

CHAPTER 3

THEMATIC DATA PREPARATION

The sources of input data and the steps in input data production to be comprehensively explained hereafter indicate that data entry and production are the most cumbersome and time consuming steps of any kinds of GIS and remote sensing techniques. The thematic data used in this thesis are to be prepared and processed below. Meanwhile, phases of natural hazard analysis in GIS-based landslide hazard zoning techniques are reviewed. However, the detailed statistic hazard analysis of the flow-flood database and the parameter maps are to be explained in the following chapter.

3.1 Phases of natural hazard analysis in GIS-based landslide hazard zoning techniques

The following phases can be distinguished in the process of a hazard analysis using GIS (Van Westen, 1993 and 1994). They are listed in logical order or sequence though sometimes they may be overlapping (Figure 3-1) as follows:

- Preliminary phase:

- Phase 1: Defining the objective of study, the working scale and the methods of analysis to be applied

- Data collection phases:

- Phase 2: Collection of existing data (collection of existing maps and reports with relevant data)

- Phase 3: Image interpretation (interpretation of images and creation of new input maps)

- Phase 4: Data base design (design of the database and definition of the way in which the data will be collected and stored)

Phase 5: Fieldwork (to verify the photo-interpretation and to collect relevant quantitative data)

Phase 6: Laboratory analysis

■ GIS work:

Phase 7: Data entry (digitizing of maps and attribute data)

Phase 8: Data validation (validation of the entered data)

Phase 9: Data manipulation (manipulation and transformation of the raw data in a form which can be used in the analysis)

Phase 10: Data analysis and modeling (analysis of data for preparation of hazard maps)

Phase 11: Presentation of output maps (final production of hazard maps and adjoining report)

Phase 12: Error evaluation and reporting (evaluation of the reliability of the input maps and inventory of the errors which may have occurred during the previous phases)

The phases of a GIS-supported landslide hazard assessment project are in part different from those in a conventional project. An ideal GIS for landslide hazard zonation combines conventional GIS procedures with image processing capabilities and a relational data base. Map overlaying, modeling, and integration with remote sensing images (scanned aerial photos and satellite images) are required, thus a raster system is preferred. The program should be able to perform spatial analysis on multiple-input maps and connected attribute data tables for map overlay, reclassification, and various other spatial functions, incorporating logical, arithmetical, conditional, and neighborhood operations, including iteration. The system should provide the use of batch files and macros for those models that require similar analysis using different parameters. Since most data sets for landslide hazard zonation project are still relatively small (mostly less than 100 Megabyte), the use of less expensive, PC-based systems is possible.

3.2 Thematic data preparation from GIS and remote sensing techniques

Remote sensing data can be readily merged with other sources of geo-coded information as a GIS. This allows the overlapping of several layers of information with the remotely sensed data, and the application of a virtually unlimited number of forms of data analysis.

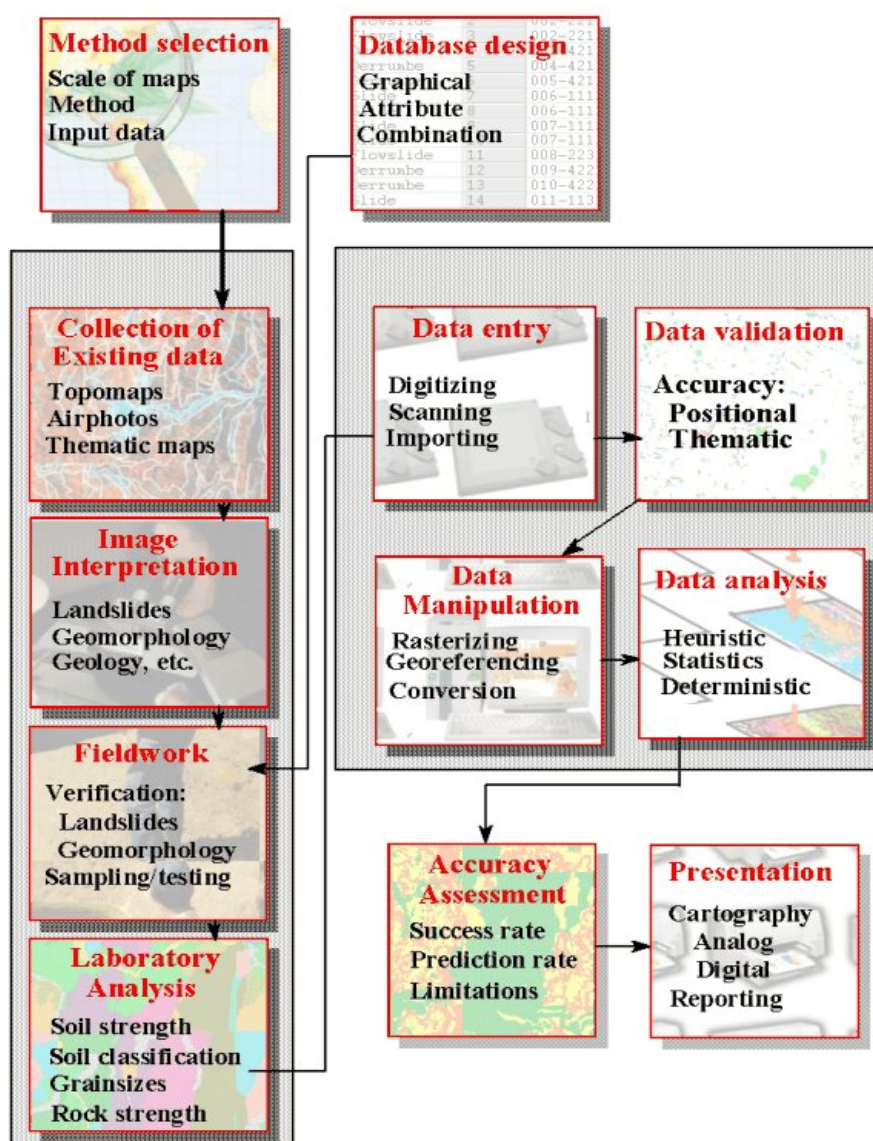


Figure 3-1 Flow chart of a GIS-based landslide hazard zonation (Van Westen, 1994).

As the use of geographic information systems increases, the availability of timely and up-to-date spatial data in digital format is required; this is expected to be easily fulfilled. Satellite imageries combined with the increased processing capability to generate meaningful data sets which represent new knowledge now available in various technologies.

GIS is suitable to meet the requirements of synthesizing the available information. The strength of a GIS lies in this capability of storing interpreted and available information as maps and linked attributes. For developing a GIS application as a landslide hazard management tool, a three-tiered GIS approach is adopted as follows; hazard assessment, vulnerability assessment and risk assessment. The parameters to be considered for assessment of the landslide hazard are vulnerability and risks being included in a landslide map, major land use/ land cover categories and topographic factors. The resulting landslide hazard management tool will help identifying the occurrence of landslides, the degree of loss as a factor of vulnerability, and will ultimately allow the assessment of risk from landslides. Therefore, risk assessments are a combination of hazard and vulnerability measurements that will assist with predicting locations where landslide events may cause damages.

The input data used for flow-flood hazard assessment in this thesis consists of several spatial data categories from the available resources (as shown in Table 3-1), being digitized from available maps and prepared from image interpretation, and from field investigation data. These input data will be further used to analyze the debris flow-flood hazard by the statistic analysis in the next chapter.

The brief techniques and thematic maps of the input data produced in this thesis, namely, elevation (digital elevation data , Digital Elevation Model-DEM, aspect, slope, landform topography), hydrology (drainage pattern, sub-catchments characteristics), geology (rock unit), soil properties (soil group unit, soil thickness), land cover, infrastructure and human settlement, flow-flood inventory of scar-scouring and

depositional locations, and meteorology of rainfall intensity are consequently presented as below.

Table 3-1 Overview of the important input data themes that were pre-processed and invented in this thesis.

Main themes	Sub-themes	Data preparation methodology
Elevation	Digital elevation data	converted from a 1:20,000 scale digital topographic map of Land Development Department (LDD)
	Digital Elevation Model (DEM)	Derived from digital elevation data with GIS
	Aspect	Derived from DEM with GIS
	Slope	Derived from DEM with GIS
	landform topography (Topographic shape)	Derived from DEM With GIS, image interpretation, field investigation
Hydrology	Drainage system	Digitized from Topographic maps, extracted from DEM with GIS
	Buffering distance to drainage-line	Digitized from topographic maps, extracted from DEM with GIS
Geology	Rock unit	Digitized from a 1:50,000 scale geological map of Department of Mineral Resources (DMR), adjusted with remote sensing imageries and field investigation
Soil properties	Soil group unit	Derived from a 1:20,000 scale soil property map of LDD and field investigation
	Soil thickness	Digitized from a 1:20,000 scale soil property map of LDD and field investigation
Land cover	Land cover	Derived from interpretation of remote sensing imageries and field investigation
Infrastructure and human settlement	Roads and villages	Digitized from geological map, adjusted with remote sensing imageries
Flow-flood inventory	Scar-scouring and depositional locations	Derived from interpretation of multi-temporal remote sensing imageries and field investigation
Meteorology	Rainfall intensity	Interpolated from existing rainfall information of the observation stations of Thai Meteorological Department (TMD)

3.3 Elevation

3.3.1 Data entry

The digital elevation data of the study area was converted from a 1:20,000 scale digital topographic map (10 m contour interval) derived from the Land Development Department (LDD). This data is composed of contour lines and points with elevation information conducted in ARC/INFO format. This map is called the color-coded contour map (Figure 3-2).

3.3.2 Input map generation

Instead of using a discrete elevation map such as contour lines, it is more advantageous to work with a continuous map. Regarding this advantage, the contour data was converted into a color-coded continuous map (Digital Elevation Model-DEM) as shown in Figure 3-2 at 10 m spatial grid resolution. DEM is used to create a slope, aspect and landform topographic shape. In order to increase visual interception of DEM, it had been chosen to convert into a color-coded DEM (Figure 3-3) and a color-draped relief model (Figure 3-4). The produced DEM would be used as the elevation input data for the elevation attributes of flow-flood.

For an accurate modeling of the topography in the study area, a watershed analysis was carried out to extract the micro-catchments, the ridge point and the drainage-lines (Figure 3-5). The process began by evaluating of elevation raster for depressions and constructs watershed polygons based on the depression recognized. This procedure identified watersheds from a raster surface image either automatically for the entire study area or for a user-defined "seed" feature image. With the automatic method, micro-catchment delineation determines the enclosed watersheds of a given raster image according to the given a real threshold.

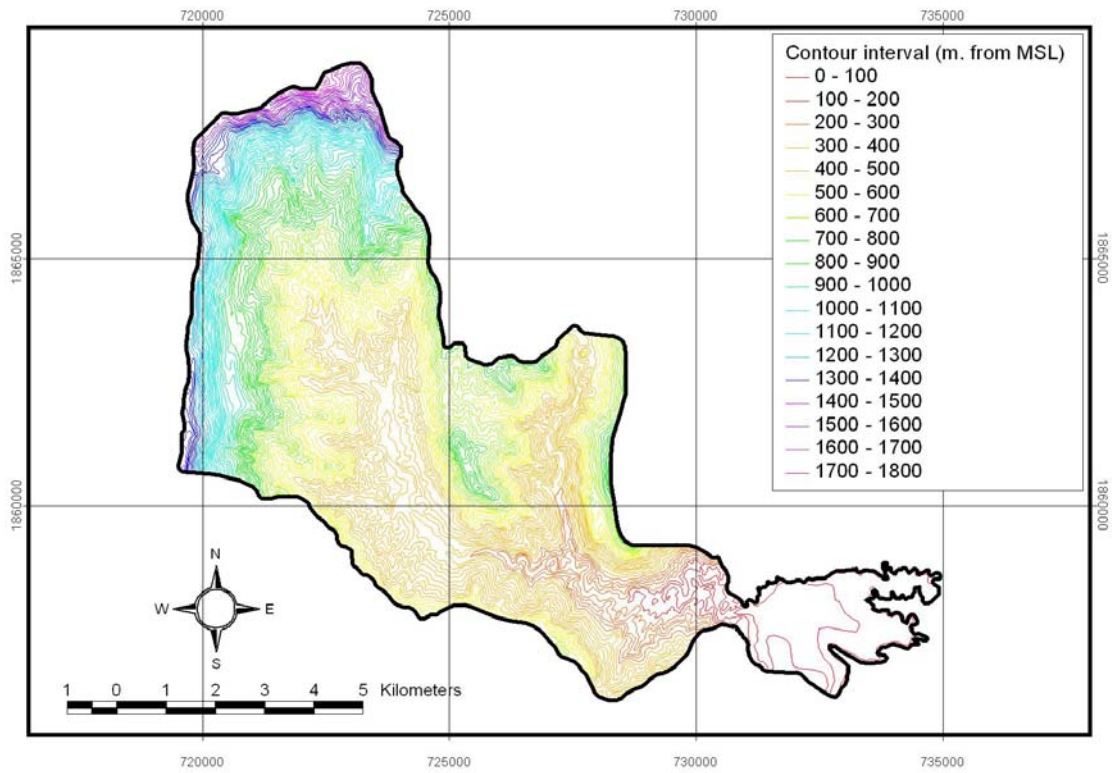


Figure 3-2 Color-coded contour map of the study area.

(Data Source: Land development Department).

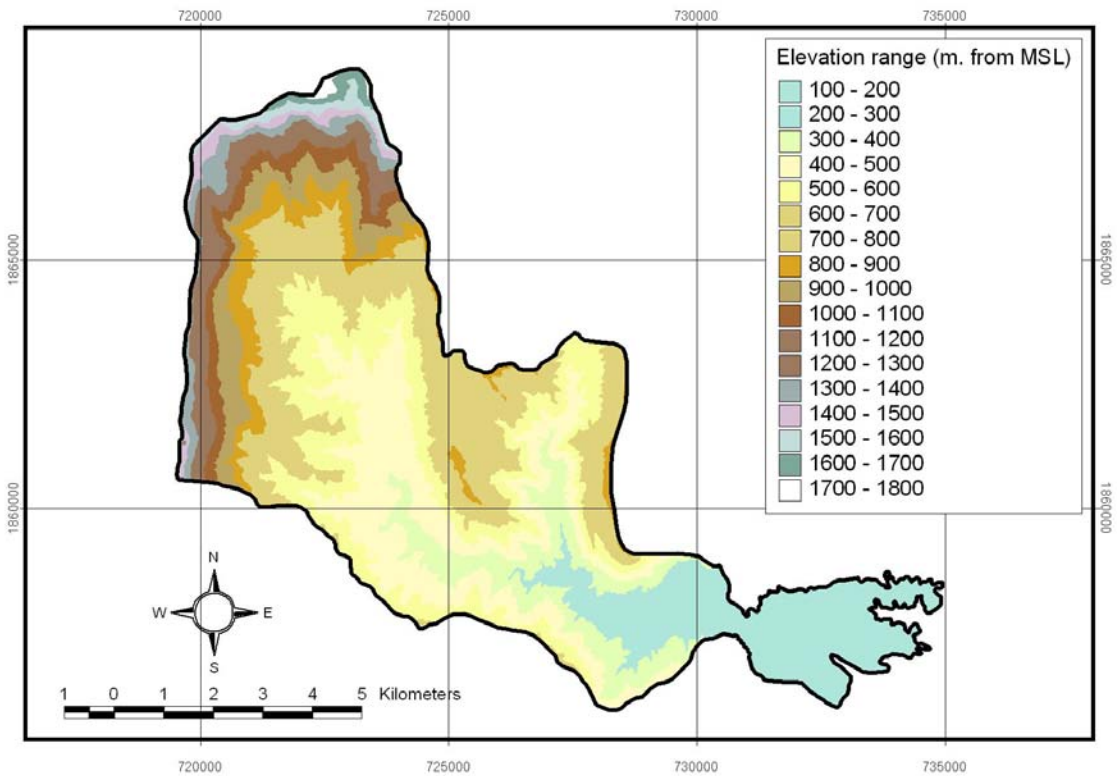


Figure 3-3 Color-coded DEM of the study area.

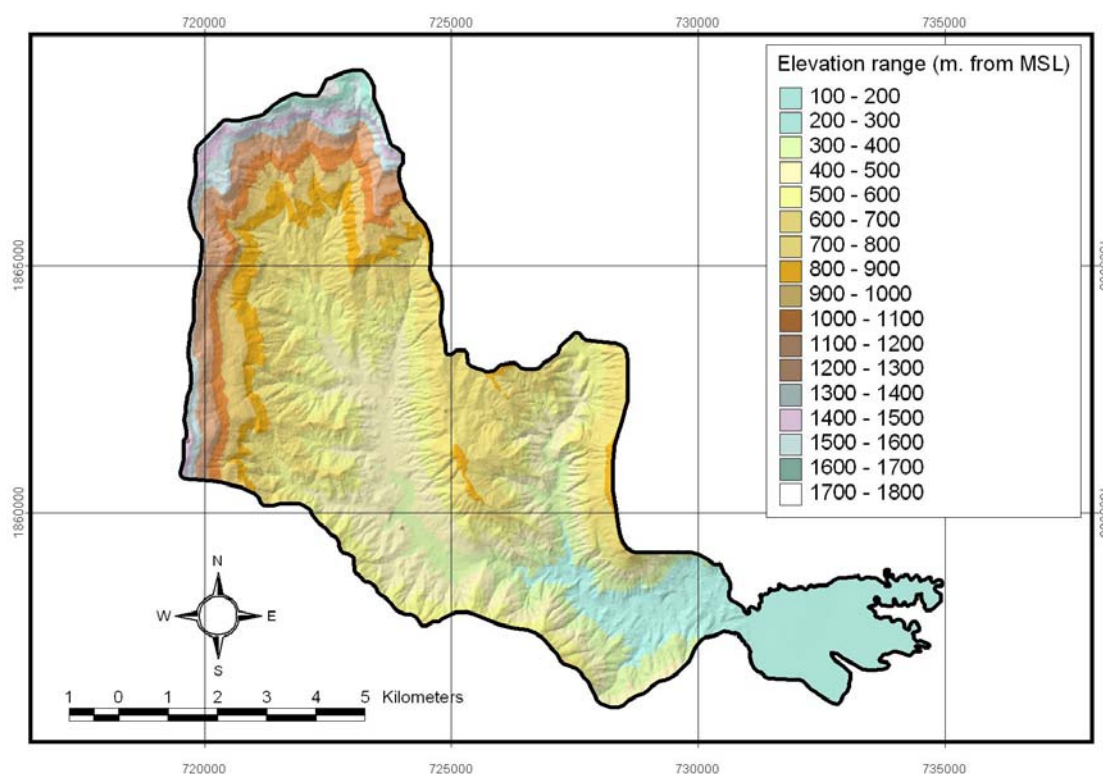


Figure 3-4 Color-draped relief model of the study area (illumination 45° , vertical exaggeration $\times 3$).

The threshold is the minimum number of cells in a watershed. Watersheds are identified if they meet or exceed this threshold number of cells and if their outlets are within the image. In the output image, each watershed was given a unique sequential identifier, starting with 1. Any pixel that was not a part of watershed as defined above was given the value 0.

The distances of every pixel regarding the drainage-lines were calculated and presented in Figure 3-6. The minimum distance of pixels was 1 meter and the maximum, 2,320.39 meters. The buffering distance to drainage-line had average value of 41.55 with standard deviation of 141.76 which means that more than 68.2 percent of Nam Ko Yai sub-catchments has distance from drainage line of basin less than 183.31 meters.

The topographical modeling of the study area also serves as an input to calculate the topographic derivatives such as aspect and slope maps of the sub-

catchment. These topographic derivatives are very crucial and contain vital information for the flow-flood development in the area.

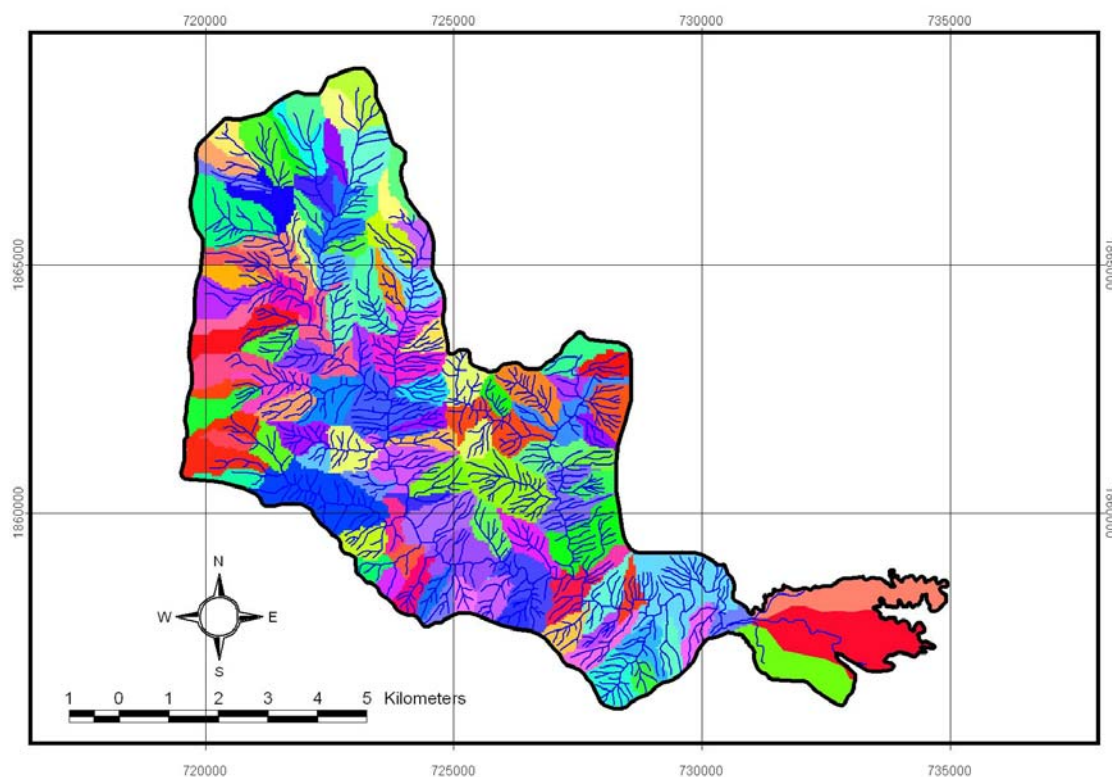


Figure 3-5 Drainage system of the study area, including micro-catchments and drainage-lines (in blue colored lines).

(Note: Color coding of micro-catchments is arbitrary)

The aspect is a measure of slope orientation and is calculated in geographic directions as the azimuthal degree from the North. Due to raster data format and the computational limits the aspect distribution has sensitivities in principle directions and in every 22.5 degrees, which is a drawback of the algorithm used. The produced and color-coded aspect map and its frequency distribution are presented in Figure 3-7. The aspect map reveals a range between -1 and 359, -1 representing the flat lying areas (no direction) and 0 as the North, and other values are the azimuth measurement from North. The minimum value is -1 and the maximum 359 degrees. The distribution has a mean value of 182.45 with standard deviation of 110 which means that most of Nam Ko Yai sub-catchments (more than 68.2 percent) has the faces to East direction (aspect faces to 65.7-112.5 degrees).

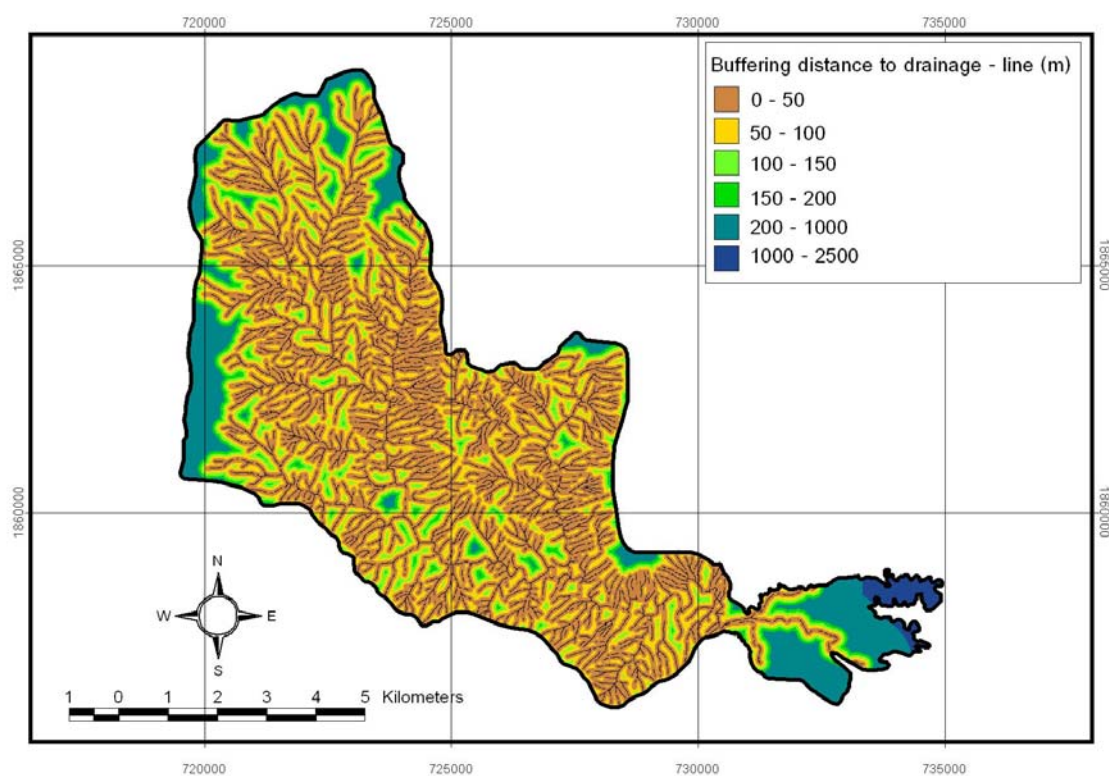


Figure 3-6 Buffering distance to drainage-line in the study area.

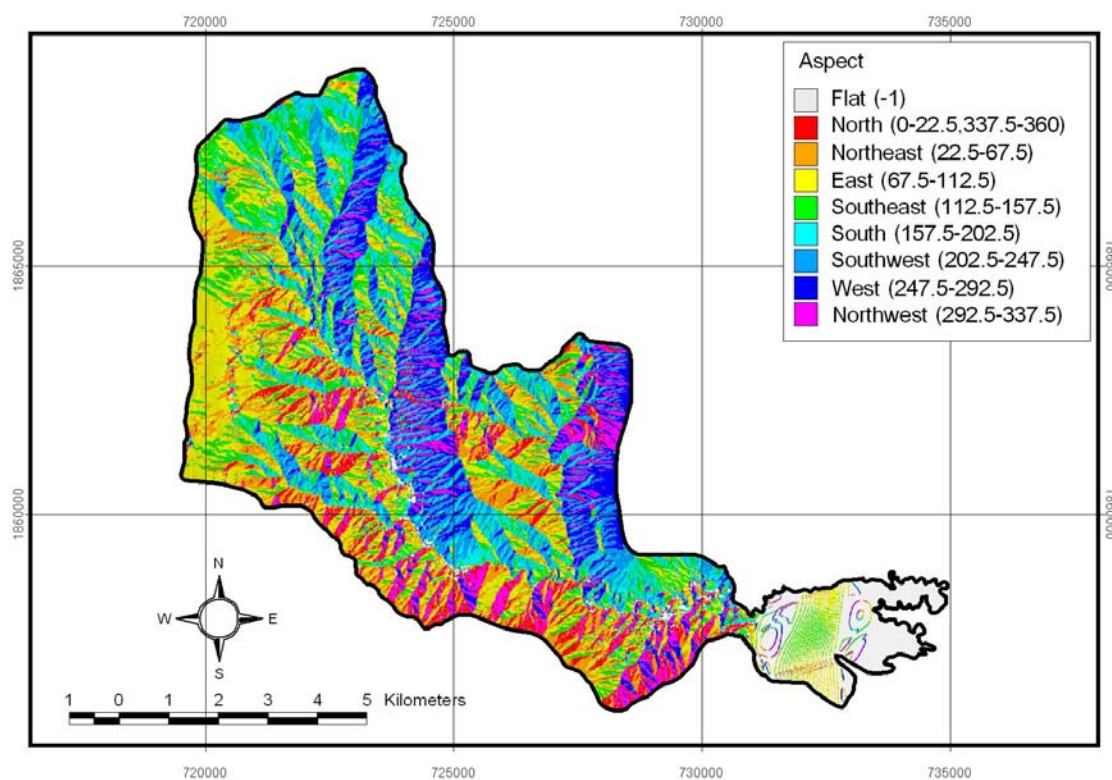


Figure 3-7 Aspect map of the study area.

The slope is a measurement of surface steepness and is calculated in degrees of inclination. The produced and color-coded slope map and its frequency distribution are presented in Figure 3-8. The slope has a range between 0 degree and 90 degrees, 0 degree representing the flat lying areas and 90 degrees as the vertical ones. Any other value indicates the inclined areas.

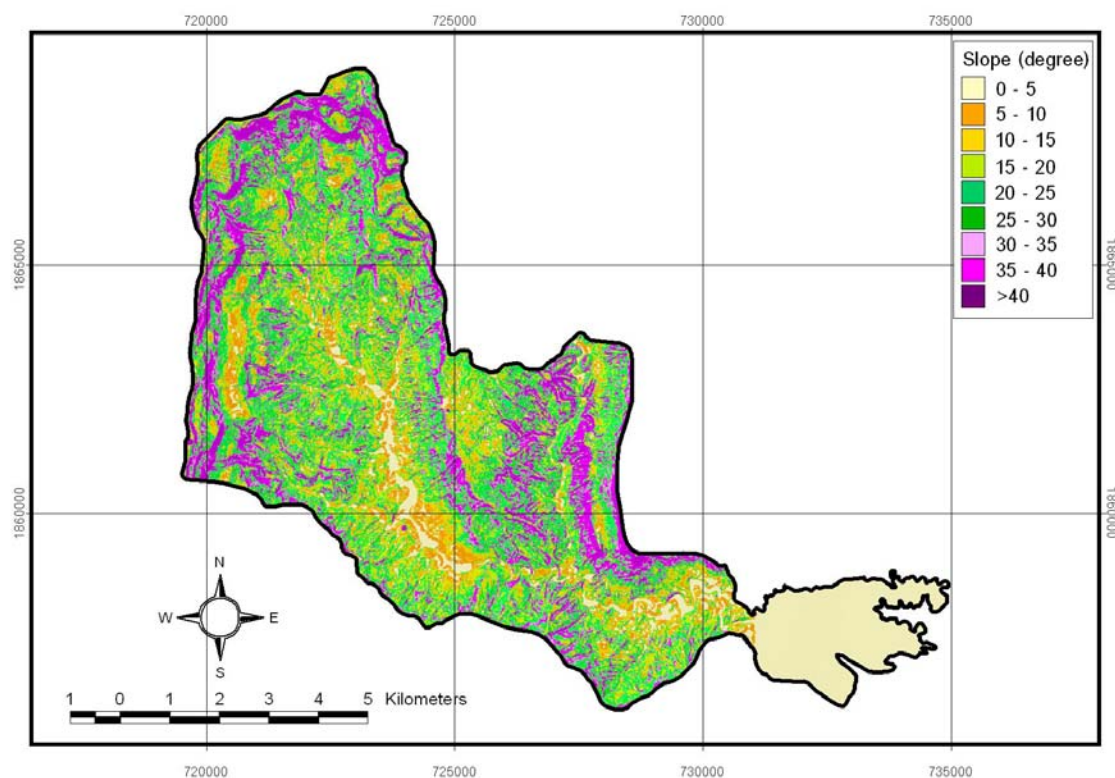


Figure 3-8 Slope map of the study area.

Hence the minimum value is 0 degree and the maximum 85.95 degrees. The distribution has a mean value of 13.28 degrees with standard deviation of 11.92 which means that most of Nam Ko Yai sub-catchments (more than 68.2 percent) has slope below 25 degrees.

From DEM data, landform topography (topographic shape) of sub-catchments was classified into 11 possible topographic features: peak, ridge, saddle, flat, ravine, pit, convex hillside, saddle hillside, slope hillside, concave hillside, and inflection hillside (Figure 3-9) by Idrisi32 software. Landform topography created by TOPOSHAPE module in Idrisi32, perform a Fourier analysis on the DEM for TOPOSHAPE results. The Fourier transform produces results that are more meaningful with better continuity of

topographic features. The procedure is performed on either a regular DEM, or on a DEM that has been transformed, filtered, and back-transformed with FOURIER and its companion modules.

The "TOPOSHAPE" module classifies each cell of an input digital elevation grid according to its topographic appearance: ridge, valley, convex slope, and eight other categories. The output is another grid of cells, each of which contains a numerical code for the topographic category.

Surface shape classification is based on polynomial surface fitting of each 3 x 3 pixel area. Eigenvalues are solved from the second directional derivative of the partial quadratic equation for a central pixel of a 3 x 3 neighborhood. The eigenvalues hold the information for the magnitude of rate of change of a tangent line along the mathematically described curve in the aspect direction of the pixel and in the direction orthogonal to aspect.

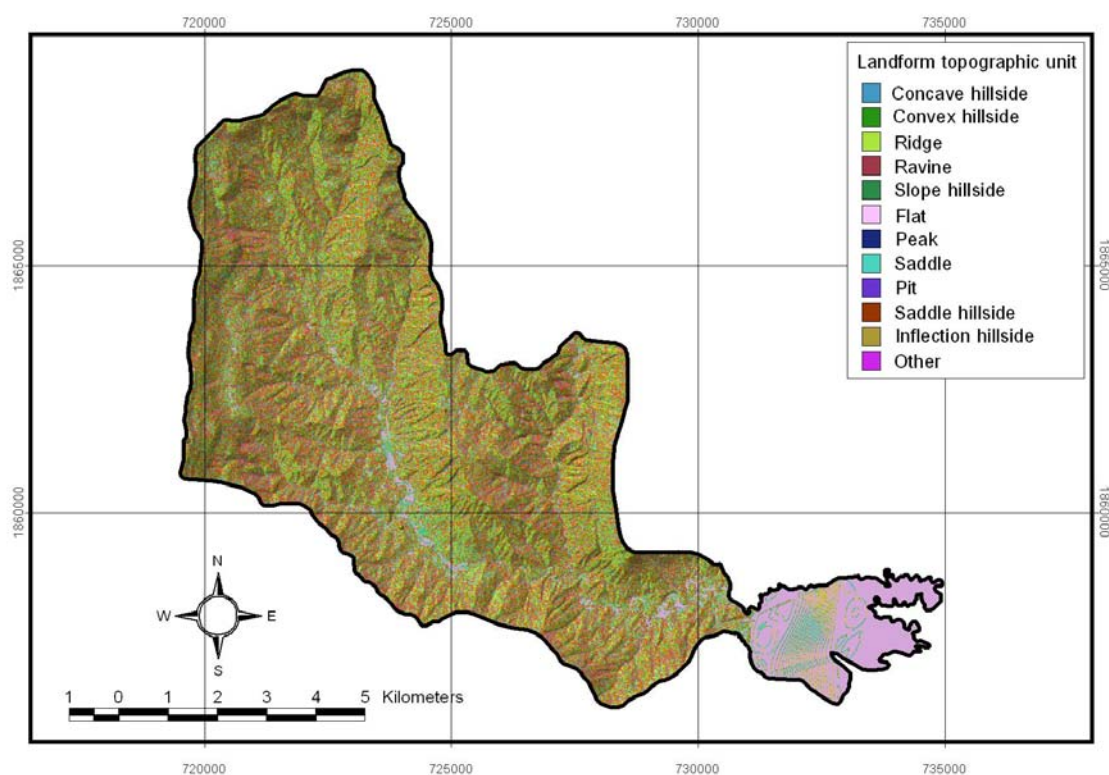


Figure 3-9 Landform topography of the study area.

3.4 Geology

3.4.1 Data entry

The previous geological map of the study area preliminary compiled from the available reports, publications, and analogue maps was prepared. The compiled analogue maps were transformed into digital image, via digitizing and edit using Arc/Info GIS software (as shown in Figure 3-10).

3.4.2 Input map generation

In the compiled geological map in this thesis, the elevation data from the previous section (e.g. slope, aspect, topographic shape, etc.), the combination of Landsat ETM+ imageries and normalized different vegetation index or NDVI (to be mentioned later) were used as additional input information for geological revision. Besides, field data from ground-truth survey was also combined to revise the previous geologic map (Figure 3-10). Finally, the compiled geologic map of the study area is proposed in Figure 3-11.

In general, various rock units ranging from the uppermost Paleozoic and Mesozoic sedimentary and volcanic rocks to the younger unconsolidated sediments were noted in the study area from the previous study by Yooyen (1985). Stratigraphically, the lowest rock unit, generally exposed in the eastern part of the study area, is Permian Lom Kao (Lk) Formation that consists of folded limestone, massive shale and slaty shale.

Unconformably above that, Triassic Lom Sak (Ls) Formation that comprises volcanic complex, siltstone, shale and slate covers the most part of the study area, especially adjacent to the central stream channel. This Ls Formation is subsequently angular- unconformably overlaid by the gently westerly-dipping Khorat Group that mainly exposed on the steepest and highest western and northern rims, to the tops of a flat highland away from the study area. This Khorat Group consists of Phu Kradung (Pk) Formation (red siltstone, conglomeratic sandstone, tuffaceous sandstone and siltstone)

and Phra Wihan (Pw) Formation (gray sandstone, tuffaceous siltstone, and red shale), both Jurassic in age, and Phu Phan (Pp) Formation (pebbly sandstone) of Cretaceous period.

The younger unconsolidated sediments of fluvial deposits (Qa1) of Quaternary age are those of mainly stream deposits, composing of river sands and gravels, silts, clays and gray soils along the drainage system here. The sediments of Quaternary age also form in the alluvial fan as alluvial fan deposits (Qa2) at the canyon mouth to the southeastern limit of the study area. It is noted that the younger unconsolidated sediments of fluvial deposits (Qa1) and alluvial fan deposits (Qa2) of Quaternary age are the new rock units that are compiled and proposed in this thesis.

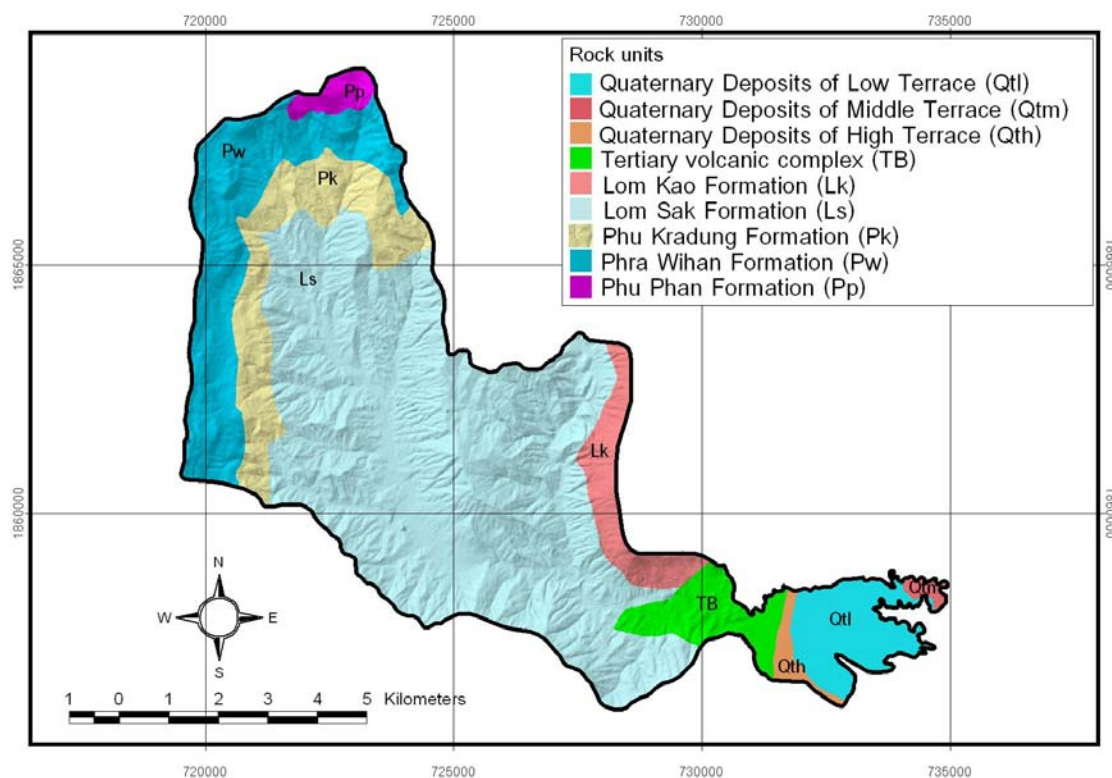


Figure 3-10 Previous geologic map of the study area (modified after Yooyen, 1985, etc.)

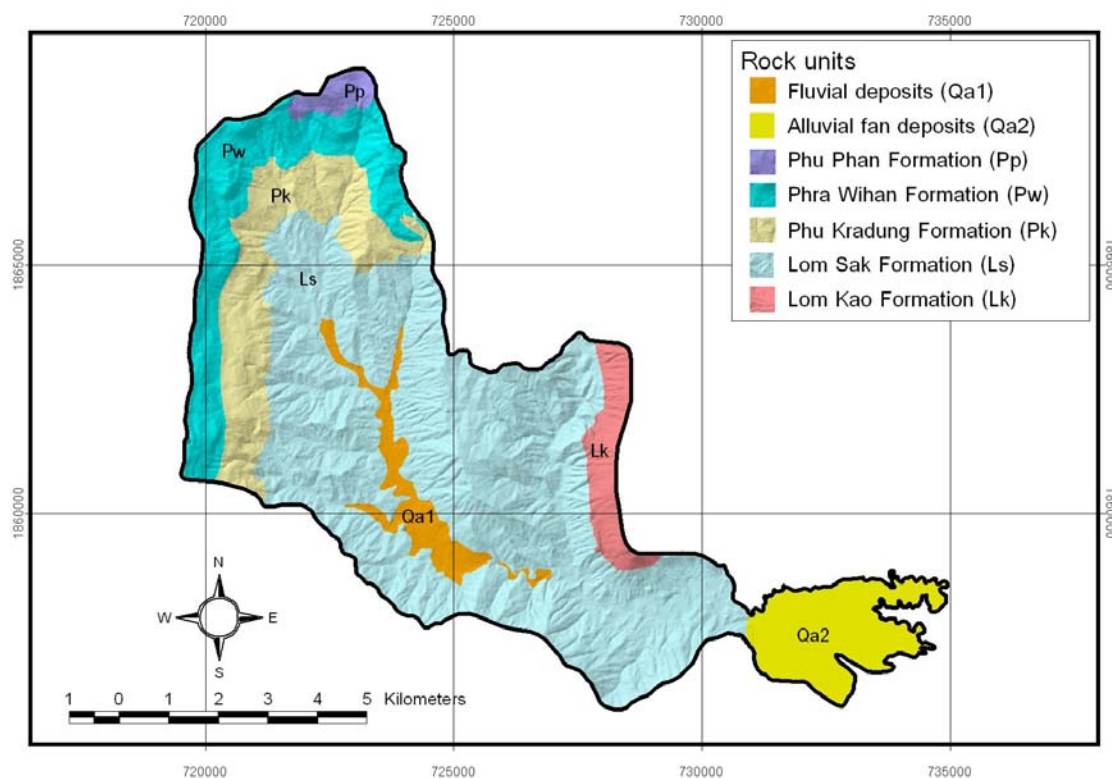


Figure 3-11 Compiled geologic map of the study area.

3.5 Soil property

3.5.1 Data entry

The soil property collected in a form of soil group unit map and soil thickness map of the study area was prepared by compiling data from the available reports, publications, and analogue map of Land Development Department (2002). The compiled analogue maps were transformed into digital image, via digitizing and edit using Arc/Info GIS software. (as shown in Figure 3-12)

3.5.2 Input map generation

The soil group unit map of the study area is further explained below.

In the study area, soils on dissected erosional surfaces are wide different according to the parent materials. Soils that derived from limestones and from basic to intermediate igneous rocks are grouped in Soil Group Unit 47 (Li, Muak Lek, Nakon

Sawan, Tali, Sob Prab and Phai Sali units). This group of soil covers 84.47 percent of total area, located on 2-20 percent (inclination) slope. Soil thickness is shallow (50-80 cm.). Soil fertility is very low with highly erosion rate. Soil pH of this group unit varied from 5.0-7.5.

As referred to Land Development Department (2002), soils that derived from sedimentary or metamorphic rocks were classified as red-yellow podzolic soils, reddish brown lateritic soils, and non-calcic brown soils and were group in Soil Group Unit 48: Ta Yang, Mae Rin, Na Cha Liaeng, Pa Yao and Nam Kun). This group of soils covers 0.189 percent of total area, located on 3-25 percent (inclination) slope. Soil thickness is very shallow (below 50 cm.), soil fertility is very low and soil fertility is very low. Soils on the high terraces were grouped in red-yellow podzolic soils, reddish brown lateritic soils, and grumusols (Soil Group Unit 55: Wang Sapung, Tab Kwang and Cha Turas). Most of these soils are well to moderately well suited for upland crops. This group of soils covers 2.30 percent of total area. Soils on the low terraces are considered to be fairly- to poorly suitable for rice growing as they are somewhat poorly drained, with coarser texture and relatively low fertility. These soils are grouped in low humic gley soils, hydromorphic and non-calcic brown soils (Soil Group Unit 28: Chai Badan, Lopburi, Wang Chompu and Samor Thod). This group of soils covers 2.76 percent of total area.

Soils on semi-recent terrace are mainly the lowland soils. The lowland soils are grouped in low humic gley soils, hydromorphic non-calcic brown soils, and grumusols (Soil Group Unit 4: Chai Nat, Ratchaburi, Tha Phon and Saraburi) and are well suitable for transplanted rice growing. The upland soils in this area are usually well suitable for upland crops. This group of soils covered 2.01 percent of total area. Soils on levees (Soil Group Unit 29: Ban Chong, Nong Mod, Mae Taeng, Pak Chong, Hang Chat, Khao Yai and Chok Chai) are well suitable for upland crops. Soils on flood plain are grouped as alluvial soils (Soil Group Unit 18: Khao Yoi, Chon Buri and Khok Samrong). They are somewhat poorly drained, with fine texture and relatively high nutrient status so they are well suitable for rice growing.

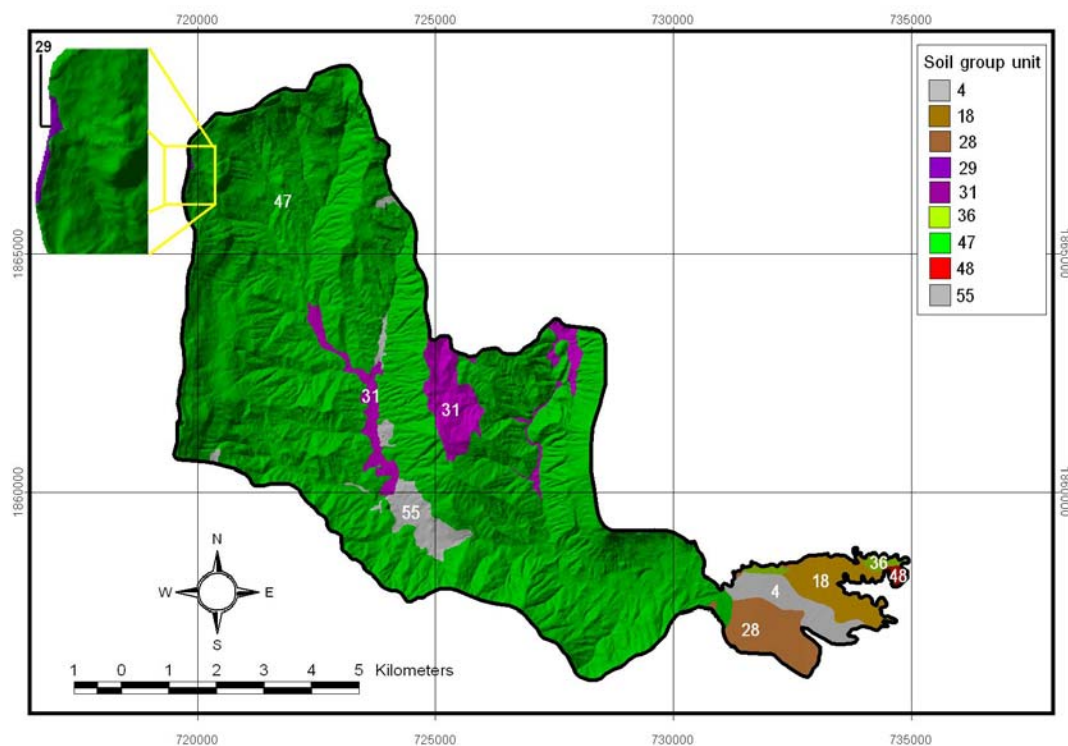


Figure 3-12 Soil group unit map of the study area (modified after Land Development Department, 2002).

Besides, soil thickness map of the study area compiled from the available reports, publications, and analogue map of Land Development Department (2002) was also prepared. The compiled analogue maps were transformed into digital image, via digitizing and edit using Arc/Info GIS software (as shown in Figure 3-13).

3.6 Land cover

About the land cover, Nam Ko Yai sub-catchment is covered by dense forests only on the western and northern high steep-slope area. Along the undulating valley floor of the stream in the central part of the sub-catchment, the deforestation is so vast to make land for agricultural usage. Erosional phenomena of many different types, ranging from sheet and rill erosion to mass movement, forming gullies and badlands, are widespread across the whole sub-catchment area. In the eastern extreme of the sub-catchment and on the alluvial fan, irrigated orchards and densely populated settlements are noted. It was noticed that the inappropriate land use in the highly deforested areas

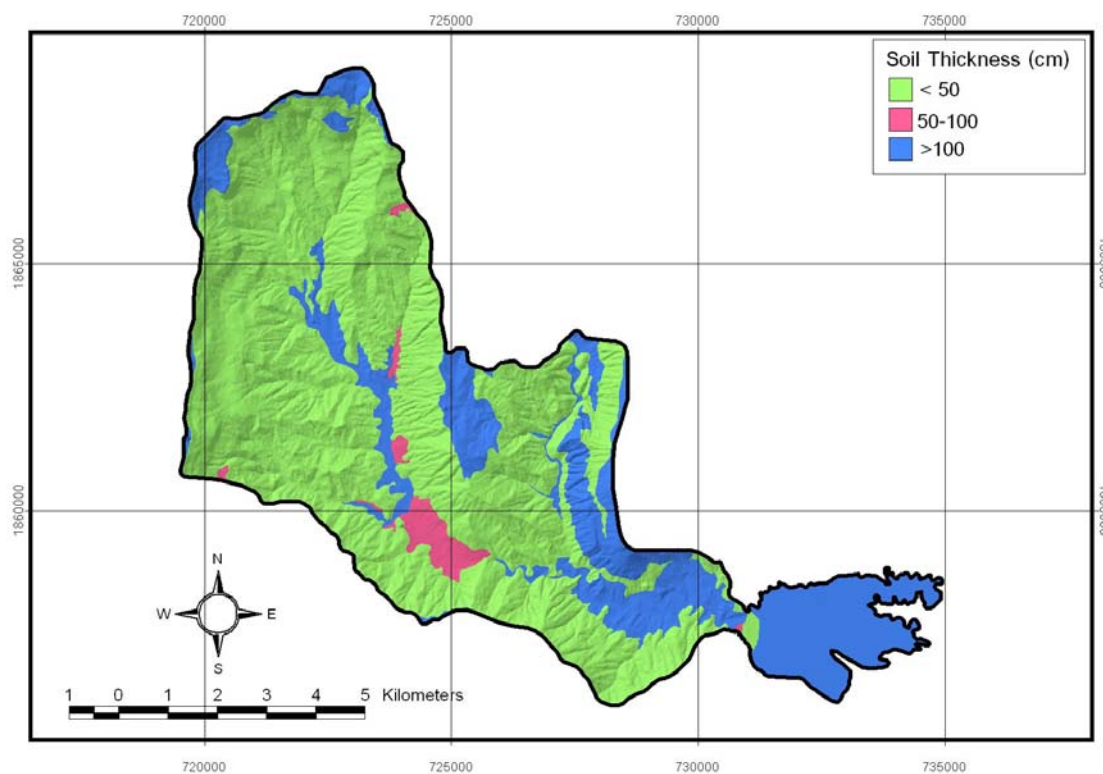


Figure 3-13 Soil thickness map of the study area (modified after Land Development Department, 2002).

of the sub-catchment was first blamed to be the major cause of this 8/11 tragedy by the publics and academics.

As a landslide hazard assessment should not only depend on the production of landslide inventory map and its analysis, a complete hazard assessment system will require also the assessment of external factors leading to instability rather than the topography and material derived properties. The land cover distribution is one of the external factors that can easily be mapped and monitored in time if needed with the aid of satellite images (Soeters and Van Westen, 1996). In landslide hazard assessment projects and in environmental and engineering studies, accurate and up to date information about land cover on a regional scale often resembles vital information to help for decision rule generation.

Basically a land cover map is obtained by classifying remotely sensed images. Typically this is performed by the spectral analysis of individual pixels and their

association with other neighboring pixels. The results of classification depend largely on the type of area, land-cover type, and image acquisition date. However, the results of the classification are directly affected by spectral confusion of land-cover types and mixed pixels. The vegetation and soil moisture conditions produce distinctive spectral responses in the electromagnetic spectrum that give the opportunity to the classifier to classify them easily.

3.6.1 Data sources

Classification of land cover in the study area was carried out using many sources of remotely sensed data. Sources of data using for this study are summarized in Table 3-2. Landsat 7 ETM+ acquired on 21st November 2001, the date close to the 8/11 event, was selected as primary remotely sensed data used for classifying the land cover in the study area in the following data processing steps.

3.6.2 Data Processing

Satellite imagery was analyzed using the program PCI Easi+ version 9.1, Erdas Imagine version 8.7 to obtain the results for land-use classification and grid interpolation and results plot. This study use ArcView GIS version 3.2a for analyzing previous secondary data and classifying results. Digital data analysis techniques employed in this study involved the following steps.

3.6.2.1 Image rectification and restoration

Since a large amount of data sets were involved, preprocessing of the data in the most efficient way is the prime concern. It is necessary to geometrically correct LANDSAT images to remove image distortion caused by variations of orbital parameters of the satellite and by imperfections of the sensor. The precise geometric correction could be applied by using the known ground control points (GCPs) from maps with a standard cartographic projection such as the Universal Transverse Mercator Projection

Table 3-2 Multi-temporal aerial photographs and satellite images that are used as primary data sources of this thesis

Image type	Acquisition date	Original		
		Format	Scale and Resolution	Source
Landsat TM5	December 12, 1995	LGSOWG	30 m.	GISTDA
Landsat 7 ETM+	March 7, 2000	LGSOWG	30 m.	GISTDA
	January 1, 2001	LGSOWG	30 m.	GISTDA
	November 21, 2001	LGSOWG	30 m.	GISTDA
IKONOS	October 31, 2002	GeoTiff	1 m.	GISTDA
Aerial Photograph	November 30, 1974	B&W	1:15,000	RTSD
	December 10, 1974	B&W	1:15,000	RTSD
	December 11, 1974	B&W	1:15,000	RTSD
	December 21, 1974	B&W	1:15,000	RTSD
	December 24, 1974	B&W	1:15,000	RTSD
	January 27, 1975	B&W	1:15,000	RTSD
	January 29, 1975	B&W	1:15,000	RTSD
	January 6, 1996	B&W	1:15,000	LDD
Orthophotographs	January 9, 2002	Color	1:25,000	MOAC
	December 18, 2002	Color	1:25,000	MOAC
	January 14, 2003	Color	1:25,000	MOAC
	January 15, 2003	Color	1:25,000	MOAC

Remarks: GISTDA Geo-Informatics and Space Technology Development Agency (Public Organization)

RTSD Royal Thai Survey Department

LDD Land Development Department

TMD Thai Meteorology Department

MOAC Ministry of Agriculture and Cooperation

ESRI Environment System Research Institute (Thailand) Co. Ltd.

(UTM). After taking known GCPs from the topographic map (scale 1:50,000), the LANDSAT satellite imageries were geo-rectified by using the first order nearest neighbor methods. The pixel size of the rectified image is 25 m x 25 m.

Transformation matrices containing the coefficients for converting coordinates were calculated from the GCPs by the least square regression methods. The best GCPs were selected and adjust until the total RMS (Root Means Square) error was less than the tolerance level (0.5 of pixel size). The first-order transformation was applied which yields adequate result because of flat terrain of the study area. Nearest neighbor interpolation was followed during re-sampling.

The formulae (Equations 3-1 and 3-2) were used to rectify the satellite image so that its geographical reference was the same as the other layers in the GIS.

$$X = X_0 + a_1 (x - x_0) + a_2 (y - y_0) \dots\dots\dots \text{(Equation 3-1)}$$

$$Y = Y_0 + b_1 (x - x_0) + b_2 (y - y_0) \dots\dots\dots \text{(Equation 3-2)}$$

Where

X, Y = satellite image coordinates

x, y = map coordinates

X_0, Y_0 = coordinates of the center (arithmetic mean) of the GCPs from satellite image

x_0, y_0 = coordinates of the center (arithmetic mean) of the GCPs from map

a_t, b_t = model constants

The values of x_0, y_0, X_0, Y_0 are dependent of the subset image and the positions of the ground control points. The model constants (a_t and b_t) are dependent of the relative orientation between the two coordinate systems.

3.6.2.2 Reduction of noise and image enhancement

Noise is a digital image that can manifest itself as either inaccurate gray level readings or missing data altogether. Noise is the result of sensor malfunctions during the recording or transmittal of information. Unlike geometric distortions and other radiometric

degradation, noise is readily identifiable, even to those unfamiliar with the scene of the image. Noise can be removed by using Fourier method; the image is then enhanced to encompass a variety of operation design to improve the visual interpretability of an image, by increasing the apparent distinction between features in a scene.

In the present study, all spectral bands excluding Thermal Infrared Channel of TM (band 6) and ETM+ (band 6L and 6H), were used. To minimize the effect of illumination difference, spectral band were normalized by an equation of total intensity below.

$$NB_i = 255 (OB_i / \sum OB_i) \quad i = 1 \text{ to } n \quad \dots\dots\dots \text{(Equation 3-3)}$$

Where

NB_i = band normalized by total intensity

OB_i = original spectral band

The constant 255 is to fit the data in a byte range of 0-255. The resulting bands have the property that the sum of some pixel values is 255 due to normalization. After normalization, the locations of these objects are projected onto a diagonal line of uniform intensity, indicating that the objects are free from intensity variation.

3.6.2.3 Image classification

The intention of classification process is to categorize all pixels in a digital image into one of several land cover classes or themes. This categorized data set will then be used to produce thematic maps of land cover presented in an image. Thematic maps provide an easily interpretable summary with which the eventual end user can make the well-informed decisions.

In conventional classification of multi spectral data, the maximum likelihood classifier is considered to provide the best results since it takes into account the shape, size and orientation of a cluster. Based on the class mean and the variance-covariance matrix, an unknown pixel is assigned to the most likely class.

Maximum likelihood classification (MLC) technique was employed to perform the classification of an unknown pixel. This technique had been found to be the most accurate procedure in quantitatively evaluate both the variance and correlation of the category spectral reflectance patterns. In this study land cover was classified into seven categories based on vegetation characteristics and field investigations.

This technique calculates the distance from each feature vector (pixel to be classified) to class means. The within-class variability is taken care of by adding a factor, which was a function of the variance-covariance matrix of that class. The formula used (Mather, 1987) is below:

$$D_i(x) = \ln |V_i| + (X-M_i)^T V_i^{-1} (X-M_i) \dots\dots\dots \text{(Equation 3-4)}$$

In which; $D_i(x)$ = distance between pixel vector X and a class means based on probabilities; X = mean pixel vector X ; M_i = mean vector of the class considered; V_i = variance-covariance matrix of the class considered; V_i^{-1} = inverse of V_i ; $|V_i|$ = determinant of V_i ; $(X- M_i)$ = distance towards a class means' and $(X-M_i)$ = transportation of $(X-M_i)$

3.6.2.4 Post-processing

During the classification process, the "island themes" may appear. These are single pixel themes that are most likely classification errors. These island themes can be assigned the same gray level as their phenomenon-surrounding theme using a mode filter. The mode filter is a filter algorithm that replaces a pixel gray value with the mode of the gray levels within the filter windows surrounding that pixel. This research applied mode filter with 5x5 windows size to reduce the island themes.

3.6.2.5 Accuracy assessment

A complete accuracy test of a classification map would be a verification of the class of every pixel. Obviously this is impossible and indeed defeats the purpose of the image classification. Therefore, representative test areas must be used instead to estimate the map accuracy with as little error as possible. Classified image accuracy consists of two accuracy

types. Firstly, overall accuracy which represents the accuracy of the entire product and secondly, user's accuracy (or map accuracy) which a map user is interested in the reliability of the map in how well the map represents what be really on the ground. Figure 3-14 illustrated the survey tracks for field data collection and land cover classification accuracy assessment.

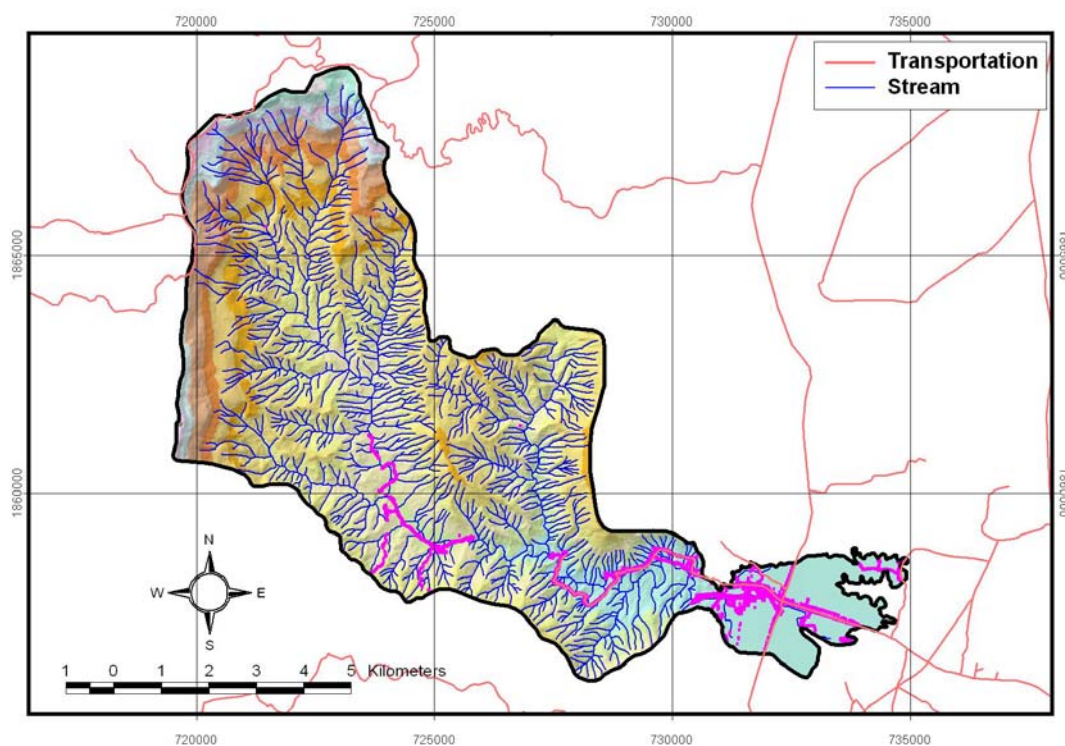


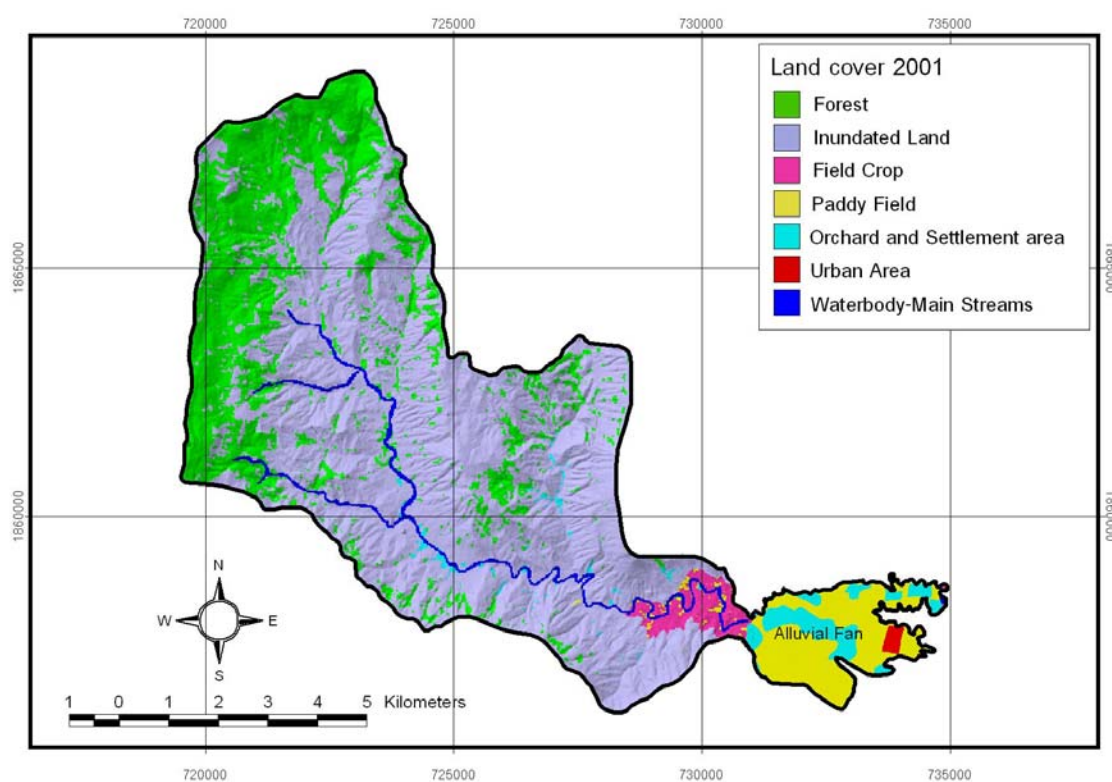
Figure 3-14 Survey tracks (pink colored dots) for field data collection and land cover classification accuracy assessment in the study area.

3.6.2.6 Classification result

The land cover classification of the study area was presented in Table 3-3 and Figure 3-15. Land cover pattern can be classified into 7 categories, namely, forest, paddy field, field crop, inundated land, urban area, water body, and orchard and settlement area.

Table 3-3 Land cover classification of the study area (on 21st November, 2001).

Land cover type	Area (pixel.)	Area (sq.km.)
Forest	183,087	18.3087
Paddy Field	46,269	4.6269
Field Crop	15,187	1.5187
Inundated Land	471,853	47.1853
Urban Area	1,832	0.1832
Waterbody (Main Streams)	15,442	1.5442
Orchard and Settlement Area	19,753	1.9753
Total	753,423	75.3423

Figure 3-15 Land cover map of the study area (classified from Landsat 7 ETM+ acquired on 21st November 2001).

3.7 Infrastructure and human settlement

The presence of several infrastructure elements such as houses, power lines and road network might contribute to the evolution of landslide in the area. The infrastructure has a mutual relationship with the flow-flood hazard as it either causes the slide or is affected from the hazard.

For the construction of infrastructure and human settlement map, the necessary features were digitized from 1:50,000 scale topographic maps of Royal Thai Survey Department (RTSD). Infrastructure (roads) and human settlement (villages) in the study area (Figure 3-16) were also mapped in addition from multi-temporal aerial photographs and satellite images as mentioned in Table 3-2.

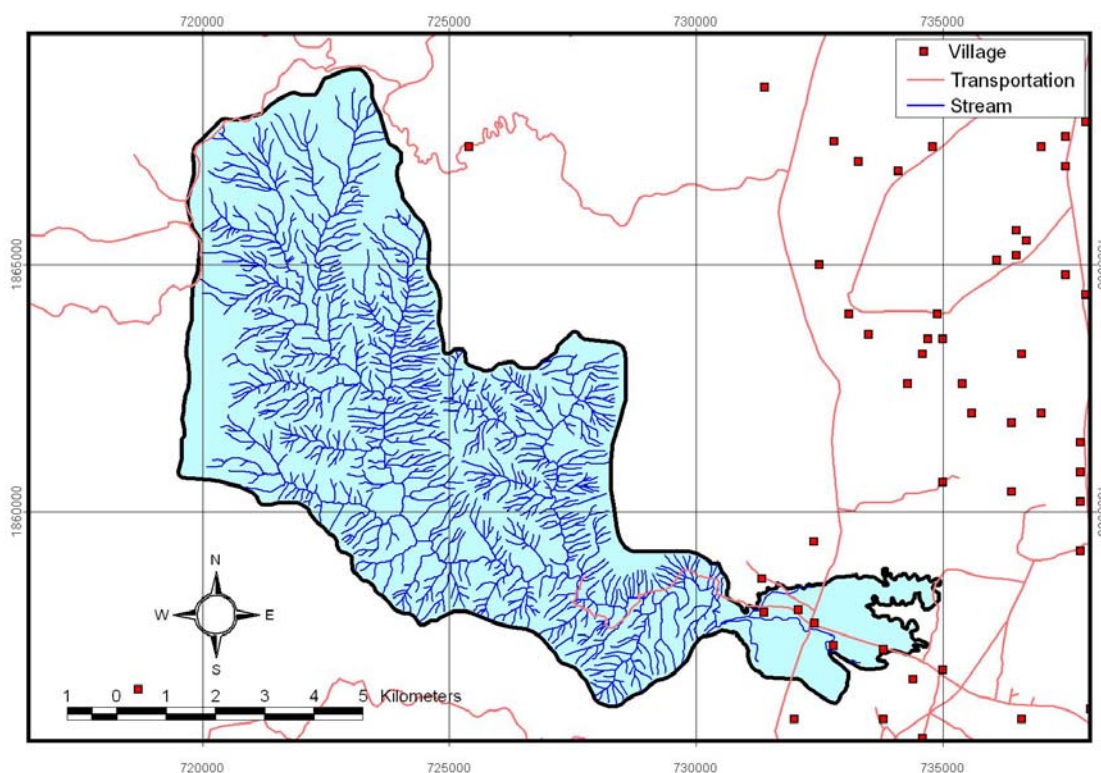


Figure 3-16 Infrastructure and human settlement map of the study area.

3.8 Flow-flood inventory: scar-scouring and depositional locations

3.8.1 Data entry

For landslide susceptibility analysis, accurate detection of scar-scouring and depositional locations is very important, by which a field survey being the most exact method. However, finding all the landslides in a large area is very difficult and is expensive in time and money. This is especially true in mountainous areas where there is no access, and visiting to the flow-flood location is difficult or even impossible. As it would be very difficult to obtain sufficient data in this way for an accurate statistical analysis, a field survey can be used to verify the results of an orthophotograph interpretation and satellite-image analysis. Fortunately, many high-resolution aerial photos and satellite images that can be used to detect the flow-flood locations are available. In this study, 1 meter-resolution IKONOS image and 0.5 meter-resolution orthophotographs were used to detect the flow-flood locations for ground truth flow-flood phenomena field check for adjusting the classification of Landsat 7 ETM+ land cover classification.

The following process describes the combination of Landsat 7 ETM+ imageries and elevation data derived from a Digital Elevation Model-DEM (e.g. slope, aspect, topographic shape, etc.) to detect and classify fresh distinctive scar-scouring and depositional locations in the sub-catchment and its alluvial fan within the study area.

3.8.2 Data Processing

The combination of Landsat 7 ETM+ imageries data and elevation data shows the characteristics of fresh scar-scouring and depositional locations clearly observable in the Landsat imageries, aerial photographs and orthophotographs. Brief field traverses were then performed at some localities as a field check to gather the ground-truth information needed to adjust the accuracy of Landsat imagery classification as well as aerial photograph and orthophotograph interpretation.

Two sets of multi-spectral Landsat imageries of different periods, one on 5 January 2001 (before 8/11) and the other on 21 November 2001 (after 8/11), were classified (as shown on Figure 3-20 and 3-21). Preprocessing of the six spectral bands (not included thermal infrared band) of these Landsat imageries involve an atmospheric correction based on the standard atmospheric-model approach. Orthorectification is accomplished using GIS vectors of road- and stream data, as well as a DEM interpolated from contour vectors (1:20,000 scales). In this data process, normalized difference vegetation index (NDVI) was used to establish a threshold of vegetated and un-vegetated pixels in the Landsat imageries for the change detection of the scar-scouring and depositional locations.

A vegetative index is a value that is calculated (or derived) from sets of remotely-sensed data that is used to quantify the vegetative cover on the Earth's surface. Though many vegetative indices exist, the most widely used index is the Normalized Difference Vegetative Index (NDVI). NDVI, like most other vegetative indices, is calculated as a ratio between measured reflectivity in the red- and near infrared portions of the electromagnetic spectrum. These two spectral bands are chosen because they are most affected by the absorption of chlorophyll in leafy green vegetation and by the density of green vegetation on the surface. Also, in the red- and near-infrared bands, the contrast between vegetation and soil is at a maximum.

NDVI is a type of product known as a transformation, which is created by transforming raw image data into an entirely new image using mathematical formulas (or algorithms) to calculate the color value of each pixel. This type of product is especially useful in multi-spectral remote sensing since transformation can be created to highlight the relationships and differences in spectral intensity across multiple bands of the electromagnetic spectrum. NDVI transformation is computed as the ratio of the measured intensities in the red (R) and near infrared (NIR) spectral bands using the following formula.

$$\text{NDVI} = (\text{NIR} - \text{red}) / (\text{NIR} + \text{red}) \quad \dots\dots\dots \text{(Equation 3-5)}$$

The resulting index value is sensitive to the presence of vegetation on the land surface and can be used to address issues of vegetation type, amount, and condition. Many satellites have sensors that measure the red- and near-infrared spectral bands, and many variations on NDVI exist. The Thematic Mapper (TM and Enhanced Thematic Mapper Plus (ETM+) in bands 3 and 4) provide RED and NIR measurements and therefore can be used to generate NDVI data sets with the following formula.

$$\text{(TM or ETM+) NDVI} = (\text{Band 4} - \text{Band 3}) / (\text{Band 4} + \text{Band 3}) \quad \text{..... (Equation 3-6)}$$

The Red and NIR images are used to calculate an NDVI value for each pixel. The NDVI equation produces values in the range of -1.0 to 1.0, where vegetated areas typically have values greater than zero, while the negative values indicate the non-vegetated surface features such as water, barren, ice, snow, or clouds. In order to maximize the range of values and provide numbers that are appropriate to display in an 8 bit image, NDVI value must be scaled. This scaling converts a number between -1.0 and 1.0 into a pixel value that is appropriate on a dark-and-white display. One example of scaling an NDVI value for display is the following equation.

$$\text{Scaled NDVI} = 100(\text{NDVI} + 1) \quad \text{..... (Equation 3-7)}$$

Thus, using this equation, a pixel with an NDVI value of 0.43 would be scaled into a gray scale value of 143. Using this technique, the NDVI computed value is scaled to the range of 0 to 200, where computed -1.0 equals 0, computed 0 equals 100, and computed 1.0 equals 200. As a result, NDVI values less than 150 represent clouds, snow, water, and other non-vegetative surfaces, and NDVI values which equal or greater than 150 represent vegetative surfaces. The resulting scaled values can be displayed on a gray tone display or even converted to a color image.

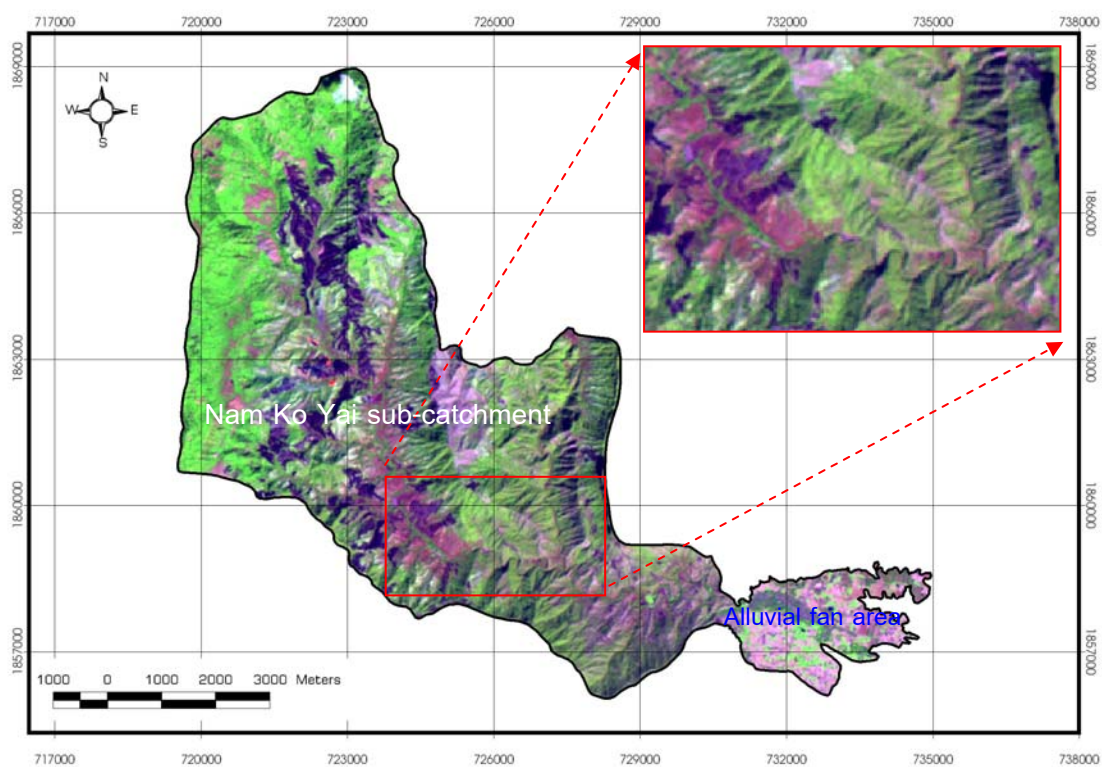


Figure 3-17 False color composite of Landsat 7 ETM+ (R=5, G=4, B=3) acquired on 5th January 2001 (before 8/11) in the study area.

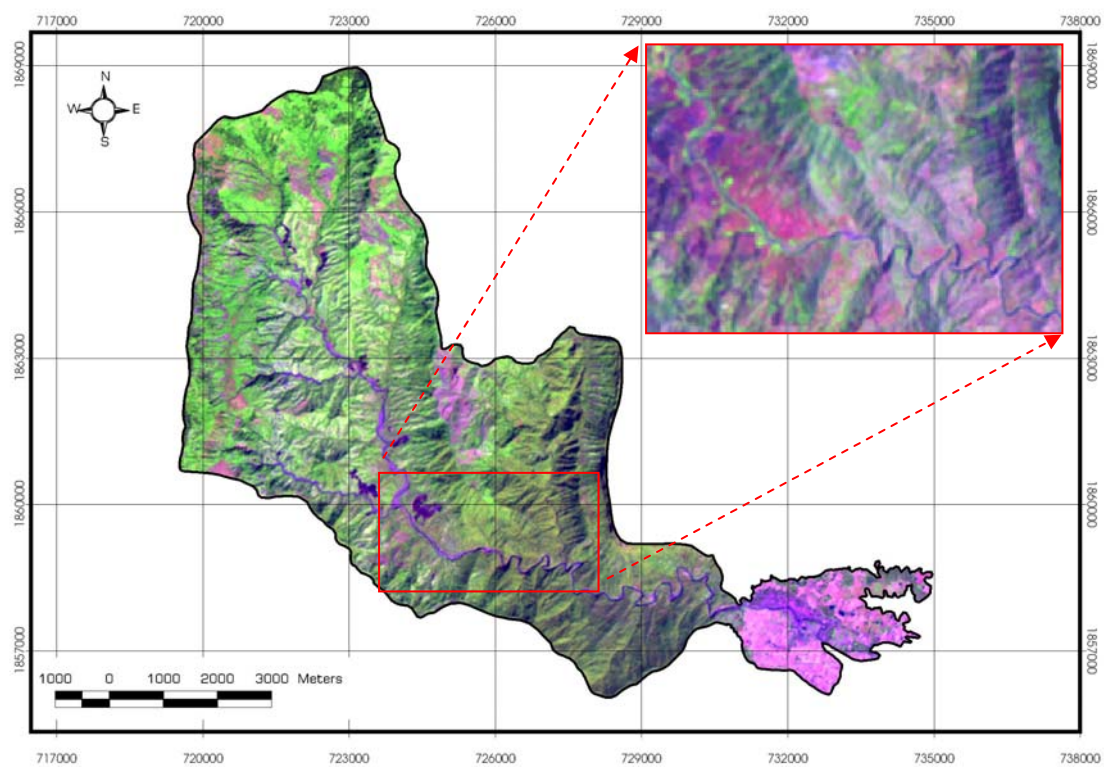


Figure 3-18 False color composite of Landsat 7 ETM+ (R=5, G=4, B=3) acquired on 21st November 2001 (after 8/11) in the study area.

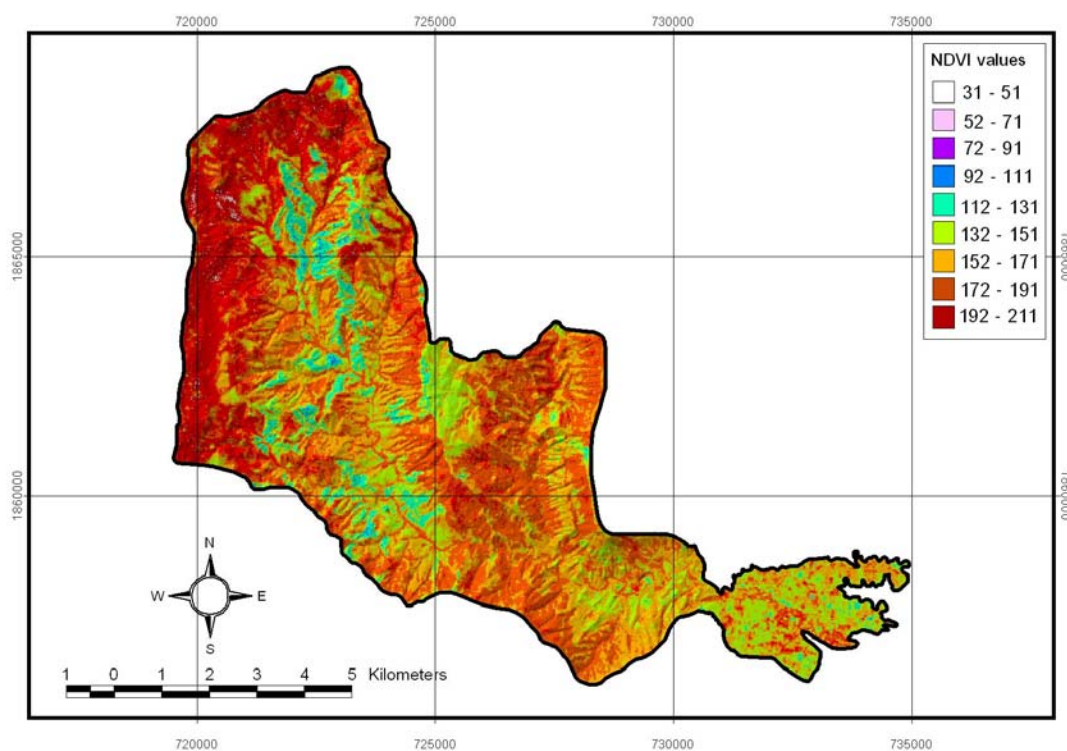


Figure 3-19 Normalized different vegetation index (NDVI) of Landsat 7 ETM+ (R=5, G=4, B=3) acquired on 5th January 2001 (before 8/11) in the study area.

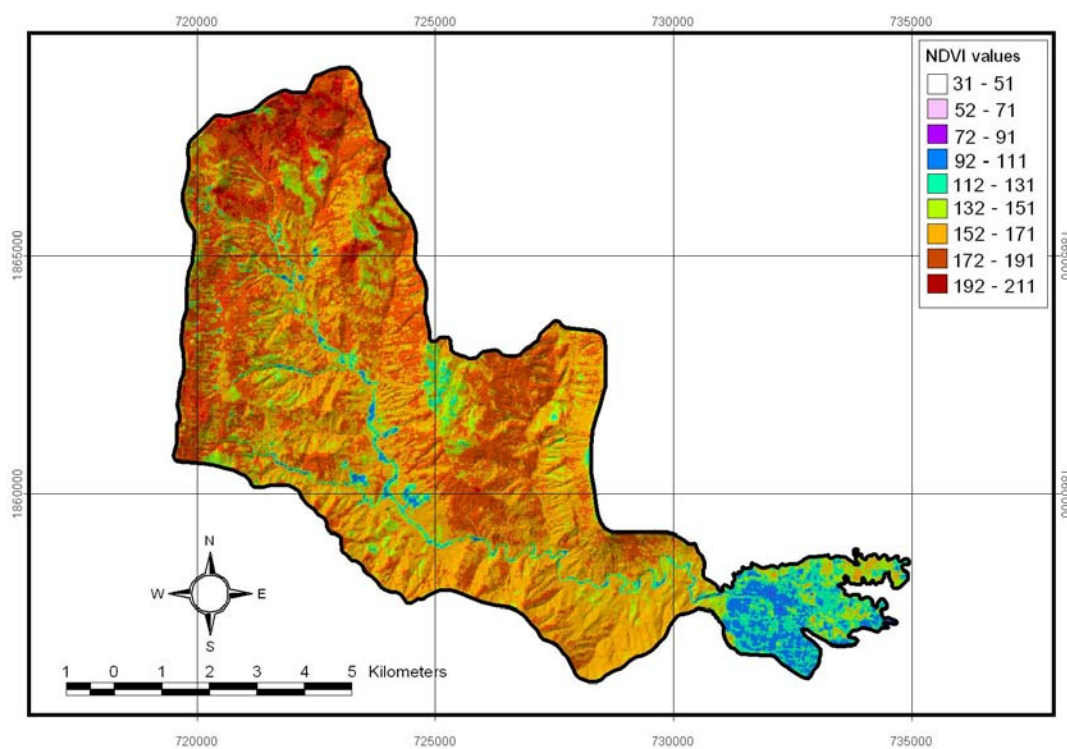


Figure 3-20 Normalized different vegetation index (NDVI) of Landsat 7 ETM+ (R=5, G=4, B=3) acquired on 21st November 2001 (after 8/11) in the study area.

The results of NDVI analyzed from the red- and infrared spectral bands of multi-spectral Landsat imageries (as shown in Figure 3-17 and Figure 3-18) are presented in Figure 3-19 and Figure 3-20. Finally, the NDVI imageries are used to establish a threshold of vegetated and un-vegetated pixels in the images for the change detection as the scar-scouring and depositional locations in the study area as shown in Figure 3-21.

The classification scheme used to detect the scar-scouring and depositional locations depended on a user-specified hierarchical structure to eliminate image objects that were not of interest. The first level was a division of the image between the vegetated and un-vegetated objects based on their NDVI value. The choice of 150.00 was empirically based on an inspection of the image objects from the ground-truth information. Those objects with NDVI value below 150.00 were considered as un-vegetated objects, and those above 150.00 as vegetated ones.

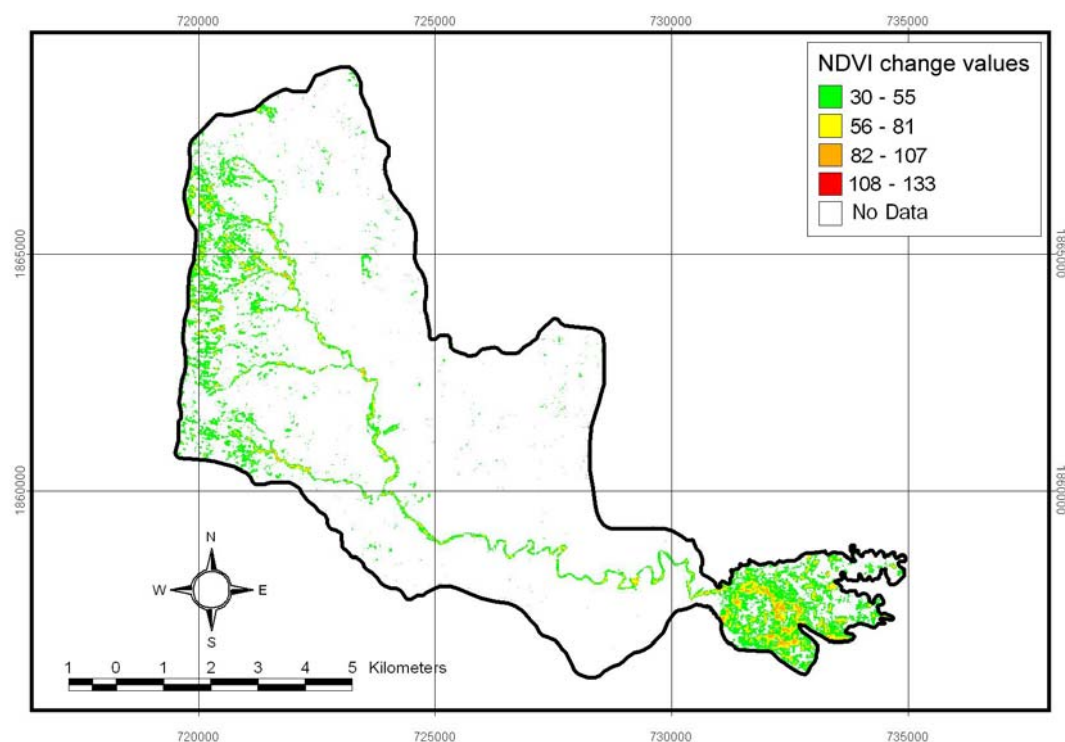


Figure 3-21 Resulted significant change detection of NDVI that used to detect the scar-scouring and depositional locations in the study area that are caused from the 8/11 flow-flood occurrence.

The scar-scouring and depositional locations were also detected and validated by an accuracy assessment. Classification accuracy was determined by comparing a sample of classified pixels with ground-truth information derived from the orthophotographs, IKONOS imagery (acquired after the 8/11 event) and the field observation (Figure 3-22 and Figure 3-23). The validity of the classified results was tested through the identified ground-truth information of the scar-scouring and depositional locations.

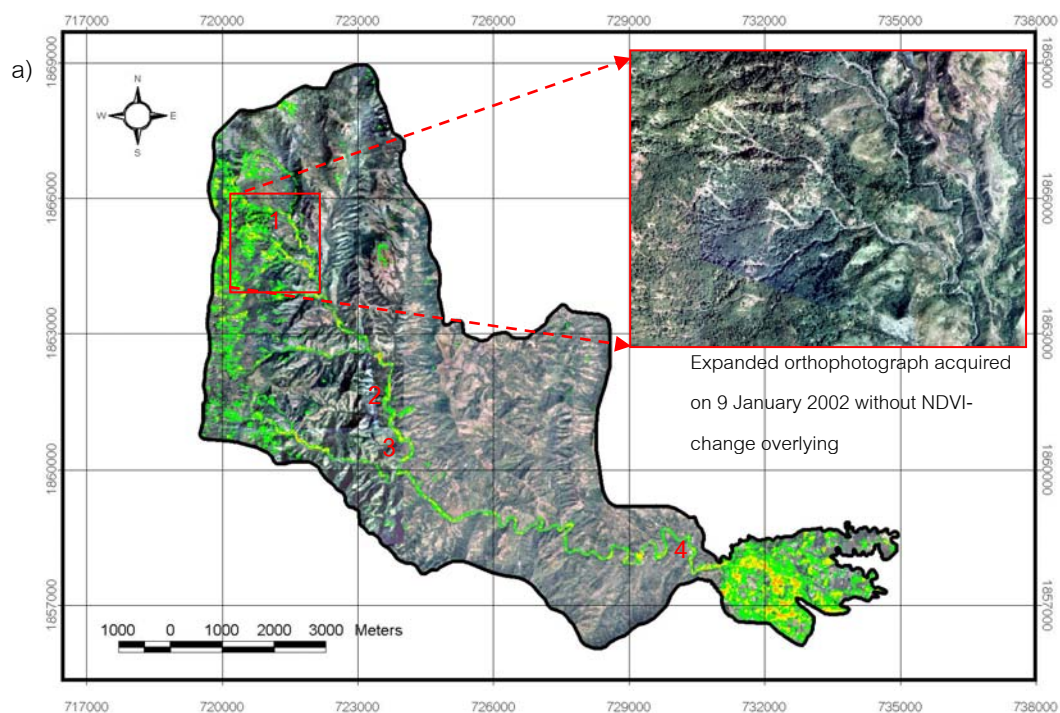


Figure 3-22 a) Significant change of NDVI (referred to Figure 3-21) overlain on the orthophotograph image acquired on 9th January 2002 (after 8/11); and b) photographs of four locations (number referred to the location in the map) taken a few days after the 8/11 event showing the ground truth evidences.



a) Orthophotographs acquired on 9th January 2002



b) Orthophotographs acquired on 29th March 2002



c) IKONOS imagery acquired on 31st October 2002

Figure 3-23 Examples of high resolution remote sensing imageries (acquired after the 8/11 event) used for classifying accuracy and validating NDVI results that related to detect the scar-scouring and depositional locations in the study area.

3.8.3 Accuracy assessment of scar-scouring delineation

This step applied the confusion matrix method to determine the accuracy of scar-scouring delineation from training area of Landsat imagery acquired on 21st November 2001 in Nam Ko Yai sub-catchment (Figure 3-24). The confusion matrix was used to compare the evaluating classification of scar-scouring locations interpreted from NDVI change (Figure 3-25) and ground truth data of scar-scouring delineation manually digitized from 1:20,000 orthophotograph acquired on January 14, 2003 (Figure 3-26). These two set data of the scar-scouring boundaries were used for calculating the accuracy assessment in 10 meter-resolution.

An error matrix was presented in Table 3-4 to determine how well a classification has categorized a representative subset of pixels used in the training process of classification. This matrix was stemmed from classifying the sampled training set pixels and listed the known ground cover types used for training (columns) versus the pixels actually classified into each classification category (rows).

Several other descriptive measures can be obtained from the error matrix. For example, the overall accuracy is computed by dividing the total number of correctly classified pixels (i.e., the sum of the elements along major diagonal) by the total number of reference pixels. Likewise, the accuracies of individual categories can be calculated by dividing the number of correctly classified pixels in each category by either the total number of pixels in the corresponding row or column. What are often termed producer's accuracies results from dividing the number of correctly classified pixels in each category (on the major diagonal) by the number of training set pixels used for that category (the column total). This number indicates how well training set pixels of the given cover type are classified.

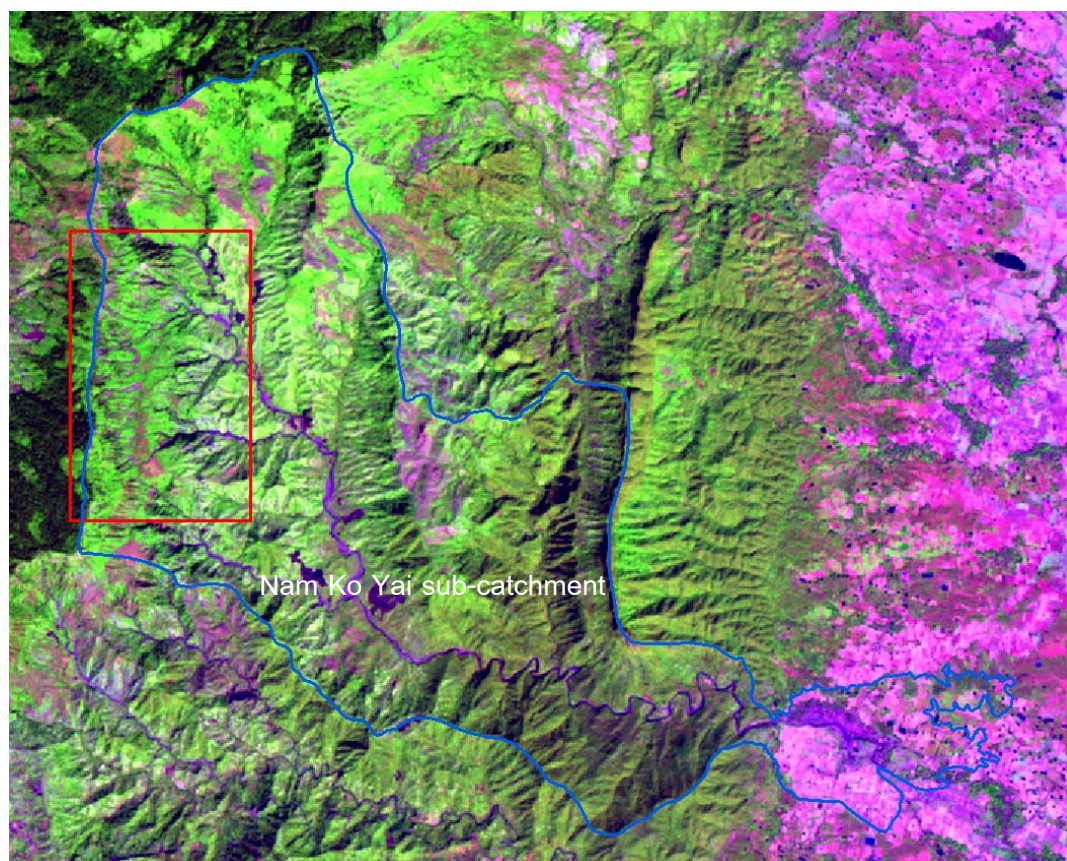
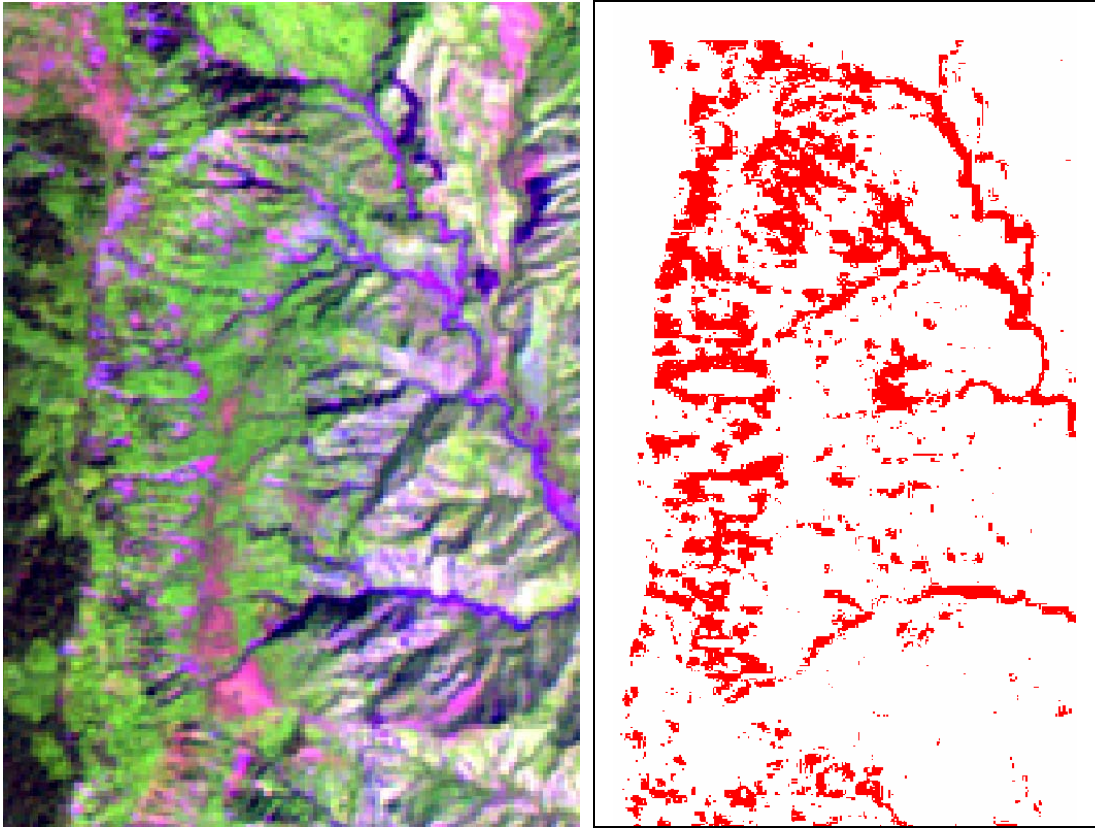


Figure 3-24 Accuracy assessment verification in training area (red box) located in Landsat 7 ETM+ (R=5, G=4, B=3) acquired on 21st November 2001 in Nam Ko Yai sub-catchment.



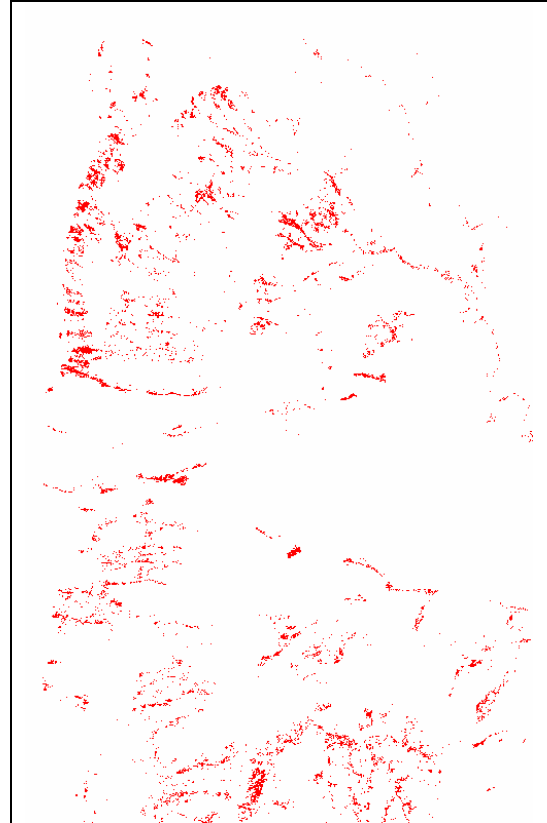
Training area (red box) referred to Figure 3-24

Scar-scouring locations (grouped in red)
interpreted from NDVI change (referred to
Figure 3-21)

Figure 3-25 Scar-scouring locations interpreted from NDVI change in training area.



Orthophotograph acquired on 14th January 2003
(1:20,000 scale) in training area of Figure 3-24



Digitized scar-scoring delineation from the
orthophotograph in training area

Figure 3-26 Scar-scouring delineation digitized from orthophotograph acquired on 14th January 2003 (1:20,000 scale) in training area.

Table 3-4 Error matrix resulting from classifying training set pixels

Categories	Ground truth verification data (from orthophotograph interpretation)		
	correctly classified pixels as scar- scourings	correctly classified pixels as no scar- scourings	Row Total
Classification data			
correctly classified pixels as scar-scourings	3,207	90,583	93,790
correctly classified pixels as no scar-scourings	11,156	458,638	469,794
Column total	14,363	549,221	563,581

According to the error matrix resulting from classifying training set pixels as presented in Table 3-4, producer's accuracy, user's accuracy and overall accuracy of scar-scouring delineation were evaluated as follows:

Producer's accuracy

$$\text{Scar-scourings} = 3,207/14,363 = 22.33 \%$$

$$\text{No scar-scourings} = 90,583/549,221 = 16.49 \%$$

User's accuracy

$$\text{Scar-scourings} = 3,207/93,793 = 3.42 \%$$

$$\text{No scar-scourings} = 11,156/469,794 = 2.37 \%$$

$$\text{Overall accuracy} = (3,207+458,638) / 563,581$$

$$= 81.95\%$$

3.9 Rainfall intensity

3.9.1 Data entry

In this upper Pa Sak region, the average annual rainfall normally exceeds 1,000 mm. The climate is a tropical kind, occasionally with tropical storms in the early and middle periods of rainy season (June-October). The tropical storm “Usa-ngi” that passed through here during the first two weeks of August 2001 was blamed to cause this tragedy.

In this study, rainfall data were received from observation stations of Thai Meteorology Department during 1st-31st August 2001 being put into database around Nam Ko Yai sub-catchment (as show in Figure 3-27).

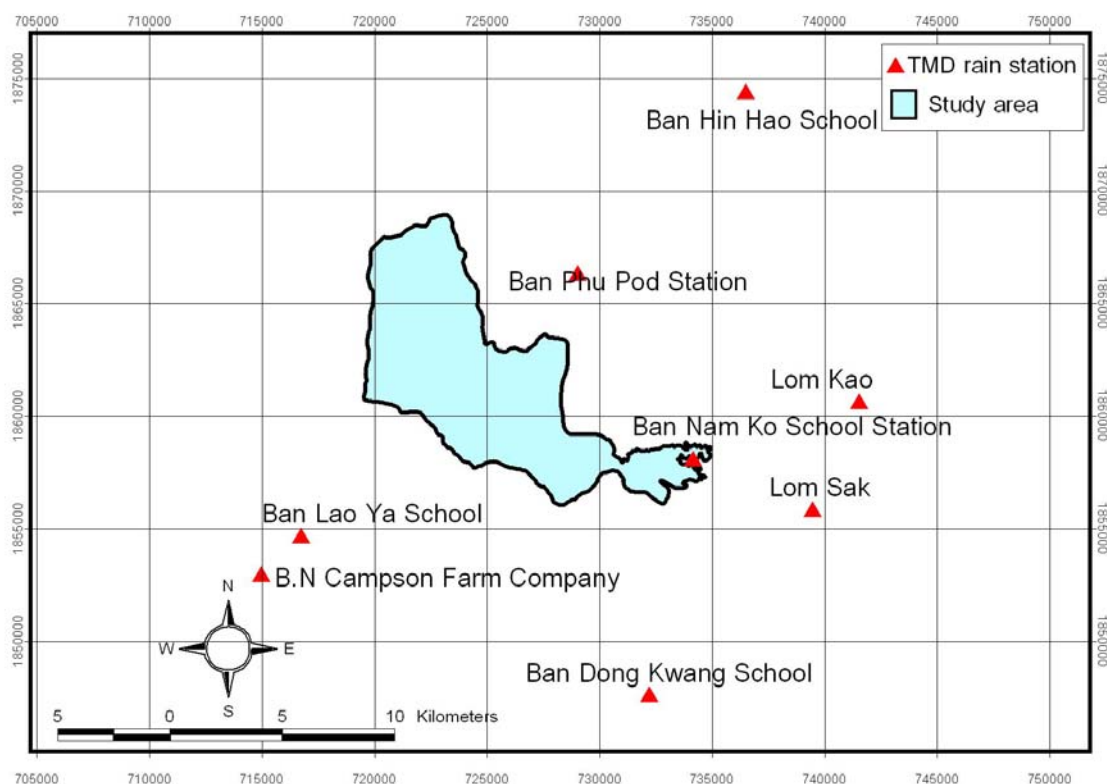


Figure 3-27 Location of seven rainfall-measurement stations of Thai Meteorological Department (TMD) near the study area.

The analysis of rainfall data (derived and interpolated from the observation stations of Thai Meteorology Department) during 1st-10th August 2001 (before 8/11 event) in relation to the configuration of sub-catchment and channel characteristics had been conducted and used as one of the most critical factors to identify the potential for the event in the study area. The graph of rainfall measurements in August 2001 from seven locations surrounding the study area (as shown in Figure 3-27) is summarized in Figure 3-28.

The pattern of rainfall during 1st-31st August 2001 recorded in most of the stations reveals the same manner that there was continuous rainfall during 2nd-14th August 2001 with the most intense raining of about 60 and 100 mm was recorded on 10th August 2001 at the Ban Lao Ya station (southwest of the study area) and Ban Hin Hao station (northeast of the study area), respectively. The average rainfall value during 1st-10th August 2001 recorded from these seven locations in Figure 3-28 is 12.98 mm whereas the average rainfall value of each station is illustrated in Figure 3-29.

For the following year 2002, the graph of rainfall measurements from eight locations surrounding the study area (seven previous locations plus a newly installed station at Ban Nam Ko Yai) is summarized in Figure 3-30. The average rainfall value during 1st-10th August 2002 recorded from the seven locations is 11.4102 mm as shown in Figure 3-31.

According to the Figure 3-29 and Figure 3-31, the average rainfall value (mm) of each station surrounding the study area during 1st-10th August 2001 (before 8/11 event) and the average rainfall value (mm) of each station surrounding the study area during 1-10th August 2002 (one year after the 8/11 event) are calculated as 12.9806 mm and 11.4102 mm, respectively. The rainfall intensity of 10 August 2001 (one day before the 8/11 event) was the highest value of this 1st-10th August 2001 duration.

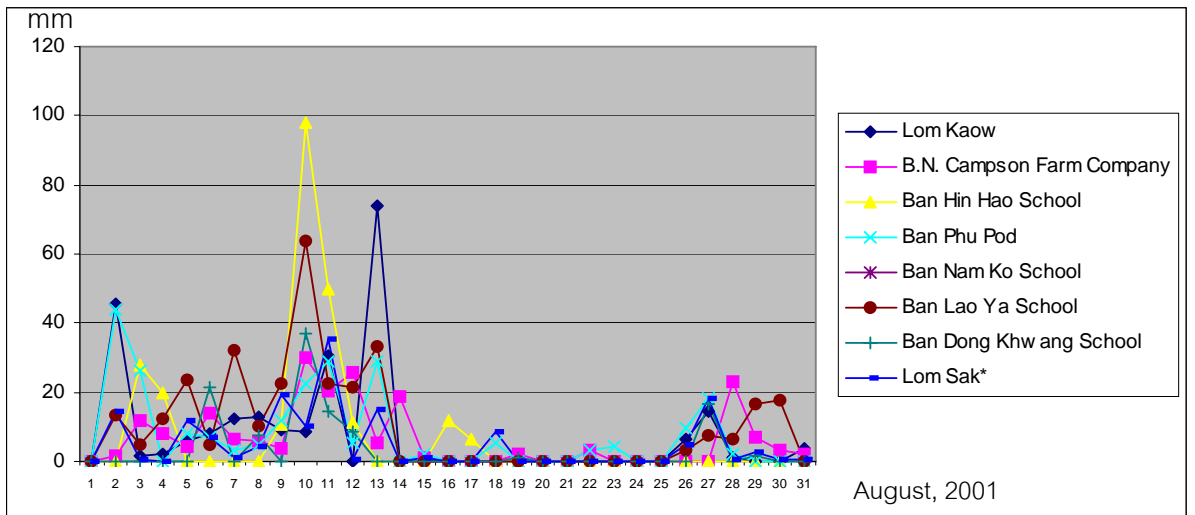


Figure 3-28 Graph showing the pattern distribution of rainfall measurements in August 2001 recorded from the seven locations near the study area.

(Note: the data of Ban Nam Ko School, the newest station in 2002, was not available during those period)

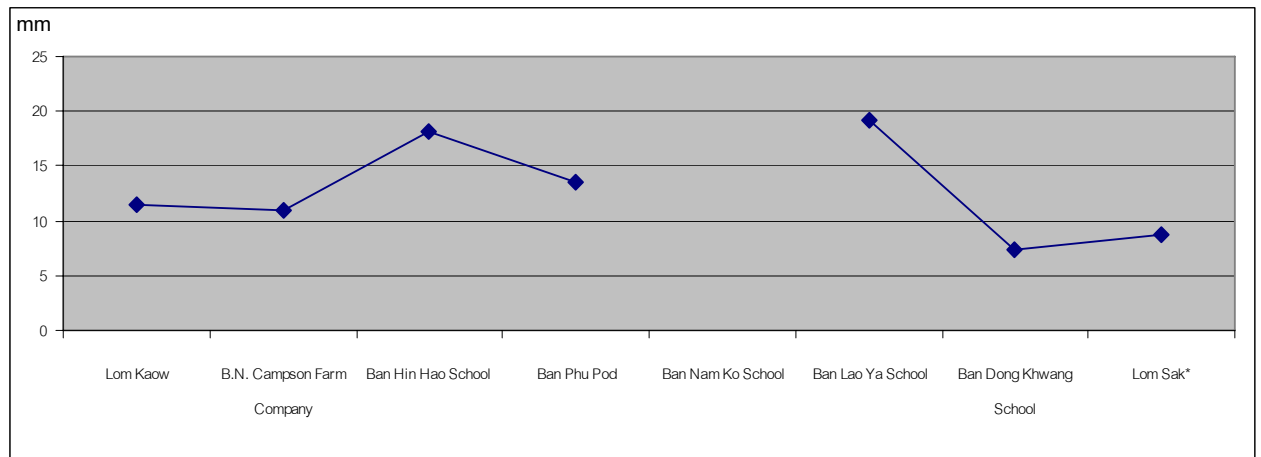


Figure 3-29 Average rainfall value (mm) of each station near the study area during 1st-10th August 2001 (before the 8/11 event).

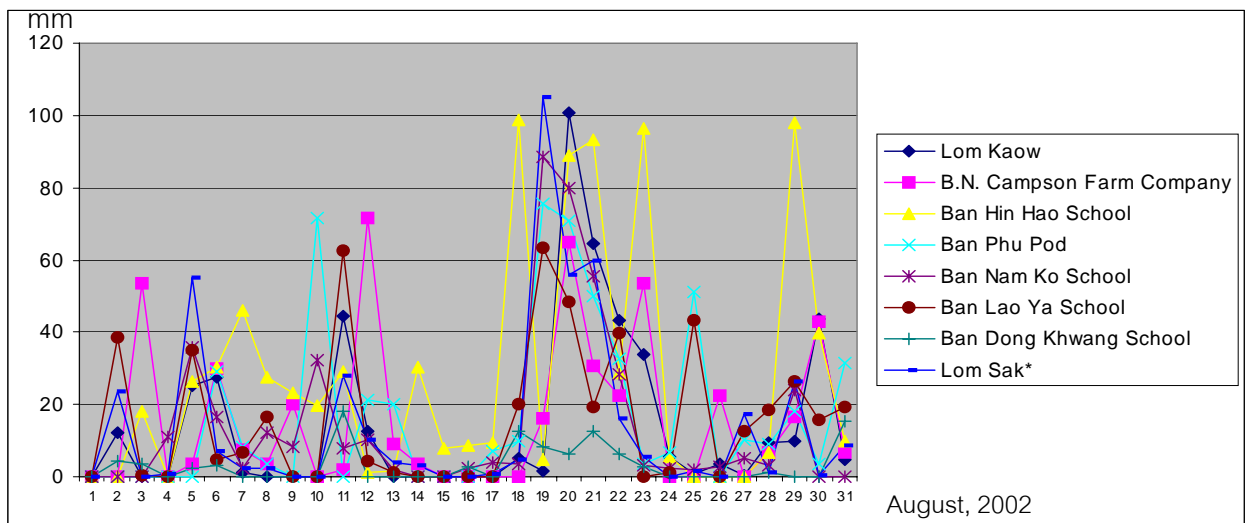


Figure 3-30 Graph showing the pattern distribution of rainfall measurements in August 2002 recorded from the eight locations near the study area.

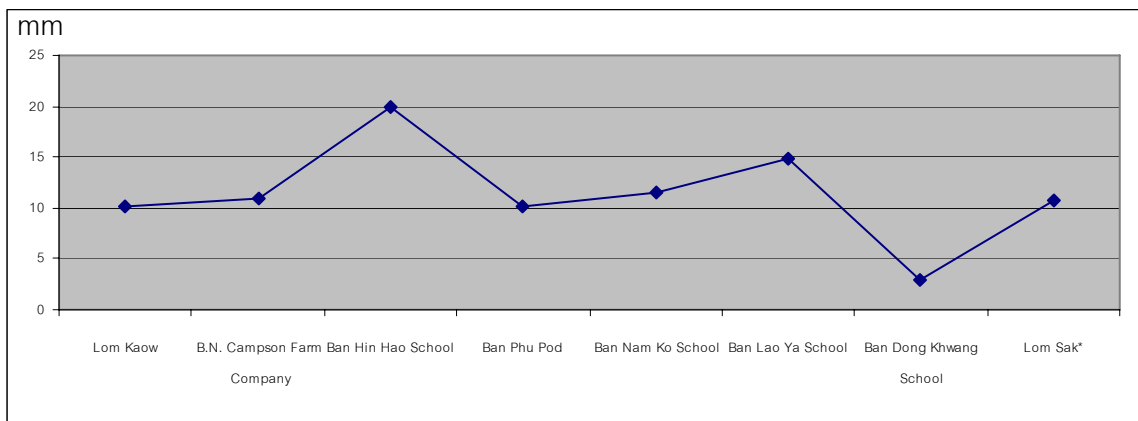


Figure 3-31 Graph showing the average rainfall value (mm) of each station near the study area during 1st-10th August 2002 (one year after the 8/11 event).

It is noted that the intensity and duration of rainfall in August 2002 as shown in the graph of rainfall measurements (Figure 3-30) were much stronger and longer than those of the year 2001 (Figure 3-28). However, there was no disastrous event of debris flow and debris flood in the study area during this time. Only a mild flash overbank-flood had occurred here throughout the whole month.

3.9.2 Input map generation

From the rainfall database mentioned above, the rainfall intensity grid during 1st-10th August 2001 (that was the strongest period of rainfall before the 8/11 event) is generated by interpolated rainfall station by Krigging procedure. The isohyte of the rainfall intensity during 1st-10th August 2001 is presented in Figure 3-31. It is remarked that the most intense raining of about 150 and 160 mm was illustrated in the western and northern steep-cliff of Nam Ko Yai sub-catchment of the study area in that duration of ten days before the 8/11 event.

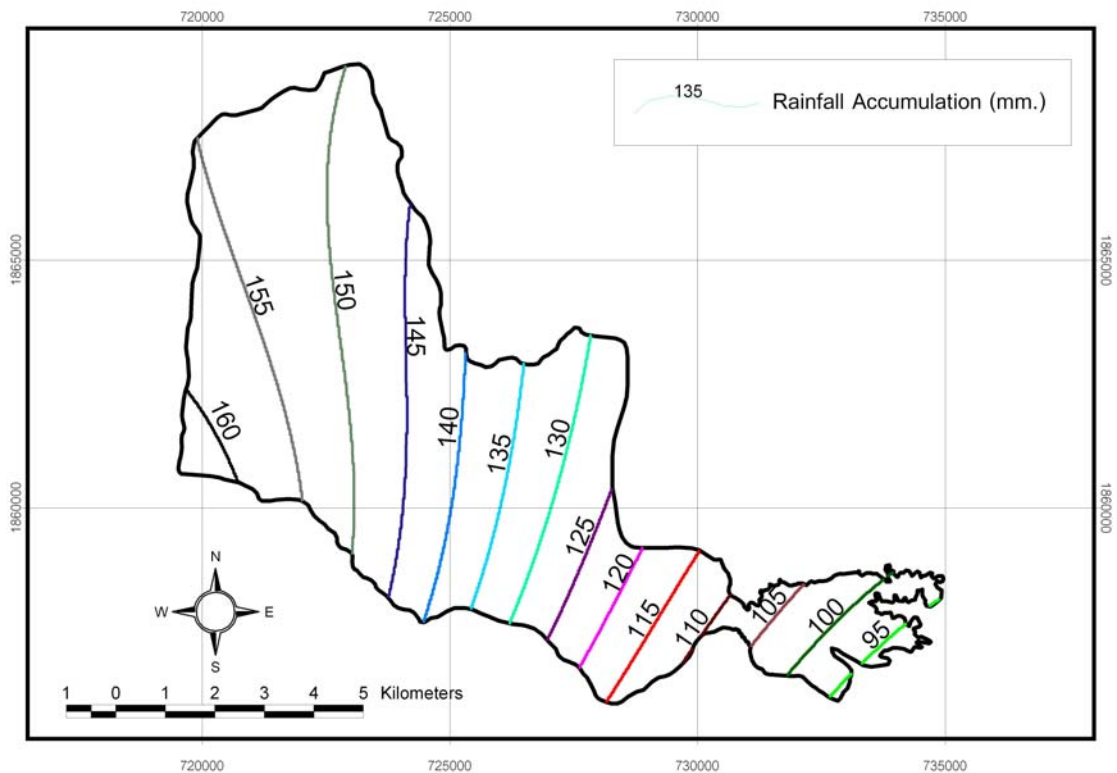


Figure 3-32 Isohyte map of rainfall intensity during 1st-10th August 2001 in the study area.