Flood Propagation Analysis in an Urban Environment; a Case Study in Yen Bai Vietnam

Niranga Alahacoon
February, 2010

Flood Propagation Analysis in an Urban Environment; a Case Study in Yen Bai, Vietnam

by

Niranga Alahacoon

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: Geo Hazard.

Thesis Assessment Board

Dr. D.G. Rossiter (Chair)
Dr. R. van Beek (External Examiner)
Dr. Dinand Alkema (First Supervisor)
Prof. Dr. V.G. Jetten (Second Supervisor)

Ms. Thi Hai Van Nguyen (Thesis Advisor)

Observer :
Drs. T.M. Loran

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Abstract

Fast growth of world population can cause the rapid development and expansion of the cities in developing countries. Furthermore, loss of life and properties in such sites is increasing during the severe flood event which can occur under a typhoon. Therefore asses the flood hazard and risk is important for future developments. 1D2D hydrodynamic model can be used as a tool for simulating severe flood events. Accurate elevation data and historic flood data are the most important parameters for such models. Rapid changes of the topography of cities may cause a problem in updating the topography such in Digital Surface Model (DSM). This thesis is focused to asses the flood risk in Yen Bai city Vietnam.

To create a Digital Surface Model (DEM), first the Digital Elevation Model (DEM) was created by integrating the different elevation data. The effect of integration of multi-source elevation data was studied through the geostatistical approach by considering the nugget value. Then the interpolation of elevation data was carried out with Kriging, Polynomial and Triangulated Irregular Network (TIN) interpolation to assess the quality of them by errors. The DEM which was generated with Kriging interpolation and the building heights integrated with that to generate the DSM for further hydrological modelling. The DSM and historic flood data was used to simulate the 5, 10, 25, 50 and 100 year return period of flood with SOBEK 1D2D flood model. Then this model was calibrated for building structure and three sets of manning's coefficient with flood extent and height data of 2008 super typhoon.

The flood hazard of the different scenarios period was assigned in to flood extent map and then categorized them in to hazard classes. Furthermore model results of 1D2D flood model was converted in to different flood parameter maps such as maximum flood depth, velocity, impulse, rate of water level rising and sedimentation. Those parameter maps was integrated with Spatial Multi Criteria Evaluation (SMCE) in ILWIS to create the Multi parameterized flood risk map.

Among the different DEM interpolations, the Kriging interpolation has given the best DEM for flood simulation. Integrated 1D2D (SOBEK) flood simulation models can be used to simulate the urban floods in complex topography. The behaviour of flood water flow is affected by the building structures of the area and the solid block representation has given the better performance for flood simulation in Yen Bai city. In the event environment experiencing with lack of flood risk data, Multi parameterized flood risk mapping approach can be applied in valuable manner.

Acknowledgements

I would like to acknowledge here the enormous strength and coverage bestowed on me by my beloved parents and other family members.

I would like to give my great respect to all those who are contribute in different ways for my studies in ITC. It is my great pleasure to express my Supervisors Dr. Dinand Alkema and Prof. Dr. Victor Jetten, who have guided me by providing invaluable knowledge, suggestion and contribution.

Special gratitude goes for Dr. D.G. Rossiter and all the staff in ITC, for shearing their experience and knowledge throughout the duration of study.

I would like to express my special thank to Dr. Jagath Gunathilaka for giving an opportunity to me to study in ITC.

Special gratitude to Dr. Hung, Miss Van, staff members of different offices in Yen Bai and all the other Vietnam friends who were helped me at all the time to fulfil the field work in Vietnam.

My special thank goes to all Geo-hazard colleagues, specially Trung, Ezra and all Sri Lankan friends who were with me at all the time giving support to finish this hard work.

Table of contents

List of figures

List of tables

List of abbreviation

DEM	Digital Elevation Model
DTM	Digital Terrain Model
DSM	Digital Surface Model
DHI	Danish Hydraulic Institute
LiDAR	Light detection and Ranging
GPS	Global Positioning System
IDW	Inverse Distance Weighted
TIN	Triangulated Irregular Network
RMSE	Root Mean Square Error
GIS	Geographical Information System
SRTM	Shuttle Radar Topography Mission
1D	One-dimensional
2D	Two-dimensional
RMSE	Root Mean Squire Error
R-Bias	Relative Bias
ARAD	Average Relative Absolute Difference
R-RMSE	Relative Root Mean Squire Error
L-RMSE	Log Root Mean Squire Error

1 Introduction

Rapid growth of the world population and migration leads to the expansion of cities all over the world causing many problems. Many of those cities are located on the flood plains of large rivers. Intense rain fall (Typhoon) generates flash flood waves causing damages to people, and their properties in those cities. Therefore flood simulation of flood events play an important role for disaster management studies.

The flood simulation modeling was challenged in the past. Development of the advanced computer processors and its processing capacity the computation capability of the complex mathematical formulas is going to be fast. Furthermore the data acquisition techniques were also being developed rapidly during last 10 years providing capability of 1D and 2D flood simulation in complex terrains (Merwade et al 2008). Hydraulic models and computing power are now sufficient to cope with large (5,000,000 grid cell) areas of inundation and large (100,000 m3s-1) peak discharge (Carrivick 2006). Capability of the flood simulation, quality and data requirement also change with the different software packages. All these are more or less proportionate to the cost of the package. (eg: SOBEK, flow 2D, PCraster). Also there are some free wares which have the flood simulation capability.

The quality of flood simulation results also depends on the quality of the terrain (topographic) and resistant data for water flowing. For the urban (complex) flood simulation, a good digital surface model (DSM) is required. Generation of a DSM is not always easy and it also depends on the data availability. Developed countries may use high resolution LIDER DEM as input, but it is not always the case for the developing countries. Therefore generation of DSM with enough details is always challenging because flood propagation of the complex urbanized area depends on the building, road (road connectivity), subways, water bearing formation, land use and breaking up of slope (mainly in the terraces), etc.

1.1 Problem statement

Flooding is a major disaster event that occurs in many parts of the world. Monsoon rain and typhoons are the major events for flood generation in tropical countries. This study is mainly focused on the flooding problem of Yen Bai city in Vietnam that suffers frequently large flood events. Yen Bai city of Vietnam has tropical monsoon climatic condition, often faces heavy rains and typhoons which caused a number of floods in last ten years.

Red River and Cuong Thinh stream system are the main drainage systems in Yen Bai city. In dry season the discharge is very low in this river system but in rainy season or Typhoon the water level

rise in Red River these courses problems in the Cuong Thinh river system. The inundation of agricultural fields and city causes loss of life and economy of the area.

From 1990 to 2005, 28 river flood events occurred together with landslides in North Vietnam with 998 deaths and missing, 698 injured, 13289 houses collapsed, 114889 houses damaged and 180000 ha of crops damages (JAT report, 2008). 27 September 2005 typhoon event was one of the biggest in recent with heavy damages of 3,339 billions of VND (Vietnam Dongs).

The complex topography and the man made structures of Yen Bai have highly influenced the flood propagation behavior. The road system, underground or surface channel system, water buffers (lakes), building, obstetrical, slope cuts, and lowland paddy fields are the major elements to be paid more attention for reconstruction of surface topography for the flood simulation. The flood simulation models for Red River flood in city of Yen Bai are very much necessary for flood hazard mitigation, preparedness and risk assessment. Then the evacuation planning is easy to carry out. Flood behavior and simulation studies could be very much useful for future development plan of the Yen Bai city to avoid the future damages for properties and lives during Red River flood events.

1.2 Main objective

To characterize the propagation of the Red River flood through the complex urban topography of Yen Bai city Vietnam.

Specific objectives

1. To reconstruct the complex urban topography required for flood propagation modeling.
2. To obtain flood information of historic events mainly extent and flood depth.
3. To create a flood hazard map, using 2D flood propagation data.
4. To generate the specific flood risk map of the area with flood hazard component.

1.3 Research Question

What are the flood propagation characteristics of Red River flood through Yen Bai?

Specific questions

1. How to reconstruct a sufficiently detail terrains, terrain model for flood propagation model with the available data?
2. What information, historic events are available and how can it be used to calibrate the model?
3. What is the flood hazard in the Yen Bai city?
4. How to integrated flood risk component to create the risk map?

1.4 Methodology

In the section on the methodology has been subdivided into four main tasks to fulfill the research objectives as follows.

Task 1

Reconstruction of Detailed surface topography

Reconstruction of the detailed surface model mainly based on the integration of elevation data from different sources. The elevation data could be coming from point, line, polygon and those have to be integrated in to ground terrain & manmade structures such as buildings, slope cuts, water bodies etc. Data sources for elevation mainly contour maps, spot heights from the topographic map, and elevation data from road, drainage and field measurements. Considering the accuracy level different interpolation techniques such as Kriging, High order terrain interpolation, TIN based interpolation can be used for DEM construction. Nearest neighbor, bilinear, bicubic resampling techniques are also available to resample the DSM for different resolution. Depending on the accuracy level, spatial error distribution and simulation time (test simulation) suitable DEM resolution will have to be selected for simulation studies.

Task2

Flood historic information

Historic flood event data of flood depth and extent are may be useful for calibration and validation the flood simulation model. Those historic data available in different places such as meteorological department, irrigation department, disaster management centers etc, but community based PGIS survey is also useful for historic data collection. Those questionnaires data can be entered on to spread sheet with coordinates to create a GIS point layer with attribute data. This basic GIS layers are used for further analysis like creation of flood extent and depth polygon.

Task 3

Flood simulation, model calibration and validation

Flood simulation mainly based on the SOBEK 1D 2D model, one dimensional channel flow and two dimensional channel overflows is represent by a 2D grid of elevation information in SOBEK model. Schematization of flood model component on the top of GIS layer is the main part in SOBEK model. Main flood model components are river cross-section, calculating points, 2D grid, connection node, boundary nodes etc. Two water discharge boundaries used for Cuong Thinh river catchment shows in Figure 2. 1. Boundary nodes represent the geographical separation of the model from the real world. Discharge information for different return periods are important parameter which can be assigned to the up and downstream boundaries. Building, surface roughness and DSM are the main input maps for simulation.

The flood model calibration is known as the adjustment model parameters to reduce the deviation between the results of the flood model with corresponds to observed flood data. The main flood

calibration parameters are flood depth and extent derived from field work and model. Different building structures (solid block or rough surface) and surface roughness maps are changed the flood propagation behavior of the system. With different sets of Manning's coefficient and building structures are used for model calibration.

Task 4

Generation of flood hazard and risk map

Flood hazard is the probability of occurrence, within a given period of time of a given area of a potentially damageable flood event. Historical flood data is necessary for hazard studies. Historical data (return period) and simulated flood data used to estimate the temporal and spatial probability of occurrence of Red River flood event. Scenarios based simulated flood data like flood depth and flow velocity can be transferred in to parameter maps such as maximum water depth, maximum duration, maximum water rise, maximum impulse …etc. the spatial distribution of Red River flood event is possible to demarcate together with indicator maps and return period.

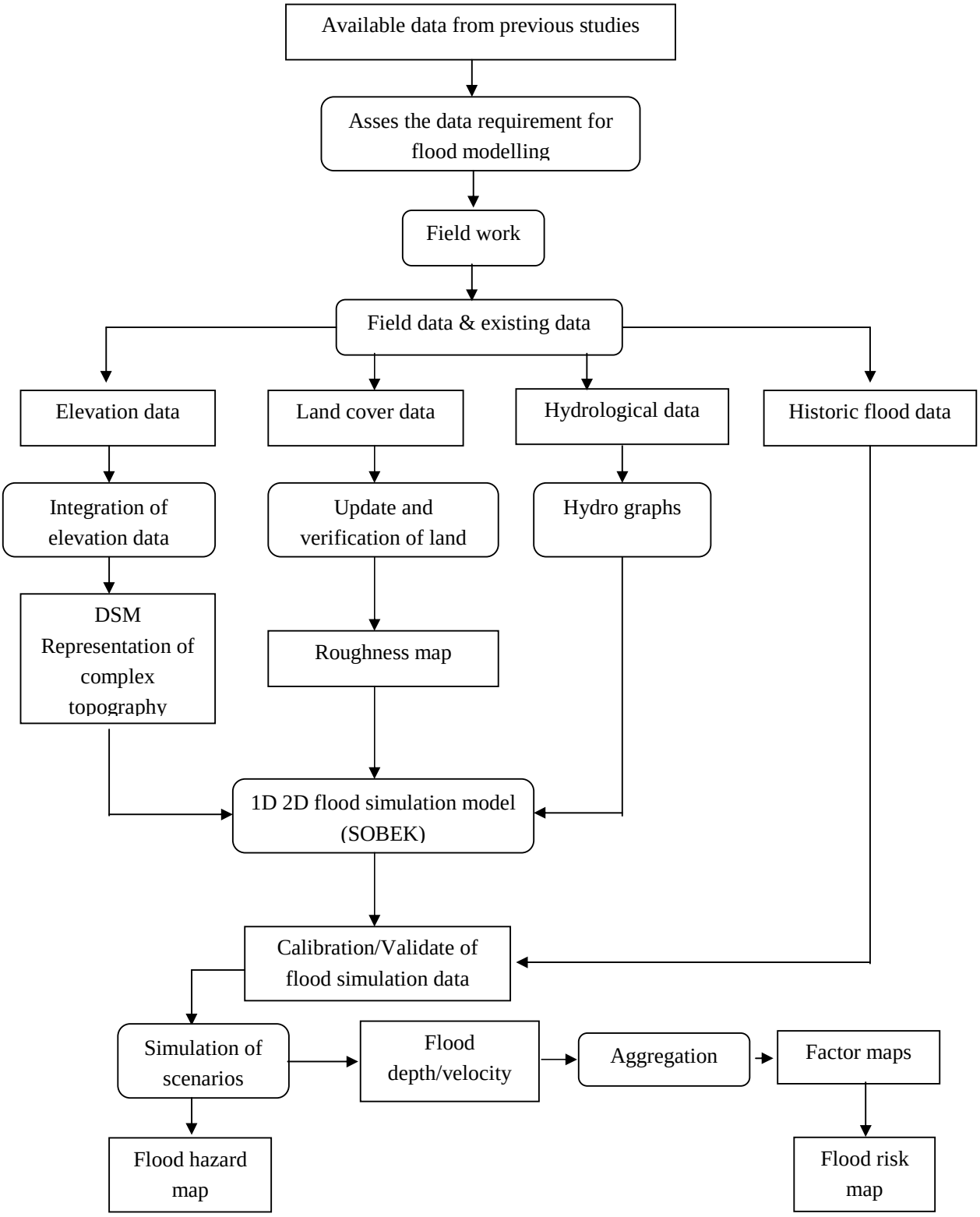


Figure 1.1: Overall methodology

1.5 Limitation of the study

Main limitation factors for the study are available data, computer hardware and software. High resolution (vertical and horizontal) LiDAR data are very much useful to represent the terrain for flood simulation. Otherwise the quality of the DSM also depends on the technique which can be used for DSM generation with integration of different data sources. The existing DEM can be updated in to DSM with the help of building heights (average floor height and number of floors) and Conventional geodetic leveling methods for different points in the study area. Amount of all above measurements depend on the time allocated for field work.

Computation capacity of the computer which is used for the flood simulation is major limitation for this study. (Modeling limitation of SOBEK is 5,000,000 grid cell and 100,000 m^3s^{-1} discharge). High resolution DSM and high amount of discharge of the system is increased the computation time for the simulation, but the model performance is to be better in this situation. Therefore selection of the appropriate DSM resolution is challenging. Simulation time is one of the major aspects to be considered due to availability of limited period of time. Limited data availability for flood depth and flood extent is the other limitation factor and it is mostly involved for the model validation and calibration procedures. Therefore limited number of points would have to be used for model validation and calibration instead of whole area studies. Building can be represented as either high Manning's permeable object or impermeable building blocks. Those two representations of the buildings are major aspects to be considered during flood simulation, because this is mainly related to flood depth, velocity and extent.

5.6 Thesis outline

Seven chapters included in this thesis

Chapter 1 is giving the overview of the study which contains the explanations on research back ground, research problems, research objectives, research questions, overall methodology and the limitation of the study.

Study area, data need assessment and data collection are explained in the chapter two. Study area is given the detailed explanation on Red River flood plain, Red River flow, topography of the area, climatic condition, geological setting and socioeconomic condition. Data need assessment covers details about the existing data, required data and the way that used to collect the data with various data categories such as elevation, land cover, flood frequency, flood depth and flood extent.

Chapter 3 is mainly focused on five main sub-topics like Digital Elevation Model, role of DEM in flood simulation, different DEM interpolation techniques, DEM quality assessment, flood modelling and Model calibration that more important for this study. Digital Elevation Model is given the detailed explanation on them and role of DEM in flood modelling explains the impotency and requirement of DEM in flood modelling. Subtopic of different DEM interpolation techniques is discussed in details which used in this research. Qualitative and quantitative approach to define the DEM quality is explained in the sub section of DEM quality assessment. Sub topic of flood modelling focused on the discussion of one dimensional and two dimensional flood modelling based on the pervious researches, it is further discussed the boundary condition and surface roughness. At the end discussion focused on the flood model calibration.

Chapter 4 is given the detailed methodology of Digital Surface Model (DSM) generation and flood frequency analysis. The DSM generation part covers the data preparation, integration, and DEM construction by interpolation, and inherent error distribution, selection of better DEM with qualitative and quantitative approach and finally creation of DSM. Flood frequency analysis is focused on detailed study of Gumbel distribution to reconstruct the return period of Red River flood.

Chapter 5 covers flood simulation of the study area with different return periods. It is also given the detailed explanation for flood model parameters such as boundary condition, building structure and surface roughness. Then it focused on model schematization with SOBEK and model calibration.

Chapter 6 is focused on flood hazard assessment and Multi-parameterized flood risk in Yen Bai city. Focused on the flood extent and return period, flood hazard was introduced. Multi-parameterized risk assessment aims the generation of risk parameter maps (maximum flood depth, velocity, impulse, sedimentation and rate of water level rising) and integration of them to create a risk map through Spatial Multi Criteria Evaluation method in ILWIS. Final part of this chapter covers the flood simulation of upstream discharge alone.

Final chapter is focused on the conclusions which gives the answers for research questions on this study as well as the recommendations for further studies.

Chapter 2 Study area, Data need assessment, data collection

2.1 Study area

2.1.1 Red River flood plain in Yen Bai

Yen Bai city is the capital of Yen Bai province of North Vietnam and it lies on the flood plain of Red River. This is the largest river of North Vietnam and the river flows through the province. The city experience several typhoons accompanied by heavy rains and floods. Further, the rapid urban growth of the area drastically changes the landscape.

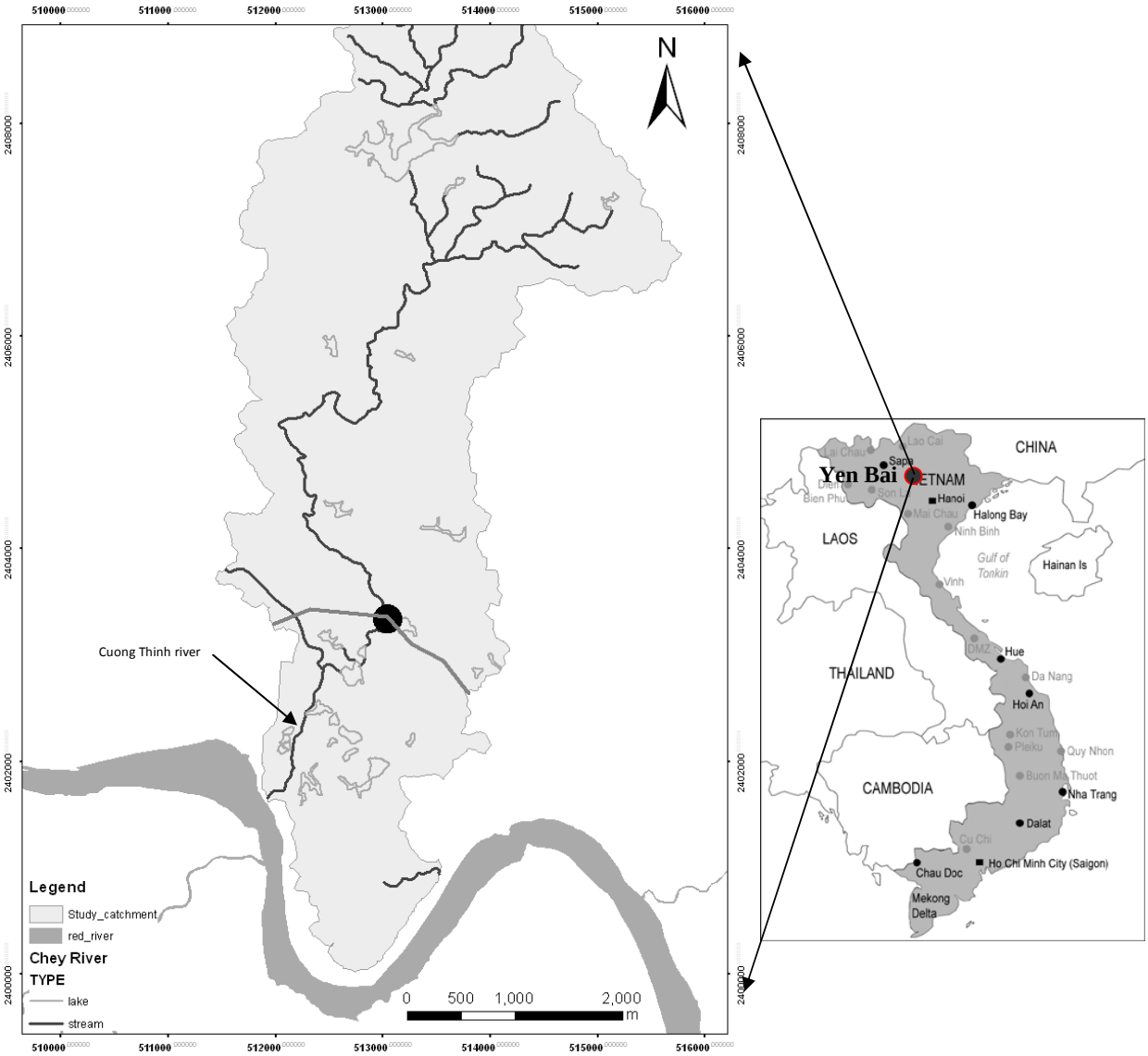


Figure 2.1: Yen Bai city in Vietnam

2.1.2 Location of the study area

Yen Bai province is located in between latitude of $21^0$40'-$21^0$46' and longitude of $104^0$50'8'-$104^0$58'15' with an area of 6,890km^2.The average elevation of the Yen Bai province is 600m above mean sea level. Yen Bai city is located much closer to the confluence of Red and Cuong Thinh River as shown in Figure 2.1.

2.1.3 Red River flow

Historic river flow data gives either discharge or water height in Red River. Discharge measurements of Red River drastically changed April to August in every year. (Discharge is changed from 950m^3/s to 5000m^3/s while changes of water heights was 23m to 30.26m in August 20008). This change is significantly high compared to average monthly discharge and water level of the river.

Figure 2.2: Red River water levels in normal water flowing situation

2.1.4 Topography of the area

Total Yen Bai province is located in the East of Hoang Lien Mountains at the transition zone of high land and mid land, it is sloping down from North-West to South east (Figure 2.3).

The Yen Bai city is also located on the low land of the province and on flood plain of the Red River. Complex topography of the area is derived from the hills, valleys, steep slopes and slope cuts. Hills are more prominent through out the city region with the height of more than 60m but the Hill's height is continuously increasing towards the North-West direction with fairly steep slopes. Slope cuts on that slopes for the roads and houses are more frequent. Valleys distribute among the hilly terrain with the 25-50m elevation variation and most of them are used for paddy cultivation (Figure 2.3).

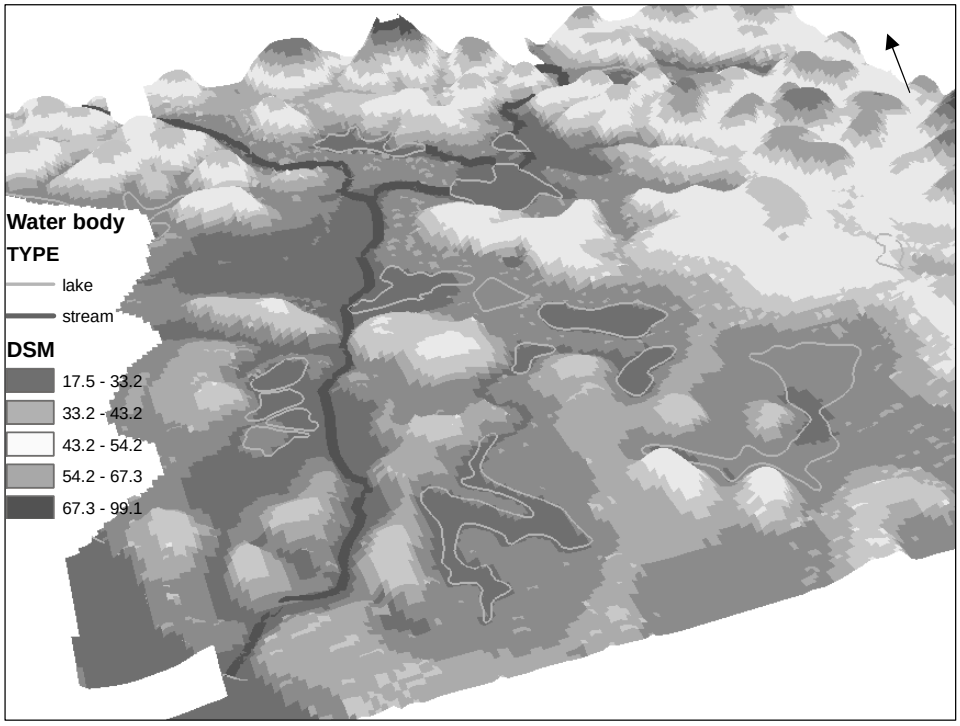


Figure 2.3: Topographical setting of the study area

2.1.5 Climate and weather condition

The Yen Bai city is situated on the tropical monsoonal climatic condition and rain occurs April and October. It is also located closer to tropical typhoon belt and experience 1-4 typhoons per year. The average temperature is 22.5^{0}C and varies from 20.5^{0}C to 27.5^{0}C through out the year. The average annual rain fall is 2057mm. The minimum and maximum rain fall ranges from 1462mm to 2705mm. the maximum daily rainfall of 349mm was recorded on 6th of September 1973. Annual average humidity is approximately 86%.

The left bank of the Red River has given the space for Yen Bai city when it is expanding due to the increase of the population. The minimum and maximum discharge of the Red River is 95m^3/s and 8400m^3/s respectively. However this could go beyond the 14000m3/s during severe typhoons. The maximum and minimum water velocity is 3.02m/s and 0.62m/s and the water depth varies from 23m to 31m above mean sea level (Red River discharge information from 1960 to 2009).

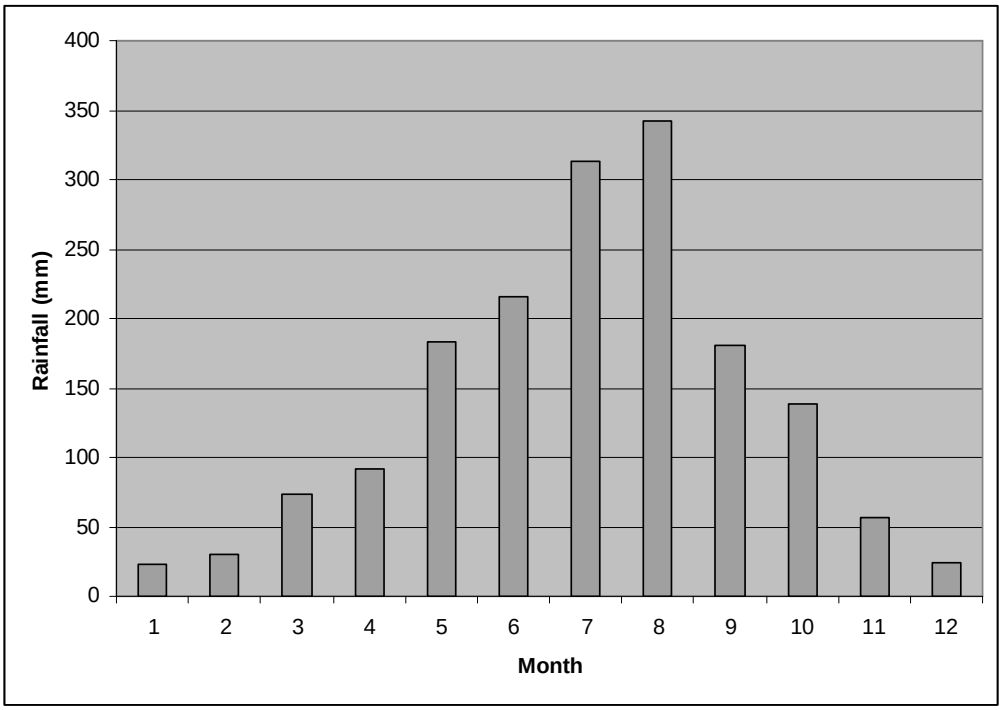


Figure 2.4: Monthly average rainfalls in Yen Bai province

2.1.6 Geological setting of the area

According to Vinh and Hoanh, (2004) seven different geological formations such as **Paleo-Preoterozoid, Paleo-Proterozoid - Meso-Proterozoid (PP-MP), NeoProterozoid (NP), Devon (D), Neogen (N), Quaternary (Q),** and **Magma complex are** existing through out the Yen Bai province. Paleo-proterozide formation is composed of biotite, sillimanite, garnet, plagioclase, amphibolites and quarts as major rock forming minerals. Those formations are also located along the left bank of the Red River closer to the Yen Bai city and the complex natural landscapes of the study area have been developed due to that formation. Quaternary formation can be observed both side of the river bank as soft sediments, pebbles, granule, clay sand and organic matter as the youngest formation in geological history.

2.1.7 Socio-Economic condition of the area

Yen Bai is a level (III) Vietnamese administrative city and it is the capital of the Yen Bai Province, the political, economical, cultural center for the North-West Vietnam. Yen Bai town is located approximately 150km away from the Hanoi (Capital of Vietnam) and total area is about 56km^2.
Approximately 30 different ethnic communities are living in Yen Bai province and Kinh is the majority of those communities in Yen Bai city. 2004 results show that the population of the Yen Bai city is about 65,000, the average population density is 1269 people/km^2. Main occupations are workers, office holders, farmers and traders. Both Governmental and Private industries are common

in the Yen Bai city, and the infrastructures are still in development stage. The main cultivations of the area are tea and paddy.

2.2 Data need Assessment and field data collection

Many different types of data available in Yen Bai province of Vietnam. The data need assessment begins with arranging the existing data in to different categories such as GIS and RS data (Checked with the different file formats, eg ArcGIS, Ilwis), Geological data, meteorological data, and Hydrological data. Then the list of required data was listed based on the existing data list.

2.2.1 Data

Data collection during field work had to be done under the supervision of thesis supervisors and with the help of my colleagues. The data collected with different sources, depending on the availability and additional requirements for the study.

Available data

- Rain fall data
- Catchment
- Contour
- DEM 10m resolution
- River network
- Road network (Road types)
- Satellite image 0.5m , 10m resolution
- Yen Bai land cover map derived from the spot image

Required data (collected through field work)

- Water depth maps for different return period
- Flood frequency data
- Hydrographs or water height for boundary conditions
- Building foot print map (updated by the field observations)
- DSM (have to be generated)
- Land use map (update by the field observations value)
- Manning's roughness map
- Flood inundation map (flood extent)

2.2.2 Preliminary analysis of existing data

Preliminary data analysis was carried out with the help of 1:10,000 topographic maps, satellite images (0.5m resolution Google Image) and landuse maps covering whole flood plain of the study area. Those maps were used to understand the catchment boundary, drainage features and topography of the area. The above mentioned preliminary work was useful for the field work plan.

2.2.3 Data collection

The field work was used to collect the required data which was listed during the data need assessment. Required data were collected from primary and secondary sources. In the beginning, field work data collected with desired precision and accuracy. But due to the limitation of the field work and availability of required data the field work plane was adjusted. All the required data was not available in the offices and therefore most of the data was collected with the field by interviewing the people and verifying. Collected data in different categories are:

1. Elevation or topographical data
2. Land use or land cover
3. Building foot prints (Urban setting)
4. Flood water depth and extent
5. Flood frequency data
6. Flood hazard

Figure 2.5: Cross-section of river and flood water height measurement

2.2.4 Elevation data collection

Elevation data was collected using different sources such as Digital contour data, elevation of buildings and measured elevation points through the field inspection. More than 100 elevation locations and small stretch with same elevation were measured during the field work to update the existing contour map. All the points located on the detailed field map (1:2000 scales) with spatial attributes (Appendix A). Some locations where plat and large areas GPS (Global Positioning System) was used to mark the locations due to the lack of field checks. GPS measurements are not convenient in the places which have the dense buildings, canopy covers, rain, and dense clouds.

2.2.5 Preliminary land covers interpretation

Remote sensing images were used for preliminary land cover interpretation of the area. Landuse classification was done with landsat images to derive the landuse map before the field work. During the field work land use data for the Yen Bai city was collected from the GIS office of the Yen Bai municipality with 1:2000 scales as a hard copy. That was prepared in 2007, and the verification and update of the existing land use map was carried out during the field visit. Also the land use type of some areas before 2007 was obtained by the discussion of the local peoples.

2.2.6 Flood frequency data

The Red River water level and discharge data was taken from the measuring station of the Yen Bai city from 1960 to 2009. Water heights and discharge was recorded with different time intervals in different years. It means that 1960 discharge and water level measured 4 times per day but in 2005 it was decreased in to 2 times per day.

2.2.7 Flood depth and extent data

The flood depth data of the 2008 super typhoon was obtained from the local people with interviews. Also considered the flood marks, which appeared on the buildings, were easy to measure during the field work. 60 locations were used to measure flood depth information through out the study area. Flood extent information was also taken from the local people and plot them on the 1:2000 scale map.

Chapter 3 Literature review

Reality is complex but we have to understand the behaviour of the processes so that it possible to predict and prevent disastrous effect in the future. Such understanding is important in flood modeling as well. The use of models as simplifications of reality enables us to gain understanding and knowledge, enhancing our ability to predict with acceptable level of accuracy.

Surface flood modeling in urban environment is a challenging task for a number of reasons: the presence of large number of obstacles of varying shapes and length scales, the storages in the buildings, the complex geometry of the city (Mignot et al., 2005). Topographic data are crucial for flood inundation modeling in urban environment and it is better to use recent and highly accurate topographic data. However, this is not always feasible given time and budget constraints of the project (Brett and Sanders, 2007). Medium resolution digital elevation data are widely available from a variety of sources, such as digitized contour data and aerial stereo-photogrammetry. However, the use of such coarse spatial descretizations may leads to poor model performance where detailed flood plain geometry, below the resolvable level of model grid, has significant on flow routing (Hilary et al., 2007).

The modeling of floods uses digital elevation models (DEM) to derive topographic attributes (Wu et al., 2008). DEM derivatives include the slope factors such as length and angle and area of catchment. For the model results to be representative enough, of the reality, the optimal DEM resolution is a critical factor, with the lowest degree of error, or the lowest level of uncertainty. DEMs are not usually reddily available and often have to be derived from a fixed set of dataset. More often, the DEMs have to be resampled so that GIS operations can be made possible using GIS software such as ILWIS. The resampling process introduces a certain degree of error or variation, making it necessary to evaluate the effects of DEMs on the flood modeling process. Rientjes and Alemseged (2009) argued that real world, concerning floods, requires schematization and conceptualization which is handled through use of models with DSMs. It can be concluded based on over role consideration of flood modeling the sensitivity of flood model prediction depends on DEM type, resolution and accuracy (Brett and Anders, 2007).

In addition, the representation of buildings is essential in flood modelling. The critical consideration is how to characterize buildings in flood models. The representation of buildings as either partially permeable with a very high friction value or as solid impermeable blocks is tested in the flood model scenarios using different DSM resolutions. (Wu et al., 2008) resampled 10m USGS DEMs to 30, 60, 100, 150 and 200m using three resampling techniques namely nearest neighbour, bilinear and cubic convolution resampling. The first was noted to have the best since the latter two lead to the unreasonable interpolated results, which in other words implies smoothening thus removing local differences which may be important in flood modelling (Wu et al., 2008) (Rientjies and Alemseged, 2009) also argued to the same effect that coarser resolution resampling obscures small scale heterogeneities.

In the urban environment the majority of the flow water is being flown through the street and junctions (Mignot et al., 2005; Chen et al.,). Flood water behavior changes in an urban environment 1D to 2D in place to place, (Weber et al., 2001), mention that the flow in the streets is mostly one dimensional except at junctions. Junctions are basically 2D flow. Furthermore two dimensional flood inundation models can be coupled to one-dimensional drainage models to account for either sub-surface flow or surface flow in channels not resolved by the 2D model (Schubert et al., 2008).

Scale issues have been noted to be critical for the flood modelling (Rientjes and Alemseged, 2009). The effects of DEM resolution were discussed in many studies including (Rientjes and Alemseged, 2009) who showed that DEM grid size significantly affects land surface representation and topographic attributes in flood modeling. This effect for example affects the rainfall runoff modelling results. They treat buildings as solid block but assign a roughness value of manning's n=1, and then two scenarios, regarding buildings as high resistance features (manning's n=1) and as solid impermeable blocks were used (using DSM). (Rentjes and Alemseged, 2009)'s third scenario for treating buildings as hollow was not considered since it modelled buildings as if they do not exist by simply assigning them the same roughness value as the surroundings. This effect removes the influence of buildings as they are similar to the other flooded areas.

3.1 Digital Elevation Model

According to (Podobnikar, 2002) surface of the earth can be described with a model and it is finite set of heights, measured with regards to mean sea level. Those models were available in analog form in early stage but now it is available in digital form as a DEM. There are three different methods for represent the terrain in the computer system according to (Weibel and Heller, 1991) such as

1. Terrain elevation represent as regular grids.
2. Triangulated irregular network (TIN).
3. Contour based models.

The grid based representation of the terrain which is called as Digital Elevation Model (DEM) widely used in many approaches such as Geomorphology, hydrological comparison. Digital Elevation Model is the digital representation of the topographical data as general term. But according to the US Geological survey a DEM is defined basically focused on GIS techniques as "digital cartographic representation of the elevation of the terrain at regularly spaced intervals in x and y direction, using z-values referenced to a common datum".

DEM is generic term which includes any digital model representing height values (Smith et al). There are different types of digital elevation representation known as Digital Surface Model (DSM) and Digital Terrain Model (DTM). DSM represents the upper surface of the physical environment including buildings, trees with the other man made structures. DTM is a bare earth model which is

represented the terrain height of the earth surface. Some filtering operations are available to create a DTM from an existing DSM.

Widespread availability of Digital Elevation Models and the development of software packages have pared the way to analyze the different geo-morphological and hydrological processes. In Runoff, erosion and flood modeling almost all the work is done based on the digital elevation model and computer based modeling approaches (Wise, 2007).

DEM resolution is the main factor to be considered in the representation of the terrain. Representation using coarse resolution Digital elevation model always lacks of terrain details but easy to analyze because of the small file size (storage capacity). Furthermore fine resolutions DEMs always represent more details for the terrain with large data storage capacity which is increased the data processing time. New Development of the raster based GIS software packages and availability of high speed computer processers have increased the capability of processing of fine resolution DEMs.

Quality and the spatial extent of DEMs have improved due to the intensive development of remote sensing techniques and availability of advanced data processing algorithms and data capturing techniques. SRTM, Airborne scanning laser altimetry (LiDAR) and Radar Interferometry InSAR data are new important data sources for terrain representation. LiDAR data is being able to map topographic heights and heights of the surface objects such as trees, building, and other man made structures with high vertical and horizontal accuracy level over the large area (Cobby et al., 2001). Ocean wave dynamics, forest mapping, mapping of dune morphology, land subsidence studies and crop height measurements are possible using LiDAR data. (Cobby et al., 2001) have described the ability of LiDAR data to be used in 1D2D flood simulation models. DEMs derived from LiDAR data are much more expensive compared to any other conventional methods such as geodetic survey. The Shuttle Radar Topography Mission SRTM is a good source for regional or global scale studies having covered 80% of the globe.

3.2 Role of DEM in flood modeling

Frequently DEMs are used to calculate the hydro-geomorphological parameters (variables) which are used for hydrological models (Le Coz, 2009).The flood simulation environment (urban, rural) and the size of the study are the main factors to be considered during the selection of the DEM resolution, because the resolution of the most accurate DEMs is too fine to run the hydrological model over the regional watersheds (Le Coz, 2009). On the other hand surface relief is more sensitive to DEM resolution, hence to represent the high and low relief area; DEMs with high and low resolution are needed (Anderson et al., 2006). Detailed DEM is the essential way to represent the terrain with the different kinds of complicated obstacles such as river embankment, buildings, roads and any other manmade structures (Werner, 2001). Representation of all the features on the terrain is not possible, because to overcome that the pixel size should be very small, however this

will increase the simulation time and the file size. In urban flood simulation environment roads and buildings are playing a big role for water flowing as barriers and as artificial channels. With considering all the factors like simulation time, data file size, and type of hazard assessment (Mark et al., 2004) suggested 1m and 5m DEM resolution are appropriate to represent the urban features with enough details. Roads and buildings are the main feature which can control the flood water behavior of an urban environment.

LiDAR (Light Detection and ranging or Air bone laser scanning) data provides the most accurate and detailed Digital Surface model at present. The RMS error for the LiDAR DEM is usually smaller than 0.15m for vertical and 0.50m for horizontal accuracy (Neelz and Pender). It is possible to represent building as either solid blocks which are not facilitating for water flow through the building or partially solid block with a high Manning's coefficient. Flood extent and depth maps are directly related to that representation type. (Tennakoon, 2004) shows that solid blocks and high rough surface can be changed the flow pattern of an urban environment. (El-Ashmawy, 2003) has given quantification for flood water height variation. If 10% of the building is represented as a solid block instead of high manning's, value the flood water height increase by 25cm. On the other hand it was suggested that if the area is less than 10% of building density there is no big influence for flood water height by the building structure.

(Tennakoon, 2004) defined the relation ship between pixel size and the application as follows.
Table 3.1: Reliable DEM resolution for different studies by Tennakoon, 2004.

Application	5m DTM	7.5m DTM	10m DTM
Urban flood hazard studies			useful
Studies of erosion and sedimentation		useful	
Small scale EIA studies	useful		
Regional level EIA studies			More than 10m

Representation of the topography and hydrological simulation results can be significantly affected by the DEM (Wechsler, 2003).

3.3 Different DEM interpolation techniques

Different interpolation techniques are used to develop continuous surface of grid based digital elevation model by filling the gaps between no data areas. This continuous surface is developed by using known elevation data in spatial extent; therefore those techniques can be defined as spatial interpolation technique. Moving average, trend surface, nearest neighbours, spline, and inverse distance weighting (IDW) are the most frequently used for spatial interpolation techniques. Same input elevation data set gives different out-puts for different interpolation techniques, therefore the DEM quality and accuracy depends on the interpolation technique. According to (Maune, 2002) IDW linearly weighted the value from near point; thus it is not valid for terrain interpolation. Spline

is good for interpolation the area with gentle slope. (Wise, 2007) found that the different DEMs produced different result in prediction of flow water, thus different interpolation techniques prodeused different flow result for same area. Since different interpolation algorithms result different surfaces; it is useful to select the proper interpolation technique depending on the application. (Podobnikar, 2002) suggested that the algorithms are the most flexible part of the digital elevation model and nobody can use the ideal, but the most suitable algorithm depending on the application.

The input data for the interpolation either contour or point height data, (Caetano et al.) studied the point interpolation methods and suggested that with lower point density the nearest neighbour method is less robust and Kriging and radial basis function are the most robust methods, and also showed that the correlation coefficient is to be different for different interpolation methods.
(Xie et al., 2003) for the contour interpolation, the input data should be in raster format, if the contour map is in vector format it has to be converted in to raster. After rasterizing the contour map discreet elevation values are entered in to grid cells not for all but for contour cells. During the process of interpolation new elevation data is assigned to all grid cells which did not have values before the interpolation. The interpolation cell size should be selected carefully, if the pixel size is as same as the adjacent contour interval this will reduced the multi value pixel. Multi value pixels occur in steep slopes and it is also difficult to avoid increasing of file size.

There are two main interpolation methods known as deterministic and probabilistic. In the deterministic method the elevation data is assigned relative to the surrounding elevation data. However in probabilistic method the assigning the elevation data is done not only considering the surrounding elevation but also consider the correlation to surrounding area. Geostatistical interpolation uses the autocorrelation function between two point pairs, and then assigns the values depending on the distance and the correlation to the predicted value and surrounding.

3.3.1 Kriging interpolation

The most popular point interpolation technique is Kriging and that method was developed by South African mining geologist called D.G. Krieg in 1963. Kriging is a spatial regression interpolation method for least square estimation of unstapled locations, which gives unrealistically smooth and continuous surface (Evane, 1993). Kriging can be controlled by the variogram models which always used to quantification of spatial variability of points.

Kriging is a probabilistic geostatistical point interpolation method which estimates the values for unknown places considering the surrounding point values. This value estimation is based on the spatial correlation which assumes that the point values of close proximity are more similar than the widely apart points. This correlation of values between two point pairs is introduced as covariance or variance and the semi-variance the degree of spatial dependence of pairs at a distance.

$$\gamma(h) = 1/2 \sum_{i=1}^{m} [(Z(xi + h) - Z(x))^2]$$

Where,
γ (h) = semi-variance at distance h
m = Number of point pairs with in h distance
Z(x) = elevation value at point i
Z(xi+h) = value at h distance from i

Experimental semi-variogram is used to analyze the elevation data set by assigning the weight for elevation sample. When the semi-variance is plot against the distance between the elevation data the result is the experimental variogram. Different mathematical functions such as Spherical, Exponential, and Gaussian can be used to model the spatial variation. Depending on that mathematical function "Nugget", "Sill" and "Range" should be determined. Semi-variance at the zero separation of the elevation data is known as the Nugget (Meijerink et al., 1994). This is the point that is not possible to explain by the spatial structure. Maximum semi variance of the semivariogram is represented by the sill. Also it represents the variability in the absence of spatial dependency. (Johnston et al., 2001) explained that the Ordinary Kriging is the most flexible Kriging method for point interpolation and it is also feasible for DEM generation.

3.3.2 TIN interpolation

Triangulated Irregular Network (TIN) is another popular GIS interpolation technique which uses the Delaunay algorithm for triangulation. This is a vector based continuous model that has the capability to integrate the line, point and polygon features for TIN generation. For point interpolation first creates circles around the point and then makes triangles using the intersection of nearest circles that are used to create the TIN surface which allows to represents the 3-dimentional plane in space. Gigged appearance and rough surface are the main disadvantage of TIN interpolation. TIN surface is good for complex environment, because the resolution can be selected according to the complexity of the terrain. (McGrath et al,) suggested that artefacts (pits and lands) will be resulted during TIN interpolation and those lead to the numerical instability in the model (Figure 3.1). Smoothing filters like Kernel are suitable to smooth the TIN surface until the numerical instability is overcome.

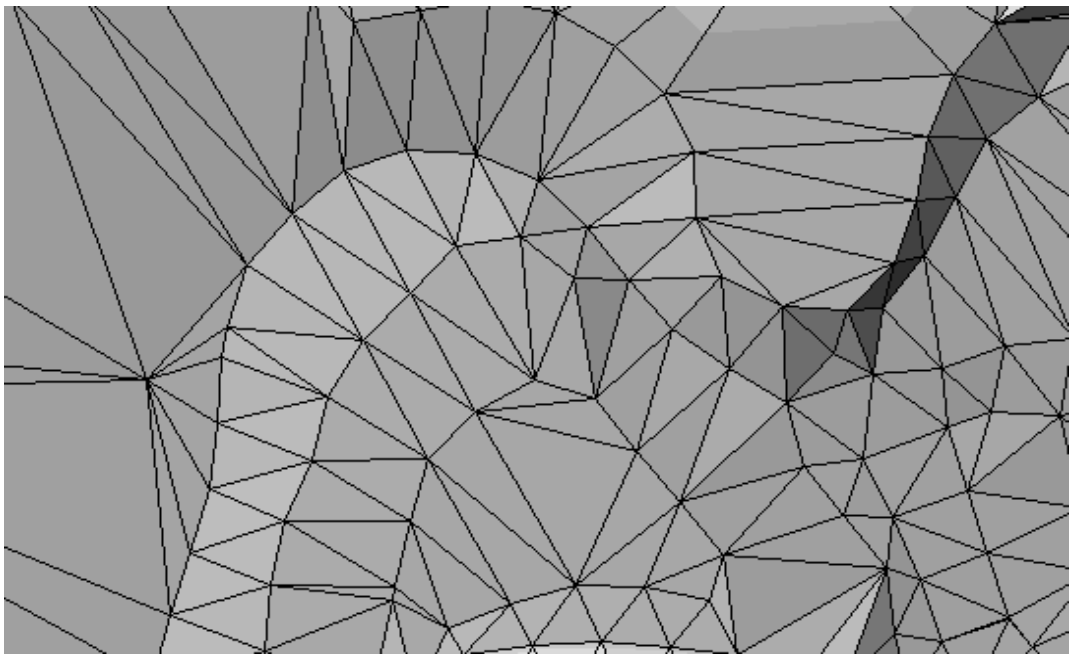


Figure 3.1: TIN interpolated surface with TIN mesh

3.4 DEM quality assessment

Even the latest technology can not guarantee an error free Geographical Information system (GIS) data base (Siska and Hung). This statement is applicable for the Digital elevation model as well. But the demand for high resolution and accurate DEMs are increasing in many sectors like telecommunication, urban planning, hydrological modelling …….. etc. Therefore new technological approaches are used to generate accurate DEMs. In the case of flood, runoff or erosion models it is important to asses the DEM quality to understand how the quality is being varied across the model. Different kind of data sources such as contour data, spot heights or any kind of data which can be collected through field verification are feasible for DEM generation. The most important part for DEM generation is to retain the quality of DEM which can also be changed with different interpolation techniques. Therefore assessing of the accuracy and quality of the DEM is more important. Accuracy is "*defined in terms of bias and precision, and is a global descriptor which is often based on the computed statistics of interest*" (Smith et al.). Different definition of the quality of available, (Butler and Lane, 1998) introduced the quality of DEM as a function of accuracy, reliability and presses of DEM generation. On the other hand suitability of DEM for a particular application is known as DEM quality.

Most of the research studies related to the DEM quality assessment have been done for the bear soil environment. Very few studies have been carried out for urban environment because of the complexity of the urban surface with verity of textures and shapes, frequent discontinuities, surface drainage (Smith, 2000). (Podobnikar, 2002) three steps were introduced to study the statistical method to asses the accuracy of a DEM as follows

1. Elimination of gross errors of reference points using reference DEM
2. statistical elevation for non-point sources with reference point
3. systematic error elimination from non point sources

Coordinates, observed during data collection may be containing blunders as well as systematic errors. Gross errors are recording failures, misinterpretation, careless observations, and undetected mismatch in image correlation. Systematic errors are system specific or procedure specific and occur in a predictable pattern and magnitude.

3.4.1 DEM validation methods

Main aim of the DEM validation is to quantification of the amount of error which has been taken palace during the different interpolation techniques. Many approaches have discussed by previous researches from different point of views. One of them is the split-sample validation and this method is capable of assessing the suitability and stability of the interpolation algorithm which has been used for DEM generation.

Through out the procedure of DEM generation from row data sources, different types of errors involves, which determine the DEM quality. Interior and exterior relative quality assessment methods have been defined by (Shan et al., 2008) for the DEM quality assessment. In this method accuracy index is calculated with row data points and the DEM quality is revealed in the local area. Then explore the relative errors are explored between different DEMs which have been acquired using different methods. Fundamental statistical parameters are used to calculate error of the DEMs as given below.

Bias is defined as the average deviation of the statistic (Wechsler, 2005)

$$\text{Bias} = \frac{\sum_{i=1}^{N} \hat{Y}_i - Y_i}{N}$$

Relative bias is defined as average percent deviation of the statistic from the original value

$$\text{R - Bias} = \frac{\sum_{i=1}^{N} \frac{\hat{Y}_i - Y_i}{Y_i}}{N}$$

Average Relative Absolute difference is expressed as average percent absolute deviation between the interpolated and true elevation data.

$$ARAD = \frac{\sum_{i=1}^{N}\left(\frac{\left|\hat{Y}_i - Y_i\right|}{Y_i}\right)}{N}$$

Root mean squire error is the most frequently used to describe the vertical accuracy of a DEM. Main assumption in the RMS error method is that the errors are normally distributed.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{Y}_i - Y_i\right)^2}{N}}$$

Relative RMSE is expressed as a percent and represents the standard variation of the estimator.

$$R - RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\frac{\hat{Y}_i - Y_i}{Y_i}\right)^2}{N}}$$

Log RMS

$$L - RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\ln\left(\frac{\hat{Y}_i}{Y_i}\right)\right)^2}{N}}$$

Where,

$Y_i^{\wedge}$ = estimator of the parameter

Y_i = Interpolated value

N = Number of points

3.5 Flood modeling

3.5.1 Introduction

Flood simulation approaches have advanced in to 1-dimenstional to 2-dimentional recent past. SOBEK is a powerful tool which was built by Delft hydraulics with enhanced capability of 1D2D flood simulation. This 2D dynamic model capability of SOBEK is given the facility to model the river bank overflow as a floodplain. On the other hand SOBEK allows integrating the 1D model with 2D. Topography of the flood plain is defined as a continuous 2D grid surface with elevation values. Finite difference method is used to calculate hydraulic parameters.

GIS was used by the large scale commercial organizations or survey department of the countries to map and organized the spatial data (Maidment, 1993). At present it has expanded due to the modern development of GIS techniques. According to (McDonnell, 1996) it is difficult to handle continues temporal changes alone with GIS environment itself. Therefore the most common way to model the temporal changes, GIS environment and Modelling environment use separately in a data base under the same file format or in file exchanging interface, which provides the flexibility for data visualization and analyzing. But there are some systems together with GIS and Modelling capabilities, but in such a system performance are less (eg SOBEK have some GIS functions, but those are not powerful).

3.5.2 1D flood modeling

The modelling approach is one dimensional; it means that the water (fluid) flow calculation is done based on the cross section of the channel by linking the channel cross section to pixels by assuming all the hydraulic parameters are uniform through out the channel (Borsanyai, 1998). Each cross-section can be used to calculate the flow velocity and water depth. 1D water flow is always parallel to the river channel. There is no any lateral flow. HECRAS, MIKE flood from DHI (Danish hydraulic Institute) are the widely used 1D hydrodynamic model. HECRAS is free ware but MIKE flood is a commercial.

On 1D modeling approach it assumes that water always flows along the free defined path and that approach has not a big impact on the flow until the water reach in to the bank full level. If the water overflows from river bank there is no any calculation for 2D domain, therefore it is problematic to understand the flood extent and depth on flat terrain. To overcome such problems coupled 1D2D models are used for model the floods nowadays.

3.5.3 2D flood modeling

2D flood models are used to be modeled the lateral water flow during the flood events. This kind of model has the capability to model not only the spatial extent and water depth, but also the water velocity and impulse according to the free defined timeframe. As a result 2D models are more popular among the hydrologists all over the World. Continuous terrain representation is the main requirement for model the 2D water flow. Hence topographical data, boundary condition and hydraulic parameters are considered to be the main inputs for 2D flood modeling. DSM are used to represent the terrain of the study area for the flood modeling, but selection of the best pixel resolution is also equally important task in 2D flood modeling. The pixel size should be selected in such that representation of the smallest feature of the study area is ensured, while representing uniform topographical variation. Increasing of the DSM resolution is the cause for increase the computation time. Therefore 2m to 10m pixel resolution is commonly used for flood simulation studies (Tennakoon, 2004).

3.5.4 Boundary conditions

Boundary conditions are represented by the nodes which can be used to define the geographical location to separate the real world from the model. Those nodes are used to assign the real world hydraulic parameters to the model either way of series of hydraulic parameter or constant value. However the uncertainty of upstream boundary can have grater impact on the model result (Pappenberger et al., 2006). On the other hand water can be entered in to the system as well as it can be flow away from the system. In general the upstream boundary used to input the upstream river water in to the system and constant or alternating water level boundary used in downstream.

3.5.5 Surface roughness

Bottom friction is common in places where the fluid is being flow. The term surface roughness is used to represent the bottom friction of any kind of underline materials. Different equations like Manning's, Chezy's, Hazen-Williams and Darcy-Weisbach are used to define the surface roughness (Haestad Methods, 2002). Surface roughness is always related to the underline materials, especially in to the land use type. The landuse information can be translated in to surface roughness information. The remote sensing data is provides better path for land use classification.
Surface roughness is important to calculate the water flow velocity. It is incorporated as Manning's roughness value.

$$V = k/n\ R^{2/3}\ S^{1/2}$$

Where,

V = mean velocity of fluid flow (m/s, ft/s)
K = 1.49 for U.S customary units, 1.00 for SI units
R = hydraulic radius (m, ft)
S = friction slope (m/m, ft/ft)
n = Manning's roughness value

In an urban environment manning's surface roughness coefficient plays the major role for water flow; due to many different landuse types. Vegetation covers, buildings, roads, underground tunnels, embankment, obstructions are the main objects to be considered when defining the surface roughness.

Table 3.2 Manning's coefficient

Landuse type	Manning's coefficient
Water (river)	0.030
Rice field	0.050
Vegetable	0.045
Build up area	0.035
Mixed tree	0.030
Mature forest	0.040
Grass and shrubs	0.025
Bear soil	0.020
Roads	0.020

3.6 Model calibration and validation

Subsequent adjustment of model parameters so that minimize the deviation arising between the observed and modeled data is expressed as model calibration. Surface runoff model can be calibrated by adjusting inputs of hydrological parameters until the computed hydrograph close enough to observed runoff data. The main approach to calibrate an urban flood models is to match the flood extent and flood depth (Mark et al., 2004). In an urban environment Manning's coefficient, DSM and building structures are important to calibrate a flood model (Tennakoon, 2004). For 1D2D flood simulation models manning's friction values should be defined for both river channel and floodplain. The manning's friction value is not only parameter which could be disturb the surface 2D water flow in urban flood plains, But a controlling factor for bulk discharge. Therefore more attention to be given while designing main channel structures specially in urban environment. For instance buildings influence water retention, but paved roads yields increases in

flow discharge by reduced surface roughness. Model validation is a process which can be used to test the calibrated model data against measured data, using model parameters which were derived during the model calibration. Flood model validation process requires comprehensive and detailed field data like flood extent, depth and water discharge in many different locations (Mark et al., 2004).

Chapter 4 DSM generation and Flood frequency analysis

4.1 DSM generation

The Digital Surface Model (DSM) is the key input layer for 1D2D hydrodynamic modelling, more accurate and the better the quality of the DSM is acting as a main model requirement. Flood propagation on an urban environment is quite complicated, because manmade structures are acting as barriers for the water flow. On the other hand roads are much easier for water flowing in urban environment. Therefore accurate terrain representation is an important requirement for urban flood simulation. There are many different ways to represent elevation data in GIS environment. According to available tools for flood simulation, DSM is the most useful elevation data representation. However accurate DEM generation is a challenge under the less data environment.

4.1.1 Data sources for DSM generation

Available elevation data in the area are:

1. 1:10,000 contour data (contour interval = 10m)
2. Measured elevation data points (Field work)
3. Digitized line with same elevation (Field work)
4. Building heights

1:10,000 scale contour data are available in digital format. Also some small elevation lines were generated using the field observations. The high resolution Google Earth image was georeferenced and the contour map (1:10,000 scales) were overlaid and zoomed in to 1:4000 scales so that all the field points are clearly visible. Then with the help of ground truth information, the new contour lines were generated on that image (chapter 4.1.2).

4.1.2 Field work data

With the help of the data need assessment (Chapter 2.2) field work planed to collect the elevation data to fill the gaps of DEM generation. Total 91 elevation data points and new contour lines were measured during field data collection. Figure 4.1 shows the distribution of the measured elevation data points throughout the field work. The observed point location was selected based on the terrain topography of the area. Point density is high in irregular topography and less in flat terrain. New contour lines were established with the help of topographical features such as ponds, slope cuts, railway cuts …..etc.

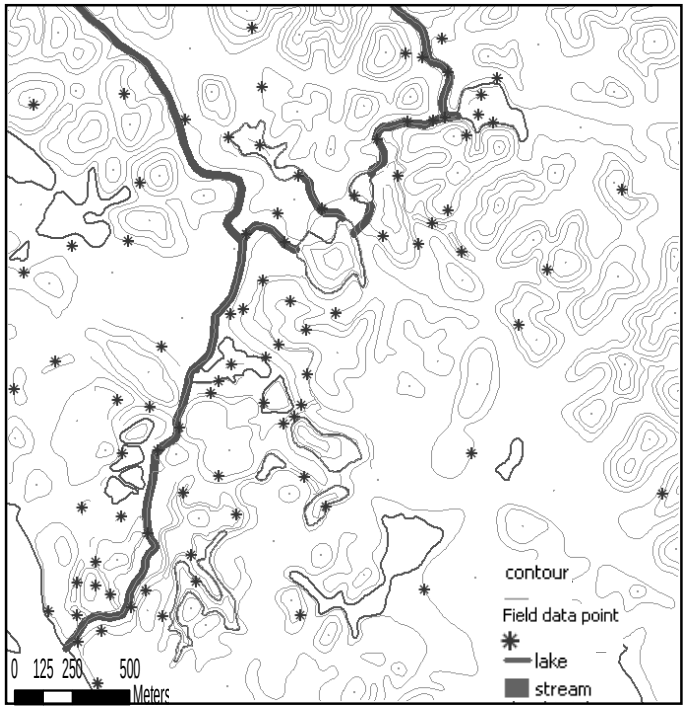


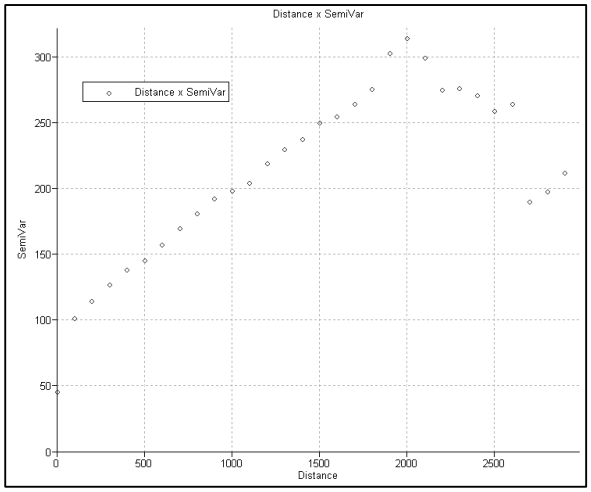
Figure 4.1: Distribution of measured elevation points with contours.

4.1.3 Method of DSM generation

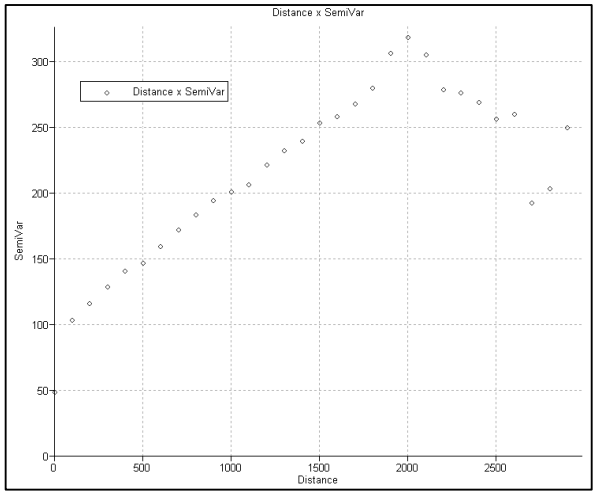
To create a DSM, a Digital Elevation Model (DEM) was first generated and then the man made structures was added. Elevation data preparation, interpolation and accuracy assessment were the three steps for the DEM generation.

In the beginning, detailed investigation of existing elevation data was carried out. Then field data collection was planned according to the data requirement for DEM generation (Chapter 2.2.1). With the help collected and available data, the data interpolation was carried out to DEM generation. To perform any interpolation techniques all elevation data should be in either form of point or line. Most of the field observed data is in the form of point; therefore the most suitable approach is to convert all the available line data in to the point data with different spacing. This conversion was done with line to point conversion with 10m, 20m, 50m, and 100m intervals. Then all the elevation data sets were combined by checking for all overlapping points. During the data combining, over lapping points within 10m radius were removed. This data combination process was used to establish the most reliable terrain elevation of the study area. Because of that data combination, the original contour elevation data variation was increased with in a short range. On the other hand complex topography of the area indicates the elevation changes within short distance. Semi-variogram analysis is a possible technique to understand those problems. If the variation of elevation values is high with in a short distance, the high nugget value can be seen in the semi-

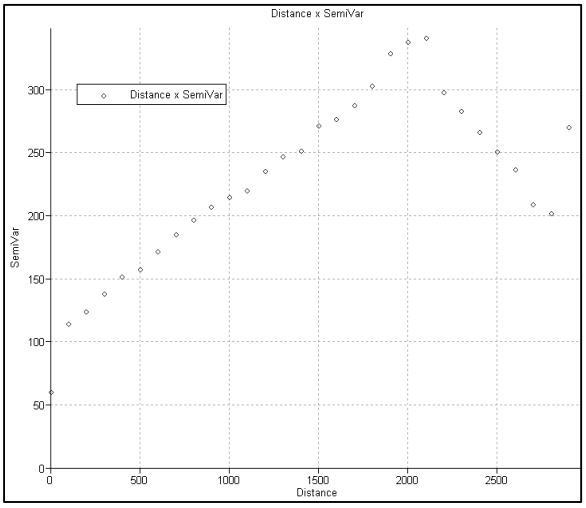
variogram. There are four different point data sets prepared after combining field measured data points with point to line conversion 10m, 20m, 50m and 100m. Therefore before the data combination, semi-variograme of the original contour data was studied to understand the spatial dependency of the data.



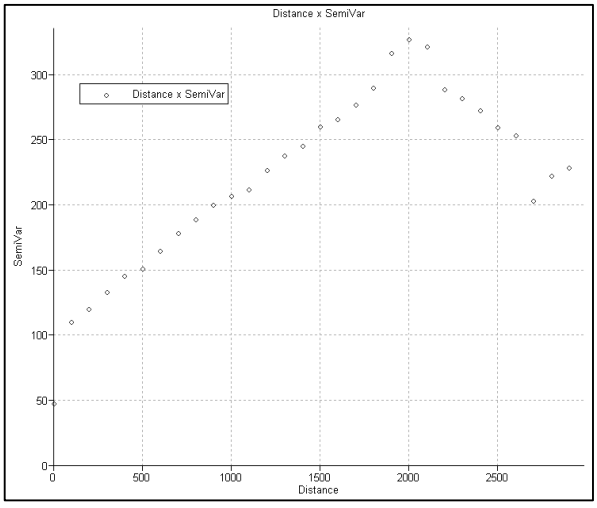
Only original contour line to pint conversion with 10m spacing



Only original contour line to pint conversion with 20m spacing



Only original contour line to pint conversion with 50m spacing



Only original contour line to pint conversion with 100m spacing

Figure 4.2: Semi-Variogram before update the contour map, Lag spacing 100m.

The above (Figure 4.2) semi-variograms which were created only with the contour line to point conversion with 10m, 20m, 50m and 100m spacing indicate high nugget values varying from 45 to 60 with 100m lag spacing. This means that the elevation variation between shorter distances is high. However by combining original contour data with the collected elevation data during the field work the elevation variation between shorter distances was minimized and nugget value was observed to be reduced (Table 4.1).

Table 4.1: Changes of nugget and sill with different lag spacing and point to line conversion

Data set	Lag	Nugget	Sill Range	Number of points	Number of Points	Model
10m contour to point conversion	**10**	**0.2**	**230**	**1000**	**11610**	**Spherical**
20m contour to point conversion	10	0.26	235	1000	6004	Spherical
50m contour to point conversion	10	0.26	242.5	1000	2647	Spherical
100m contour to point conversion	10	0.28	245.5	1000	1519	Spherical
Data set	Lag	Nugget	Sill	Range	Number of points	Model
10m contour to point conversion	20	1.25	332	2000	11610	Spherical
20m contour to point conversion	20	7.92	337	2000	6004	Spherical
50m contour to point conversion	20	4.31	350	2000	2647	Spherical
100m contour to point conversion	20	3.04	370	2000	1519	Spherical
Data set	Lag	Nugget	Sill	Range	Number of points	Model
10m contour to point conversion	50	16.8	338	2900	11610	Spherical
20m contour to point conversion	50	17	337	2900	6004	Spherical
50m contour to point conversion	50	26.4	350	2900	2647	Spherical
100m contour to point conversion	50	22.12	365	2900	1519	Spherical
Data set	Lag	Nugget	Sill	Range	Number of points	Model
10m contour to point conversion	100	36.74	335	2900	11610	Spherical
20m contour to point conversion	100	38.7	336	2900	6004	Spherical
50m contour to point conversion	100	45	361	2900	2647	Spherical
100m contour to point conversion	100	45	361	2900	1519	Spherical

The Yen Bai city region is quite complicated in topographically. Also the lag spacing is given as 10m because of the complexity of the area. This complex topography may be the reason for the deviation of the semi-variogram model from the Gausian to spherical. This semi-variogram is fitted at 1000m range with low spatial dependency (Figure 4.3). The spherical model is the only model that can be used to fit the available elevation data points. On the other hand different lag spacing was studied with semi-variogram to get the understanding about the elevation data variation through the nugget effect of the elevation data (table 4.1). Finally the combined point cloud which was prepared from field elevation points and 10m line to point conversion was given the smallest nugget value with 10m lag spacing. Consequently, it was selected as the best elevation data set for DEM interpolation.

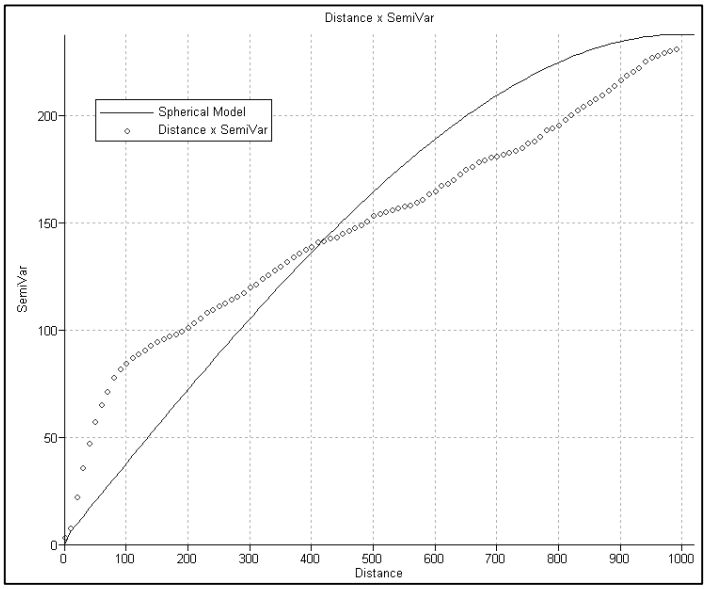


Figure 4.3: Semi-variogram

4.1.4 Terrain interpolation

After the selection of the best point data set, the main focus was on the generation of Digital Elevation Model (DEM) through the point interpolation. The interpolation should be done to represent the surface topography as accurate as possible while minimizing the errors. Kriging, polynomial trend surface and TIN based grid are the main available DEM generation methods and the best DEM was selected after considering the error distribution different outputs (chapter 4.1.5).

Fitting up of a mathematical function for elevation data is the basic approach for the polynomial interpolation. Higher order mathematical function is used to generate a complex terrain with large number of data points. Low order polynomial mathematical function is used to produce a smooth surface with lesser number of elevation data points. Lower and higher order interpolations are used for smooth and complex terrain interpolation respectively.

Triangulated Irregular Network (TIN) based grid interpolation is a vector based interpolation technique and depending on the complexity of the terrain sample point density can be defined. Lesser number of elevation data points is adequate for smooth or less complex terrain and dense elevation data adequate for complex terrain. However TIN interpolation is an exact interpolation method which means that the interpolated surface always following on the input elevation point. Both TIN and polynomial interpolation mainly consider the distance between two elevation points. But in Kriging interpolation distance as well as the spatial correlation are taken in to consideration. This method is more sensitive to the errors in the input data set.

4.1.5 Error distribution of Kriging interpolation

With the semi-variogram analysis the point data with the lowest nugget value was selected for the interpolation (Table 4.1). But that data set is not the best due to the shape of the Semi-variogram model is not Gausian and it is not best fit with data range because of the complexity of the terrain. Therefore the error distribution with different interpolation was analyzed with reference elevation point (91 reference point) to select the best DEM for further flood modelling. Figure 4.4 shows the error distribution of Kriging interpolation with different line to point conversions. The contour interpolation having the highest number of points produced the highest error as 42 points have shown error of more than 1.5m (table 4.2). The best result is obtained from 10m line to point conversion data set with Kriging interpolation. 60 points have given less than 0.5m error, and it is about 65% of the total. Therefore this point data set was good enough for the DEM generation.

Table 4.2: Number of points with error

	Number of points with error		
Interpolation	>1.5m	1.5-0.5m	<0.5m
contour	42	21	27
Kriging_10m	**5**	**28**	**60**
Kriging_20m	7	31	55
Kriging_50m	7	31	55
Kriging_100m	33	8	50

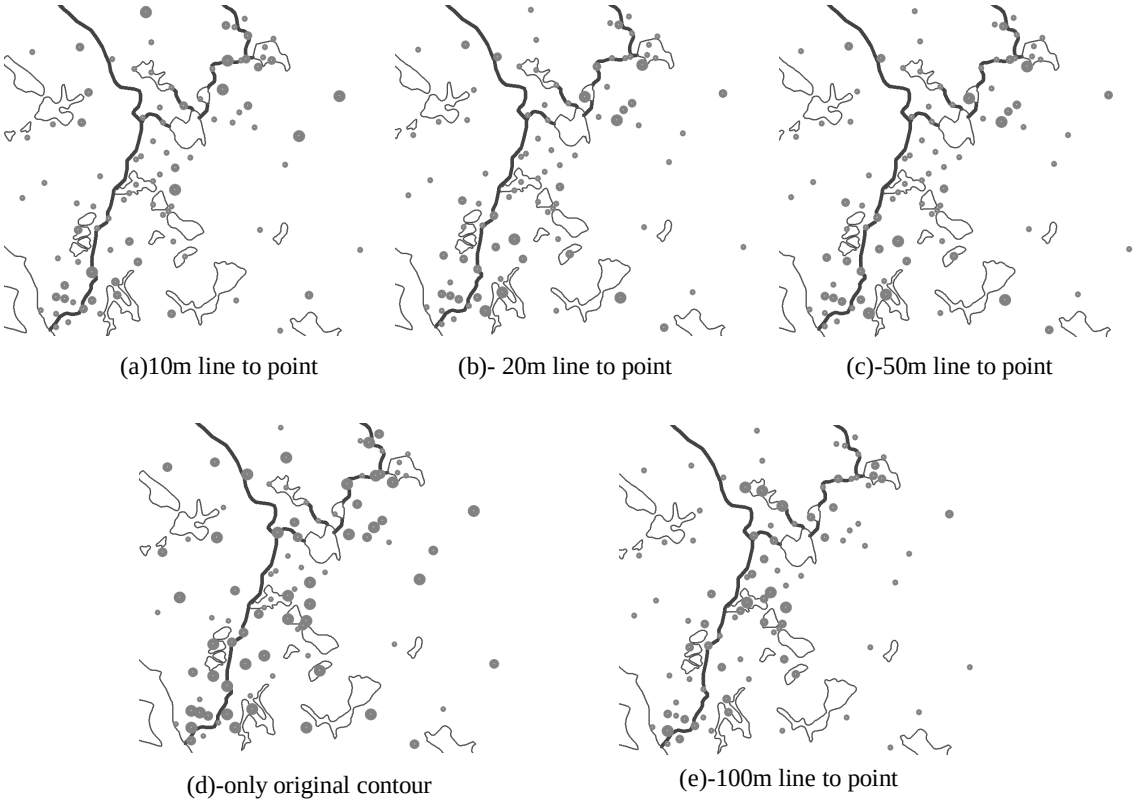


Figure 4.4: Error distribution of Kriging interpolation with different line to point conversion and contour interpolation

The statistical parameters such as Bias, R-Bias, ARAD, RMSE, R-RMS and L-RMSE were used to express the quality of a DEM (Chapter 3.4.1). According to the table 4.3 when the spacing of line to point conversion was increased R-Bias, ARDA, RMSE, R-RMSE and L-RMSE have increased. But Bias shows slightly different changes; it shows that the average deviation of statistic is better in 50m line to point conversion. However, when consider the whole statistics 10m line to point conversion data interpolation has given the best overall performance.

Table 4.3: Error of Kriging interpolation

Line to point conversion	Bias	R-Bias	ARAD	**RMSE**	R-RMSE	L-RMSE
10m	**-0.12**	**-0.004**	**0.015**	**0.861**	**0.023**	**0.024**
20m	-0.13	-0.004	0.016	**1.012**	0.028	0.029
50m	-0.10	-0.005	0.027	**1.469**	0.042	0.044
100m	-0.11	-0.007	0.044	**2.277**	0.066	0.068

4.1.6 Error distribution of polynomial interpolation

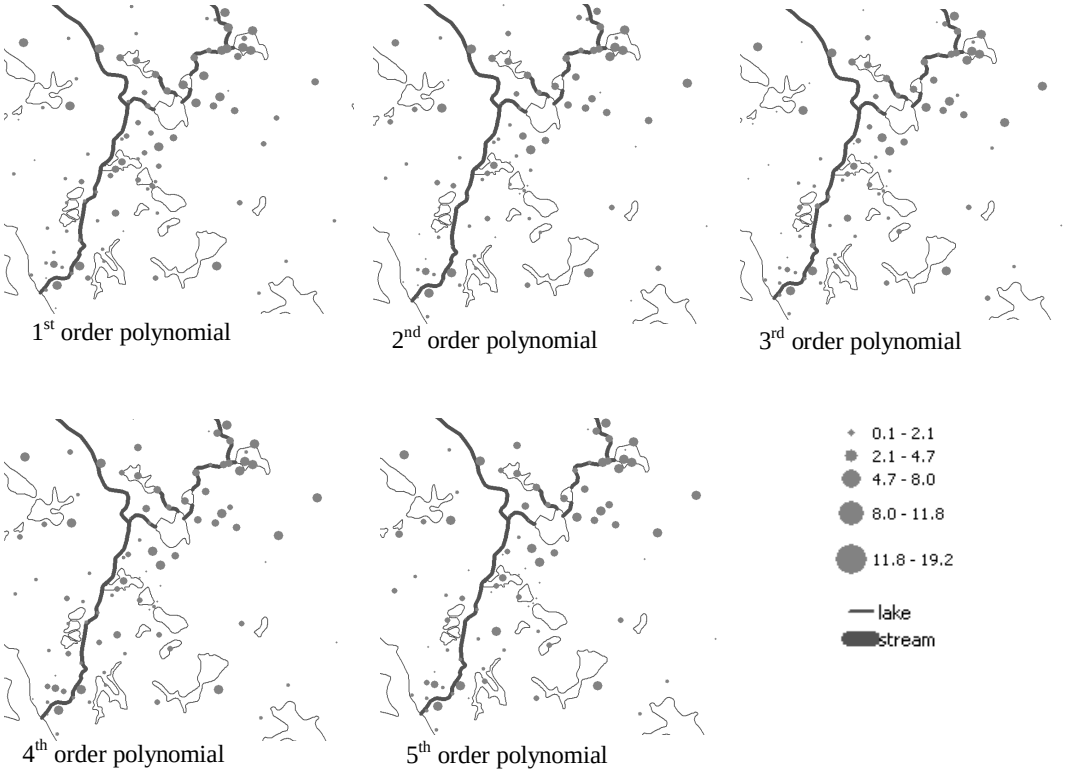


Figure 4.5: Error distribution of different order of polynomial interpolation

Table 4.4: Number of points with error

	Number of points with error		
Interpolation	>1.5m	1.5-0.5m	<0.5m
1st	83	5	3
2nd	80	7	4
3rd	79	8	4
4th	76	10	5
5th	**70**	**10**	**11**

The polynomial interpolation with 10m line to point conversion point data set depicts the lowest accuracy of the final DEM (Table 4.4). When the order of the polynomial interpolation was increased, the accuracy of DEM was observed to be increased but is much less compared to the Kriging interpolation (Table 4.2). The number of reference points having the error les than 0.5m is about 10% of the total number of points in polynomial interpolation.

Table 4.5: Error of polynomial interpolation

Order of polynomial	Bias	R-Bias	ARAD	**RMSE**	R-RMSE	L-RMSE
1st	-4.66	-0.09	0.25	**12.90**	0.30	0.34
2nd	-3.42	-0.16	0.27	**10.91**	0.34	0.29
3rd	-3.00	-0.07	0.20	**10.05**	0.24	0.279
4th	-2.48	-0.06	0.19	**9.44**	0.23	0.25
5th	-1.99	-0.05	0.19	**8.98**	0.23	0.24

The error distribution of the DEM which created by the polynomial distribution is much higher compared to the Kriging interpolation and TIN interpolation. The best value for polynomial interpolation was obtained for the 5th order, but appears much higher. The DEM which was created with the TIN also produced less error compared to 5th order polynomial interpolation.

4.1.7 Error distribution of TIN interpolation

Table 4.6: Number of points with error

	Number of points with error		
Interpolation TIN	>1.5m	1.5-0.5m	<0.5m
contour	48	19	21
10m	**25**	**14**	**52**
20m	34	11	46
50m	43	9	39
100m	55	8	28

The DEM generated from the TIN interpolation gives a higher accuracy compared to the polynomial interpolation but it is less than the Kriging interpolation (table 4.6). 10m line to point conversion elevation is still giving the highest accuracy which is 57%. It is still les compared to the 65% given by the Kriging interpolation.

Table 4.7: Error of TIN interpolation

Line to point conversion	Bias	R-Bias	ARAD	**RMSE**	R-RMSE	L-RMSE
10m	-0.20	-0.006	0.019	**1.352**	0.032	0.033
20m	-0.26	-0.006	0.026	**2.321**	0.049	0.047
50m	-0.18	-0.008	0.032	**2.861**	0.081	0.086
100m	-0.24	-0.009	0.056	**3.357**	0.102	0.106
Original contour	-0.86	-0.060	0.213	**6.256**	0.265	0.272

TIN based interpolated DEM also given the less error compare to the polynomial interpolation but it is little bit higher compare to Kriging errors (Table 4.7).

4.1.8 Visual interpretation of DEM quality

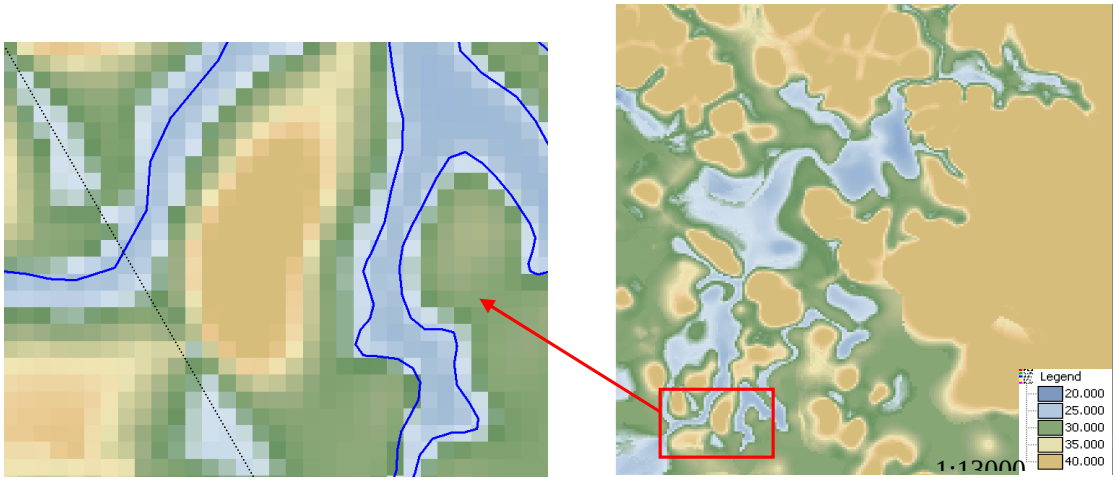


Figure 4.6: DEM generated through Kriging interpolation

The DEM which was generated from the Kriging interpolation has produced the better topographical representation. Most of the surface features like river, lake can be observed on the DEM. Figure 4.6 clearly shows the railway line, ponds and river in the DEM. Finally the DEM which was created through the Kriging interpolation with 10m line to point conversion was found to be the best DEM for flood simulation under this condition.

4.1.9 DSM

With the field experience it was understood that the Building density of the study area is little bit less compare to large city. The area can be classified as a semi-urbanized area. Houses and roads can be identified as the main manmade structures of the area. Therefore the buildings were superimposed on the DEM to create a DSM. The road network was not included in the DSM but it will be introduced in the flood model with low Manning's value.

In general, the widths of the road in Yen Bai city region varies from 5 to 15m, but in the flood simulation area it is 5 to 10m width. Either less than 5m width roads or 10m width roads which are not running parallel to the DSM grid, therefore it is not easy to represent in the 10m DSM without having a discontinuous road pixels. According to (Thennakoon, 2004) the road width should be greater than or equal to 2 pixel resolution to make sure that the continuous grid connected side by side. The minimum road width that can be used for the flood simulation is depends on the pixel size of the DSM. For the small roads it should have to represent at the minimum width by buffering the

existing road layer. Small pixel size (2m, 5m) represents the road network better in an urban environment. But, when the pixel size is small the number of pixels which has to be used for flood simulation is going to be increased and either limiting study area or increasing the simulation time drastically.

On the other hand, there is an impact on the representation of the building foot print map with respect to different pixel sizes. When the 10m pixel DSM was used the minimum area of the building should be equal or more than $100m^2$. Representation ability of a building on top of digital terrain model of any place is also increased with the reduction of pixel size, if pixel size is 5m, it is possible to represent a building of $25m^2$. There was no any building having an area less than $25m^2$. But when consider the available terrain information and flood simulation time, the DEM resolution for flood model was 10m. Therefore 10m pixel size was also used for building representation. Due to this representation some of the building details were removed from the final DSM.

4.2 Flood frequency analysis

4.2.1 Introduction

Extreme flood events are experienced in many parts of the world and the magnitude of an extreme event is inversely related to its frequency of occurrence (Gabriel Parodi, 2004). It means moderate events occur more frequently than the severe events. To understand the relationship between magnitude of the extreme events with frequency of occurrence, flood frequency analysis is extensively used.

Return period is widely used in flood frequency analysis and according to the (Kidson and Richards, 2005), the return period is the average recurrence interval of a given magnitude and it can be explain as follows,

Return period = number of years of river records+1/rnak of a given flood

Table 4.8 Red River discharge data from 1960 to 2009

year	1960	1970	1980	1990	2000
0	5600	4262	6739	5756	3373
1	4605	13402	2426	6129	6796
2	3476	6335	4060	8756	6601
3	3840	7049	4954	4665	4143
4	4567	5782	3532	8157	4396
5	2064	4302	4861	8764	6994
6	6275	3961	7931	11632	4887
7	4182	4187	8634	5060	5655
8	10071	5553	4020	4908	10761
9	5109	7452	8290	5678	3476

Table 4.8 shows 50 years flood data of Red River from 1960 to 2009 in Yen Bai city. There are four occurrences in exceeding discharge 10000m^3s^{-1}. 50 total numbers of events were recorded. Then return period of the 10,000m^3s^{-1} annual maximum discharges exceeded 4 times. Therefore return period is equal to 50/4=12.5 years.

All the Flood frequency analysis is done with interpolation and fitting up of a model with existing flood parameters such as discharge or water heights. Kidson and Richards (2005) has summarized the flood frequency analysis as form of a risk analysis. Flood frequency analysis is a statistical approach for flood hydrological parameters. The first approach is extreme value analysis, using reliable past recorded data to establish a probability distribution which is used to describe the flooding process. The second approach is used when exact flood information is not available. Here hydrological model is used to establish the flood parameters based on the meteorological data such as rain fall.

Selections of data, suitability of model and parameter estimation are the main steps for flood frequency model fitting. Data should be either maximum flood parameters (discharge, water height) or series of data over threshold value. It can be continues records of gauged or discrete historical data of a hydrological model. Limited numbers of models are available for flood frequency analysis but there are many number of variant of them.

4.2.2 Extreme value distribution

Depending on the application criteria the statistical distribution is selective to certain extent. As an example log-Pearson distribution is widely used in hydrological parameters. In extreme value distribution, Gumbel distribution and lognormal distribution are commonly used. 2-Parameter functions are used in Gumbel distributions for statistical fitting. Two parameters represent the

mean and variance of the sample population. Mean, variance and skewness are 3-parameters for log-normal distribution. (Kidson and Richards, 2005)

4.3.3 Gumbel Extreme value Distribution

Gumbel extreme value distribution can be used to construct a relation between the probabilities of the occurrence of event with its return period. Generally hydrological event's probability distribution is right skewed which gives the concave curves on linear-normal probability paper. Straighten out of right skewed cumulative distribution by extreme value transformation or double exponential transform is called the Gumbel transformation.

The five steps approach to create a gamble distribution was discussed (Gabriel Parodi, 2004)

- Rank the yearly maximum flood values from low to high
- calculate the left side probability for each observation

$$P_L=R/N+1 \quad \text{.....(4.1)}$$

- determination of the return period (T)

$$T=1/P_R=1/1-P_L \quad \text{.....(4.2)}$$

- determination of plotting position of each observation

$$Y=-\ln(-\ln P_L) \quad \text{.....(4.3)}$$

- construction of the probability graph

Where,

P_L=Left side probability
R=Rank
N=Number of observation
P_R=Right side probability
P_L=Left side probability

Y axis of the Gumbel probability distribution is related to the return period (T_r). Calculation of return period is not an easy task but the extra table which have been developed by equation 4.4and 4.5, can be used to calculate the return period. Appendix A is given the value range for the calculation of return period.

$$P_L = e^{-e(-y)} \quad \text{..... (4.4)}$$

$$T = 1/P_R = 1/1-P_L \quad \text{..... (4.5)}$$

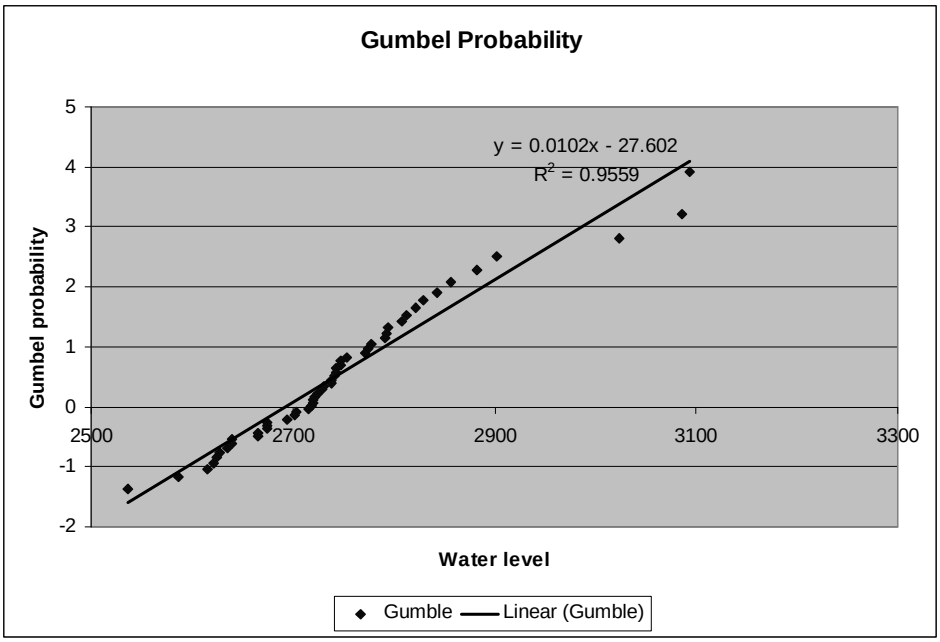


Figure 4.7: Gumbel distribution

Gumbel probability was calculated through the procedure discussed in chapter 4.3.3, and the data shows in Appendix B. Values of the y-axis are directly related to the probability of the return period. Then corresponding return period was calculated using the equation (4.4).

Table 4.9: Water level changes with different return period

return period	Water level above datum (m)
5	28.63
10	29.21
25	30.20
50	30.88
100	31.57

From 1960 to 2009 Red river water level changes and discharge measurements are available. Then to describe the relationship between discharges and water levels, a rating curve was established. As the downstream boundary condition for the flood modelling the water heights were used (chapter 5.2). The relation between the water height and the discharge is given the positive linear correlation 0.68 (Figure 4.8).

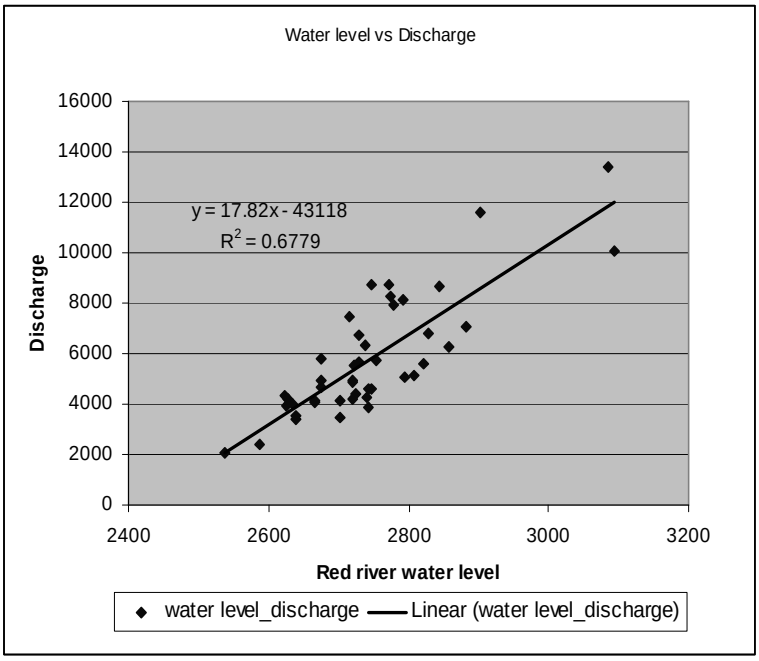


Figure 4.8: Water level discharge relation ship of Red river

To estimate the frequency of occurrence of water levels above a certain level, a Gumbel time series analysis was performed.

Chapter 5 Flood modeling

5.1 Introduction

Different flood scenarios were simulated with SOBEK 1D2D model for the study area. SOBEK 1D2D model integrates one dimensional channel flow with two dimensional overland flow. The overall methodology for flood modelling is given by (Figure 5.1).

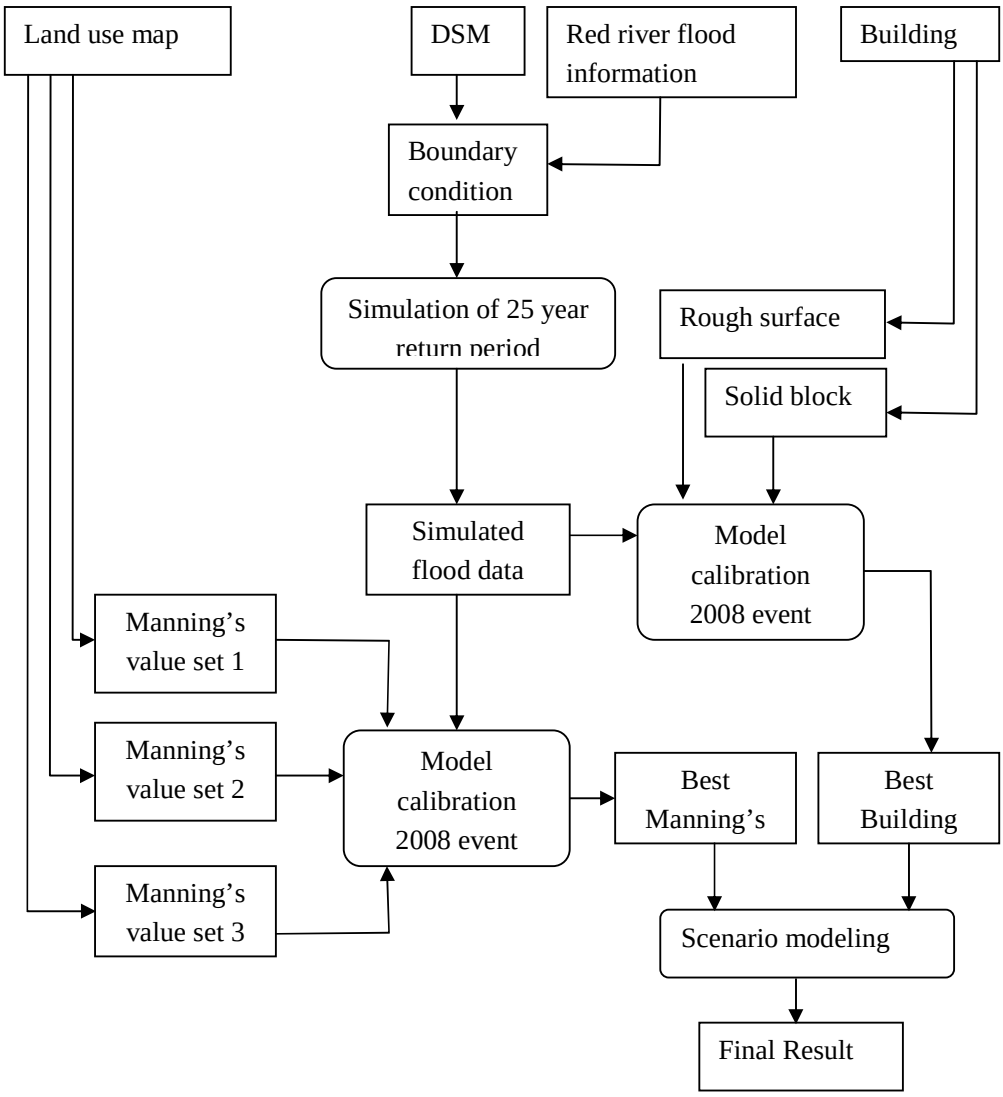


Figure 5.1: Methodology for flood modelling

Two dimensional (2D) elevation grid is used to represent the topography of the study area, the width of the finest river branch of the study area is grater than 10m, and therefore 10m resolution is sufficient to represent the river network with details. Then the boundary conditions were defined for the flood simulation based on the field observations. Then simulated and field observed flood data of super typhoon 2008 was used for model calibration and validation. Model calibration was done by step by step approaches because the simulated flood depths are influenced by the building structures and surface roughness of the underline land use. However two sets of building structures and 3 sets of manning's coefficient was used for model calibration. This was the methodology to select the final manning's coefficient and best building structures for the final flood scenario simulation.

5.2 Boundary condition

The two upstream boundaries and one downstream boundary were used for model schematization. Downstream data are based on daily water level reading of the Red River from 1960 to 2009 (Figure 5.2). The water from upstream watershed also continuously flows in to the system but measured discharge data is not available. Then the upstream boundary condition was defined with constant discharge which assumes to be varied with return period (Table 5.1). That estimation was done based on the measured diver data.

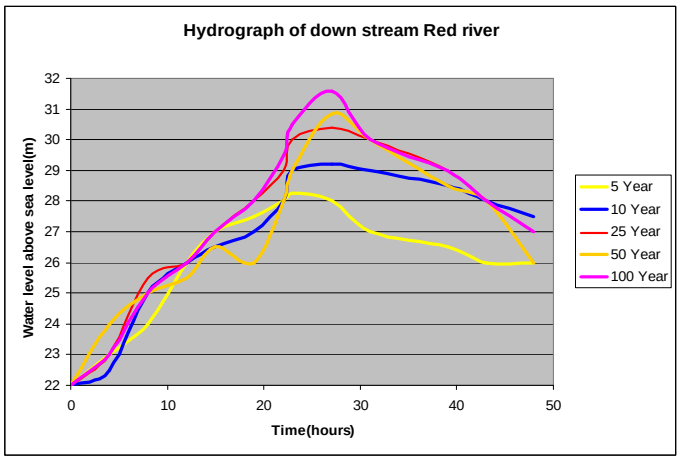


Figure 5.2: Boundary condition for downstream Red River

Table 5.1: Upstream boundary data for flood simulation (figure 5.2)

Return Period	Upstream Discharge	
	Boundary 1	Boundary 2
5	5m3/s	5m3/s
10	10m3/s	7m3/s
25	15m3/s	10m3/s
50	20m3/s	15m3/s
100	25m3/s	20m3/s

5.3 Surface roughness

Manning coefficient is related to the corresponding land uses. Therefore suitable manning's coefficient ware derived from the literature (http://www.flowsizer.com/content/mannings-n-values-literature). The landuse map of the study area is presented in figure 5.3.

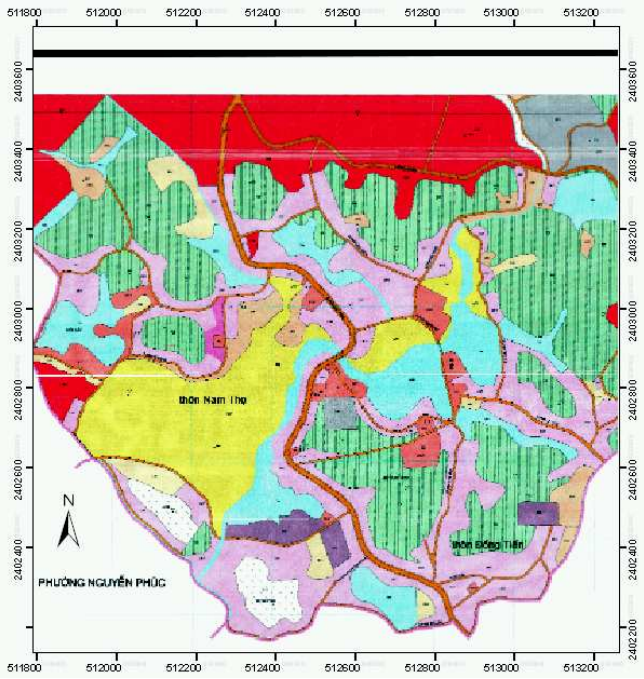


Figure 5.3: Landuse map of study area

Three different sets of Manning's coefficient values were used for model calibration process and are presented in table 5.2. The set 1 derived from literature and then the set 2 derived from set 1-0.005 and the set 3 derived from set1+0.005.

Table 5.2: Different sets of Manning's coefficient used for model calibration

Landuse type	Set 1	Set 2	Set 3
Water (river)	0.030	0.025	0.035
Rice field	0.040	0.035	0.045
Vegetable	0.035	0.030	0.040
Build up area	0.045	0.040	0.055
Mixed tree	0.030	0.025	0.035
Mature forest	0.050	0.045	0.055
Grass and shrubs	0.025	0.020	0.030
Bear soil	0.020	0.015	0.025
Roads	0.015	0.10	0.020

5.4 Building representation

Theoretically buildings can be represent as solid block, partial solid block which allows water to enter in to the building and rough surface. In this study buildings were represented as solid block and rough surface. The partially solid block for the building representation was introduced in such a way that the first 40cm level of the building represent as solid block which not allow entering water in to the pixel, then rest represent with manning coefficient. The simulation was done with different approaches and results were evaluated and compared with field observed data to determine the best building representation.

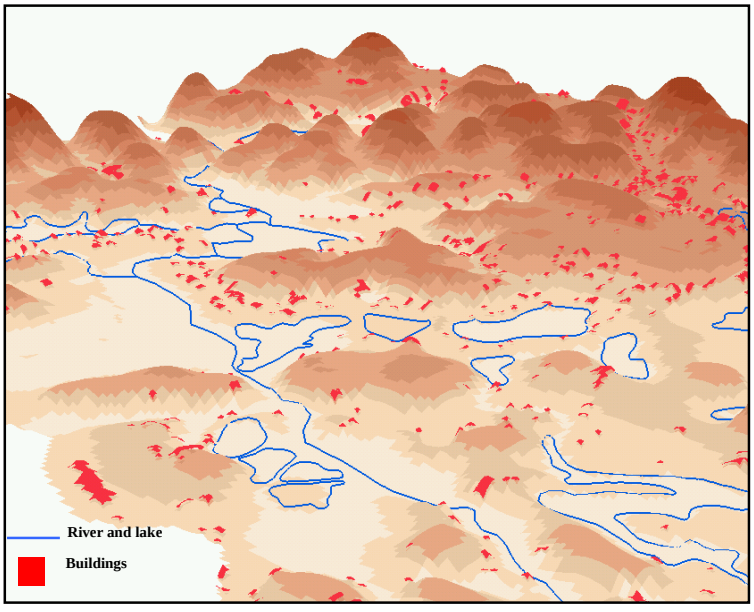


Figure 5.4: Building structure of solid block

5.5 Model Schematization

Network editor in SOBEK is used to schematize the model components on the topographical layer. Schematization and attribute editor are the main two editing modes in NETTER. River network, river cross-sections, boundary nodes, 2D grid, connection nodes, historic points, calculation nodes are the major components which can be defined through the process of schematization (Figure 5.6). The attribute editor is used to edit their attributes.

Starting point of the schematization river network was imported as GIS layer. Then the DSM used to introduce the terrain information for flood simulation model in ASCII format. DSM resolution is always influenced the model computation time which means high resolution DSM needs large computation time compared to coarser resolution. Application type, details of flood simulation results have to be considered to select the suitable DSM resolution. Terrain values can be checked and changed manually with the help of grid cell editor.

Surface roughness is an important parameter for flood simulation models and SOBEK is given the facility to represent the surface roughness either single value for whole area or one value for each pixel. River bed elevation is excluded from the DSM and represented as a river cross-section in the 1D flow. 1D flow model is representing the water flowing system along the river channel. Flow cross-sections (Figure 5.5), bed level, surface level and roughness are the main parameters which determine the flow characteristics.

The river network for the 1D flow is defined with reaches and connection nodes. Then the edit reach vector is used to reshape the straight river line. 1D flow is representing the river flow in the system and its characteristics determine by the cross-section, roughness, surface level and bed level of the channel. The river cross-section for 1D was defined with the type of trapezium (Figure 5.5) through the field measured values. 1D and 2D boundary conditions are introduced during schematization on the edge of the river network as upstream and down stream. 2D boundary condition is only available in downstream edge of the river. The study area represents several ponds therefore initial water levels were introduced during schematization. Finally historic flood stations were placed at the selected areas which are useful to get the water level and water velocity.

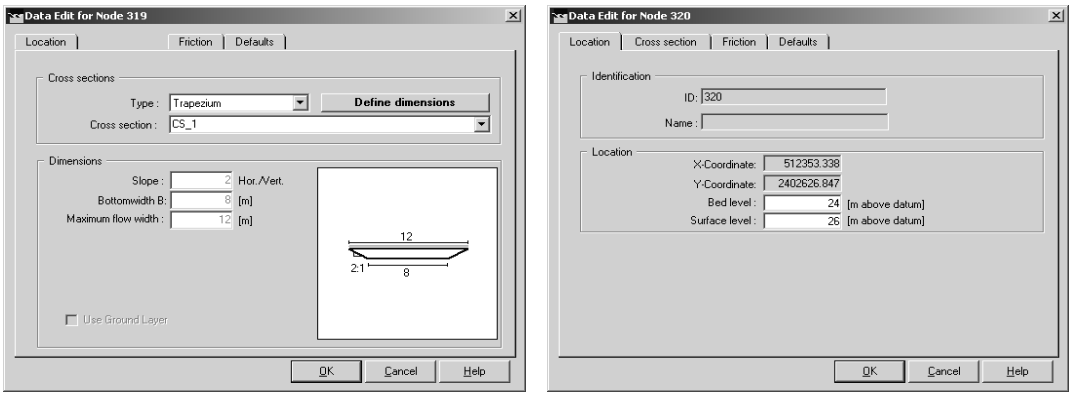


Figure 5.5: Parameters for 1D flow cross-section

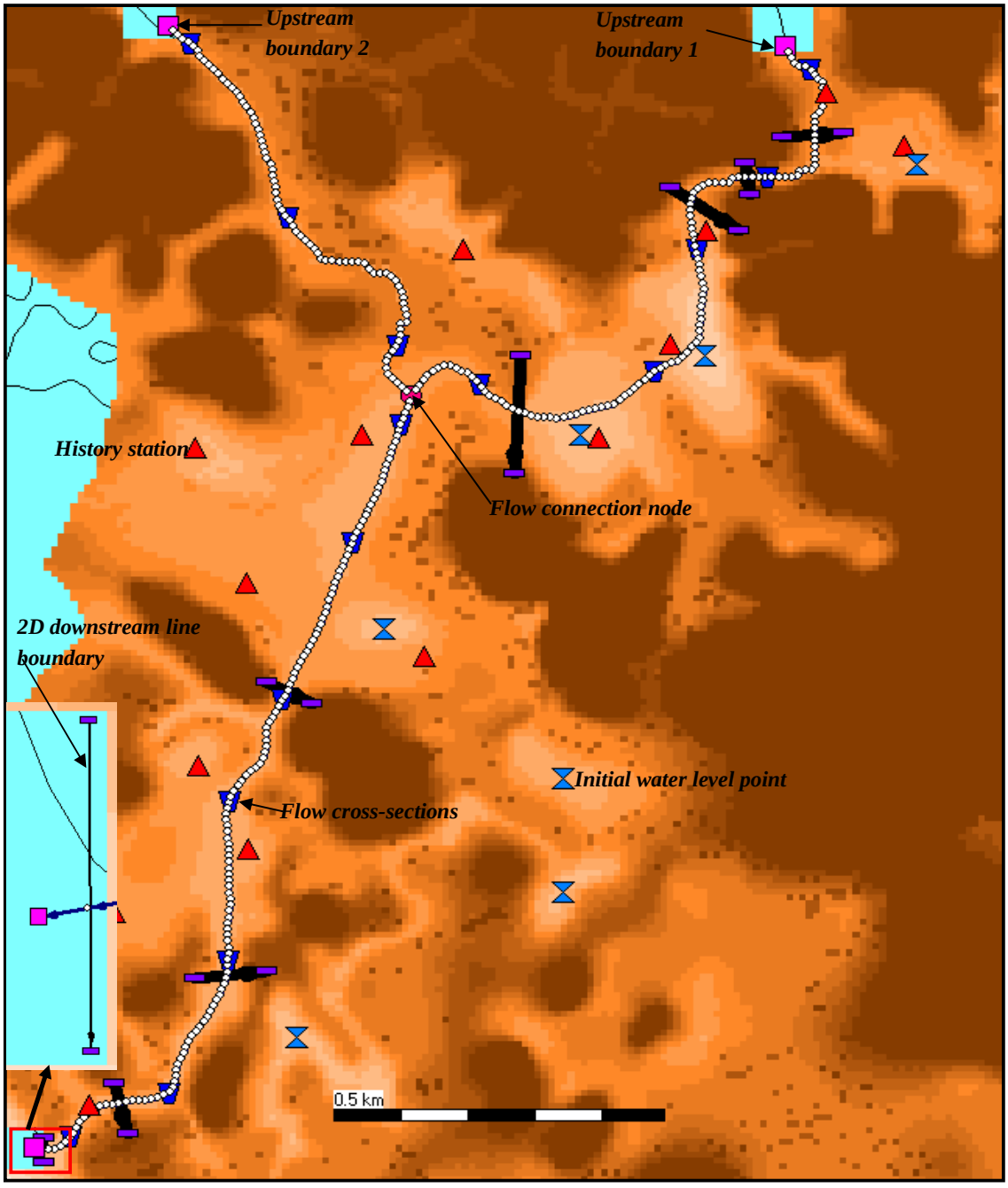


Figure 5.6: Model schematization of SOBEK 1D2D flood modeling in Netter network editor

5.6 Model calibration

Flood simulation model is a reconstruction of a flood event through a computer program. There is no any flood simulation model which gives 100% fit with nature. This is because many different simplified parameters have to be used for flood simulation. Therefore model calibration is simply defined as tuning the flood model parameters so that gives the lowest deviation between modeled and measured parameters. As an example Manning's coefficient can adjusted with different model runs. According to (Mark et al., 2004) flood data of one event is not good enough to use for model calibration in large flood event. But only one flood event data is available as flood extent and depth information which are the most common parameters used for model calibration. In general flood depth has local variation, because the topography and the land use type, and building structures are playing a major role for flood depth changes.

Due to the lack of data availability it is difficult to define proper boundary data for upstream (chapter 5.2). Upstream boundary data was derived from constant discharge. Whole process of model calibration was done through two main steps. First step is concerned on the representation of the building as solid block and high manning surface (n=1). The second step consider with different manning's value with same landuse type. Model information, flood depth and extent compared with observed value during field work.

Typhoon of 2008 data used to check the prediction accuracy of the model. 52 flood information points and flood extent information were collected by interviewing people during field work. Measured water height distribution of field samples is given in Figure 5.7.

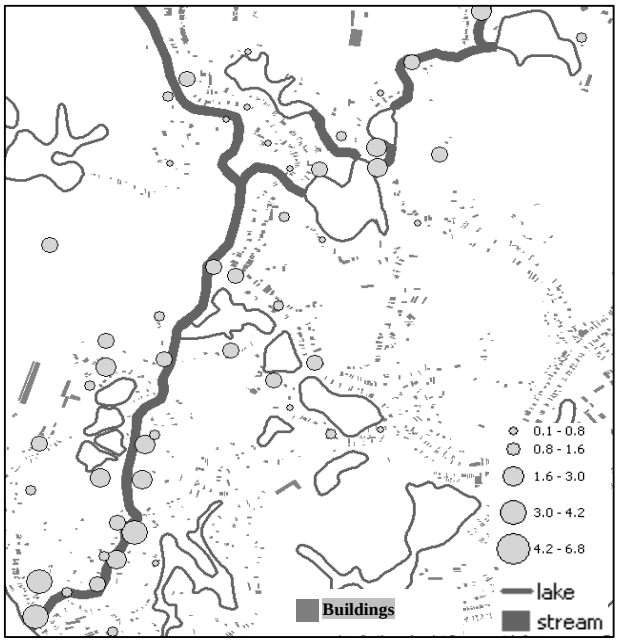


Figure 5.7 flood depth distribution of the area measured through field work

Simulated flood depth information against both building blocks and rough surface compared with field observed flood depth information and then difference (Field measured flood depth – simulated flood depth) was obtained. According to two histograms (Figure 5.8) which represent the difference of flood depth for both building representations, there are only 3% of sample of both building structure with the difference between 1m to 2m and all of them are given over-predicted values of simulated flood. This difference may be due to local deviation of DSM. 85% of data can be observed the difference between 0 to 0.5m levels which is not bad. When the building representation is solid block, the difference between measured and simulated flood depth distribution shows that increase the under prediction compare to solid block representation.

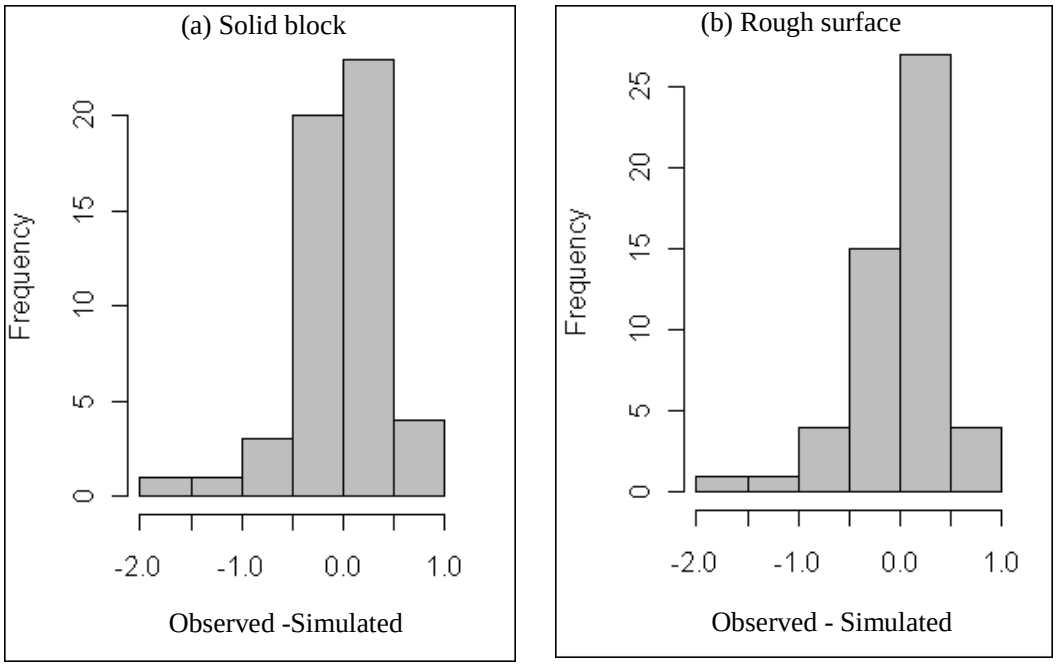


Figure **5.8: Histogram of maximum water depth prediction error for building representation as solid block (a) and rough surface (b)**

Mean and Standard deviation of the difference between measured and simulated flood depth data for both building structure of solid block and rough surface was not changed significantly (Table 5.3). Mean of solid block representation is closer to 0 compare to Rough surface.

Table 5.3: Mean and standard deviation of solid block and rough surface

Building Structure	Mean	Standard deviation
Solid block	-0.32	0.423
Rough Surface	-0.36	0.421

Then compare the flood extent with different building structures against field observed data. The visual interpretation (Figure 5.9) of flood extent clearly indicates that flood extent of solid building structure representation is comparative to the field observed data.

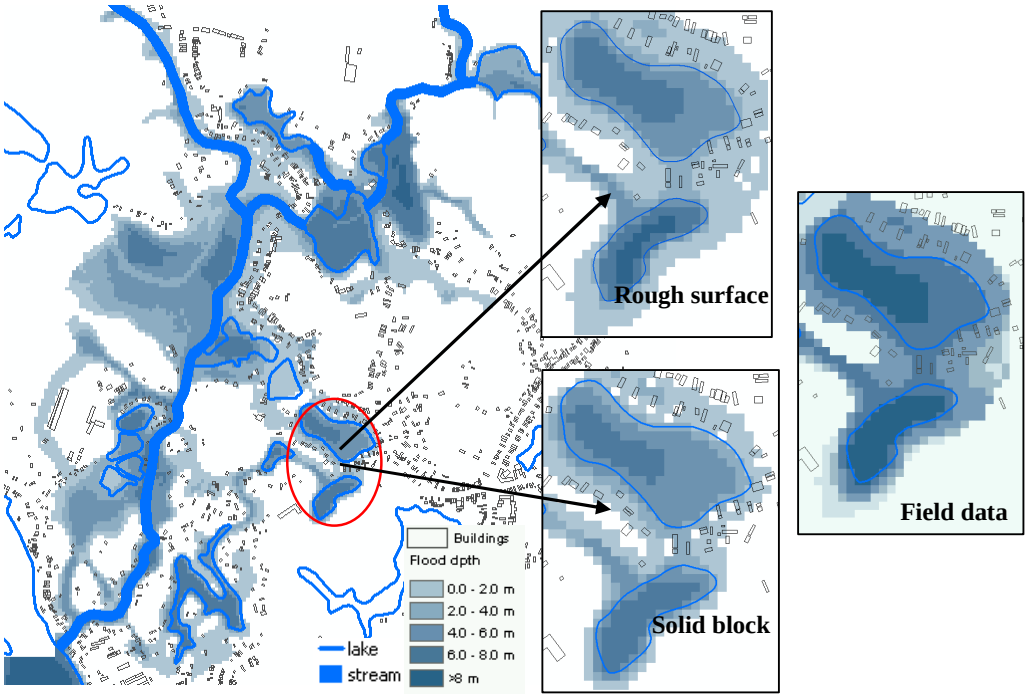


Figure 5.9: Different flood depth and flood extent for 25 year return period of flood with the build representation of solid block and rough surface against Field observed data

Also the water velocity changes in study area were studied with both solid block and rough surface of building structures. The water velocity is comparatively high in rough surface representation of building compared to the solid block representation (Figure 5.10). It also gives a slightly different flow pattern. The water velocity of solid block representation decreases because building is acting as a barrier while water not flowing through them. But rough surface representation allows to water to flow faster and it also causes to increase the flood extent. When compare with field observed data (fig 5.8) the flood depth is over estimated.

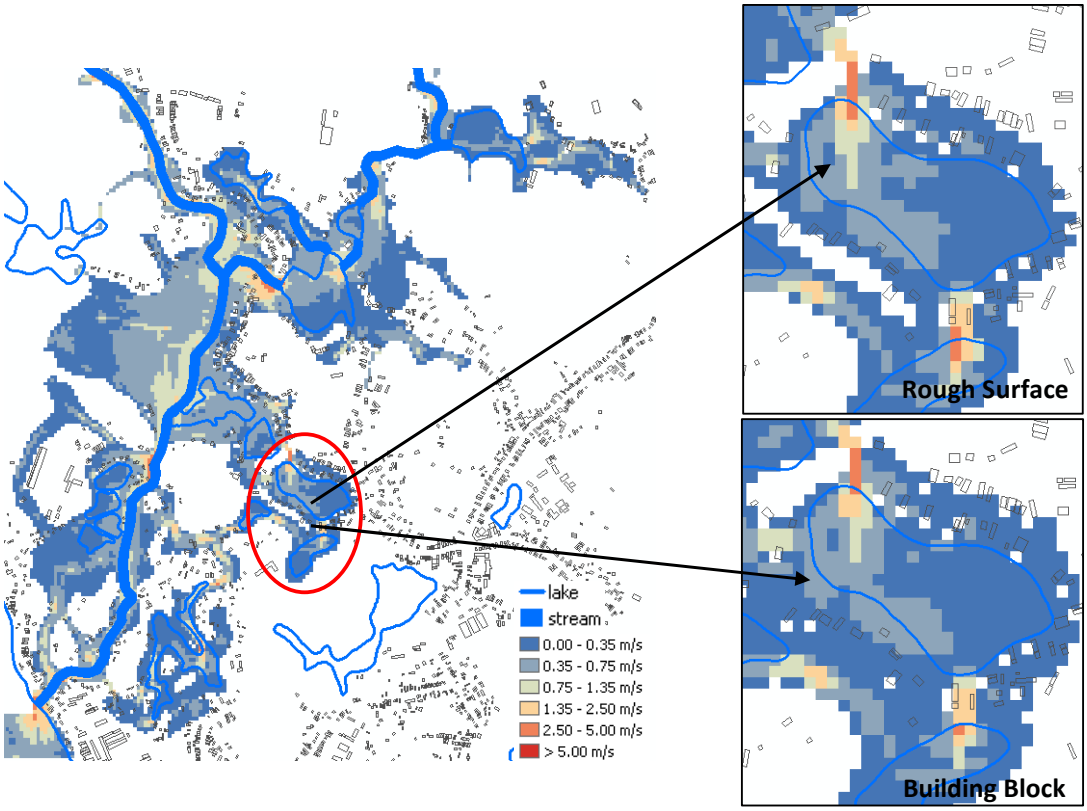


Figure 5.10: Different flood extent and water velocity for 25 year return period of flood with the build representation of solid block and rough surface

The next approach was to calibrate the Manning's coefficient values for further flood simulation. As explained in chapter 5.3; three sets of manning's coefficient have been used to flood model calibration with 25 year return period of flood. Individual simulations for all three Manning's values carried out and the flood depth information used for understand the error distribution.

According to the histograms of maximum flood depth error distribution (Figure 5.11) there is no large deviation of prediction error in all three sets of Manning's coefficient. It means that the maximum water depth for the three different Manning's coefficients is not much different. The initial run was given the minimum percentage of deviation of error from more than 0.5m level which is equal to 17%. Therefore initial manning's value (set 1) was used for flood simulations for scenarios.

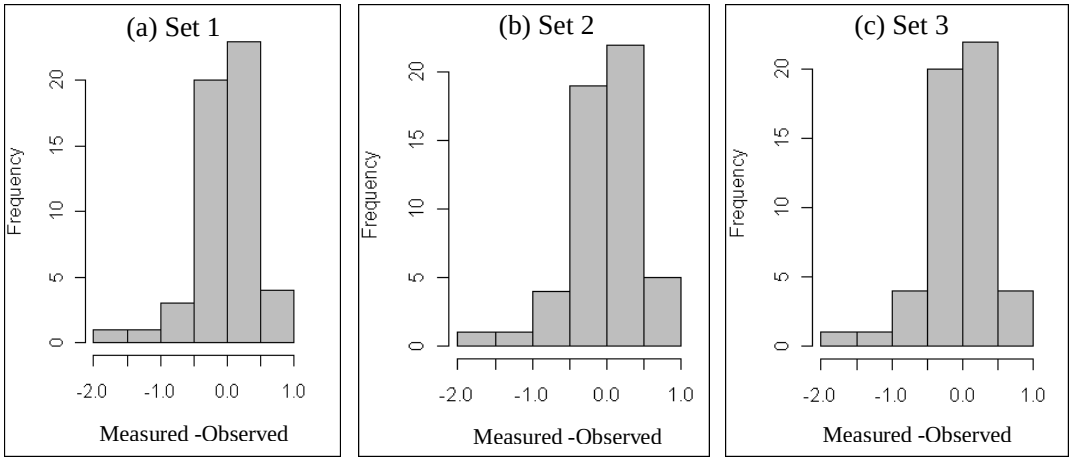


Figure 5.11: Histograms of maximum water depth prediction error by flood modelling with initial manning coefficient Set 1(a), Set2 (b) and Set 3 (c)

Mean and Standard deviation of difference between measured and simulated flood depth data for three sets of Manning's values are closer to each other (Table 5.4). The differences are close to millimetre level. Manning's set 1 gives the value closer to 0 compare to the set1 and set2. Finally considering all the above discussed factors Manning's values set 1 can be used for further flood simulation for scenarios.

Table 5.4: Mean and standard deviation of three sets of Manning's values

Manning values	Mean	Standard deviation
Set 1	**-0.036**	**0.423**
Set 2	-0.044	0.426
Set 3	-0.040	0.420

Chapter 6 Flood hazard and risk in Yen Bai city

Flood hazard in Yen Bai city can be expressed as the probability of an occurrence of a flood event with in a given period. Therefore to develop a flood hazard map always need to have flood events with different return periods. Flood hazard assessment is based on the five return periods such as 5, 10, 25, 50 and 100 year. Those scenarios were derived from the historical data analysis of Red River by Gumble distribution (Chapter 4.2). With the calibrated flood model parameters (Chapter 5.6), five scenarios were model with SOBEK 1D2D flood model.

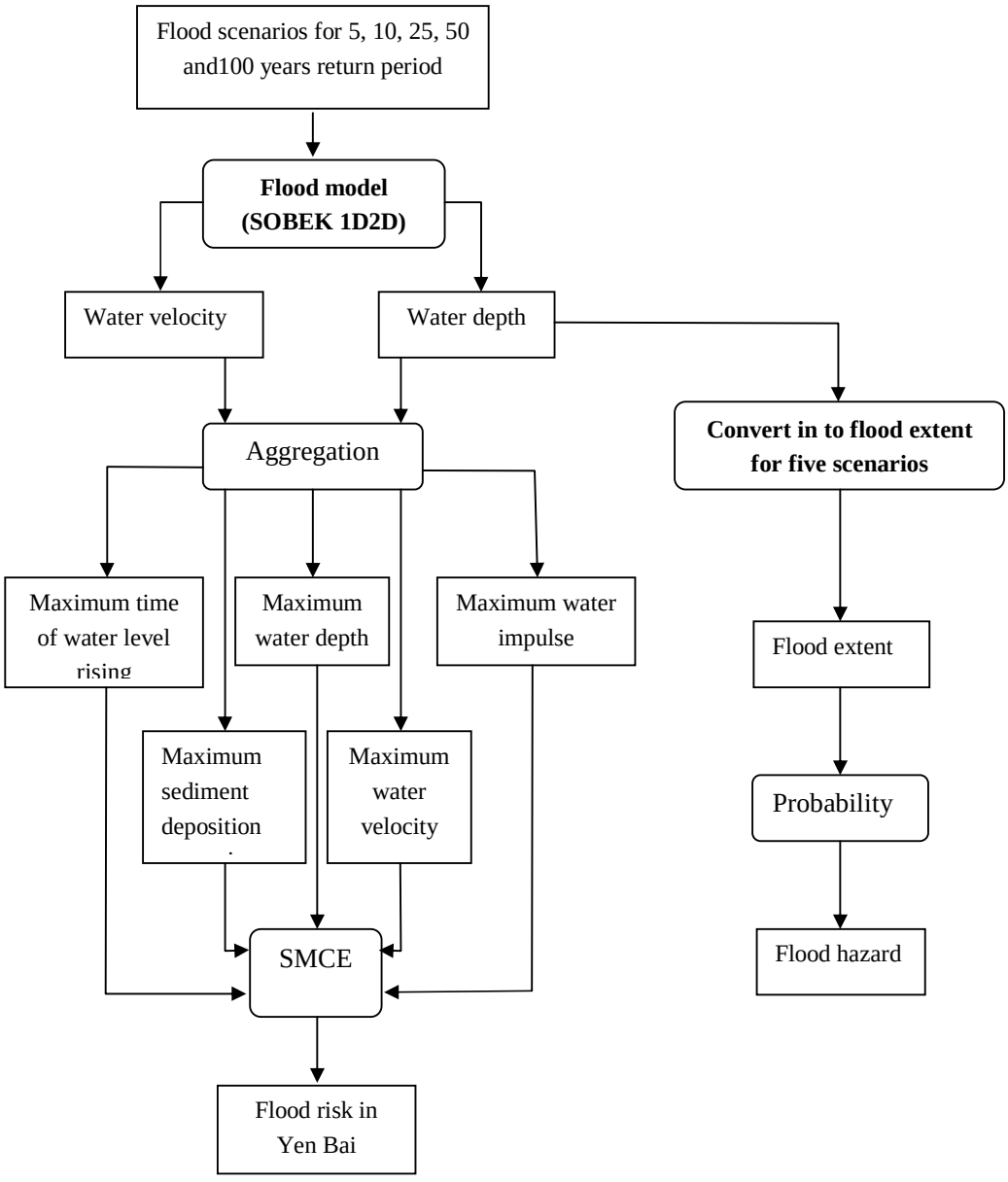


Figure 6.1: Overall methodology to create flood hazard and risk

6.1 Hazard assessment

People have defined many methods to categorise the flood hazard (Geological society of Australia, 2005). However to generate a flood hazard map for any place, should have to understand the flood propagation characteristics of that area. In principle flood extent, depth, velocity for any flood event is the important parameters for flood hazard mapping. However generation of those parameter maps for scenarios is difficult only through the field observation itself. Therefore 1D2D flood models can be used to simulate the scenarios to over come that problem.

Five flood events with different return period (5, 10, 25, 50 and 100 year return periods) were simulated with 1D2D SOBEK flood model and studied the flood characteristics. Considering the flood extent data from all those simulated flood parameter, the flood hazard map was created for Yen Bai city. On the other hand multi-parameterized flood risk of the area was studied with other parameter maps.

To generate a flood hazard of the area in beginning, the probability of occurrence of different flood events was calculated (Table 6.1). Then probability values assigned to the flood extent map of five different return periods. Finally all those maps were integrated to get the flood hazard of the area.

Table 6.1: Probability values for different return periods

Return period (year)	Probability of occurrence	Red River water level (m)	Discharge of Upstream boundary 1 (m^3/s)	Discharge of Upstream boundary 2 (m^3/s)
5	0.2	28.63	5	5
10	0.1	29.21	10	7
25	0.04	30.00	15	10
50	0.02	30.20	20	12
100	0.01	30.88	25	15

Due to the combining of all those maps, the probability values which vary from lowest (0.01) to highest (0.2) can be observed in final map (Figure 6.2).

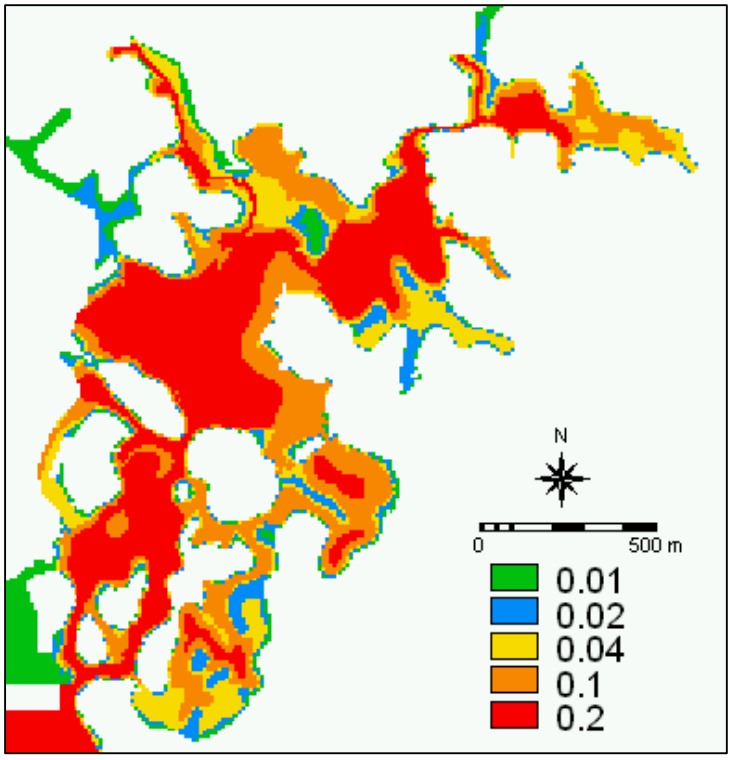


Figure 6.2: Probability value distribution for different return period

Then the probability values classify in to three categories considering the probability values (Table 6.2).

Table 6.2: Probability values for different hazard classes

Hazard Class	Probability value range
Low Hazard	0.01.0.02
Medium Hazard	0.02-0.1
High Hazard	0.1-0.2

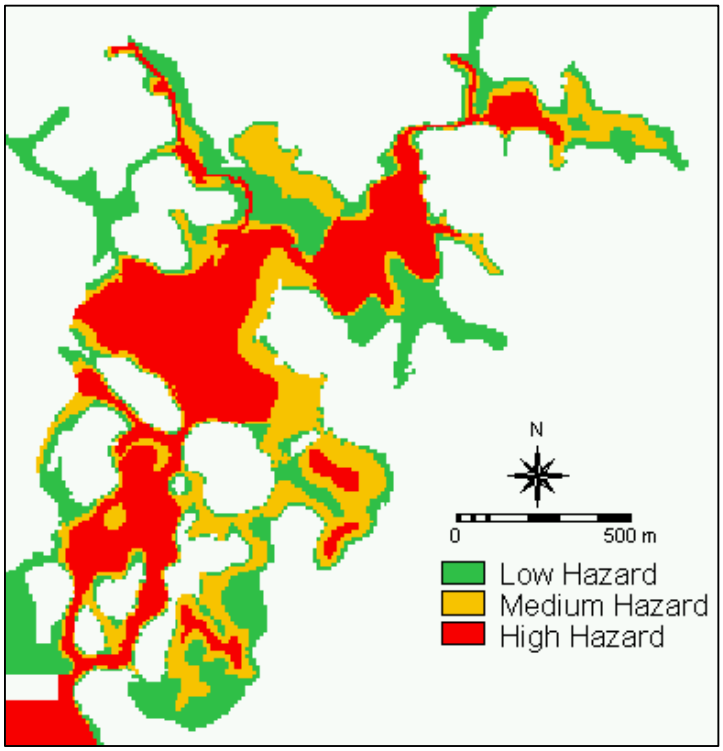


Figure 6.3: Hazard class in Yen Bai city

Then flood hazard can also be explained by the combination of flood depth and flow velocity explained by the Geological Society of Australia (2005). The flood depth and velocity are contributing for damage the properties and people. It is in between hazard and risk. Therefore considering all the flood risk parameter maps such as flood velocity, depth, impulse, rate of rising of water and sedimentation can be used altogether to create the specific risk for the area by using decision making tool known as (SMCE) Spatial Multi Criteria Evaluation (chapter 6.2).

6.2 Multi parameterized Flood risk

6.2.1 Parameter maps for flood risk

With the simulation time frame every one hour was given flood depth and water velocity map during simulation (Chapter 5). Five scenarios like 5, 10, 25, 50 and 100 year return period were simulated for 48 hours and each of those scenarios created 48 hourly maps for flood depth and velocity. Hourly water depth and flow velocity map were transformed in to five different parameter maps: Maximum water depth, maximum flow velocity, maximum impulse, maximum speed of water level rising and sedimentation map.

The lowest and highest water depth and inundation area was given in 5 year and 100 year return period of flood (Figure 6.4). There is no large difference in flood extent for 10-100 year compare to 5 year flood but more than 75% of the area of 100 year flood, the depth is grater than 1.5m.

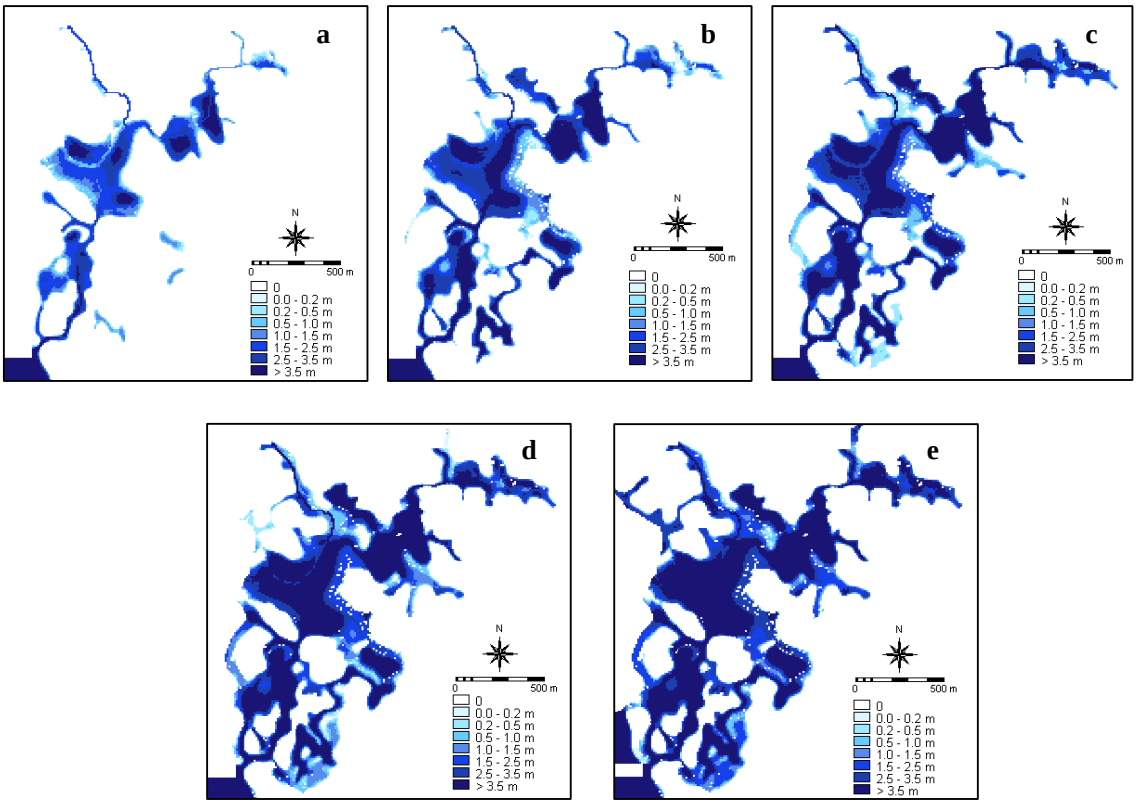


Figure 6.4: Parameter maps of maximum flood depth in different return periods 5 year return period (a) 10 year return period (b) 25 year return period (c) 50 year return period (d) 100 year return period (e).

Always high water velocity can be observed along the main stream and it is grater than 2m/s closer to the bottle necks (Figure 6.5). Water velocity pattern in 10, 25, 50 and 100 year floods bit comparable but 50 year flood shows large area (x) with high velocity compared to the others. That is implied from the input hydrograph (chapter 5.2) of flood simulation.

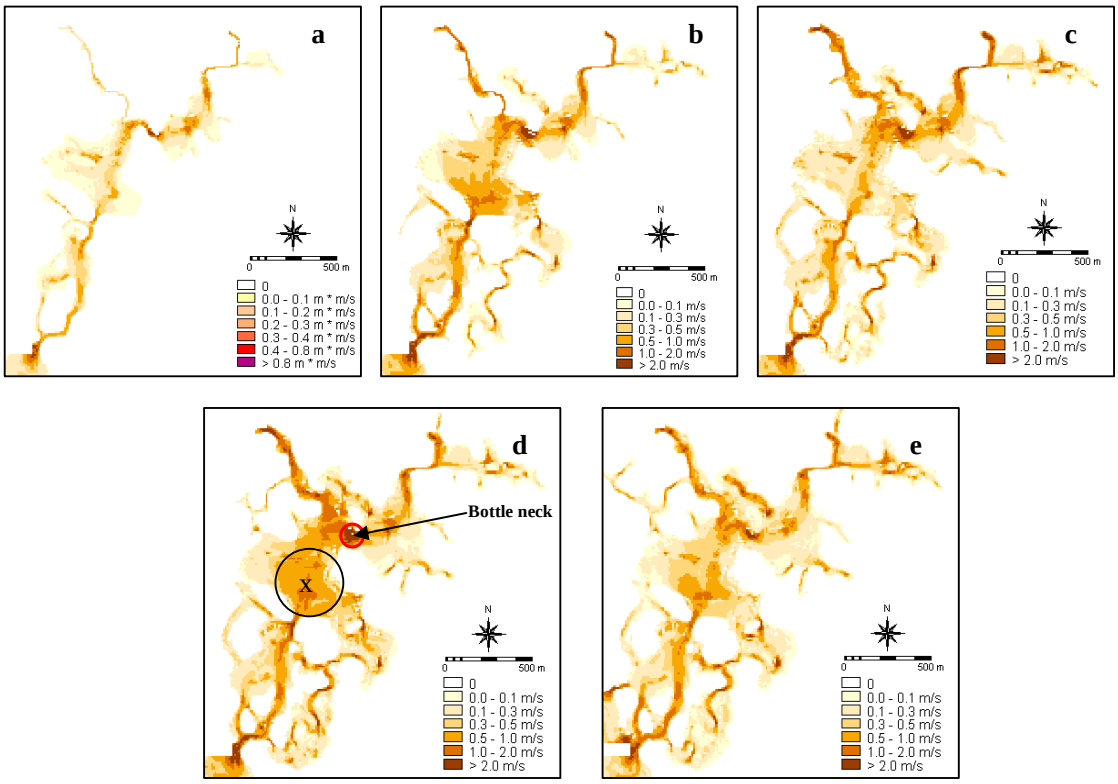
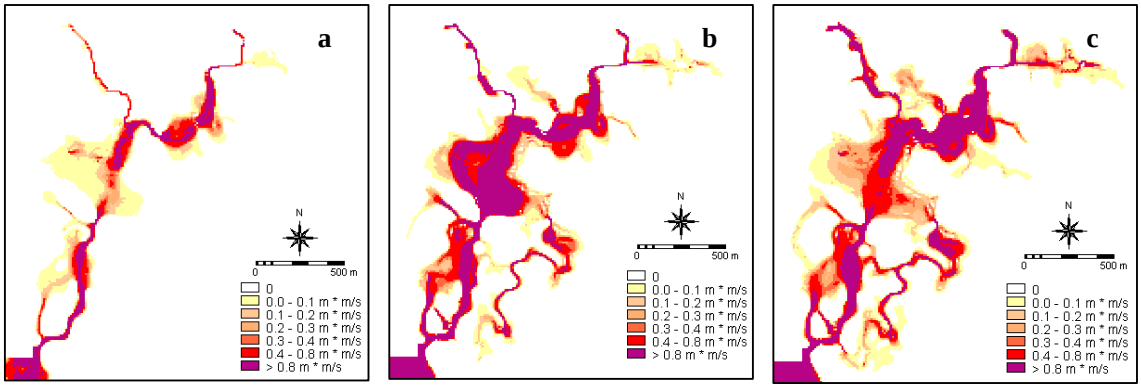


Figure 6.5: Parameter maps of maximum water velocity in different return periods 5 year return period (a) 10 year return period (b) 25 year return period (c) 50 year return period (d) 100 year return period (e).



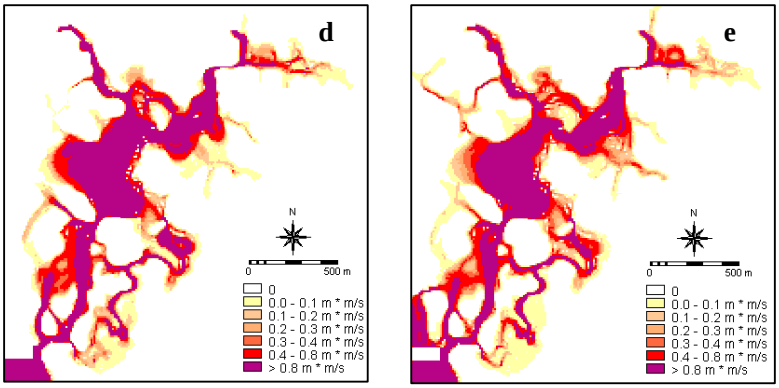


Figure 6.6: Parameter maps of maximum impulse of different return periods 5 year return period (a) 10 year return period (b) 25 year return period (c) 50 year return period (d) 100 year return period (e).

The maximum impulse (water depth * water velocity) for different floods (fig 6.5) was derived from the maximum water depth and velocity. The pattern of impulse of 50 and 100 year flood is more comparable than the other three (Figure 6.6 a,b,c).

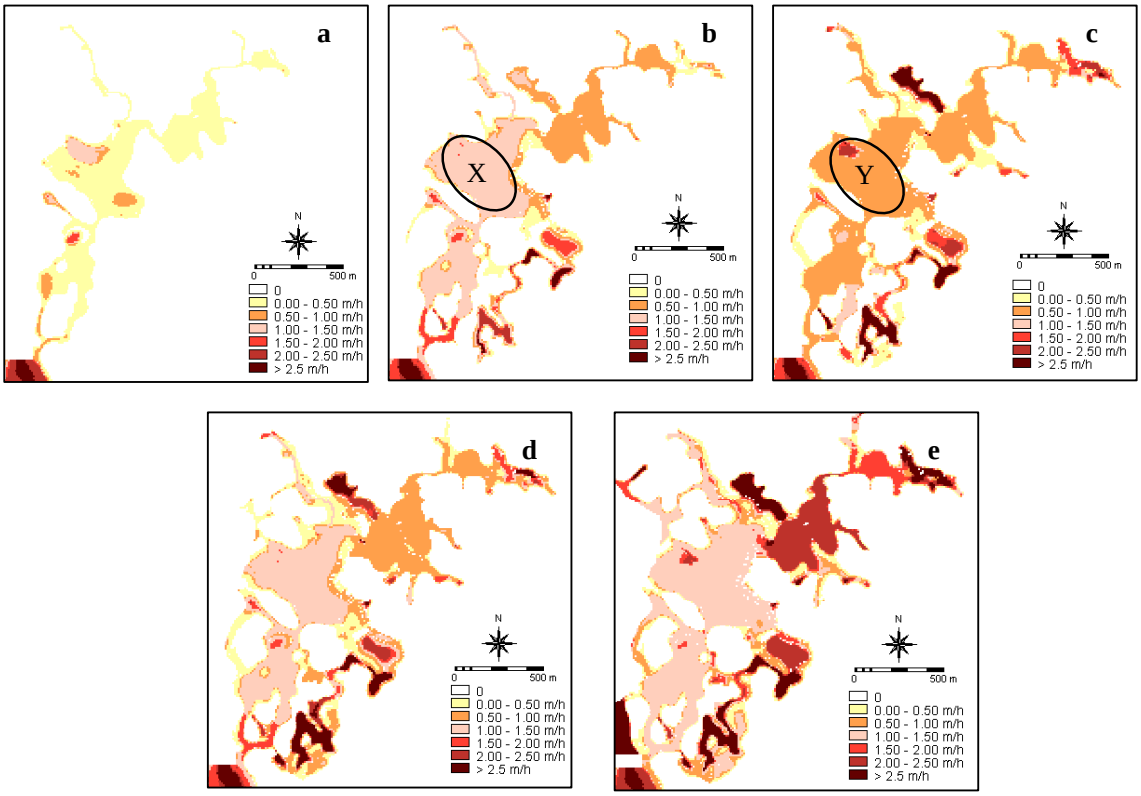


Fig 6.7: Parameter maps of maximum rate water level rise in different return periods 5 year return period (a) 10 year return period (b) 25 year return period (c) 50 year return period (d) 100 year return period (e).

As same as the other parameters, the highest rate of water level rising (Figure 6.7) was observed for large spatial extent from 100 year flood compared to the others and it is lowest in 5 year flood. Almost same spatial extent and quite different values for rate of water level rising was observed for 10 (X) and 25 (Y) year flood with high and low values. This is because the maximum water level of 10 year flood reached at 21 hour and 25 year reach at 28 hour (chapter 5.2) during the flood simulation.

Severe potential erosion was occurred close to the bottle necks of the river and severe deposition was occurs in places where the low velocity and high water depth has (Figure 6.8).

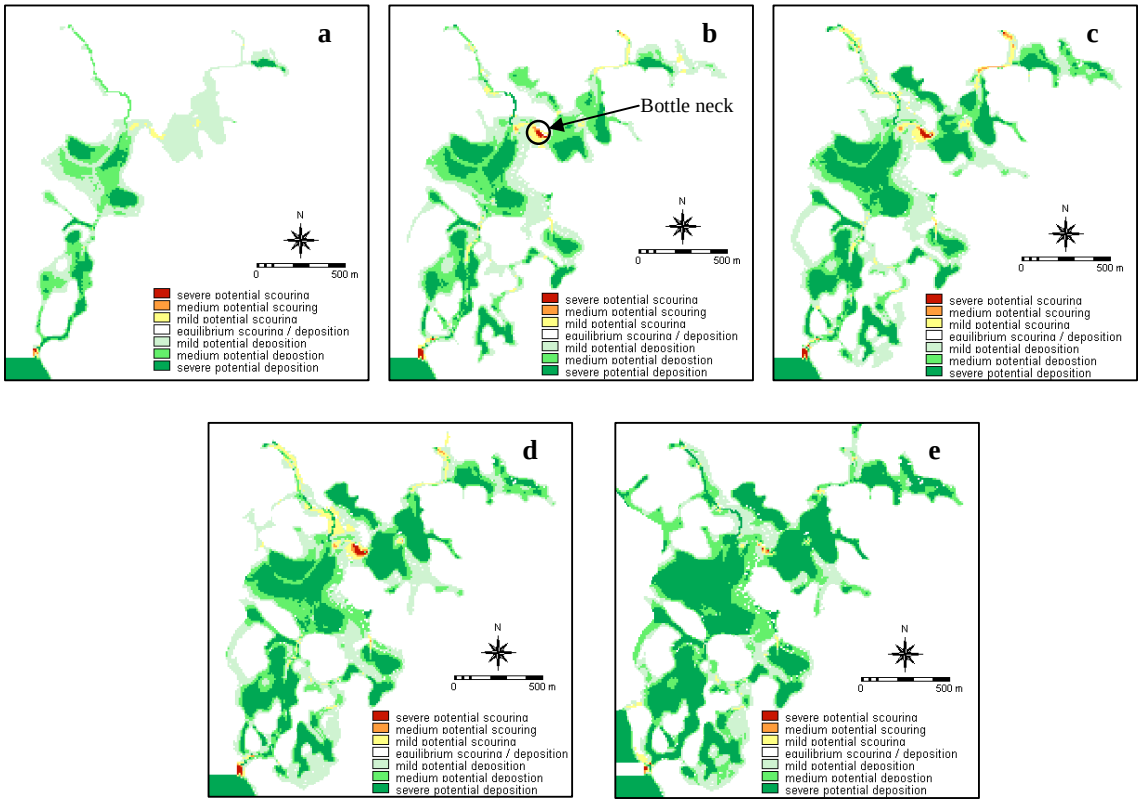


Figure 6.8: Parameter maps of maximum sedimentation or erosion in different return periods 5 year return period (a) 10 year return period (b) 25 year return period (c) 50 year return period (d) 100 year return period (e).

6.3.2 SMCE (Spatial Multi Criteria Evaluation)

To perform a complete risk assessment additional information like vulnerability and element at risk are needed. But the study mainly concerned about the specific risk of the area by considering flood parameter maps. Maximum water depth, velocity, impulse, rate of water rising and type of sediment (either erosion or sedimentation) are used as flood parameter maps. Those are the main factors to damage the flooded area either people or structures. Therefore decision making tool can be used to evaluate the risk of the area by aggregating all the parameter maps. The Spatial Multi Criteria Evaluation (SMCE) method can be used for decision making.
SMCE is the process of transformation of multidimensional spatial information in to single parameter map. This transformation has to be done through the four steps: 1) identification of parameter map as cost or benefits; 2) standardization of the parameter map; 3) identification of importance of criteria acceding to dissection problem; 4) establishment of aggregation procedure.

Cost and benefit parameter

The main aim is to assess the flood risk in the study area with flood parameter maps. Parameter maps which have high values of (flood depth, velocity, rate of water level rising and impulse) are corresponding to high risk. Therefore high values of them are considered as benefit. Cost parameters are inverse of benefit, therefore high cost value is correspond for low risk. Here there is no use of cost parameter map like warning time; thus there is no pay of more attention to high cost parameters. In case of parameter map of sediment, only used the sedimentation which has the positive value as the main consideration is risk on building structures. If more sediment added in to flood it is cause for risk, it means that is benefit.

Standardization

Standardization is conversion of parameter map's information in to parameter score through the value function in to 0 to 1 scale. There are different kind of value functions are available in ILWIS-SMCE but the function should be selected depending on the objectives of the decision making studies. This study used the direct standardization method for value function because classified flood parameter maps was used for decision making. On the other hand the first run was done with direct standardization and further research could be carried out to evaluate the pair wise standardization to understand what would be more reasonable.

Prioritization

Prioritization was done by the assigning weights to the criteria tree (Figure 6.11). Criteria tree is a tool that can be used for SMSC in ILWIS. The weights for the criteria tree assign in three phases explaining bellow. First depends on the return period and second depends on the prioritization of indicator map and finally it depends on classes (eg flood depth class). When higher the return period, the probability to occurrence of that event is less. Therefore it is less susceptible for occurrence a flood and then assigns the low priority value for it (Figure 6.9). According to the basic understanding of the flood behaviour, high velocity and water depths are the higher contributing factors for damaging of buildings or peoples. If both water depth and velocity is high, it is more damaging. Likewise the factor maps can be prioritized (Figure6.9). Then the classes of parameter maps (eg. velocity) prioritized based on their values. If the class value is high it is more causative for damages (Figure 6.10).

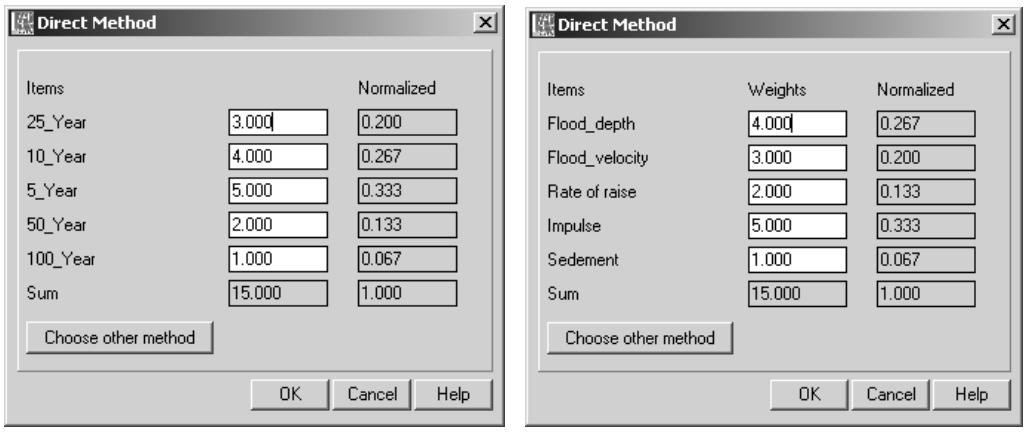


Figure 6.9: prioritized values flood scenarios and parameter maps

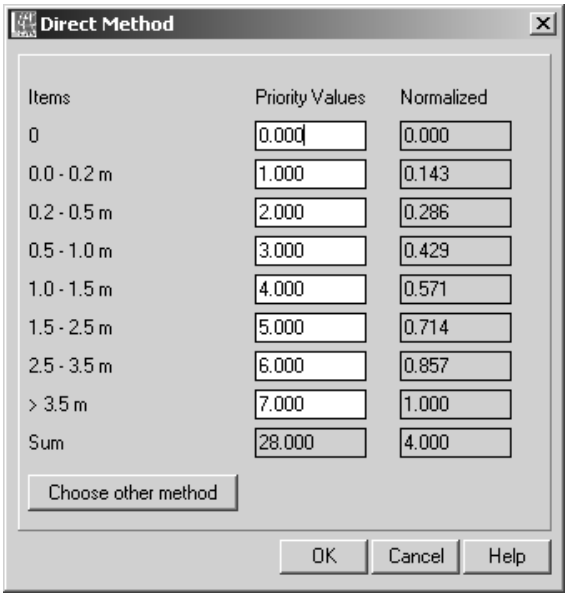


Figure 6.10: prioritized values for classes of parameter maps

Aggregation

The aggregation method was weighted with linear combination. ILWIS SMCE criteria tree (Figure 9.11) was used to aggregate all the factor maps. Aggregated specific risk map can be used to evaluate the specific risk on the exposed element by overlaying them (Figure 6.12). To evaluate the total risk of the area extra information like number of people living, value of the building, building type have to be considered.

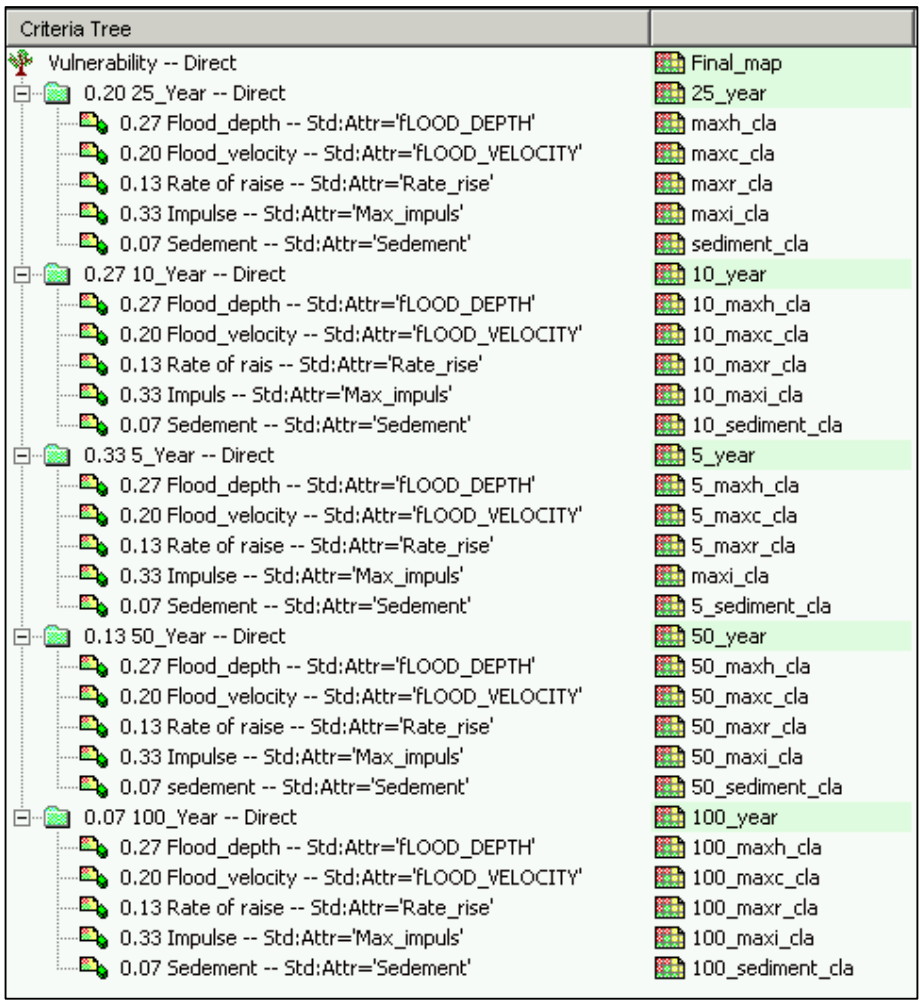


Figure 6.11: Criteria tree for aggregation of factor maps with direct standardization

Buildings which are located closer to high values of risk are in high flood risk.

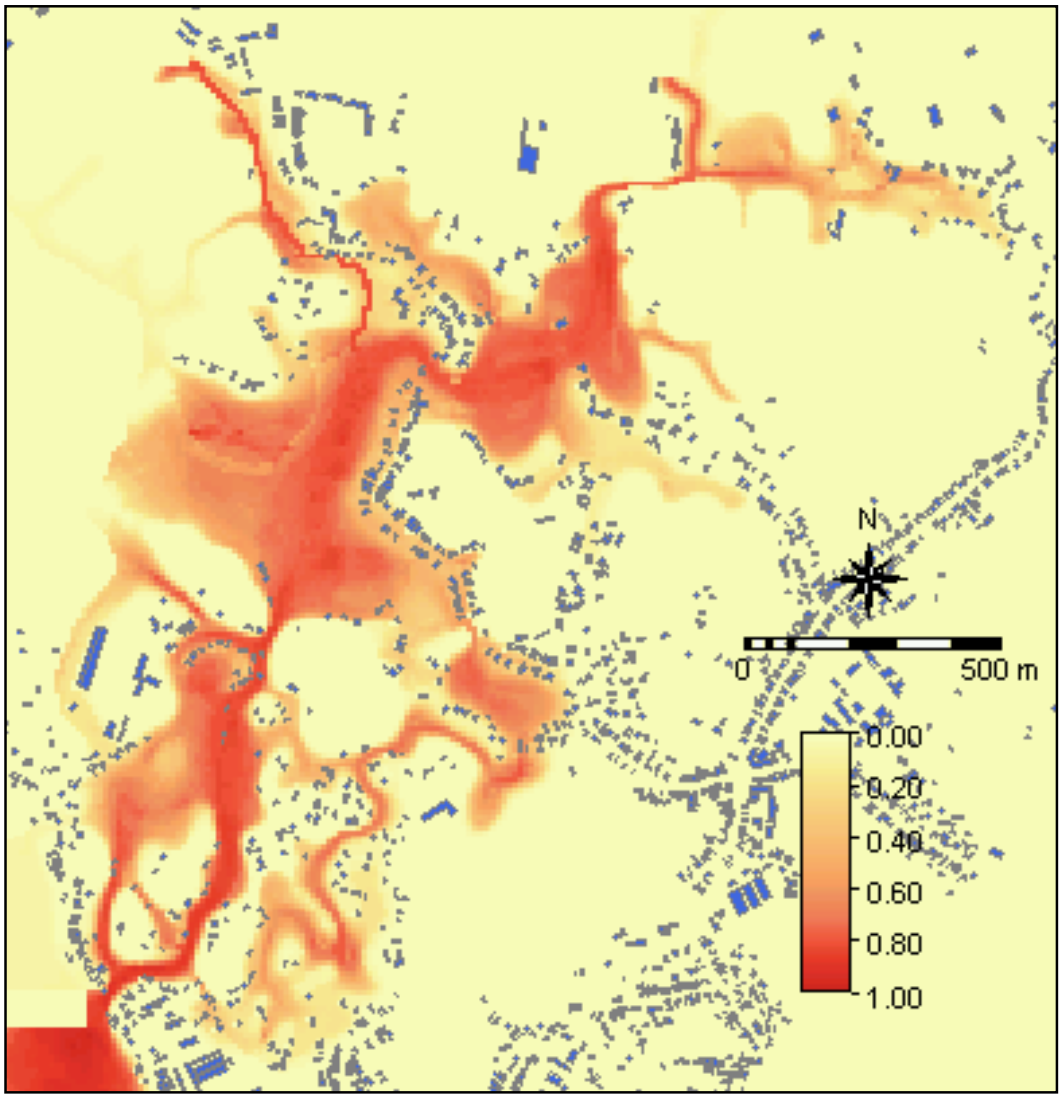


Figure 6.12: Multi-parameter risk map for Yen Bai city, Vietnam

6.3 Flood from Numkun river catchment

With the study of upstream of Cuong Thinh catchment surface runoff model and 1D2D flood model, two different hydrographs (Figure 6.13) were generated by (Pedzisai, 2010). Depending on both hydrograph the flood was simulated with SOBEK 1D2D flood model and compared flood behaviour of those two scenarios (Figure 6.14). Those two scenarios gave less flood extent compared to the Red River flood.

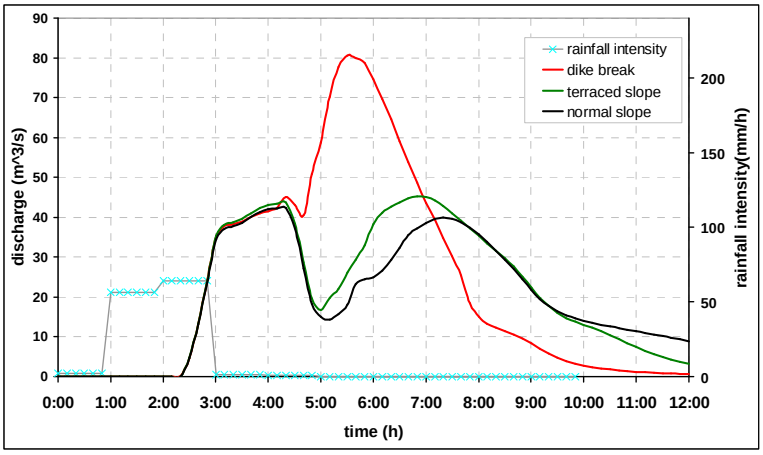


Figure 6.13: Input hydrographs for Upstream boundary 1(Figure 5.6), (Pedzisai, 2010)

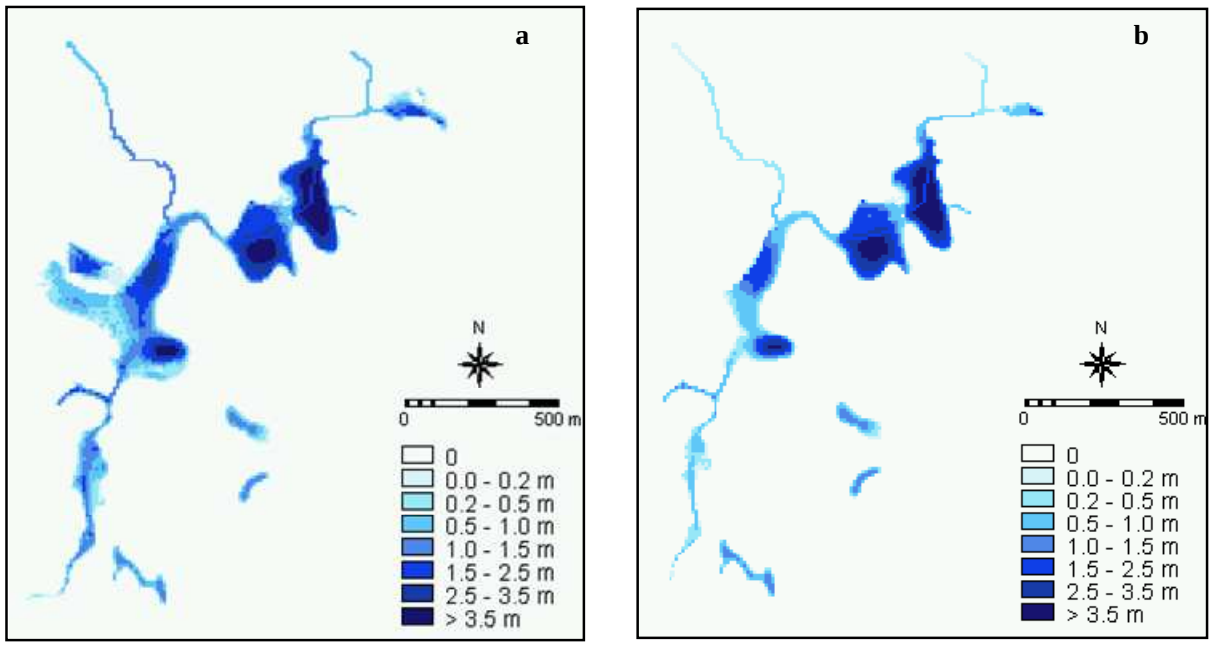


Figure 6.14: flood depth and extent with two different scenarios with dike (a) without dike (b)

Chapter 7 Conclusions and Recommendations

This chapter will provide the answers to the research questions which were formulated in the chapter 1.3 to address the research objectives in this research.

7.1 DSM Creation

How to construct a sufficiently detailed terrain for flood propagation model with the available data?

The elevation data which was derived from the different sources (1:10,000 contour line, field measured elevation points and building heights) have been integrated and interpolated to produce the final DEM and DSM which was used for detailed flood modeling. Most part of the study area consists of complex terrain features like hills, valleys, slope cuts and water bearing features (lakes); those features give the high variation of the terrain elevation with in a shorter distance. Therefore the most suitable elevation data set for DEM generation was selected so that reducing the variation of elevation data within a shorter distance with respect to the nugget effect from the integrated elevation data set. Furthermore the averaging of the elevation data set with in the 10m distance some of the detailed terrain information will be eliminated. Integrated elevation data of 10m line to point conversion and field measured elevation data points was selected as the best point data for further interpolation.

After selection of best point data set for interpolation, after studying the capability of commonly used terrain interpolation techniques, a Digital Elevation Model was generated. Numerical and visual interpretation methods were used to determine the suitability of terrain model. TIN interpolation is sensitive to the errors and noise of the input dataset and it is possible to create the artificial pits. These artificial pits are not suitable for hydrological modelling applications. Polynomial interpolation yields the biggest error compared to the other two interpolation techniques, as such polynomial interpolation also not accurate enough to give the detailed landscape. Consequently Kriging interpolation was selected as the best method for terrain interpolation because of the lowest error.

According to the field observations 10m resolution is enough to represent the terrain information of the study area. Therefore final DEM resolution was selected as 10m by considering the field observations and flood simulation time. Finally DSM was created by overlaying the manmade structures on the DTM. This approach can be used as an economical way for DSM generation.

7.2 Flood modelling

What information, historic events are available and how can it be used to validate model result?

Flood of the study area was simulated with two upstream boundaries and one downstream boundary condition. According to the field data collection conceived the idea of flood behaviour of that area. The area is flooding with two sources, when the Red River water level increases, the backward water flow cause for flooding. On the other hand the high discharge of the Cuong Thinh river catchment during the heavy rain causes floods. Those two scenarios are being given the completely different flood behaviour in to the area. The flooded area due to the Red River water level increase much higher compared to the flood of Cuong Thinh discharge.

With the field observed data of 25 year return period of Red River flood, both building structures and Manning's coefficient was used for model calibration. When the building structures treated as solid block and rough surface have given the changes of water velocity, depth and extent. The water velocity was reduced when the building is treating as a solid block because it is acting as barrier for water flow. It is also cause to increase the flood depth locally where the building density is high. The maximum water depth difference of solid block and rough surface representation is 2.0m, but rough surface under estimate the water depth compare to solid block. Furthermore, when the building treated as a rough surface, the flood extent was overestimated (Figure 5.9).Therefore solid block was used as the best building representation for this study.

Then the analysis was carried out based on the effects of the Manning's coefficient value to the maximum flood depth. When study the distribution of difference of measured and simulated maximum water depth for three sets of Manning's values, no significant changes observed (Figure 5.12). When use the manning's set 1 (Table 5.2), mean value of the differences of measured and simulated flood depth close to zero compare to the manning's set 1 and set 2. Finally the manning's set 1 was used for further flood scenarios modelling.

7.3 Flood hazard in Yen Bai city

What is the flood hazard in Yen Bai city?

Hazard is defined as the probability of occurrence of a flood event in Yen Bai city. By assigning the probability values to the flood extent map of different return periods the hazard was defined. The aggregated map of all probability values used to indicate the hazard of the area (Figure 6.2) and the hazard class such as Low Hazard, medium hazard and high hazard zones defined based on the probability of occurrence of an event. 5 year return period flood occurs more frequently than 100 year flood. Thus 100 year flood susceptible for low hazard compare to 5 year flood. Then the hazard was categorized in to three classes (Table 7.1)

Table 7.1: Probability values range for hazard classes

Hazard Class	Probability value range	Return period
Low Hazard	0.01.0.02	>50 years
Medium Hazard	0.02-0.1	50 year-10year
High Hazard	0.1-0.2	<10year

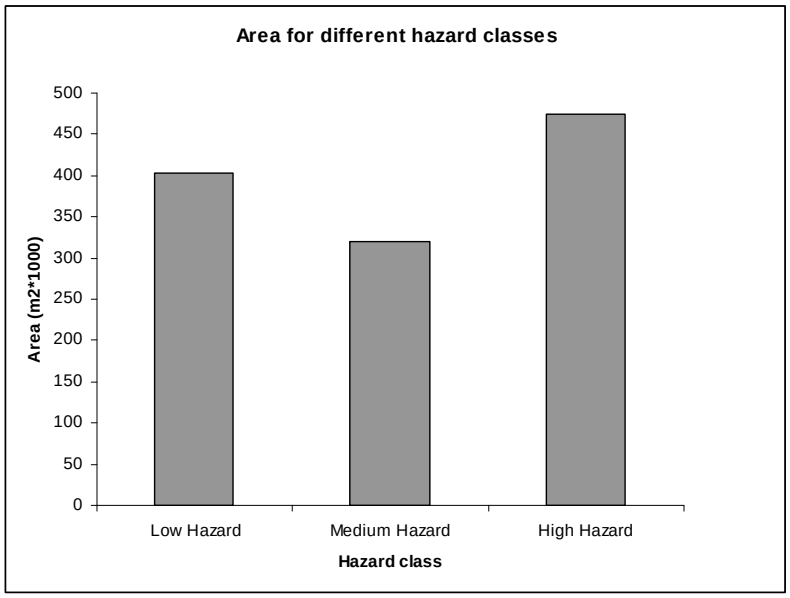


Figure 7.1: Area distribution for different hazard classes

Depending on (Figure 7.1) the classes of hazard, the area of high hazard class is greater than compared to low and medium hazard classes. The lowest area is given by medium hazard. But Number of buildings in high hazard zone is less compared to low and medium hazard zone, it is only 12 houses (Table 7.2) of total number of building in flooded area. Therefore maximum numbers of buildings are located in low hazard zone i.e 175. When increased the return period of Red River flood, the flooded area is also increased (Figure 7.2).

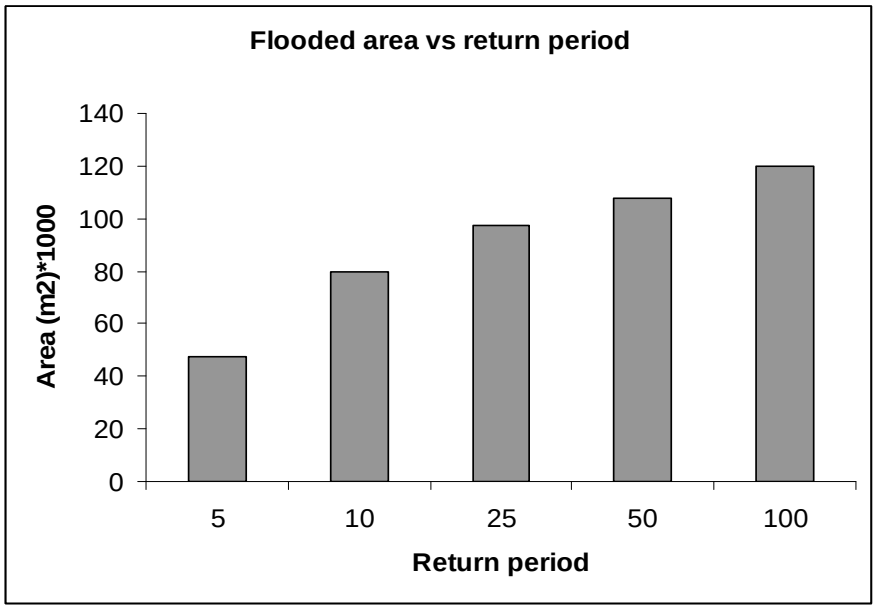


Figure 7.2: Flooded area with different return period

Table 7.2: Number of buildings locates in different hazard class

Hazard class	Number of buildings
High Hazard	12
Medium Hazard	175
Low Hazard	40

Furthermore the flood simulation was carried out for two hydrographs of upstream Cuong Thinh catchment which was generated by (Pedzisai, 2010), based on with and without dikes on paddy field (Chapter 6.3). Then it was revealed that if the flooding associated with the dike scenario, the flood extent is higher compared to without dike scenarios. For instance, six buildings were affected by dike Scenario model and only two buildings affected for without dike scenario. Hence it can be concluded that the study area is more severally affected by Red River flood compare to the flood of Cuong Thinh catchment.

7.4 Spatial Multi Criteria Evaluation (SMCE) for flood risk

How to integrate the flood risk component to create the risk map?

To describe the flood behaviour and risk of the study area different flood parameter maps had to be used. With Spatial Multi Criteria Evaluation (SMCE) method same flood event can be used to create the different flood risk maps. This is mainly depends on the way of defining the vulnerability. The vulnerability can be defined by standardization, prioritization and aggregation method in SMCE. According to that vulnerability, the priority area for flood risk can be identifiable with the risk values which vary from 0-1. 0 indicate the no risk and 1 indicate the highest risk. By overlaying the building foot print map top of the risk map, the building which undergoes high risk or low risk can be easily demarcated (Figure 6.12).

7.5 Recommendations for further studies

Recommendations for future studies as flows:

1. Further studies are needed to simulate the upstream (Cuong Thinh Catchment) flood and to evaluate the flood behaviour with field observed data. Then the resultant flood which can be occurred due to upstream discharge and Red River can be studied in details. Therefore the Upstream discharge estimation is more useful for simulate similar floods.

2. Total inundation area in Yen Bai city is much higher compared to the area covered by this research. Development of a flood risk map for entire city is better for future urban development plan.

3. With this study of the two different flood Scenarios of Red River flooding and Cuong Thinh catchment flooding, the high hazard can be observed from the Red River flooding. Therefore to mitigate the flood in Yen Bai city, due attention should be given to the Red River flood rather than the flood of Cuong Thinh catchment. Study the pertinent flood mitigation measures to reduce the flood impact of Yen Bai city and development of impact assessment.

4. The Multi Parameterized risk assessment was carried out with direct standardization of SMCE and further research focused on the different standardization methods will be useful to evaluate the applicability of SMCE for future flood risk assessments.

5. A large underground drainage system exists in the central part of the Yen Bai city. Therefore the influence of drainage system towards flood problems is needed to be analysed by including drainage network to the model schematization.

Chapter 8 References

Alkema, D., 2009. *Digital Surface Model (DSM) Construction and Flood Hazard Simulation for Developement Plans in Naga City, Philipines*

Anderson, E.S., Thomppson, J A., Crouse, DA. and Austin, RE., 2006 *Horizontal resolution and data density effects on remotely sensed LIDAR based DEM in Geoderma* 132:406-415.

Maidment, D., 1993. Developing a spatially distributed unit hydrograph by using GIS. In Applications of GIS in hydrology and water resources. Procedings of Vienna conf, April 1993, 221: 181-192.

Borsanyi, P., 1998. *Physical habitat modelling in Nidelva, Norway*. The Norwegian University Science and Technology.

Butler, J.B. and Lane, S.N., 1998. *Assessment of DEM quality for characterising surface roughness using close range photogrammety*. Photogrammetric Record, 16(92):271-291.

Carrivick, J.L., 2006. *Application of 2D hydrological modeling to high-magnitude outburst floods: An example from Kverkfjoll, Iceland.* Journals of Hydrology 321:187-199.

Chen, J., Arleen, A. Hill and Lensyl D. Urbano., 2009. *A GIS-based model for urban flood inundation.* Journal of Hydrology 373:184-192.

Cobby, D.M., Mason, D.C. and I.J. Davenport., 2001. *Image processing of airborne scanning laser aultimetry data for improved river flood modelling.* ISPRS Journal of photogrametry and Remote Sensing 56:121-138.

Caetano, M. and M.Painho., *Analysis of interpolation errors in urban digital surface models created from Lidar data.* 7th international symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Science.

El-Ashmaway, N.T.H., 2003. *Refined model for Flood Extent Predictions Using Laser Scanning. Master*, International Institute for Geo-Information Science and Earth Observation.

Evan, J.E., 1993. *Spatial Simulation*: Environment application. Environmental Monitoring Systems Laboratory U.S. Environmental Protection agency Las Vegas.

Haestad Methods, 2002. Computer Application in Hydraulic Engineering, Basic Hydraulic Principle.

Hillary, K. McMillan and Brasington, J., 2007. *Reduced complexity strategies for modeling urban floodplain inundation:* Geomorphology 90:226-243.

Joint rapid assessment flood disaster after Kammuri storm in Yen Bai Province August, 2008, Yen Bai assessment team.

Kidson, R. and Richards, K.S., 2005. Flood frequency analysis: assumptions and alternatives. Progress in Physical Geography 29,3: 392-410.

Le Cos, M., Delclaux, F., Jenthon, P. and P.Favreau.,2009. *Assesment of Digital Elevation Model (DEM) aggregation methods for hydrological modelling: lake chad basin, Africa.* Computers and Geosciences 35:1661-1670.

Mark, O., Weesakul, S., Apirumanekul, C., Aroonnet, S.B. and Djordjevic, S., 2004. *Potential and limitations of ID modelling of urban flooding.* Journal of Hydrology 299:284-299.

Maune, D.F. 2001. *Digital elevation model technologies and application:* the DEM user manual The American society for Photogrammetry and Remote Sensing (ASPRS), Bethesda.

McDonnell, R.A., 1996. *Including the spatial dimension: Using Geographical information system in hydrology*. Physical Geography 20(2):159-177.

McGrath, J.K., O'Kane, J.P., Barry, K.J. and Kavanagh, R.C., *Channel-adaptive Interpolation to Improved Bathymetric TIN.* Department of Civil and Environmental Engineering, National University of Ireland.

Meijerink, A.M.J., de Brouwer, J.A.M., Mannaerts, C.M. and Valenzuela, C.R., 1994. *Introduction to use of geographic information systems for practical hydrology.* IHP-IV M 2.3 ITC, Enschede.

Merwade, V., Cook, A. and Coonrod, A., 2008. *GIS technique for creating models for hydrodynamic modeling and flood inundation mapping;* Environmental Modelling & Software 23:1300-1311

Mignot, E., Paquier, A. and Haider, S., 2005. *Modeling floods in a dense urban area using 2D shallow water equations:* Journal of Hydrology:186-199.

Neelz.,S. and Pender, G., *The influence of error in Digital Terrain Model on flood flow routes:* School of the Built Environment, Heriot-Watt University, Edinburgh, UK.

Pappenberger, F., Matgen, P., Beven, K.J., Henry, J.B., Pfister, L. and Fraipont de, P., 2006. *Influence of uncertain boundary conditions and model structure on flood inundation predictions.* Advanced in water resources 29:1430-1449.

Parodi, G., 2004. *Flood frequency analysi*s, Design precipitation hydrographs from IDF relationship, ITC, Enschede.

Pedzisai, E., 2010. *Rainfall Runoff Modelling for Flash Floods in Cuong Thinh Catchment; Yen Bai Province, Vietnam.* MSc thesis, International Institute for Geoinformation Science and Earth Observation, Enschede.

Podobnikar,T., 2002. *Various Data Sources of Different Quality for DTM production*. 5th AGILE Conference on Geographical Information Science.

Rientjes, T.H.M. and Alemseged, T.H., 2009. *The effectiveness of high - resolution LIDAR DSM for two dimensional hydrodynamic flood modeling in an urban area*. In: Quality aspects in spatial data mining. Ed. by. A. Stein, W. Shi and W. Bijker. Boca Raton : CRC, 2009. ISBN 978-1-4200-6926-6 pp. 221-238.

Sander, B.F., 2007. *Evaluation of on-line DEMs for flood inundation modeling:* Advances in Water Resources 30 (2007) 1831-1843.

Schubert, J.E., Sanders, B.F., Smith, M.J. and Wright, N.G., 2008 *Unstructured mesh generation and landcover-based resistance for hydrodynamic modeling of urban flooding:* Advanced in Water Resources 31:1603-1621.

Shan, Z.Y. and Tseng, Y.H., 2008. *Quality Assessment for Digital Elevation Models*. Journals of photogrammetric and Remote sensing 13:206-215.

Siska, P.P. and Hung I.K., *Assesing the Kriging accuracy in the GIS environment*.

Smith, S.L., Holland, D.A. and Longley, P.A., *Quantifying Interpolation Errors in Urban Airborne Laser Scanning Models*. Centre for Advanced Spatial Analysis, University of college, London.

Tennakoon, K.B.M., 2004. *Parameterisation of 2D hydrologic models and flood hazard mapping of Naga city, Philippines.* MSc thesis, International Institute for Geoinformation Science and Earth Observation, Enschede.

The Geological Society of Australia, 2005. Case Study- johnstone River Flooding (Flood Hazard Categories).

Vinh, N. and Honah N.V., 2004. *Geology and Mineral Resources, Yen Bai sheet*. Department of Geology and Mineral resources.

Wechsler, S.P., 2003. *Perception of Digital Elevation Model by DEM users*. URISA Journal 15:57-64.

Wechsler, S.P., 2005. *Digital Elevation Model (DEM) Uncertainty:* Elevation and effect on topographic parameters.

Weibel, R. and Heller, M., 1991. *Digital terrain modelling. Geographical Information Systems:* Principle and applications. London: Longman: 269-297.

Werner, M.G.F., 2001. *Impact of grid size impact in GIS based flood extent mapping using a 1D flow model*. Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere 26:517-522.

Wu, S., Li, J. and Huang G.H., 2008. *A study on DEM-derived primary topographic attributes for hydrologic applications: sensitivity to elevation data resolution in Applied geography* 28:210-223.

Wise, S.M., 2007. *Effects of differing DEM creation method on the result from a hydrological model*: Computer and Geosciences 33:1351-1365.

Xie, K., Wu, Y., Ma, X., Liu, Y., Liu, B. and Hessel, R., 2003. *Using contour line to generate digital elevation model for steep slope areas: a case study of the Loess Plateau in North China*. Catena 54:161-171.

Appendix

9.1 Appendix A (values for Y axis, and probability for return period calculation)

Y	Probability	Tr (return period)
4	0.981851073	55.09968
3.8	0.977877598	45.20305
3.6	0.973046194	37.10051
3.4	0.967177474	30.46688
3.2	0.960057401	25.03593
3	0.951431993	20.58969
2.8	0.941001954	16.94971
2.6	0.928417664	13.96993
2.4	0.913275261	11.53074
2.2	0.895114927	9.534245
2	0.873423018	7.900331
1.8	0.847640317	6.563416
1.6	0.817179487	5.469846
1.4	0.781455585	4.575729
1.2	0.739934055	3.845179
1	0.692200628	3.24887
0.8	0.638056167	2.76286
0.6	0.577635844	2.367625
0.4	0.511544834	2.047271
0.2	0.440991026	1.78888
0	0.367879441	1.581977
-0.2	0.294816321	1.41807
-0.4	0.224961794	1.290259
-0.6	0.161682814	1.192866
-0.8	0.108008978	1.121088
-1	0.065988036	1.07065

Appendix B (Calculated Y value for Gumbel probability distribution)

Year	Water Level	Sort	rank	left probability	TR	y=-ln(-ln(left_prob)
1960	2822	2537	1	0.01960784	1.02	-1.369103856
1961	2747	2586	2	0.03921569	1.04081633	-1.175165361
1962	2702	2616	3	0.05882353	1.0625	-1.041411525
1963	2743	2622	4	0.07843137	1.08510638	-0.934339379
1964	2743	2625	5	0.09803922	1.10869565	-0.842595846
1965	2537	2627	6	0.11764706	1.13333333	-0.760836746
1966	2857	2635	7	0.1372549	1.15909091	-0.686080009
1967	2720	2639	8	0.15686275	1.18604651	-0.616473507
1968	3093	2639	9	0.17647059	1.21428571	-0.550777448
1969	2808	2666	10	0.19607843	1.24390244	-0.48811398
1970	2741	2666	11	0.21568627	1.275	-0.427833304
1971	3086	2674	12	0.23529412	1.30769231	-0.369436456
1972	2739	2675	13	0.25490196	1.34210526	-0.312528045
1973	2883	2675	14	0.2745098	1.37837838	-0.25678589
1974	2674	2695	15	0.29411765	1.41666667	-0.201940696
1975	2622	2702	16	0.31372549	1.45714286	-0.147761953
1976	2625	2703	17	0.33333333	1.5	-0.094047828
1977	2627	2715	18	0.35294118	1.54545455	-0.040617693
1978	2722	2719	19	0.37254902	1.59375	0.01269357
1979	2715	2720	20	0.39215686	1.64516129	0.066040065
1980	2729	2720	21	0.41176471	1.7	0.119568534
1981	2586	2722	22	0.43137255	1.75862069	0.173421465
1982	2666	2724	23	0.45098039	1.82142857	0.227739827
1983	2720	2729	24	0.47058824	1.88888889	0.282665606
1984	2639	2730	25	0.49019608	1.96153846	0.338344257
1985	2719	2738	26	0.50980392	2.04	0.394927186
1986	2778	2739	27	0.52941176	2.125	0.452574378
1987	2843	2741	28	0.54901961	2.2173913	0.511457286
1988	2635	2743	29	0.56862745	2.31818182	0.571762102
1989	2775	2743	30	0.58823529	2.42857143	0.633693595
1990	2753	2747	31	0.60784314	2.55	0.697479696
1991	2666	2748	32	0.62745098	2.68421053	0.76337711
1992	2772	2753	33	0.64705882	2.83333333	0.831678317
1993	2675	2772	34	0.66666667	3	0.902720456
1994	2793	2775	35	0.68627451	3.1875	0.976896805
1995	2748	2778	36	0.70588235	3.4	1.054671882
1996	2903	2791	37	0.7254902	3.64285714	1.136601674
1997	2795	2793	38	0.74509804	3.92307692	1.223361309
1998	2675	2795	39	0.76470588	4.25	1.315783759
1999	2730	2808	40	0.78431373	4.63636364	1.414915347
2000	2639	2813	41	0.80392157	5.1	1.522097745
2001	2829	2822	42	0.82352941	5.66666667	1.639093245
2002	2791	2829	43	0.84313725	6.375	1.768284083
2003	2703	2843	44	0.8627451	7.28571429	1.913005502
2004	2724	2857	45	0.88235294	8.5	2.078137249
2005	2813	2883	46	0.90196078	10.2	2.271239187
2006	2695	2903	47	0.92156863	12.75	2.504970212
2007	2738	3024	48	0.94117647	17	2.803054168
2008	3024	3086	49	0.96078431	25.5	3.218742468
2009	2616	3093	50	0.98039216	51	3.921940658