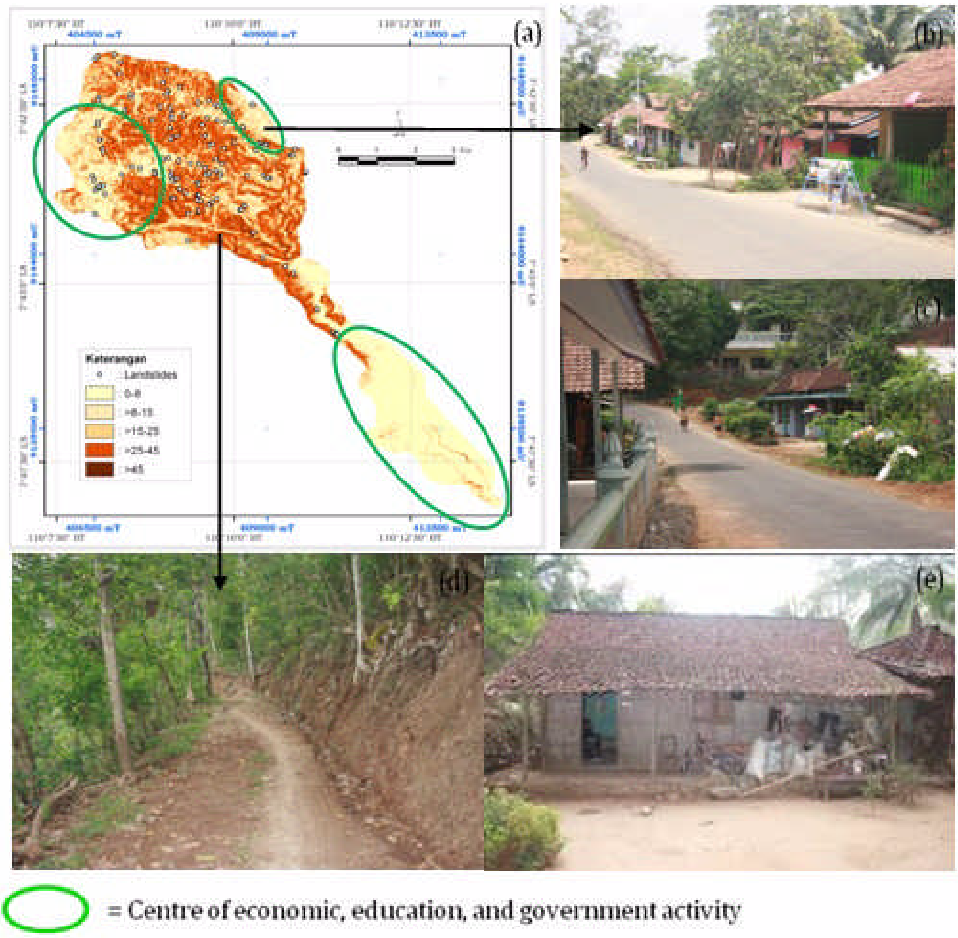

Figure 4.1 (a) Slope Map of Kayangan Catchment (b) Good Road showing High Accessibility (c) Centre of Education (d) Bad Road Condition Showing Low Accessibility (e) High Vulnerability Degree or a Household in Hilly Area

Therefore, the preferences of people to build settlement will influence the distribution of the settlement location. Whynne-Hammond (1979) classified rural settlement/village into loose-knit villages, clustered villages, linear villages, open space village and double village. The distribution of settlement in Kayangan Catchment is likely similar with the combination of loose-knit village and linear village. In the loose knit village, settlement are scattered irregularly. This pattern represented the settlement that was built by the people living under poverty line. The poor tend to build settlement in remote area and hilly area. It caused irregularity of settlement distribution. Linear village represent the preference of the people that build settlement near to the road (have high accessibility). It was built by the people with high prosperity. A new house building was usually built by excavating the slope without reinforcement. It may also lead to the reduction of the instability of natural slopes and can cause landslides. It was usually built in the steep area and far from the centre of government, education and economic activity.

Thus, there was spatial tendency between topography, economy, village distribution and vulnerability. People living in hilly area and low accessibility were the poor. In consequences, it would increase the vulnerability degree either physical, social or economic vulnerability. High vulnerability degree was located in the hilly area and low degree of accessibility. Meanwhile, low vulnerability degree was located in the more gentle area and high accessibility degree (near to the centre of economic, education, and government activity). There were traditional market, village office, and school in this place. Thus, multinomial logistic regression was developed in order to determine the spatial characteristic of vulnerability.

4.2. Pre Processing and Image Fusion

The use of high resolution image in many applications was increasing significantly due to the availability of high to very high resolution. Unfortunately, the availability of very high resolution satellite imagery is very rare in a less developed country such as Indonesia. Thus, it is a critical to use a kind of technique which is directed to improve both spatial and spectral resolution of an image cost effective. Image data fusion has been applied widely to minimize those backward. Image fusion is the combination of two or more different spatial or spectral resolution of images to form a new image by using certain algorithm (Pohl and van Genderen, 1998). It was directed to achieve improved information for decision making. Image fusion sometimes is confused with pan sharpening. Pan-sharpening is an image fusion method in which high resolution panchromatic data is fused with lower resolution multispectral data to create a colorized high spatial resolution dataset.

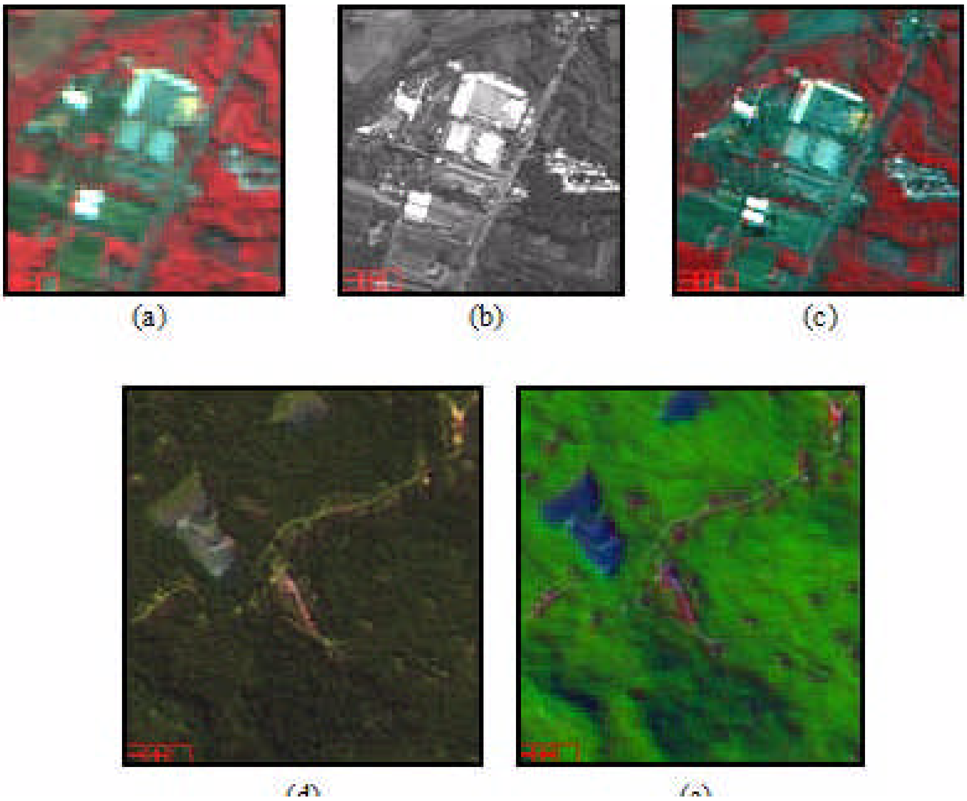

Figure 4.2 Multispectral SPOT Image 10 m (a) Panchromatic SPOT 2.5 m (b) Pan-sharp Multispectral SPOT 2.5 m (c) HIS Pan-sharpening (d) Gram-Schmidt Pan-sharpening (e)

In this research, orthorectified multispectral SPOT data with ith the orthorectified

with 2.5 m spatial resolution (Figure
. The result of orthorectification showed that RMS Error
pectral and 3.2 for panchromatic. The

ification has been matched with vector topography

Formerly, Gram-Schmidt pan-sharpening was used in this

is typically more accurate

evaluated qualitatively (Figure 4.2). Thus, HIS (Hue-Saturation Index) was applied in this research.

4.3. GEOBIA-Multinomial Logistic Regression and Its Performance

Recently, remote sensing technique has been significantly increasing due to the availability of very high resolution whether spatial, temporal and spectral of satellite imagery. It has also been supported by high capability of computer machine. Thus, the use of remote sensing techniques in many applications can be less time consuming and cost effective. One of the relatively new applications of remote sensing is object based image analysis. The application of object based image analysis in many applications is still develop in the last decade.

The emergence of object based image analysis an opposed to the single pixel based analysis is due to the limitation of single pixel based analysis. (Baatz et al., 2004) defines that single pixel based image analysis have limitation related to the rich information of very high resolution data. It includes the inconsistent salt-and-pepper classification and having limited capabilities of extracting object of interest. The limitation of single pixel based classifier, then, has been addressed by object based image analysis. Object based image analysis assumes that the semantic information needed to analyze an object is not merely available in the pixel, but image objects and the relation among the object (Baatz et al., 2004). Thus, in object based classification homogeneous image are firstly extracted and then classified into labelled image.

The main procedure of object based image analysis is segmentation as the basic processing unit. It divides image objects into homogeneous image segment. According to Marpu et al. (2008) the procedure of object based image analysis can be classified into pre processing, segmenting image, collecting sample, identifying optimum feature, calculating final threshold, and finally classification. In this research, those procedures were run in SPRING Software (Camara et al., 2006) in order to segment element at risk of landslide (settlement).