Lund University GEM thesis series nr 14

Dam site selection using an integrated method of AHP and GIS for decision making support in Bortala, Northwest China

Xinyi Dai

2016

Department of Physical Geography and Ecosystem Science

Lund University

Sölvegatan 12

S-223 62 Lund

Sweden

LUND UNIVERSITY

UNIVERSITY OF TWENTE.

ITC

FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Dam site selection using an integrated method of AHP and GIS for decision making support in Bortala, Northwest China

By

Xinyi Dai

Thesis submitted to the department of Physical Geography and Ecosystem Science, Lund University, in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation for Environmental Modelling and Management

Thesis assessment Board

First Supervisor: *Dr., Abdulghani Hasan* (Lund University)
Co-supervisors: Prof. Petter Pilesjö

Exam committee:
Andreas Persson
Hongxiao Jin

Disclaimer

This document describes work undertaken as part of a program of study at the University of Lund. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Course title: Geo-information Science and Earth Observation for Environmental Modelling and Management (GEM)

Level: Master of Science (MSc)

Course duration: January 20*16* until June 2016

Consortium partners:

The GEM master program is a cooperation of departments at 5 different universities:
University of Twente, ITC (The Netherlands)
University of Lund (Sweden)
University of Southampton (UK)
University of Warsaw (Poland)
University of Iceland (Iceland)

Abstract

Constructing dam is a possible solution for remitting the scarcity of irrigation water resource in Bortala, Northwest China. Considering the complexity of dam site selection, Analytic Hierarchy Process (AHP) was chosen in combination of Geographic Information System (GIS) as the method of Multi-Criteria Decision Making on selecting dam site. Precipitation, slope, geological layer, soil type, land cover and drainage order were considered as criteria in this study. A suitability map of constructing dam in Bortala was generated, based on which *8* possible dam sites with relatively high suitability were proposed. According to the suitability map, West Bortala, where the elevation is relatively high, was found to be the densest area with high and very high suitability for constructing dam while Northwest and Central Bortala were found as low or extremely low suitability area for constructing dam. In a percentage describing, *11.8%* area of Bortala dropped in the range of high suitability and very high suitability in which only *0.0001%* was very high suitability area; *47.6%* area of Bortala was found in the range of moderate suitability; *34.1%* of Bortala has got low or extremely low suitability. The profile of proposed dams containing cross section and characteristics of dam such as height, width, and volume were generated. Among the proposed dams, there were *2* small dams, *5* medium dams, and *1* large dam proposed with storage capacities varied from $122{,}506\ m^3$ to $5{,}033{,}652\ m^3$.

Key words: GIS; AHP; MCDM; site selection; Bortala; Northwest China.

Acknowledgement

I would like to send my thankfulness to my supervisors **Abdulghani Hasan** and **Petter Pilesjö** for proposal and technical support during thesis writing. Thanks to my GEM family **Gunjan Sharma**, **Jose Ignacio Diaz Gonzalez**, **Tanyaradzwa Joy Nyarai Muswera**, **Maria Angela Dissegna**, **Chao Yang**, **Jialong Duanmu**, and **Sida Zhuang** for your great friendship and company in the whole MSc program. I would also like to show my gratitude to my friend **Feng Li** who showed me the beautifulness of Xinjiang and let me fall in love with this place. I'm sure this two-year program would be impossible to finish by me without the mental and academic support I got from ITC, Lund and home.

Dedication

This thesis is dedicated to:

My grandmother, who raised me up but passed away before I finish this MSc program.

Table of Contents

Table of Figures

Table of Tables

1. Introduction

Water is a renewable but finite resource which is essential for human life and human activities by supporting agriculture, sanitation, and economic development, etc. As an unevenly distributed resource, in both temporal and spatial scales, the different amount of water resources has been causing flood and drought events at the same time in different places in the world. For instance, flood event occurred at Southeast Spain in *1997* (Hooke & Mant 2000), while drought event occurred at Papua New Guinea (PNG) in *1997* (McVicar & Bierwirth 2010); flood event occurred at South France in *2003* (Maillet et al. 2006), while drought event occurred at South Africa in *2003* (Mondlane 2004); and of course in different years such as flood event at Northeast Iberian Peninsula in *2000* (Llasat et al. 2003) and drought event at the Amazon River basin in *2005* (Chen et al. 2009).

In China, the total amount of freshwater resource has a huge storage which is around *2.8 trillion* m^3, and the ranking of it is at 6th place in the whole world (Piao et al. 2010). However, from the report of Ministry of Water Resource (MWR) of China, freshwater resources per capita in China is *2100* m^3, which only hits *28%* of global average, way below the average number *7831* m^3 per capita (Ministry of Water Resources 2016).

Hence, water resources scarcity is still a problem in China, especially in the area of Northwest China, where large areas covered by arid regions. Northwest China contributes *33%* of the total land area of China and it consists of *6* provinces, while only *7%* of the total water resources (Ministry of Water Resources 2004). In addition, Northwest China is one of the most important snowfield and glaciers area in China with the biggest glacier, which has an area of *1.74* km^2, located here (Zhu et al. 2003). That is to say, a huge amount of freshwater resources are stored in solid state and difficult to be utilized, and this makes the water scarcity even more severe in Northwest China.

From previous studies, Northwest China has been proved as one of the area with highest water resource pressure in a global scale. In this region, the Budyko aridity index (BAI),which is used to evaluate the degree of dryness of the climate, is larger than *3.0* which means this region is one of the most dry place among basins in the world (Shen & Chen 2010). The annual precipitation in Northwest China varies from *40 mm to 600 mm* (Xie et al. 2005)(Fan et al. 2005), while the annual potential evaporation reaches a high point at *1500-3000 mm* (Xie et al. 2005). As an important region of producing wheat and cotton, under the influence of low amount of precipitation and high amount of evaporation, *41.58%* of the total amount of water consumption is spent on agricultural irrigation in this area (Geng et al. 2006).Hence, for Northwest China, irrigation water supply is essential considering both food security and economy development. Xinjiang Uygur Autonomous Region, as the biggest province in Northwest China, the agricultural development here totally relies on irrigation (Liu et al. 2009).

According to a study focused on the time period of *1989 to 2010* (Shen et al. 2013), in the aspect of temporal variation of irrigation water demand, July and August are the months with maximum irrigation water demand. However, the most critical period of water supply, comparing between water demand and water availability, appeared in April and May during *1989 to 2010*. The temporal distribution of irrigation water demand and water availability leads to a severe irrigation water scarcity (Shen et al. 2013).

However, it's extremely difficult for human beings to change the total amount of accessible freshwater. Instead, efforts on improving water usage efficiency can be done as a powerful alternative. Under all above reasons and for a more reasonable and efficient water resource utilization, one of the main approaches is constructing dams, which not only collects and redistributes water for activities like irrigation, residents' consumption, industrial use, and aquaculture but also makes use of hydropower as a source for electricity.

Since Chinese government has a strong influence on lots of issues in China, it is also important to know if water projects such as dam construction are supported by the government or not. According to Chinese government hydrological projects work report of *2014* (Ministry of Water Resources 2014), the investment on water project construction showed an increasing trend from *108.82 billion* CNY in *2008* to *408.31 billion* CNY in *2014*. In *2014, 88.4%* of the total investment comes from central government and local government. That is to say, Chinese government has attached great importance to water resource management. And dam construction in suitable location in responding to water usage demand is supported by the government.

Besides the support from the Chinese government, water resources availability is also an important foundation for dam construction. As an increasing trend of precipitation in Northwest China has been showed in the past *50* years (Shi et al. 2007) and predictions for future precipitation under different scenarios also showed an increasing trend in Northwest China (Wan et al. 2015), the environment is becoming more suitable for surface water harvesting in the aspect of precipitation amount.

However, Northwest China is a vast area that can only be described by massive amount of data which may not be possible to process within a short study time. Hence, a smaller area with a suitable size around *25000* km^2 called Bortala region within Northwest China is chosen as the study area of this research.

1.1. Aim

The aim of this study is to find suitable dam sites in Bortala region, Northwest China for helping decision makers choosing the best site to build dam(s). The suitable dam sites will be generated mainly based on the environmental conditions while irrigation water supply is put as the main usage of these dams.

Besides the aim of this study, there are several objectives which need to be achieved in this study:

1) Compute the suitability of dam construction according to different criteria and generate a suitability map.
2) Propose several dam sites with relatively high suitability.
3) Compute the cross sections and some other characteristics, including volume of reservoir, height and width of dam, of proposed dam sites.

1.2. Study Area

The study area of this research is Bortala Mongol Autonomous Prefecture (Bortala for short). The area is around *27,000* km^2 with *0.4 million* population, in Xinjiang Uyghur Autonomous Region (Xinjiang for short), Northwest China. The location of study area is shown in *figure 1*.

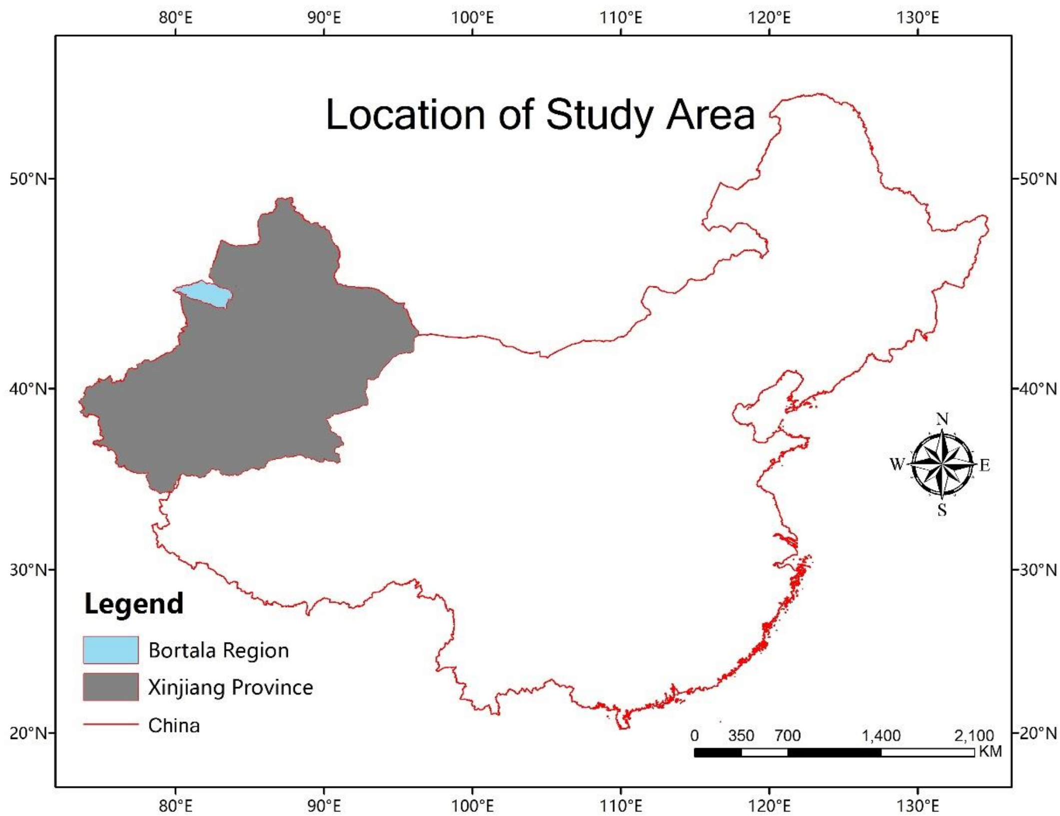

Figure 1 Location of study area with respect to Xinjiang Province and to whole China

Bortala is located in the southwestern part of the Jungar Basin where there is a mountain range at the Northwest Bortala and another mountain range at the Southwest Bortala. As *figure 1* shows, Bortala is a border region in Northwest China. It borders on Kazakhstan in the west and north with an international border of *385 km* (Wikipedia 2016). There are *2* large closed lakes, Ebi-Nur Lake with salt water and Sayram Lake with freshwater, inside Bortala.

In the aspect of climate, Bortala is considered as an area with arid climate, which also called desert climate according to the Köppen-Geiger climate classification (Climate-Data.Org 2016). And there are *271* records out of *431* cities are classified as desert climate in Xinjiang Province which means the desert climate is the dominate climate type for Xinjiang Province (Climate-Data.Org 2016). Hence, Bortala is representative for Xinjiang Province on the aspect of climate type. In Bortala, the average annual temperature is *7.4 ℃* and the average annual precipitation is *192 mm* in the period of *1982 to 2012* (Climate-Data.Org 2016). According to a study on climate change in Bortala in the period of *1960 to 2006*, the annual temperature in the prefecture has been increasing *from 1960 to 2006* and the annual precipitation in the prefecture has been increasing since *1980* (Yu et al. 2006).

The main crop cultivated in Bortala is cotton and it contributes around *54%* of the total crop production in the prefecture. According to a research on irrigation water requirement and supply risk in the arid region of Northwest China *1989 to 2010*, the irrigation water requirement showed an increasing trend in the whole study and it’s mainly caused by increasing cultivated area of cotton, which has larger irrigation water requirement compared to other crops (Shen et al. 2013).

2. Background

2.1. Dam site selection

Dam site selection is a complex problem since it is affected by, as well as it influences, many different factors including both environmental factors and human society factors. Meanwhile, dams can be built for different purposes which can also lead to a wider diversity of factors and influences. Many papers have been published in the field of dam site selection using different factors as criteria, as well as using different method to measure the influences from each criteria.

In *2013*, a research was carried out in Central Iran using Geographic Information System (GIS) to choose the suitable sites of underground dam, which is used to trap groundwater and store it below the surface, combined with aqueducts (Alaibakhsh et al. 2013). At the same year, another research group published a paper about a case study in Western Iran aiming on site selection of earth dam using Analytic hierarchy process (AHP), which is one of the most flexible and simplest methods, but powerful at the same time, for solving multi-attributes decision-making (MADM) problems (Yasser et al. 2013). In *2015*, from the same research group focusing on the same study area, a decision making support system for dam site selection was proposed and applied which contains fuzzy AHP and VIKOR method for Multi-Criteria Decision-Making (MCDM) (Yasser et al. 2015). At the same year, another research is carried out in Northwest Saudi Arabia using a combination of Remote Sensing (RS) techniques and Geographic Information System (GIS) on dam site selection (Abushandi & Alatawi 2015).

Besides the variety of methods used in the field of dam site selection, the factors considered also vary a lot between each individual case. In the Northwest Saudi Arabia case, four criteria was chosen for best dam site location, in which catchment slope, land cover type, soil type and soil infiltration rate are involved (Abushandi & Alatawi 2015). In the Western Iran case from *2013*, *9* categories of criteria and *11* sub criteria are considered as most important attributes for locating the earth dam site in this study. Specifically, criteria used in this study are: topographic conditions; economic development; health dam site; river flow regime; annual yield (volume of water); volume of reservoir; annual volume of sediment; probable max flood; average annual evaporation; access to materials and facilities; overall cost; water quality; dam body and reservoir damage; probable dam break; environmental impact; social impact; political impact (Yasser et al. 2013). In the western Iran case from *2015*, 18 criteria are chosen for the best dam site selection which involve:

Health of dam site; Overall cost; Annual yield; Topographic conditions; access to materials and facilities; economic development; water quality; damage of dam body and reservoir; volume of reservoir; river flow regime; water diversion and transfer; annual volume of sediment; water diversion and transfer; annual volume of sediment; probability of dam break; probable maximum flood; average annual evaporation; environmental impacts; social impacts; political impacts (Yasser et al. 2015). While all studies mentioned above put more attention on influences on constructing dam, there are studies focused more on influences of constructing dam. For example, a research from *2003* proposed *13* indicators for dam site selection on the purpose of sustainable development and these indicators involve: reservoir surface area; water retention time in reservoir; biomass flooded; length of river impounded; length of river left dry; number of downriver tributaries; likelihood of reservoir stratification; useful reservoir life; access roads through forests; persons requiring resettlement; critical natural habitats affected; fish species diversity and endemism; and cultural property affected (Ledec & Quintero 2003).

It is difficult to define a certain number of criteria for dam site selection generally as the objectives of decision makers can vary from strongest hydropower utilization, like the world's largest power station Three Gorges Dam, to ministrant usage in small villages such as irrigation or aquaculture. Meanwhile, the specific local conditions, in both of environment and human society, and decision makers' tendencies between high speed development or sustainable development also brings difficulties in a general datasets of criteria for dam site selection. However, there are still some criteria related to dam safety and essential requirements that are always important for dam construction, such as slope and hydrological conditions.

2.2. Multi-Criteria Decision Making

Making decisions in real world sometimes can be complex because of the complexity of reality. One of the possible solutions is Multi-Criteria Decision Making (MCDM) which refers to making decisions in the presence of multiple criteria that usually conflicting each other (Zavadskas et al. 2014). Recent foundations for MCDM were developed in the 1950s and 1960s. In late 1970s, the abbreviation of MCDM was proposed for the first time (Zionts 1979). By using MCDM, complex problems can be divided into a certain number of smaller parts which are easier to measure or make judgement on. There is no unified MCDM method can be followed stepwise. Instead, a big variety of different methods are proposed for MCDM. One of the most used ones is called Analytic Hierarchy Process (AHP).

2.2.1. Analytical Hierarchy Process

Analytical Hierarchy Process (AHP) is a simple and powerful decision making model as well as one of the most used MCDM method. From *2000 to 2014*, nearly *32.6%* of published papers used AHP as a MCDM method out of *9* MCDM methods in *393* MCDM related published papers in total (Mardani et al. 2015). Both tangible and intangible criteria can be measured with absolute scales by using AHP (Saaty 2013). A typical AHP hierarchy is shown in *figure 2*.

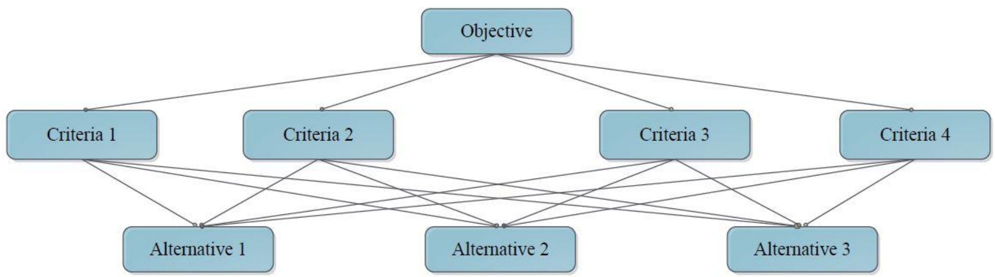

Figure 2 A typical AHP hierarchy

The procedure of AHP can be divided as three parts which include identifying a hierarchy of objectives, criteria and alternatives; pairwise comparison of criteria; an integration with result from pairwise comparison as relative importance over all levels of hierarchy (Saaty 1988). According to Saaty, for the consistency of comparison between different criteria, a scale with continuous 9 absolute numbers which indicates the intense of relative importance between two criteria was proposed. In this scale, a value of *1* indicates "equal importance" and a value of *9* indicates factors with "extreme importance" over another factor (Saaty 1977). *Table 1* shows the fundamental scale of absolute numbers with brief explanation.

Table 1Scale for pairwise comparison (Saaty 2008)

Intensity of Importance	Description	Explanation
1	Equal importance	Two activities contribute equally to the objective
2	Equal to moderate importance	
3	Moderate importance	Experience and judgement slightly favor one activity over another
4	Moderate to strong importance	
5	Strong importance	Experience and judgement strongly favor one activity over another
6	Strong to very strong importance	
7	Very strong importance	An activity is favored very strongly over another; its dominance demonstrated in practice
8	Very strong to extremely strong importance	
9	Extremely strong importance	The evidence favoring one activity over another is of the highest possible order to affirmation
Reciprocals of above	If activity i has one of the above non-zero numbers assigned to it when compared with activity j, the j has the reciprocal value when compared with i	

Applications of AHP with Geographic Information System (GIS) has been operated a lot since the beginning of 21st century (Mardani et al. 2015). The integration of AHP with GIS gives an efficient and user friendly way for solving complex problems as it is a combination of decision making support method and tools with powerful capabilities of mass data computation, visualization, and mapping (Marinoni 2004). The implementation of AHP in ArcGIS, which is one of the most used GIS software globally, can be summarized as following procedure: definition of objective; identification of criteria; data collection and preprocess; digitization of criteria and convert all data into raster data; classification of raster datasets; creation of preference matrix; determination of criteria weights according to calculation based on preference matrix; weighted summation of criteria raster datasets as result (Marinoni 2004).

2.3. Drainage Network Extraction

Drainage networks are areas of land from which all surface water converges and be transported to other locations through fluvial process. For selecting location for constructing dams, it is necessary to have drainage networks extracted.

The basic idea of popular drainage networks extraction method comes from *1984*. By using digital elevation data as input, F. O'Callaghan and M. Mark proposed an effective method for extracting drainage networks. In their work (O'Callaghan & Mark 1984), the proposed method consists of a sequence of procedures, which contains elevation smoothing (optional), flow direction assignment, drainage feature assignment, drainage basin labelling, interior pit removal, drainage accumulation, drainage feature assignment, and channel link recovery. In *1988*, the idea of extracting drainage networks using digital elevation data from O'Callaghan and Mark was carried forward by Jenson and Domingue with more types of conditions and corresponding solutions involved.

2.3.1. Sink fill

Almost every DEM contains sink area that stands in the way of flow routing. These sinks can be caused by systematic data error like limitation of resolution as a source of systematic data error or natural landform like Karst Topography. The process of filling sink is called elevation smoothing or filling depressions. The main purpose of elevation smoothing is to reduce the number of artificial depression generated by data collection system (O'Callaghan & Mark 1984). In O'Callaghan and Mark's work, a weighted sum of a cell in the center of a 3 * 3 matrix was used as the method of elevation smoothing. The most used procedures of sink fill, taken as the official algorithm in the most wide-used GIS software ArcGIS (ESRI 2016), were developed in *1988* by Jenson and Domingue. In their study, depression cells are raised to the nearest elevation of connected neighborhood cells (Jenson & Domingue 1988).

2.3.2. Flow direction

One of the most important idea of O'Callaghan and Mark's work is the flow direction assignment, also known as "Deterministic 8" and of which the modified method (Jenson & Domingue 1988) is still used in the most wide-used GIS software ArcGIS as flow direction. The data used for a flow direction calculation is elevation value. In O'Callaghan and Mark's work, the flow direction for a cell was calculated within a 3 * 3 neighborhood matrix. The direction is from central cell to the cell with steepest downslope in the 8 neighbors and this minimum elevation must be smaller than the elevation of central cell. Considering the path length is larger in the diagonal directions, the elevation values for the 4 neighbors in the diagonal directions are adjusted by a 1/√2 factor. The values in the flow direction matrix are defined by *equation 1*.

$$DDIRN(i,j) = 2^{d-1} \quad for\ d = 1,2,\dots,8,$$

$$= 0 \quad for\ an\ undefined\ direction,$$

Equation 1 Flow direction

In which, d is the flow direction. The first one of flow direction is northeast neighbor and it goes clockwise. So the coded direction in the 3 * 3 neighborhood matrix of the central cell (c) is as the following matrix (O'Callaghan & Mark 1984).

$$\begin{matrix} 64 & 128 & 1 \\ 32 & c & 2 \\ 16 & 8 & 4 \end{matrix}$$

In later work of Jenson and Domingue, the concept of flow direction was carried forward with solutions for 2 more condition than the original method of O'Callaghan and Mark. The first new condition is when two or more cells have the same value as largest distance-weighted drop. And the second new condition is all 8 neighborhood cells are larger or equal in elevation value compared to central cell (Jenson & Domingue 1988). In ArcGIS, the direction coding is slightly changed compared to the ideas from these 2 pieces of works. The direction coding starts at east and also continues in clockwise. The coded direction is as following (Jenson & Domingue 1988).

$$\begin{matrix} 32 & 64 & 128 \\ 16 & c & 1 \\ 8 & 4 & 2 \end{matrix}$$

2.3.3. Flow accumulation

Flow accumulation is a process that assign every cell with a value equals to number of cells flow into it (O'Callaghan & Mark 1984). To be more specific, the process of flow accumulation is based on the data of flow direction, and rules below are followed.

1) Starting from the first cell, track cells along the flow direction until a stop point (lowest elevation in *3*3* matrix) of flow or the edge of DEM;
2) Cells along the tracking route gain 1 accumulated value;
3) When cross with other tracking routes, the accumulated value of the other tracking route is added to current tracking route.
4) Calculation stops when every cell in flow direction data matrix is calculated.

Catchment area can be calculated by multiplying value in accumulated flow matrix with the area of a single cell.

2.3.4. Drainage network identification

The drainage channels are defined as cells with accumulated flow exceeding a user-defined threshold (O'Callaghan & Mark 1984). Different methods for quantitative description of drainage networks are studied by many researchers. The most used ones are Strahler method (Strahler 1957) and Shreve method (Shreve 1966). In both of these two methods, drainage network is idealized as a tree with a strong root and slimmer branches. But the identification of different levels of branches are not the same within these two methods.

Strahler method follows the rules bellow (Strahler 1957).

1) Level *1* is assigned to those streams without any branch;
2) Level *2* is assigned to downstream where two of the level *1* upper streams meet;
3) Level *n* is assigned to downstream where two of the level *n-1* upper streams meet;
4) If two upper streams with different levels meet, then the downstream is assigned with the higher level in the two upper streams.

Shreve method also follows the same rule for the level *1* streams assignment. But for other levels, the rules are different. When a level *n* upper stream meets a level *m* upper stream, the downstream is assigned with level *n+m*. That is to say, when all streams ideally flows to the same outlet, then the level value of the last downstream, where outlet is, equals to the total amount of level *1* streams (Shreve 1966).

3. Data and methodology

3.1. Data

3.1.1. Digital Elevation Model (DEM)

The DEM used in this study is Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global DEM, a product of The Ministry of Economy, Trade, and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA). And the data was generated in 2006 with resolution of *30 m*. The dataset is acquired on http://reverb.echo.nasa.gov/reverb/#utf8=%E2%9C%93&spatial_map=satellite&spatial_type=rectangle .

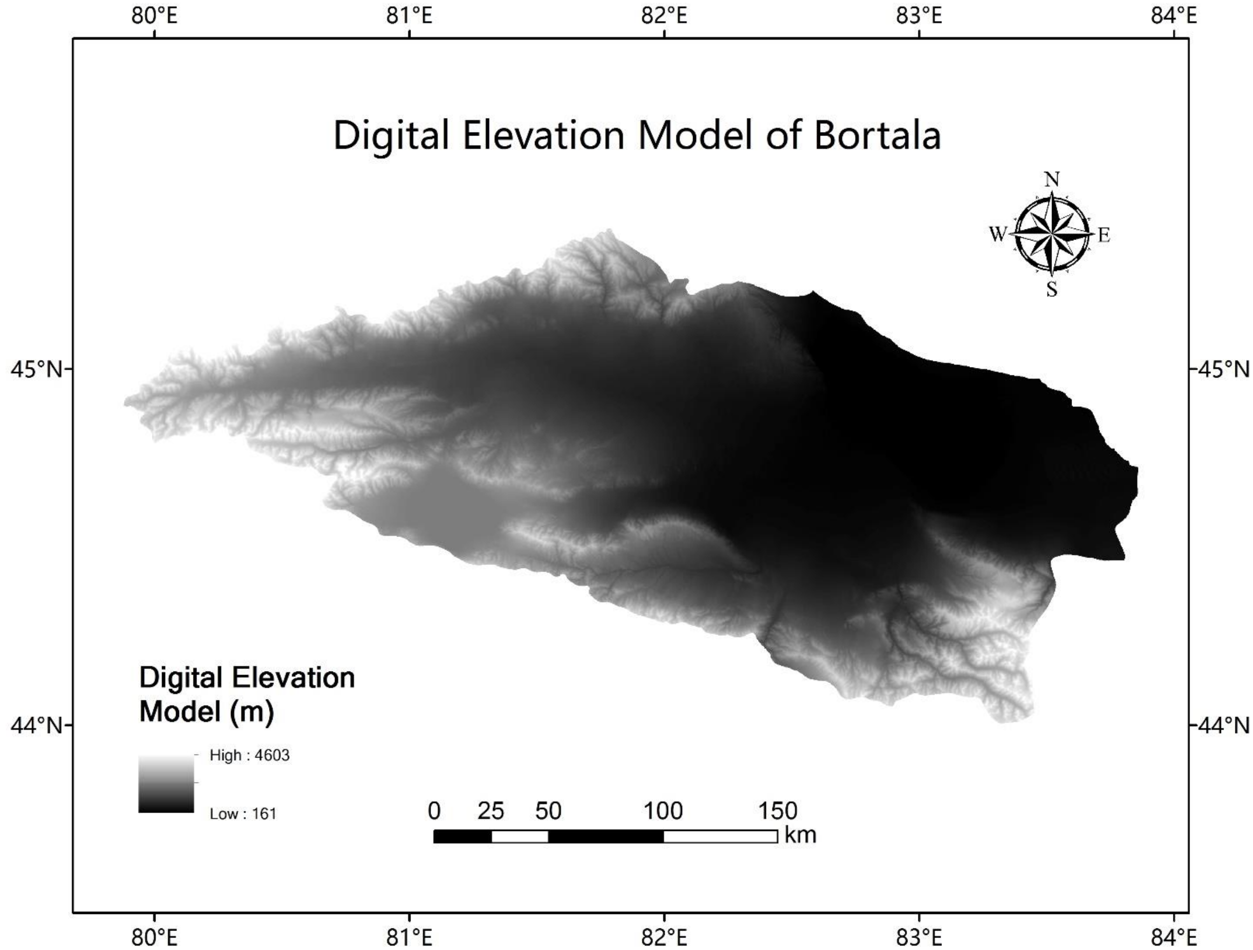

Figure 3 Digital Elevation Model of Bortala in 2006

According to the DEM shown in *figure 3*, the topography of Bortala is low in Northeast and high in other parts with the elevation range of *161 m to 4603 m*. Northwest Bortala and Southwest Bortala are mountainous as well as Southeast Bortala.

3.1.2. Soil type

The soil data used in this study is ISRIC-WISE global dataset of derived soil properties on a *5 by 5 arc-minutes* (*9 km*) grid from 2012. This dataset is acquired from ISRIC/WDC-Soils, http://www.isric.org/data/data-download/.

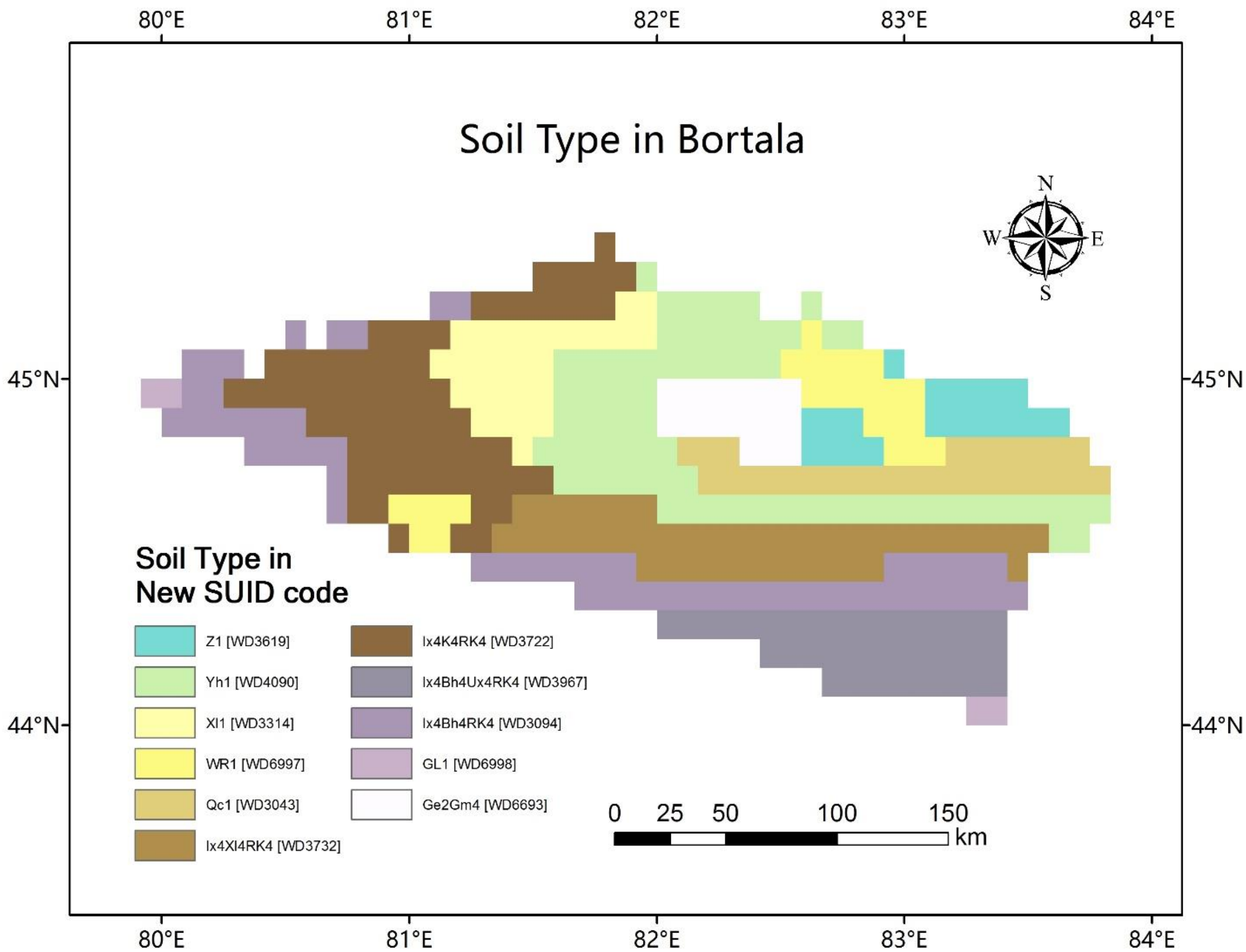

Figure 4 Soil type data of Bortala in 2012

In *figure 4*, the New SUID code for soil type contains soil properties information including composition percentage which can be used for classification of soil texture later.

3.1.3. Land cover

Land cover data used in this study is from Global Ecological Land Units map at *250 m* resolution released by United States Geological Survey (USGS) and ESRI in 2014. This dataset is acquired from http://rmgsc.cr.usgs.gov/outgoing/ecosystems/Global/.

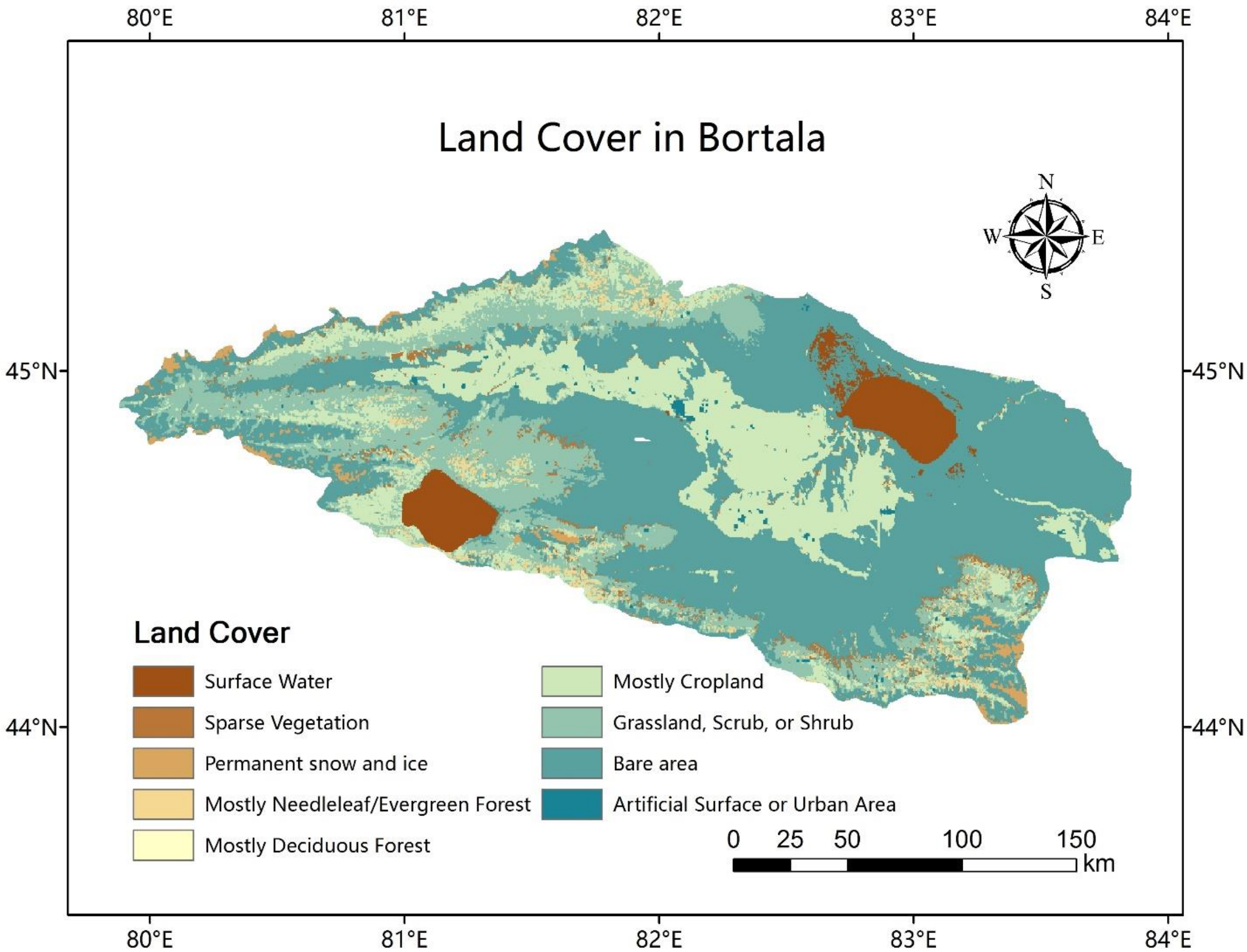

Figure 5 Land cover data of Bortala in 2014

According to the land cover data shown in *figure 5*, more than half area of Bortala is covered by farmland or bare land and these area are located in Northeast and Central Bortala. Northwest, Southwest and Southeast Bortala are covered by mixture of forest, shrub, grassland and cropland with small areas of permanent snow and ice.

3.1.4. Geological layer data

Geological data used in this study is Global Ecological Land Units map at *250 m* resolution for *2014* and it is acquired from USGS http://rmgsc.cr.usgs.gov/outgoing/ecosystems/Global/.

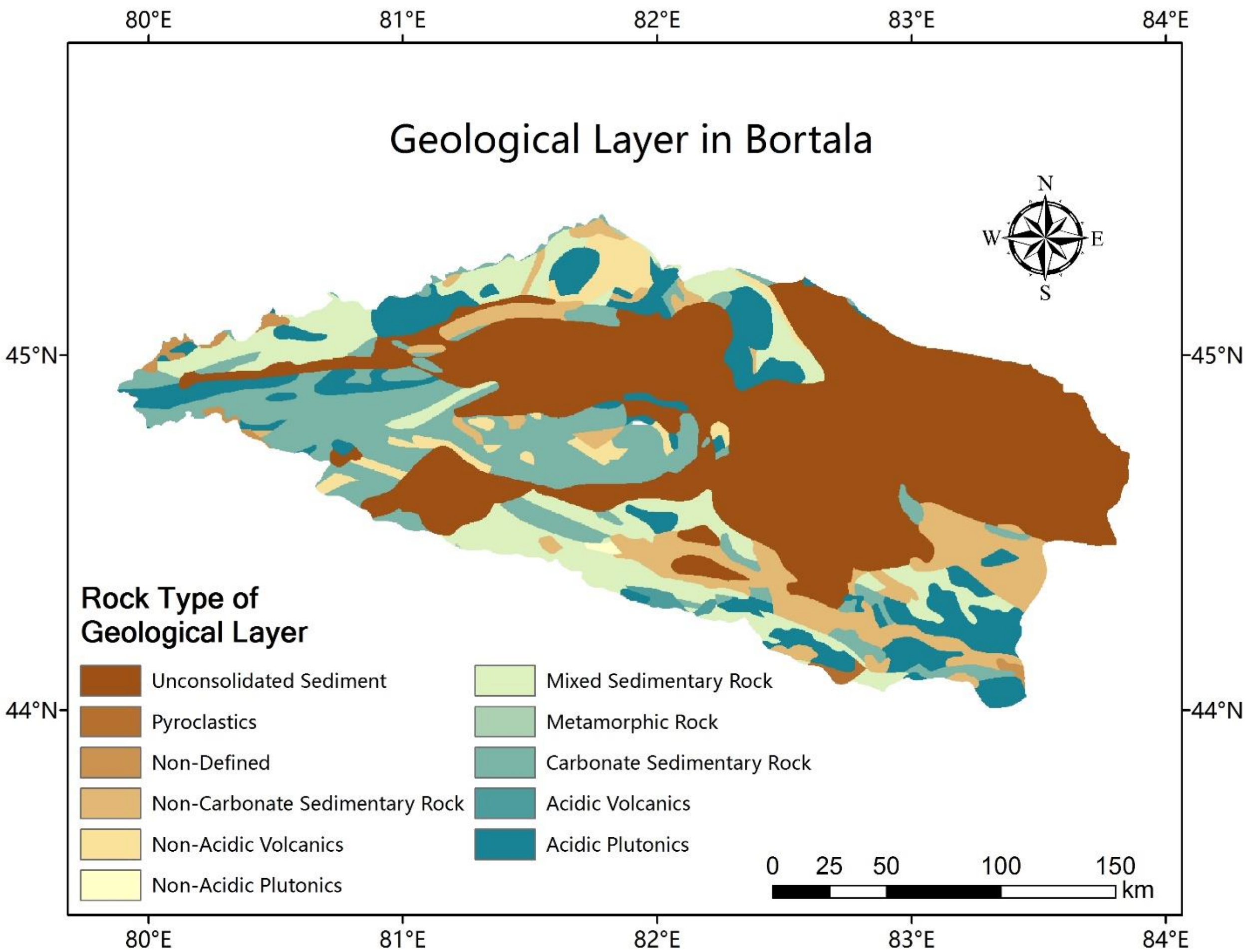

Figure 6 Geological layer data of Bortala in 2014

According to the geological layer data shown in *figure 6*, Northwest Bortala and most of the Central Bortala are covered by unconsolidated sediment which is the most common rock type of geological layer in Bortala. The second common rock type of geological layer in Bortala is carbonate sedimentary rock, which is mainly exist in Southwest Bortala. The other types are more mixed and exist in Northeast and Southwest Bortala.

3.1.5. Predicted precipitation data in responding to climate change

Based on the availability of data, predicted precipitation data used in this study is Downscaled IPPC5 (CMIP5) data using Global Climate Model (GCM) CCSM4 under scenario Representative Concentration Pathway (RCP) 6, which is the aim scenario for sustainable development on greenhouse gas concentration. There are two period of time with available predicted precipitation in CMIP5 dataset including *2041 to* 2060 and *2061 to 2080*. And the average annual precipitation data for *2041* to *2060* is chosen since it is the available data for nearest future. The spatial resolution for this dataset is at *30-seconds* (of a longitude/latitude degree) which is around *900 meters*.

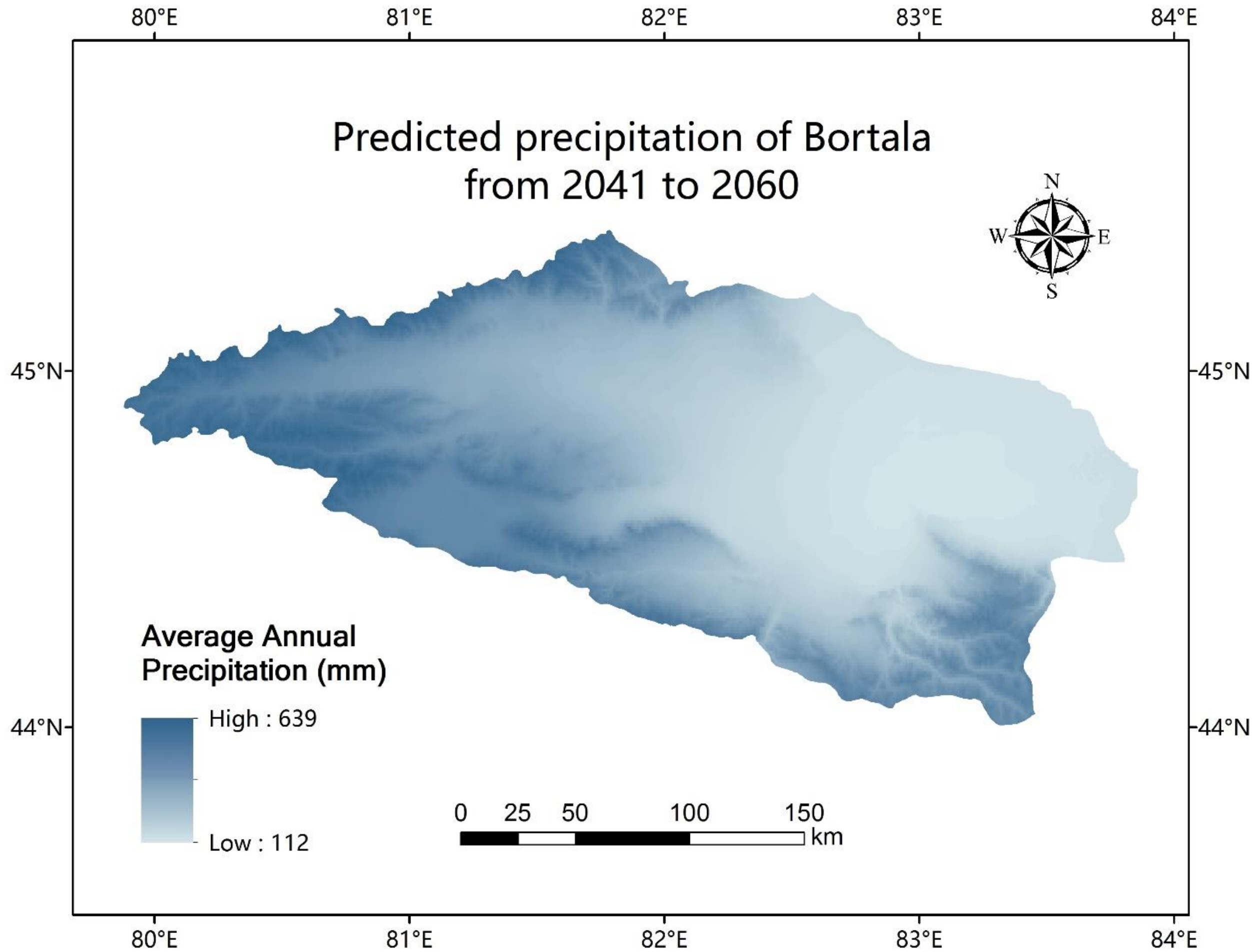

Figure 7 Predicted annual precipitation map of Bortala for 2041 - 2060

According to the predicted annual precipitation data map in *figure 7*, the predicted average annual precipitation of Bortala for *2041 to 2060* ranges from *112 mm to 639 mm*. The general distribution trend is low in Northeast and Central Bortala and high in Northwest, Southwest and Southeast Bortala.

3.2. Methodology

ArcGIS10.3.1 is used for processing imagery datasets including DEM, soil, geological layer, land cover, precipitation and products from these datasets. Excel 2013 is used for simple statistical analysis and creation of tables and histograms.

3.2.1. Select the criteria

Based on literature review of previous studies, specific conditions of Bortala region, and data availability, *6* criteria considered as main factors are chosen for this study including: level of drainage network, precipitation amount, topographic conditions (slope), soil type, land cover, and resistance of geological layer.

Drainage network provides necessary runoff water for dam function, different levels of drainage network indicates different amount of runoff water when the relation of streams are upper stream tributaries and downstream main streams.

Precipitation is the main source of runoff water recharge. In another word, precipitation has a positive influence on runoff amount. Hence, precipitation has a positive influence on dam function when there is no natural disaster caused by extremely strong precipitation such as landslides or floods.

Slope is one of the main factors that influence the safety of dams, since large degree of slope has a higher risk of landslides and gives more pressure on foundation of buildings.

Soil types can be divided by the texture of soil, which leads to different soil infiltration rate hence different influences on runoff amount. Therefore, soil type has an influence on dam function in the aspect of runoff amount.

Another factor that influences dam safety is the resistance of geological layer, which is decided by the rock type of geological layer.

Since the focus of this study is on dams that provides agricultural irrigation water, the priority of changing agricultural land into dam construction area is lower than constructing dam on shrub or forest cover.

3.2.2. Data preprocess

3.2.2.1. DEM mosaic

Bortala region is covered by *9* pieces of ASTER DEM. For a more efficient analysis, raster mosaic is implemented to make these *9* individual DEM into one.

3.2.2.2. Projection & Reshape & Uniform resolution

The original projection of most data used in this study is GCS_WGS_1984, which is a geographical coordinate system uses degree as angular unit. To change geographical coordinate system (with spherical coordinates) into projected coordinate system (with planar coordinates), reprojection is implemented. The new projection is WGS_1984_World_Mercator, which uses meter as linear unit. All datasets are reshaped from a national or global scale into a regional scale of study area. For applying raster calculation in later procedure, all datasets of criteria are adjusted into a uniformed resolution which is the same as the resolution of DEM data which is *30 m*.

3.2.2.3. Precipitation

3.2.2.4. Predicted precipitation

In Bortala region, the predicted average annual precipitation for *2041 to 2060* varies from *112* mm to *639* mm. A classification is carried out to make individual values into categories related to preferential level of constructing dam on. The classes indicate the average annual precipitation in the level of *112 mm – 200 mm, 200 mm – 300 mm, 300 mm – 400 mm, 400 mm – 500 mm, and 500 mm – 639 mm*. *Table 4* shows the preference value for different precipitation classes.

3.2.2.5. Soil

The soil properties data for Bortala is derived from a global soil properties image of *2012*. There are *11* classes of soil properties in this region. The classes are in a global unique SOTER code through which the percentage mass of components, which are sand, silt and clay, can be searched. According to the different percentage of this three types of soil, a simplified classification is done for tracing the different soil infiltration in study area.

From a reference of a textural triangle describes the relative proportions of sand, silt and clay in various types of soils (University of Hawaiʻi - College of Tropical Agriculture and Human Resources n.d.), in combination of basic infiltration rates for different types of soil (Brouwer et al. 1990) and the soil composition data for Bortala, 4 classes are defined for soil type in Bortala. These classes are sand, sandy loam, loam, and clay loam. According to the basic infiltration rates of each type of soil (Brouwer et al. 1990), the preference values for constructing dam are assigned to soil types which can be found in *table 8*. According to the reclassified soil texture map in *figure 8*, soil with lowest soil infiltration, which is clay, are distributed mainly in Northwest, Southwest and Southeast Bortala while a smaller area in Central Bortala also has this type of soil.

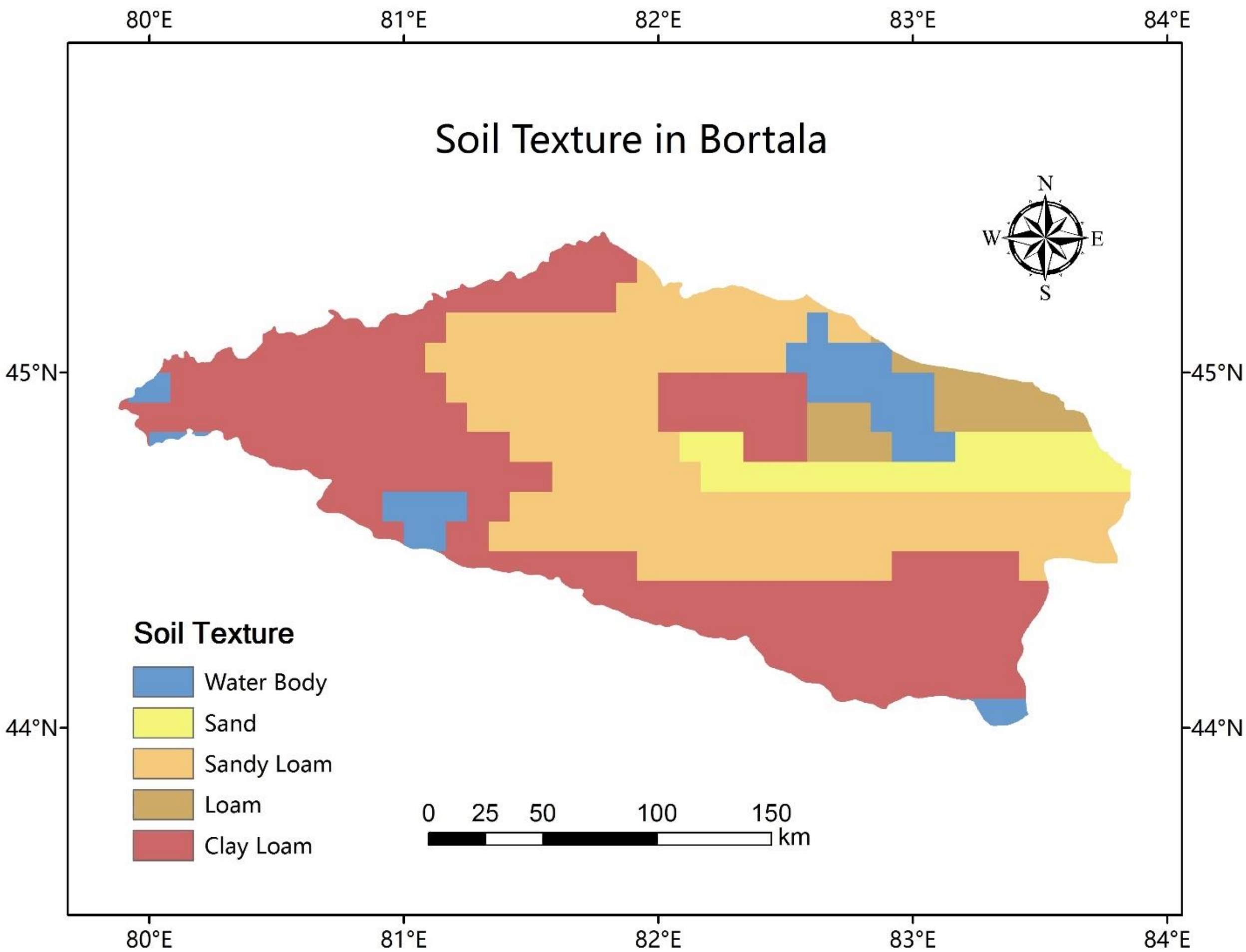

Figure 8 Soil texture map of Bortala in 2012

3.2.2.6. Slope

Slope can be described in two different ways. One is degree of slope, which is used in this study, indicates the angle between ground surface and horizontal plane. The other one is percentage slope which indicates the percentage ratio of elevation change on horizontal distance change. Slope can be generated according to DEM using the Pythagorean Theorem. As dealing with slope in each cell individually leads to highly mixed area, which can cause confusion in suitability map and dam site selection, an aggregated generalization, which use mean value for generalized value, is carried out on the generated slope layer. Several matrix scales for aggregation are tested and compared with profile graph of valleys based on DEM, and finally the aggregation is applied with the mean value of neighbor cells in a *7 * 7* matrix that makes the resolution rise to *200 m* scale.

For constructing dam, different slope thresholds has been chosen in previous studies such as less than *10* percent which equals to *5.71* degree (Singh et al. 2009), less than *3* percent which equals to *1.72* degree (Abushandi & Alatawi 2015). In a research of watershed resource prioritization, slope is categorized as gently sloping (*5* degree), moderately to steeply sloping (*5* to *18* degree), and steep to very steep sloping (more than *18* degree) (Adinarayana et al. 1995). In this study, slope is classified into *6* classes, in which every *1* degree under or equal to *5* degree stands for a category and slope larger than *5* degree is defined as another category as shown in *figure 9*. Classes and preference value are shown in *table 5*.

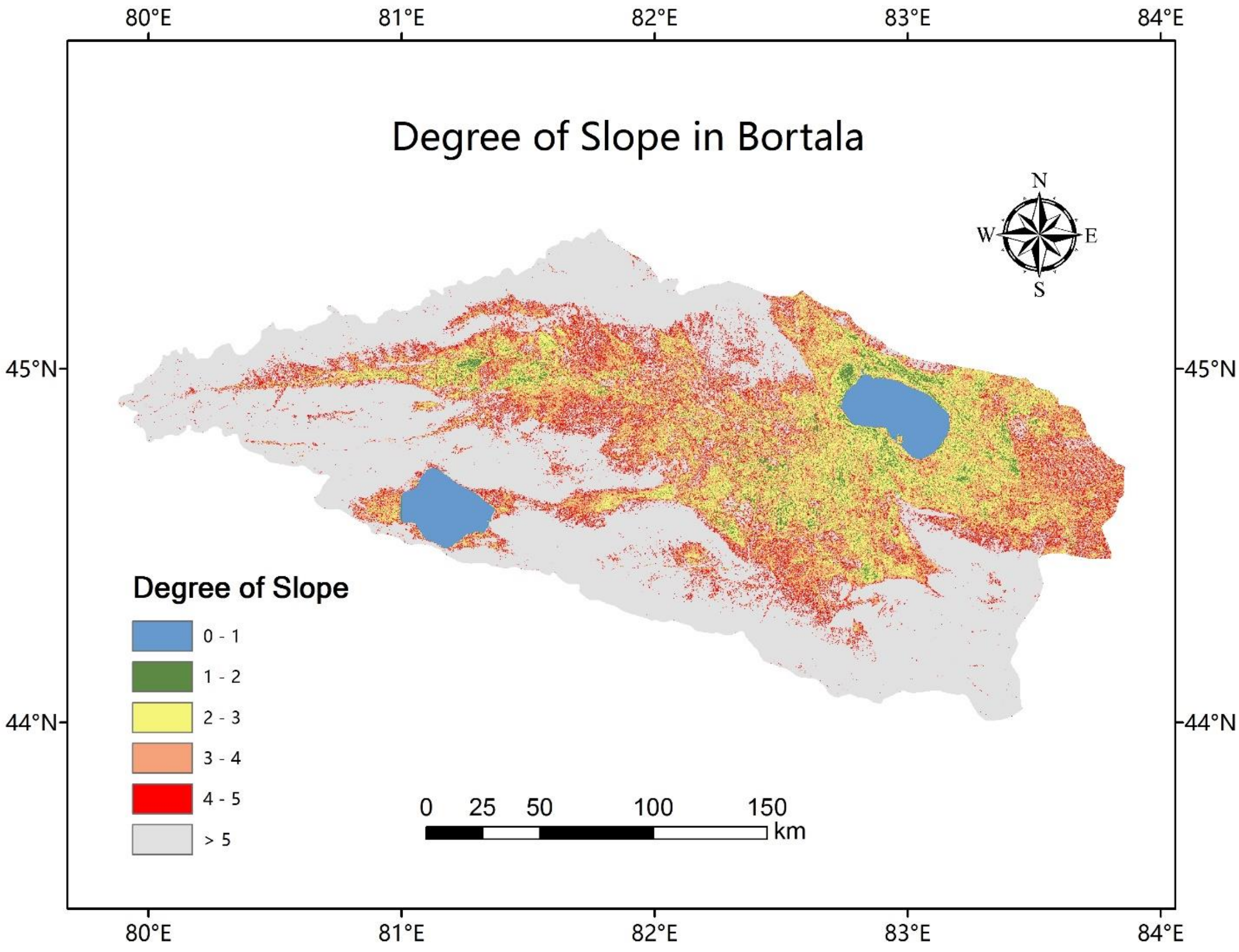

Figure 9 Slope map of Bortala in 2006

3.2.2.7. Geological layer

Among various natural factors that affect dam constructing, none are more important than the geological ones (Lashkaripour & Ghafoori 2002). In a summary of dam performance statistic, foundation problems are found out to be the most common causes of dam failure (Wyllie 2003). Competent rock foundations are rocks with relatively high resistance to erosion, percolation and pressure (US Army Corps of Engineers 2005). For a desirable foundation, igneous rocks such as granite; metamorphic rocks such as quartzite; sedimentary rocks such as thick-bedded sandstones, flat-lying sandstones, and limestones are among the most satisfactory materials (Baban & Wan-Yusof 2003). According to the characteristics of different classes of rocks, which can stand for the preference for constructing dam on, the original geological data is reclassified from *11* classes in to *4* classes as shown in *table 2* and *figure 10*. The preference values for each class are shown in *table 6*.

Table 2 Geological layer reclassification

Original Classification	Modified Classification
Non-Defined	Low Resistance
Pyroclastics	
Unconsolidated Sediment	
Mixed Sedimentary Rock	Slightly Low Resistance
Carbonate Sedimentary Rock	
Non-Acidic Volcanics	
Acidic Volcanics	Moderate Resistance
Non-Carbonate Sedimentary Rock	
Non-Acidic Plutonics	
Acidic Plutonics	High Resistance
Metamorphic Rock	

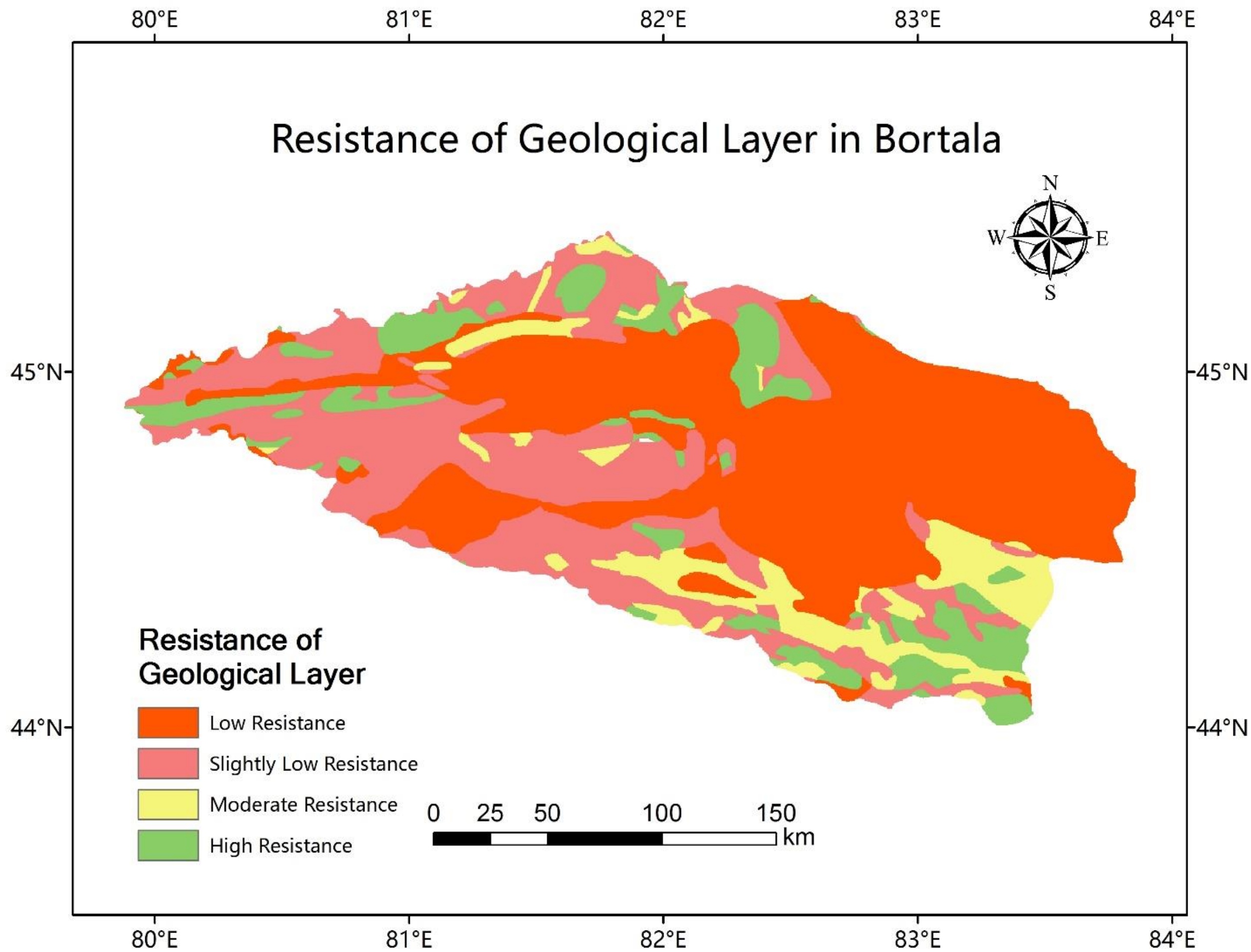

Figure 10 Resistance of geology layer map of Bortala in 2014

3.2.2.8. Land cover

Land cover plays several roles in dam site selection. Firstly, land cover greatly modifies the effect of rainfall which gives land cover a place in influencing soil erosion (Adinarayana et al. 1995), and high soil erosion area makes a weak foundation for constructing dam (Baban & Wan-Yusof 2003). On the other hand, constructing dam leads to expropriation of land which have different economical cost according to land cover type (Baban & Wan-Yusof 2003). In this study, land cover is reclassified from *9* classes into *6* classes, as shown in *table 3* and *figure 11*, according to the preference of constructing dam on. In addition, a Boolean classification is done for excluding dam sites on water or settlement. The preference values of land cover classes are shown in *table 7*.

Table 3 Land Cover reclassification

Original classification	Modified classification
Mostly Needle leaf/Evergreen Forest	Forest
Mostly Deciduous Forest	
Grassland, Scrub, or Shrub	Shrub and Herb
Sparse Vegetation	
Mostly Cropland	Farmland
Bare Areas	Bare Land
Surface Water	Water
Permanent Snow and Ice	
Artificial surfaces or Urban Area	Settlement

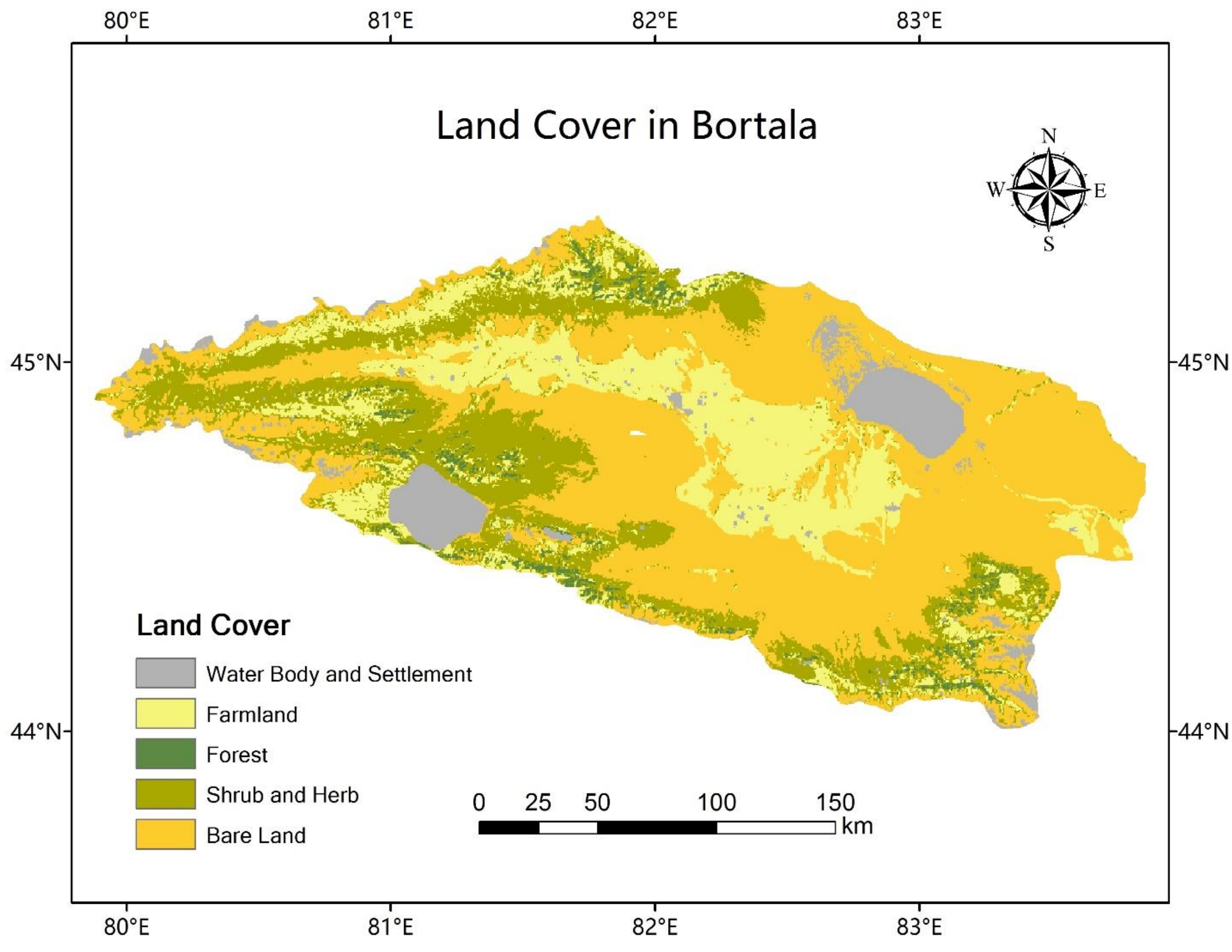

Figure 11 Land cover map of Bortala in 2014

3.2.3. Sink fill

DEM is a topographical model with elevation records of cells in a certain size. However, there is still a potential of existence of sunk areas because of data error or landform like Karst Topography. Data error is mainly caused by the resolution limitation of DEM on both vertical and horizontal direction or system error during the generation of DEM (Hao et al. 2010). Due to the existence of these sinks, unreasonable flow direction may be generated during the calculation. If these sinks are not filled by technical process, then the generated drainage network will not be continuous.

The first step of sink fill is generating flow direction. Flow direction of a cell is the direction in which water flows out from this cell. In ArcGIS, flow direction is determined by the elevation of a cell itself and the elevation and position of 8 neighborhood cells. The flow direction is always from each cells to the steepest downslope neighbor. To calculate which neighbor should be the direction target (the steepest downslope neighbor), a distance weighted elevation difference is used. It means when the neighbor cell is located at a due direction of the source cell, then the distance would be 1 times cell size, while diagonal directions are calculated as $\sqrt{2}$ times cell size.

Sinks are computed according to the generated flow direction. A sink is defined as the area where contains unreasonable flow direction. To be more specific, a sink is a cell or set of spatially connected cells whose flow direction cannot be assigned as one of the eight valid values in a flow direction raster. This can occur when all neighboring cells are higher than the processing cell or when two cells flow into each other, creating a two-cell loop. If a cell or a set of continuous cells with exactly same elevation has higher elevation value comparing to surrounding cells, then the water inside this cell or set of cells cannot flow out. The elevation of these sinks are reset to the most proximate elevation in their neighboring cells. After these sinks are filled, new sink calculation is carried out for the possible existence of new sinks. The progress of sink fill is an iteration progress. It stops when all the sinks are filled and no more new sinks can be found.

3.2.4. Compute drainage network

In one hand, drainage network, as the channels water flows through, is essential for dam construction; in the other hand, the stream level, which indicates capacity of transporting water, is one of the important criteria of dam site selection.

Flow direction calculation based on filled DEM is carried out as the foundation of flow accumulation. Accumulated flow calculation is based on the idea that water in a cell always flows to the comparatively steepest downslope neighbor cell out of 8 surrounding neighborhood cells.

Since every cell is with a certain amount of water inside, which can be defined as 1 unit of water amount. Accumulated flow of a cell can be derived from water of how many cells are flow into this cell. According to a user-defined threshold, cells with a flow accumulation value higher than threshold are marked as *1*, otherwise value *0* is assigned. The threshold here works as a divider between flows considered as streams and small flows that more likely to disappear due to natural process like infiltration and evapotranspiration. A series of thresholds are tested by comparing generated stream network with the water features on world topographic map provided by ESRI on ArcGIS Online. After the comparison, a threshold value of *30000* accumulated unit is chosen as the stream definition threshold in this study. All cells with accumulated flow value equal or more than *30000* accumulated units are considered as streams. Then, streams are segmented according to flow direction to generate stream links. Stream link records structure information of a drainage network. In a stream channel, all streams can be considered as links, and the points of intersection between link sections are considered as junctions. To compute stream link, a raster layer of stream network and a raster layer of flow direction are used.

After stream link is generated, flow direction can be adjusted with stream link and a polygon layer of lakes. This process is to make cells within a lake flows towards the closest stream (ESRI 2011) in that lake so that drainage points, which are generated in later procedure, in lakes are located more accurately. Based on the adjusted flow direction, several previous steps are operated again including flow accumulation, stream definition, and stream segmentation.

Then, drainage order is assigned to each section of streams. Drainage order is a numeric order assigned to a linear stream network that indicates different level of flow accumulation. The more the flow accumulation of a stream is, the higher level of drainage order is assigned to this stream. In another words, higher drainage order indicates more tributaries are flowing into the stream which gives a higher possibility to harvest more water. Drainage order is used in later process as a criteria on dam site selection. Since the whole Bortala region is processed as one integrated image, Shreve method (Shreve 1966) will lead to large number of drainage order levels. Hence, Strahler method (Strahler 1957), which will lead to few different drainage order levels, is chosen in this study. By using Strahler method (Strahler 1957), drainage order with 5 order levels is generated.

To give a more direct view on flow conditions, flow length is generated along downstream direction which means the flow length value indicates how many cells has the flow from one certain cell passed before the flow stops or reaches the edge of dataset. Based on a converted vector drainage order layer and a raster flow length layer, a drainage network map is presented in *figure 12*.

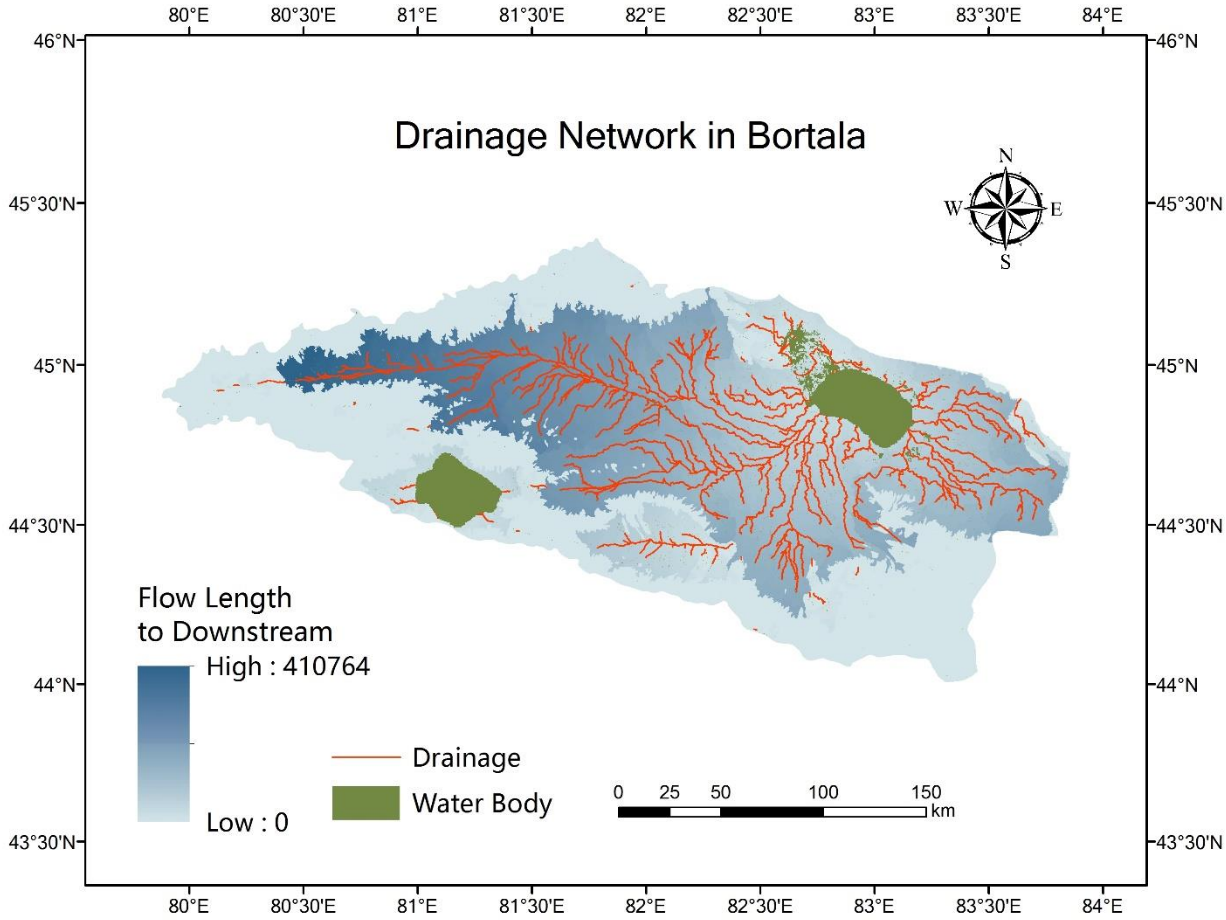

Figure 12 Drainage network and downstream flow length map of Bortala

3.2.5. Generate potential dam sites

3.2.5.1. Catchment delineation & vectorization

Catchment basin, also called drainage basin, is the extent of an area where all water flows to a single point and pours to another water body. The delineation of catchment assigns each cell with a value indicating which catchment the cell belongs to (ESRI 2011).

Delineated catchment is then converted into a vector layer of catchment polygon to meet the format requirement of later procedures. Then HydroIDs are assigned to features in catchment polygon layer. The reason of assigning HydroID is because that for data not generated in Arc Hydro tools, HydroID should be assigned to these data (ESRI 2011).

3.2.5.2. Generate pour points

Pour points, also called drainage points, are outlets of catchment. These pour points have the potential of being dam sites since they are the end of each section of streams where considered having maximum flow volume. For generating pour points, flow accumulation data and catchment data are needed when running drainage point processing tool in ArcGIS.

3.2.5.3. Narrow down potential dam sites

A raster calculation is applied to extract suitability value of cells on stream link. Cells located in lakes or built up area are excluded using a Boolean land cover layer. As part of the result, *41* sites are excluded out of *518* potential sites, which are the drainage outlets. *Figure 13* shows the drainage outlets of each watershed in Bortala.

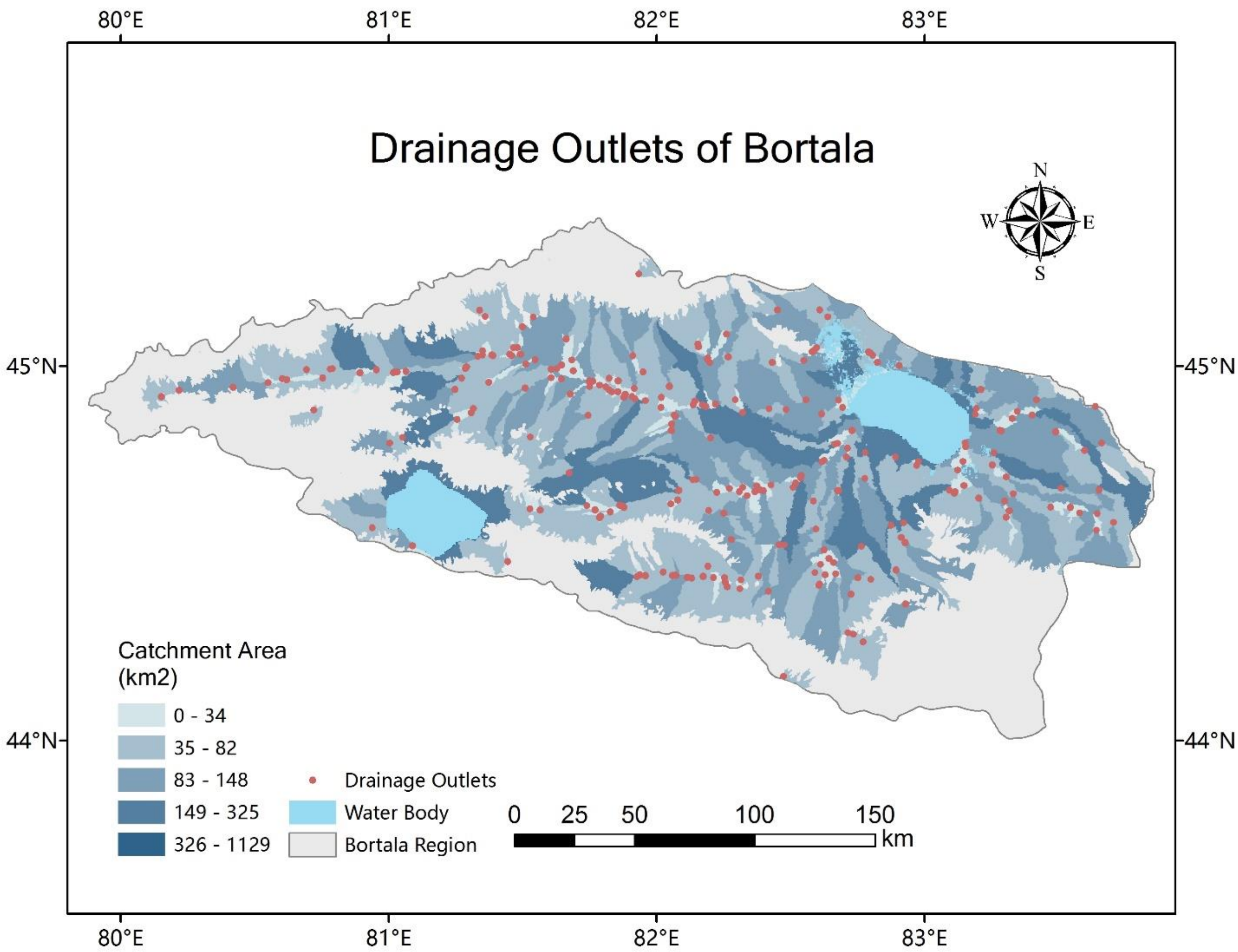

Figure 13 Drainage outlets of small catchments in Bortala

3.2.6. Suitability Evaluation

3.2.6.1. Data preparation

Data used for ranking dam sites are preprocessed as described in *part 3.2.1*. Reclassification is done to group classes with same preference for dam constructing. However, for different layers that stands for different digitized criteria, number of classes are different which makes the preference value comparison of different criteria difficult. Hence, a unified standard for preference value is applied. A linear function is used to assign preference value to different classes of all criteria. When relative preferences of pairwise classes, number of classes and maximum preference value are certain, preference value for every class can be easily calculated. The maximum preference value used for every criteria is *100* in this study. Individual preference values are adjusted accordingly. The unified preference value for classes in every criteria are shown in *table 4* to *table 9*.

Table 4 Unified preference value for predicted precipitation classes

Predicted Precipitation (mm)	Preference Value	Unified Preference
112 to 200	1	20
200 to 300	2	40
300 to 400	3	60
400 to 500	4	80
500 to 639	5	100

Table 5 Unified preference value for degree of slope classes

Slope (degree)	Preference Value	Unified Preference
0 to 1(include)	5	100
1 to 2	4	80
2 to 3	3	60
3 to 4	2	40
4 to 5	1	20
Higher than 5	0	0

Table 6 Unified preference value for geological layer classes

Geological Layer	Preference Value	Unified Preference
Low Resistance	1	25
Slightly Low Resistance	2	50
Moderate Resistance	3	75
High Resistance	4	100

Table 7 Unified preference value for land cover classes

Land Cover	Preference Value	Unified Preference
Settlement	0	0
Water	0	0
Farmland	1	25
Forest	2	50
Shrub and Herb	3	75
Bare Land	4	100

Table 8 Unified preference value for soil type classes

Soil Type	Preference Value	Unified Preference
Sand	1	25
Sandy Loam	2	50
Loam	3	75
Clay Loam	4	100

Table 9 Unified preference value for drainage order classes

Drainage Order	Preference Value	Unified Preference
1	1	20
2	2	40
3	3	60
4	4	80
5	5	100

3.2.6.2. Pairwise comparison

Pairwise comparison is applied on all criteria using the fundamental scale shown in table *1* which was proposed as part of AHP by Saaty in *1977*. Intensity of importance is assigned to criteria *i* when compared to criteria *j* and the reciprocal value is assigned to criteria *j* as intensity of importance. When comparison between all possible criteria pairs is done, the weight of criteria *i*, which is used in later analysis for suitability analysis, is calculated using *equation 2* (Saaty 1977).

$$W_i = \sum_{j=1}^{n} P_{ij} / (\sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij})$$

Equation 2 Calculation of criteria weight

In which P_{ij} indicates relative importance in pairwise comparison of criteria *i* comparing to criteria *j*.

The preference matrix of is shown in table *10* and the explanation of criteria identifier is shown in table *11*.

Table 10 Pairwise comparison preference matrix

	C1	C2	C3	C4	C5	C6
C1	1	2	3	4	5	6
C2	0.5	1	2	3	4	5
C3	0.333	0.5	1	2	3	4
C4	0.25	0.333	0.5	1	2	3
C5	0.2	0.25	0.333	0.5	1	2
C6	0.167	0.2	0.25	0.333	0.5	1

Table 11 Criteria identifier and weight

Criteria No.	Criteria	Weight
C1	predicted precipitation	0.343
C2	slope	0.254
C3	geological foundation	0.177
C4	soil type	0.116
C5	land cover	0.070
C6	drainage order	0.040

3.2.6.3. Evaluation of matrix consistency

The value of pairwise comparison relies on subjective judgement which might lead to arbitrary result with bias. Hence, an evaluation is needed. A numerical index, called Consistency Ratio (CR), is used for evaluating the consistency of pairwise comparison matrix following the proposal in AHP (Saaty 1977). This index indicates the ratio of the Consistency Index (CI) to the average consistency index, which also called Random Index (RI), as shown in *equation 3*.

$$\mathrm{CR} = \mathrm{CI}/\mathrm{RI}$$

Equation 3 Consistency ratio calculation

The value of RI can be found in a table, prepared along the proposal of AHP, according to number of criteria involved as shown in *table 12*.

Table 12 Values for Random Index (Saaty 1977)

Number of criteria	2	3	4	5	6	7	8
Random Index	0	0.52	0.9	1.12	1.24	1.32	1.41

The value of CI can be calculated from the preference matrix according to *equation 4*.

$$\text{CI} = \frac{\lambda_{max} - n}{n - 1}$$

Equation 4 Calculation of Consistency Index

In which λ_{max} indicates the greatest eigenvalue of preference matrix and n is number of criteria involved.

The detailed procedure of calculating λ_{max} is as following steps.

i. Multiply the weight of criteria *i* with columnwise sum of criteria *i* in preference matrix;
ii. Calculate for all criteria as step i described;
iii. Calculate the summation of result values from step ii.

The value of CR is compared to *0.1* which is the maximum CR value for an acceptable pairwise comparison (Saaty 1977). The CR value of *table 10* is *0.0587* which is less than the acceptable maximum CR value recommended in AHP. Hence, the consistency of pairwise comparison is accepted.

3.2.6.4. Suitability calculation

Suitability of dam sites is calculated as weighted summation of different criteria layers using ArcGIS Raster Algebra tool. Weight of different criteria is shown in *table 11*. For exclude water body and built up area, a Boolean layer of land cover, which has the class of water and built up area as value *0* and other classes with value *1*, is multiplied with the summation result. Then, a reclassification is done for differentiating different levels of suitability.

Figure 14 is a flowchart briefly indicates the working steps of suitability calculation in this study. Parallelograms contain data and rectangles contain operations. Some preprocessing working steps are not included in the flowchart to keep the presentation of main steps more clear.

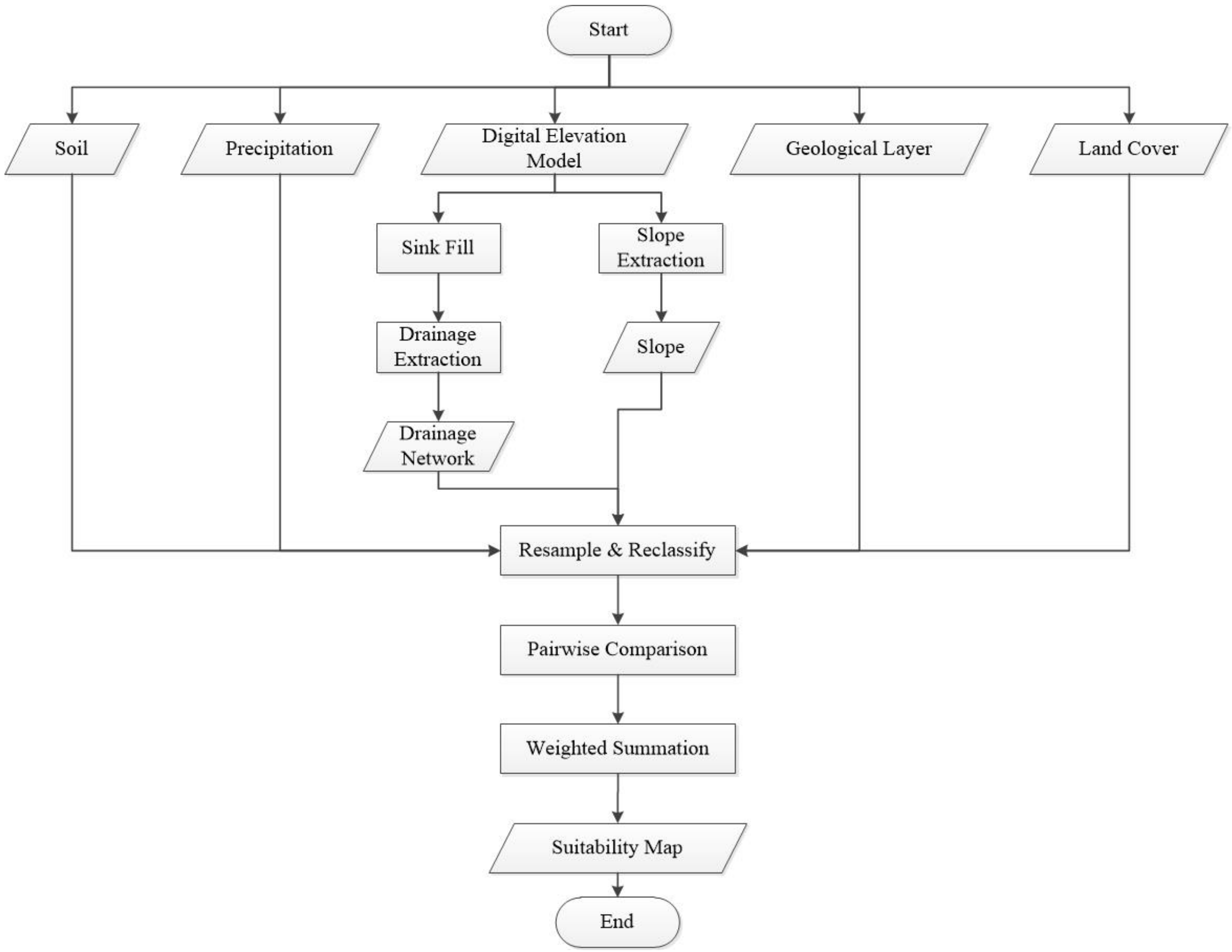

Figure 14 Flowchart of generating suitability map for dam construction

3.2.7. Select and Generate profile of top ranking dam sites

3.2.7.1. Choose possible dam sites

Drainage outlets are used as potential dam sites since these points are the end points for each stream and highest runoff amount of each stream appears at these points. Possible dam sites are chosen based on an integrated interpretation on suitability layer, drainage network layer, drainage outlet layer, Triangulated Irregular Network (TIN) generated from DEM and profile graph based on DEM. When choosing possible sites, check 0 suitability cells near drainage outlets with relative high suitability value and highly mixed with high suitability cells. If a cell with relatively low suitability is caused by slope, then check regional TIN and profile graph to see if it’s acceptable to involve this cell in a dam site. A *20 km* buffer layer of possible dam sites is used for further selection. Dams located within *20 km* of other dam are compared to each other until only one possible site left in a *20 km* range buffer zone. *Figure 15* shows the process of dam sites proposal.

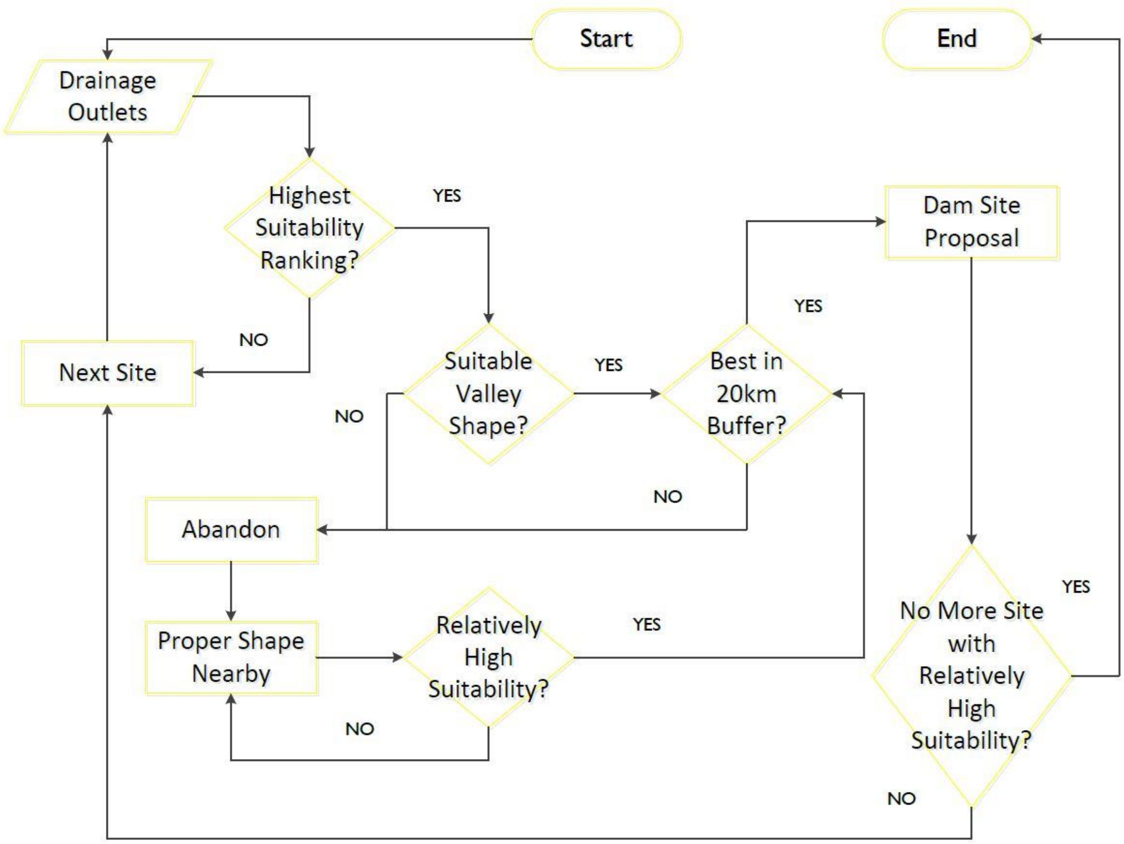

Figure 15 Flowchart of generating proposed dam sites

3.2.7.2. Profile of proposed dam sites

For every proposed possible dam, *5 meters* interval contour layer is generated from DEM and the base height is selected according to elevation of the dam. For drawing interpolation line, which is essential to get cross section profile, and calculating volume of reservoirs, Triangulated Irregular Network (TIN) and *5 meters* interval contour line layer are used. Hence, possible heights of dams are identified at a *5 meters* interval. Besides, when an increased height of dam does not lead to an increased volume of reservoir due to possible leakage on the catchment border, *1 meters* interval contour layer is generated to calculate the height of dam with maximum volume of reservoir. When dam height, location of dam and ground surface (TIN) is certain, storage capacity and area of reservoir surface can be calculated by using 3D Analyst Tools in ArcGIS. *Figure 16* shows the structure of profile processing.

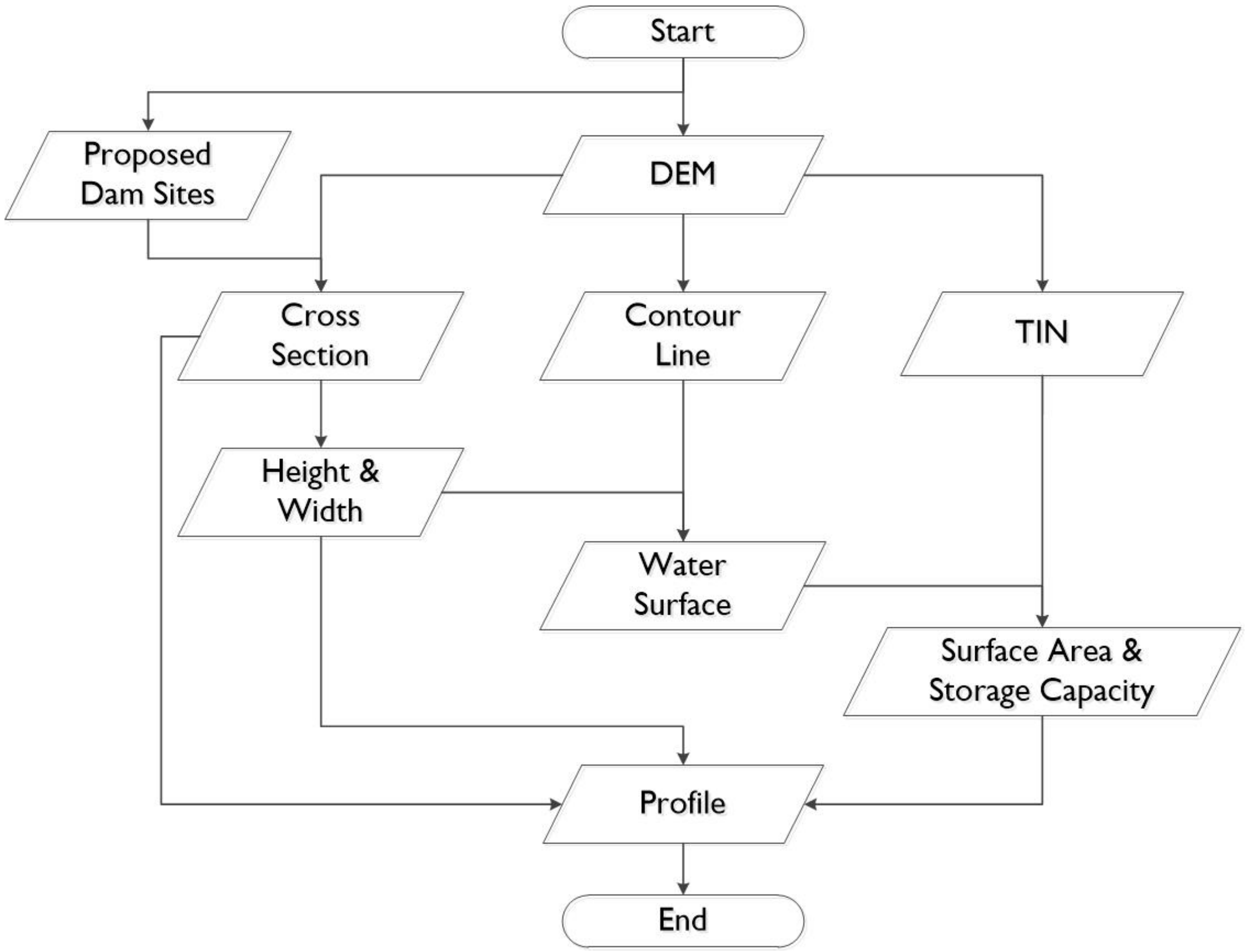

Figure 16 Flowchart of dam sites profiles generation

4. Results & Discussion

4.1. Dam construction suitability map

A suitability map for constructing dam in Bortala, shown in *figure 17*, was produced as well as a histogram showing area and percentage of different suitability levels which is shown in *figure 18*.

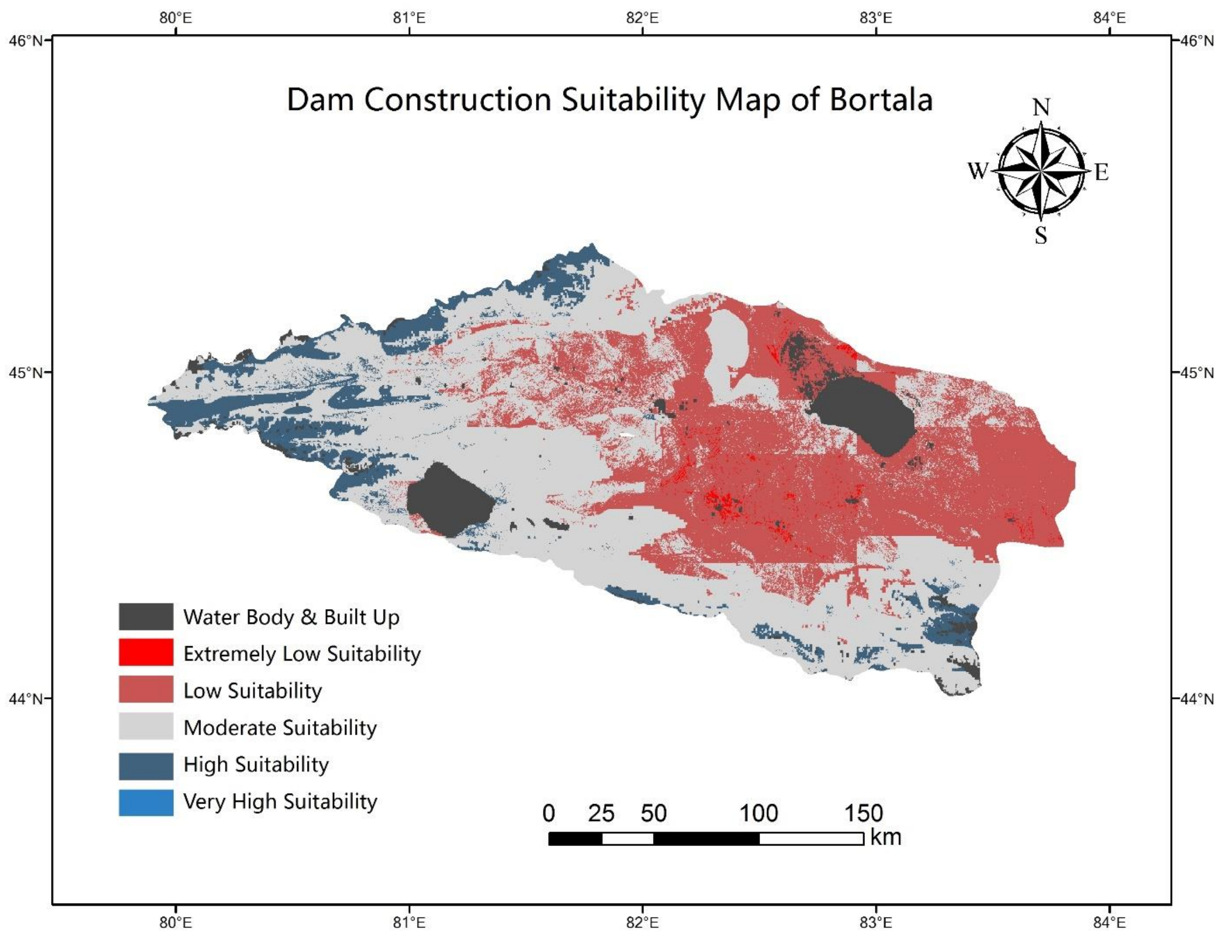

Figure 17 Dam construction suitability map of Bortala

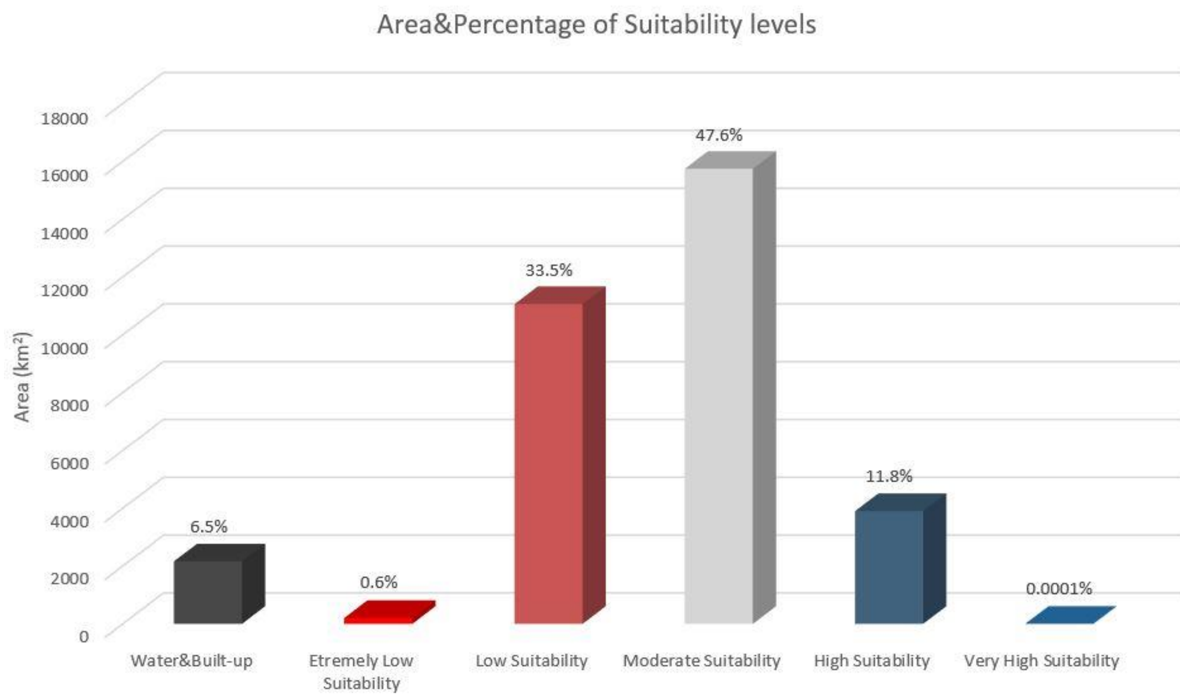

Figure 18 Area and percentage of different suitability levels of Bortala

As the suitability map showing, Northeast and Central Bortala showed a general low suitability for constructing dam and different levels of low suitability are highly mixed in this area which was probably caused by slope change; from Northeast and Central Bortala to other directions, an increasing trend was shown on suitability at a macro scale; most of the area dropped in the range of moderate suitability and high suitability lays within West Bortala. In some part of the map, clear boundary occurred due to the low resolution of original soil data. In general, *11.8%* area of Bortala dropped in the range of high suitability and very high suitability in which only *0.0001%* are very high suitability area; *47.6%* area of Bortala dropped in the range of moderate suitability; *34.1%* of Bortala has got low or extremely low suitability; and *6.5%* area of Bortala was covered by water body or built up area which identified as not suitable for constructing dam.

By comparing suitability map with elevation map, all high and very high suitability area has been found located in the area with higher elevation and most of them has been found with an elevation higher than *1850* m. This is related to but not mainly caused by the positive correlation between elevation and precipitation.

By visual interpretation of suitability map and comparison between suitability map and slope map, cells with different suitability have been found highly mixed in a micro scale and it's mainly caused by change of slope, which is also highly mixed in study area. When the shape of a valley is suitable for constructing dam, which requires valley with tenacious high walls (Yasser et al. 2013), there is a high possibility that the bottom of the valley is with flat slope which defined as suitable slope and the walls of the valley is with steeper slope which defined as unsuitable slope. To alleviate this problem, a generalization was carried out on slope. However, though the slope data used was aggregated to avoid incontinuous slope and suitability value within the area which can be covered by a dam, which would be a severe problem if dealing with individual slope value for each cell at small scale resolution, the problem concerned was alleviated but still exist. The problem of incontinuous slope and suitability value has brought part of the uncertainty in dam site selection since it can be subjective when choosing a dam site covering cells with different suitability value caused by slope. More investigation on generalization scales can be done in further research.

Generally, resolution of datasets affects accuracy of the suitability map because of the resolution differences between datasets and limitations caused by resolution. The limitation caused by resolution can be understand as following: Value of a cell can represent most area covered by the cell, but it does not mean area covered by this cell have a uniformed value in reality. In this study, resolution of datasets varies from *30 meters* (DEM) to *9 km* (soil type), therefore resamples to a uniformed resolution was needed. Still, the different resolution of data sources caused considerable uncertainty and some mosaic area on the suitability map. For further studies, effects of resolution difference should be considered during data collection and resamples should be treated with caution especially on DEM since the topographic attributes extracted from DEM varies a lot with change of resolution (Wu et al. 2008).

Another source of uncertainty in this study is the subjective judgment on functions of preference value of criteria. The preference value used for all criteria classes had linear functions which probably correct for some criteria but incorrect for others. Although the judgements on relative preference within each criteria were based on previous studies, the judgements on absolute preference value within each criteria were subjective and difficult to validate in simple tests.

For validation of the suitability map, similar studies on Bortala region were searched for but not found. Therefore, as an alternative, mid-products of this study, drainage network and flow accumulation, are compared with hydrographical mapping product called HydroSHEDS (Lehner et al. 2008). The similarity in the mainstreams is satisfying while there are first level tributaries missing in this study comparing to HydroSHEDS dataset. This difference is caused by generating flow direction from DEM in a large scale which causes huge variation on flow accumulation value between first level tributaries and the biggest main streams. And this correspond to the result of a previous study that different sizes (number of cells) of DEM can lead to different flow accumulation value generated and the total flow accumulation increases in respond to increase of DEM size (Wolock 1995). Hence, more investigation on suitable DEM size for extracting hydrological attributes should have been done in this study as well as generating flow accumulation separately for sub catchments in smaller DEM size, and these are recommended for further studies.

4.2. Proposed dam sites

Possible dam sites were chosen based on an integrated dataset composed of drainage outlet layer, suitability layer, drainage order layer, Triangulated Irregular Network (TIN) generated from DEM and profile graph based on DEM. As a result, 8 possible dam sites were proposed. The location of proposed dams along with a DEM layer are shown in *figure 19* and the coordinates and elevation of proposed dams are listed in *table 13*.

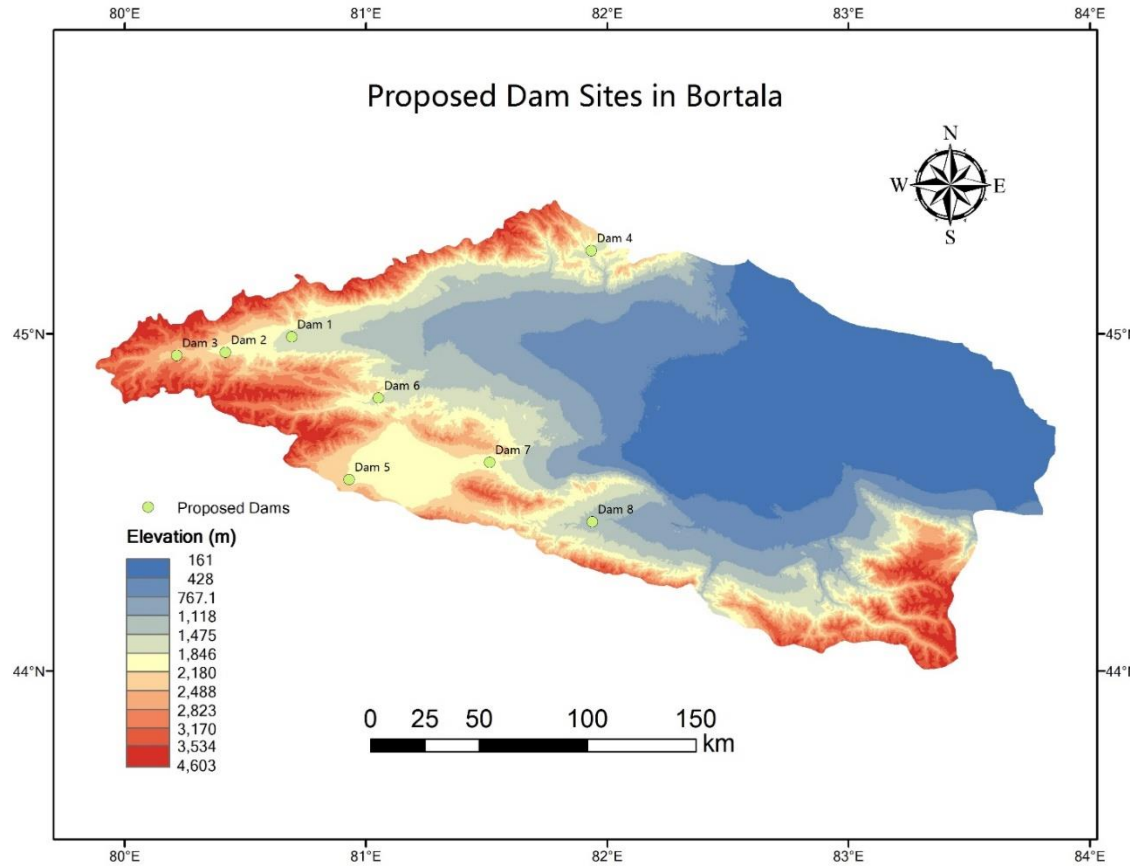

Figure 19 Location of 8 proposed dam sites in Bortala

Table 13 Coordinates and elevation of proposed dams

Rank	Dam 1	Dam 2	Dam 3	Dam 4	Dam 5	Dam 6	Dam 7	Dam 8
Longitude (E)	80°41'34"	80°41'3"	80°12'56"	81°56'2"	80°55'49"	81°3'6"	81°30'46"	81°56'17"
Latitude (N)	44°59'27"	44°56'41"	44°56'12"	45°14'36"	44°34'11"	44°48'38"	44°37'15"	44°26'41"
Elevtion (m)	1730	2072	2342	1571	2152	1694	1951	1188

The 8 proposed dams were located at area with elevation fairly distributed in the range of *1180 m* to *2342 m*. All dams were located in or close to mountainous area where proper valley shape and higher precipitation value exist. However, on the aspect of drainage order, 6 dams were located at 1st drainage order and 2 dams were located at 2nd drainage order. The reason why all proposed dam sites were located in lower level of drainage network is probably caused by multiple factors as follow:

i. Comparing the proposed dam sites with suitability map, all proposed dams are on high elevation which is with relatively high suitability, as the suitability map shown, and streams in these area are more likely to be defined as upper stream since the relatively high elevation and close to data edge;
ii. Comparing the drainage network with maps of criteria, most of the higher level drainage are located in the area of relatively low preference value in almost all criteria except for slope. To be more specific, most of the higher level drainage are located in the area with relatively low predicted precipitation, sand or sandy loam soil texture, low resistance geological layer, and farmland. All of these characteristics would lead to a relatively low suitability;
iii. Comparing between the suitability map, drainage network and TIN, it is easy to find out that area with lower elevation, which higher level drainage network exist, are flatlands that gives these area lower priority for proposing dam sites in terms of proper valley shape;
iv. When generating drainage network, data from whole Bortala region was processed as one whole input which leads to a multimillion level variety of accumulation value. Plus, the flow accumulation model does not consider water consumption during water transportation.

To get a knowledge of the influence, from threshold value of defining drainage, on the drainage order of proposed dam sites, a quick test of changing threshold for defining streams from *30,000* to *8,000* is done. As the result of changing, *1* higher drainage order class appeared and 5 dams has scored a higher drainage order out of all *8* sites (shown in *table 14*).

Table 14 Dam sites drainage order change in responding to stream definition threshold change

Dam	Dam1	Dam2	Dam3	Dam4	Dam5	Dam6	Dam7	Dam8
Elevation (m)	1730	2072	2342	1571	2152	1695	1951	1180
Drainage Order (30k)	2	1	1	1	1	1	1	2
Drainage Order (8k)	3	2	1	1	1	2	2	3

Most of the new streams defined are located in lower elevation area which are still defined as low suitability area for constructing dams. Hence, the threshold changing on defining streams does not have a significant effect on dam site selection in this study. However, more accurate stream defining is still an important part for dam site selection and it should be considered for further studies as well as the accuracy of flow accumulation. The difficulty of drainage definition according to flow accumulation in a relatively large scale is that the accuracy of tributaries simulation and mainstreams simulation cannot be proved high at the same time when comparing to topographic map. Meanwhile, when using Shreve method (Shreve 1966) to define drainage order in a large scale, the list of drainage order level is relatively long, in this study *206* levels of drainage order on streams defined by threshold value of *30,000* and *681* levels of drainage order when taking *8,000* as threshold. A possible solution would be breaking filled DEM into small parts and generating flow accumulation for every sub catchment then make a mosaic dataset of flow accumulation from all sub catchments. But this possible solution requires massive time for data processing and it also leads to an underestimation of flow accumulation in biggest mainstreams. Further studies can be done in balancing the accuracy of flow accumulation in tributaries and mainstreams.

Besides the threshold defining and the procedures for computing drainage order, the preference of constructing dam on different level of drainage order is also a considerable problem. In one hand, higher drainage order, which means large main stems of rivers, should be preferred while seeking for a higher runoff. In the other hand, lower drainage order, which means upper tributaries, should be preferred while seeking for a more environment benign alternative (Ledec & Quintero 2003). Hence, the priority of different classes within a criteria category as well as different choices on involved criteria and the weights of different criteria are all very dependent on decision makers' purposes and environmental conditions in each individual case.

In this study, a subjective judgment on minimum distance between two proposed dam sites has occurred during dam site selection. It is necessary to have a value of minimum distance between two proposed dam sites to avoid redundancy and the value can be subjective depending on decision makers' purposes. For further investigation on dam site selection in Bortala, high resolution runoff data in both temporal and spatial scales can be used to improve the accuracy of flow accumulation thus the accuracy of stream defining. Meanwhile, more information about artificial infrastructure and human activities, as well as the specific purposes of decision makers, can be involved as criteria for analysis of suitability. Another thing can be considered is the special location of Bortala. Since Bortala is a border region between China and Kazakhstan, both the usage of resource and construction of engineering program might be politically sensitive. Further investigation can be done to get more information on aspect of resource and cultural conflict that may affect the site selection of dam.

4.3. Profile of proposed dams

Profile of proposed dams were generated according to location of dams in combination with DEM, Triangulated Irregular Network (TIN), 5 meters and 1 meter interval contour layers. Profile of dam consists of cross section of dam and some other characteristics of dam, which contains elevation of dam, height of dam, width of dam, elevation of reservoir surface, maximum storage of reservoir and maximum surface area of reservoir. Generally, possible heights of dams were identified at a 5 meters interval unless the maximum volume of reservoir appear at a height that lays out of 5 meters interval. Widths of dams were generated according to cross profiles and heights of dams. Cross section of proposed dam sites are shown in *figure 20* to *figure 27* and profile of proposed dams are shown in *table 15*.

Among all 8 proposed dams, the maximum height of dams, corresponding to cross section of dam locations, varied from *10 m* to *35 m*; maximum width of dams varied from *148 m* to *461 m*; the maximum storage capacity of reservoirs, corresponding to distribution of topographic conditions in surrounding area, varied from *122,506* m^3 to *5,033,652* m^3; the maximum surface area of reservoir varied from *49,154* m^2 to *414,713* m^2. There are many different way to classify dams, but no documented standardization of dam size classification found on related Chinese governmental website. In an alternative source, which is documented standardization for dam safety from Vermont USA, *12.2 meters* and *30.5 meters* (converted from *ft*) in height or *1,233,482* m^3 and *61,674,092* m^3 (converted from *ac-ft*) in storage are the border lines between small, medium, and large dam (Government of Vermont 1990). According to this classification standardization, there were *2* small dams, *5* medium dams, and *1* large dam among proposed dams.

For dam 1 and dam 6, the storage capacity reached a limitation while there was still space for a larger dam. This phenomenon was caused by high elevation at the top of both sides of mountain slope in the valley while some part of the surrounding catchment boundary was with a lower elevation than maximum dam height. Since there is a possibility to construct auxiliary dyke to close the leaking part, possible maximum dam height and width of these two dams are still available.

While generating cross section for dam sites, it is flexible to change the angle between dam body and stream direction for a cross section that can support higher possible dam height. The uncertainty of this angle choosing operation may lead to an unrecorded maximum dam height which is larger than the analysis result of this study. For application of proposed dam sites, fieldwork should be done for further information.

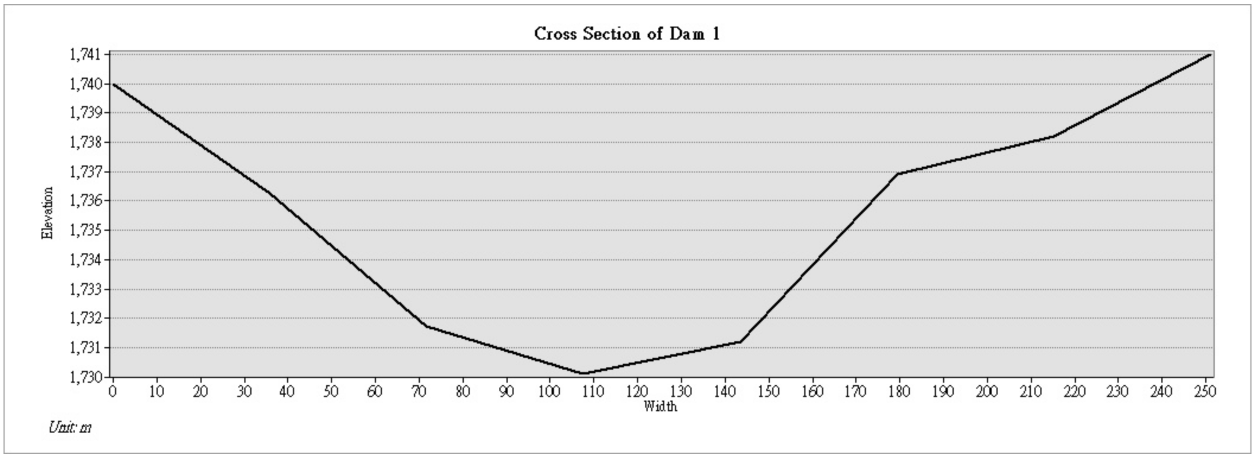

Figure 20 Cross section of proposed dam 1

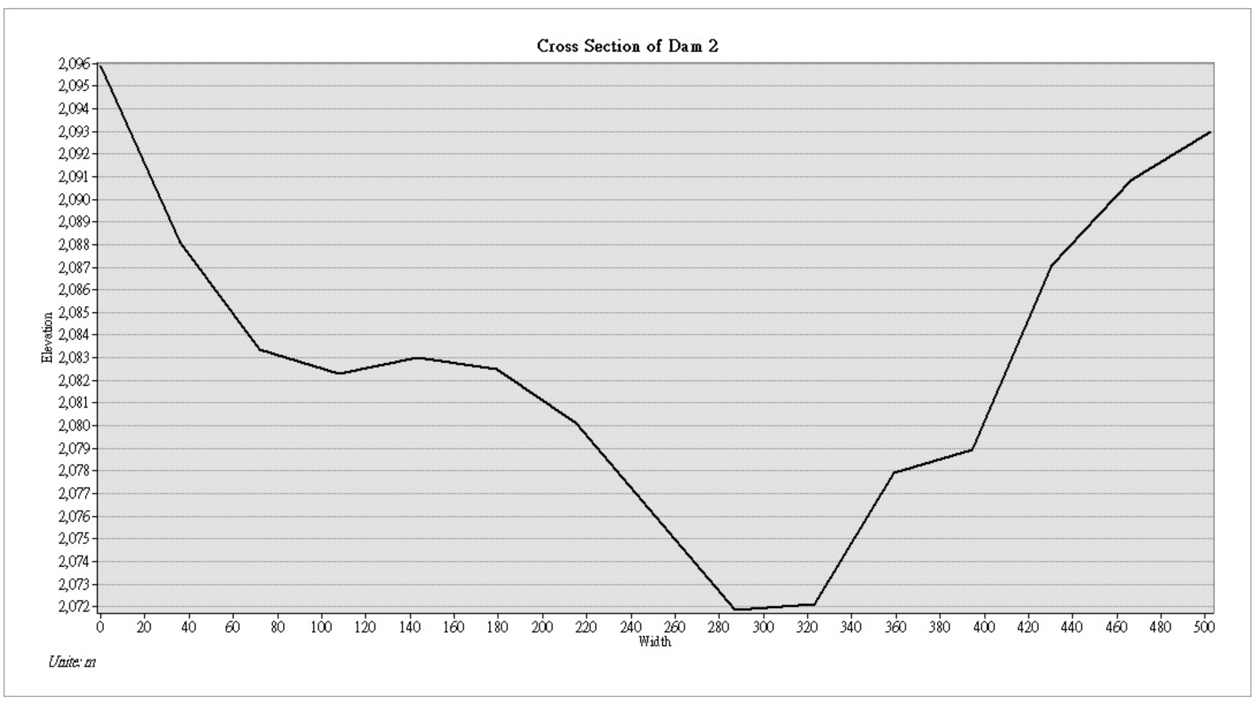

Figure 21 Cross section of proposed dam 2

Figure 22 Cross section of proposed dam 3

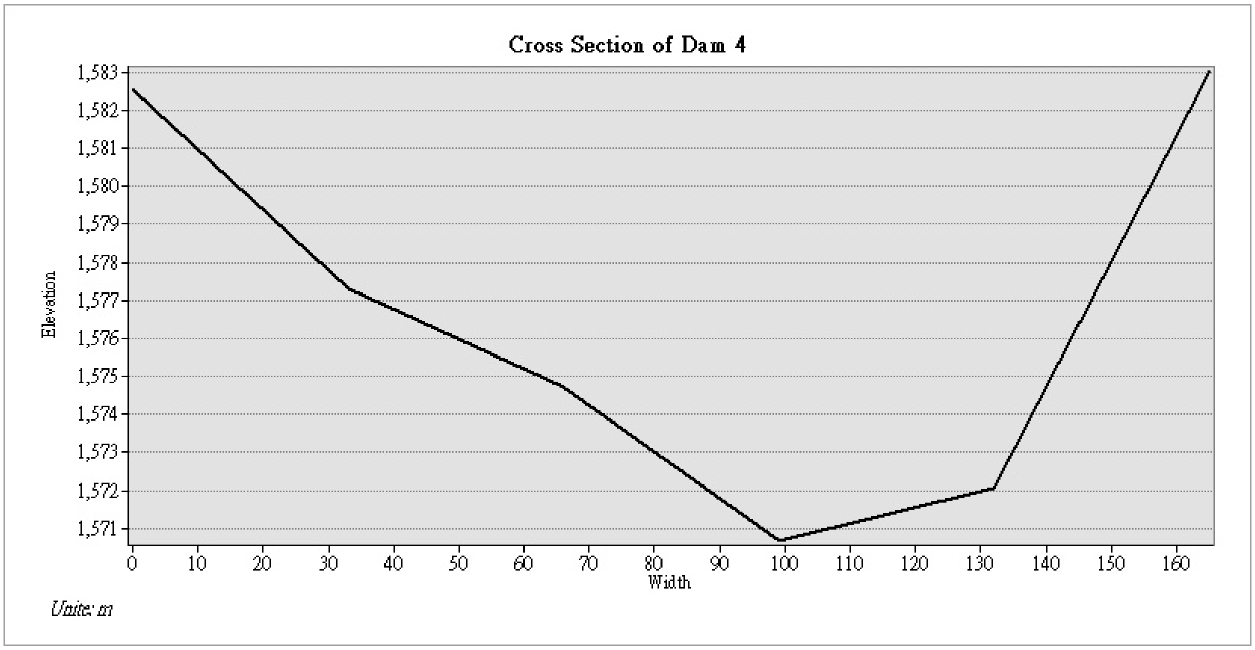

Figure 23 Cross section of proposed dam 4

Figure 24 Cross section of proposed dam 5

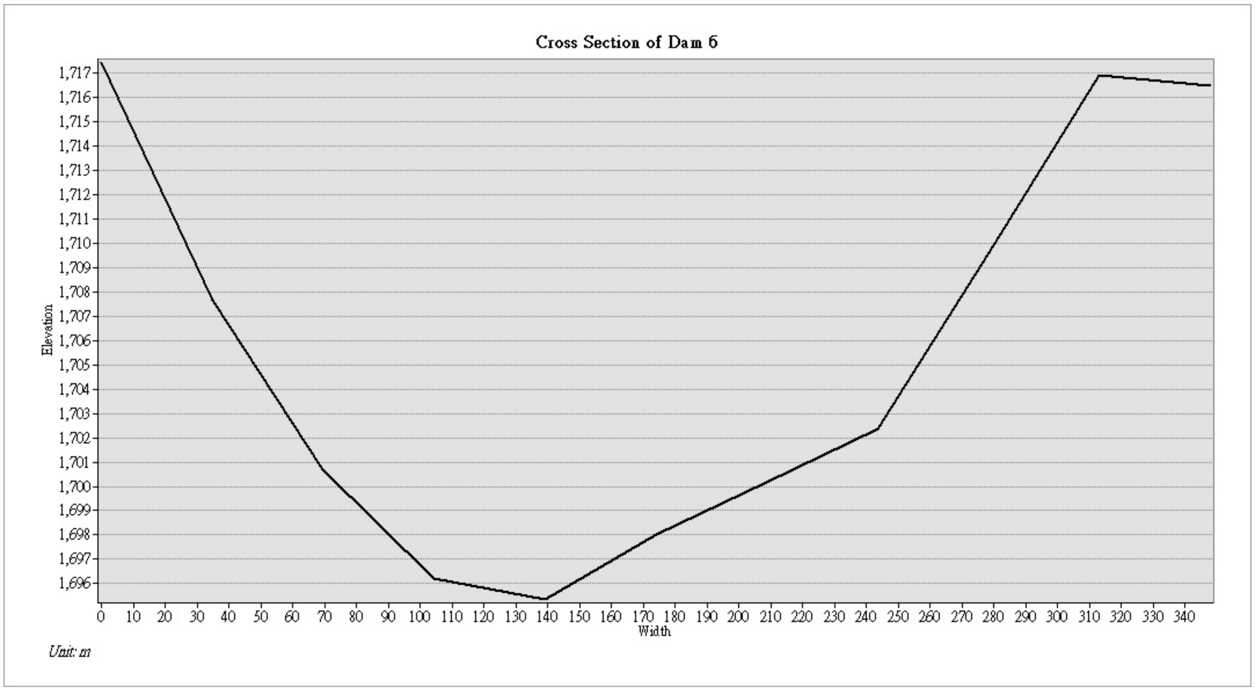

Figure 25 Cross section of proposed dam 6

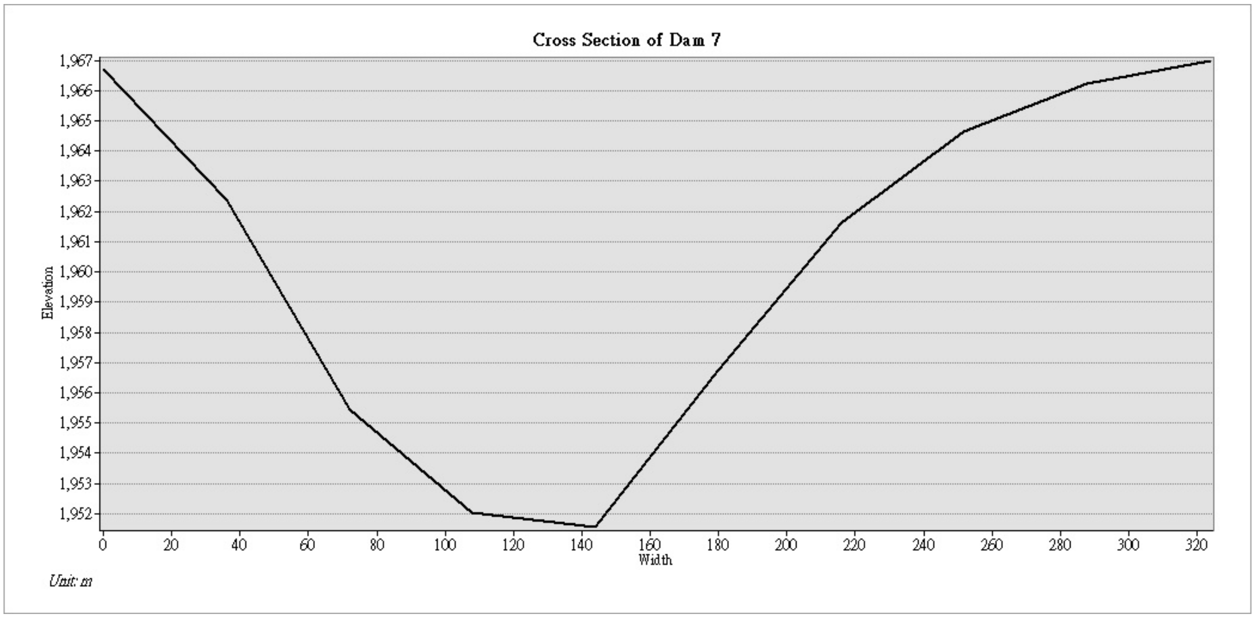

Figure 26 Cross section of proposed dam 7

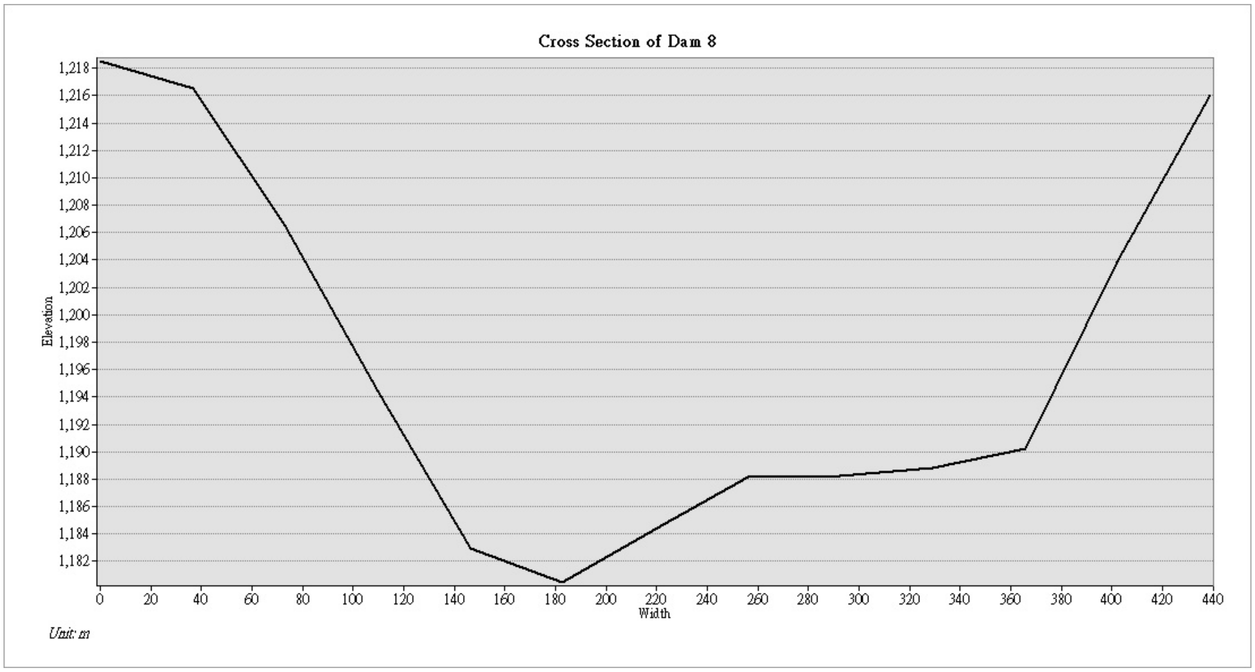

Figure 27 Cross section of proposed dam 8

Table 15 Characteristics of proposed dams

Proposed Dam	Elevation (m)	Dam Height (m)	Dam Width (m)	Max Storage (m3)	Max Surface Area (m2)
Dam 1	1730	0	0	0	0
	1735	5	122	122,506	52,770
	1740	10	238	122,506	52,770
Dam 2	2072	0	0	0	0
	2077	5	110	17,349	5,891
	2082	10	220	70,762	16,960
	2087	15	383	750,458	180,500
	2092	20	461	2,108,118	314,668
Dam 3	2342	0	0	0	0
	2347	5	76	10,686	3,903
	2352	10	168	77,539	14,097
	2357	15	255	263,110	31,913
	2360	18	310	490,431	49,154
Dam 4	1571	0	0	0	0
	1576	5	94	111,058	37,669
	1581	10	148	366,657	66,084
Dam 5	2152	0	0	0	0
	2157	5	174	359,891	127,143
	2162	10	314	1,371,397	281,712
Dam 6	1695	0	0	0	0
	1700	5	142	14,867	10,078
	1705	10	210	174,395	47,931
	1710	15	252	538,628	110,448
	1711	16	264	672,631	140,023
	1715	20	294	672,631	140,023
Dam 7	1951	0	0	0	0
	1956	5	105	14,336	6,071
	1961	10	174	160,677	50,258
	1966	15	278	620,540	130,911
Dam 8	1180	0	0	0	0
	1185	5	125	3,328	1,789
	1190	10	230	72,368	32,255
	1195	15	272	318,679	60,623
	1200	20	298	760,168	124,703
	1205	25	332	1,710,511	244,952
	1210	30	360	3,116,614	315,831
	1215	35	395	5,033,652	414,713

5. Conclusion

As a responding to water resource shortage in Northwest China, constructing dam has been a possible solution considered by both the central government of China and local governments in Northwest China. In this study, a suitability map for constructing dam and several proposed dam sites with profile were generated for Bortala region Northwest China. Due to the complexity of dam site selection on aspect of number of factors, a Multi-Criteria Decision Making (MCDM) method called Analytic Hierarchy Process (AHP) was used in this study. As a powerful and flexible tool, AHP provides an integrated measurement on tangible factors with different priority by pairwise comparison. For a more visualized measurement and presentation, Geographic Information System (GIS) was used for implementation of AHP. As one of the most used GIS software globally, ArcGIS was chosen as the software tool for this study.

Raster layers of all criteria were integrated by a weighted summation using the weights generated by pairwise comparison to generate a suitability map showing different levels of suitability for dam construction. Another land cover layer in Boolean classes, where water body and built up were classified as *0*, was used for excluding water body and built up area from potential dam sites. The suitability map is useful for decision makers and whoever interested in this topic to have a quick idea on suitability distribution and determine the potential area for dam sites. In general, West Bortala was found to be the densest area with high suitability for constructing dam while Northwest and Central Bortala were found as low or extremely low suitability area for constructing dam. For percentage describing, *11.8%* area of Bortala dropped in the range of high suitability and very high suitability in which only *0.0001%* was very high suitability area; *47.6%* area of Bortala was found in the range of moderate suitability; *34.1%* of Bortala has got low or extremely low suitability. Most of the high suitability area in Bortala was found located on higher elevation. Fieldwork should be done when choosing dam sites due to the limitation of data resolution as well as mixture of different suitability within same area.

Besides the suitability map, *8* possible dam sites were proposed according to analysis on the suitability map, drainage network, drainage outlet, Triangulated Irregular Network (TIN) generated from DEM and profile graph based on DEM. Along with the location of proposed dam sites, a profile of each dam was generated including a cross section of the dam site, possible dam heights, possible dam widths, elevation of the dam, possible maximum storage capacity and surface area of the reservoir coming with the dam. All proposed dam sites were located in relatively high elevation area in Bortala. According to a dam size classification of Vermont USA, there were *2* small dams, *5* medium dams, and *1* large dam among proposed dams. The maximum height of dams, corresponding to cross section of dam locations, varied from *10 m* to *35 m*; maximum width of dams varied from *148 m* to *461 m*; the maximum storage capacity of reservoirs, corresponding to distribution of topographic conditions in surrounding area, varied from *122,506* m^3 to *5,033,652* m^3; the maximum surface area of reservoir varied from *49,154* m^3 to *414,713* m^3.

For further studies, high resolution sources of all related datasets are recommended to improve the accuracy of flow accumulation, thus the accuracy of stream defining, as well as the accuracy of suitability map. Meanwhile, more information about artificial infrastructure and human activities, as well as decision makers' purposes, can be involved as criteria for analysis of suitability. Also, more information about determining preference values of criteria classes and effects from resolution of datasets as well as processing sizes of datasets can be considered. Another possible factor can be considered is the special location of Bortala as a border region between China and Kazakhstan that both the usage of resource and construction of engineering program might be politically sensitive. Hence, further investigation can be done to get more information, which may affect the site selection of dams, on aspect of resource and cultural conflict between these two countries.

References

Abushandi, E. & Alatawi, S., 2015. Dam Site Selection Using Remote Sensing Techniques and Geographical Information System to Control Flood Events in Tabuk City. *Hydrology Current Research*, 6(1), pp.1–13.

Adinarayana, J., Krishna, N.R. & Rao, K.G., 1995. An Integrated Approach for Prioritization of Watersheds. *Journal of Environmental Management*, 44(4), pp.375–384.

Alaibakhsh, M., Azizi, S.H. & Kheirkhah Zarkesh, M.M., 2013. Water resource management with a combination of underground dam/qanat and site selection of suitable sites using GIS. *Water Science and Technology: Water Supply*, 13(3), pp.606–614.

Baban, S.M.J. & Wan-Yusof, K., 2003. Modelling optimum sites for locating reservoirs in tropical environments. *Water Resources Management*, 17(1), pp.1–17.

Brouwer, C. et al., 1990. Irrigation Water Management : Irrigation Methods, Training manual n°5. , (5), p.140.

Chen, J.L. et al., 2009. 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models. *Journal of Geophysical Research*, 114(B5), p.B05404.

Climate-Data.Org, 2016. Climate: Bortala. Available at: http://en.climate-data.org/location/2220/.

ESRI, 2011. ArcHydro Tools Tutorial. , (October), p.184.

ESRI, 2016. How Fill works. Available at: http://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/how-fill-works.htm.

Fan, T. et al., 2005. Grain yield and water use in a long-term fertilization trial in Northwest China. *Agricultural Water Management*, 76(1), pp.36–52.

Geng, Y.H., Min, Q.W. & Cheng, S.K., 2006. Water demand of crop plantation in northwest China - A case study of the Jinghe watershed. *Journal of Ecology and Rural Environment*, 22(4), pp.30–34.

Government of Vermont, 1990. Safety of Hydroelectric Dams. , pp.1–8. Available at: http://psb.vermont.gov/sites/psb/files/rules/OfficialAdoptedRules/4500_Dam_Safety.pdf.

Hao, Z. et al., 2010. *Extraction of catchment characteristics based on DEM in "Theory and Method of Distributed Hydrological Models"* 1st ed. J. Shen, ed., Beijing: Science Press.

Hooke, J.M. & Mant, J.M., 2000. Geomorphological impacts of a flood event on ephemeral channels in SE Spain. *Geomorphology*, 34(3-4), pp.163–180.

Jenson, S.K. & Domingue, J.O., 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric Engineering and Remote Sensing*, 54(11), pp.1593–1600.

Lashkaripour, G.R. & Ghafoori, M., 2002. The engineering geology of the Tabarak Abad Dam. *Engineering Geology*, 66(3-4), pp.233–239.

Ledec, G. & Quintero, J.D., 2003. Good dams and bad dams: environmental criteria for site selection of hydroelectric projects. *Latin America and the Caribbean Region:Sustainable Development Working Paper No. 16*, 16(November), p.21.

Lehner, B., Verdin, K. & Jarvis, A., 2008. New global hydrography derived from spaceborne elevation data. *Eos*, 89(10), pp.93–94.

Liu, Yu; Wang, Lin; Ni, Guangheng; Cong, Z., 2009. Spatial distribution characteristics of irrigation water requirement for main crops in China. *Transactions of the Chinese Society of Agricultural Engineering*, 25, pp.6–12(7).

Llasat, M. del C., Rigo, T. & Barriendos, M., 2003. The "Montserrat-2000" flash-flood event: a comparison with the floods that have occurred in the northeastern Iberian Peninsula since the 14th century. *International Journal of Climatology*, 23(4), pp.453–469.

Maillet, G.M. et al., 2006. Morphological changes and sedimentary processes induced by the December 2003 flood event at the present mouth of the Grand Rhône River (southern France). *Marine Geology*, 234(1-4), pp.159–177.

Mardani, A. et al., 2015. Multiple criteria decision-making techniques and their applications – a review of the literature from 2000 to 2014. *Economic Research-Ekonomska Istraživanja*, 28(1), pp.516–571.

Marinoni, O., 2004. Implementation of the analytical hierarchy process with VBA in ArcGIS. *Computers and Geosciences*, 30(6), pp.637–646.

McVicar, T.R. & Bierwirth, P.N., 2010. Rapidly assessing the 1997 drought in Papua New Guinea using composite AVHRR imagery. *International Journal of Remote Sensing*.

Ministry of Water Resources, 2004. *Pilot Experiences of Establishing a Water-saving Society in China*, Beijing, China.

Ministry of Water Resources, 2016. *Water Resources in China*, Available at: http://www.mwr.gov.cn/english/mainsubjects/.

Ministry of Water Resources, P.R. of C., 2014. *2014 Statistic Bulletin on China Water Activities*,

Mondlane, A.I., 2004. Floods and droughts in Mozambique–the paradoxical need of strategies for mitigation and coping with uncertainty. In *4th International Conference on Computer Simulation in*

Risk Analysis and Hazard Mitigation (RISK 2004) 4th.

O'Callaghan, J.F. & Mark, D.M., 1984. The extraction of drainage networks from digital elevation data. *Computer Vision, Graphics, and Image Processing*, 27(2), p.247.

Piao, S. et al., 2010. The impacts of climate change on water resources and agriculture in China. *Nature*, 467(7311), pp.43–51.

Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15(3), pp.234–281.

Saaty, T.L., 2008. Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 1(1), p.83.

Saaty, T.L., 1988. Mathematical Models for Decision Support. In G. Mitra et al., eds. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 109–121.

Saaty, T.L., 2013. The analytic network process. *Decision making with the analytic network process*, 195, pp.1–40.

Shen, Y. et al., 2013. Estimation of regional irrigation water requirement and water supply risk in the arid region of Northwestern China 1989-2010. *Agricultural Water Management*, 128, pp.55–64.

Shen, Y. & Chen, Y., 2010. Global perspective on hydrology , water balance , and water resources management in arid basins. *HYDROLOGICAL PROCESS*, 135(September 2009), pp.129–135.

Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data. *Proceeding ACM '68 Proceedings of the 1968 23rd ACM national conference*, pp.517–524.

Shi, Y. et al., 2007. Recent and future climate change in northwest China. *Climatic Change*, 80(3-4), pp.379–393.

Shreve, R.L., 1966. Statistical Law of Stream Numbers. *The Journal of Geology*, 74(00221376, 15375269), pp.17–37.

Singh, J.P., Singh, D. & Litoria, P.K., 2009. Selection of suitable sites for water harvesting structures in Soankhad watershed, Punjab using remote sensing and geographical information system (RS&GIS) approach- a case study. *Journal of the Indian Society of Remote Sensing*, 37(1), pp.21–35.

Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. *Transactions, American Geophysical Union*, 38(6), p.913. Available at:

University of Hawaiʻi - College of Tropical Agriculture and Human Resources, Soil Texture and Soil Structure. Available at: http://www.ctahr.hawaii.edu/mauisoil/a_factor_ts.aspx.

US Army Corps of Engineers, 2005. “general considerations”, in General design and construction considerations for earth and rock-fill dams. , (July). Available at: http://www.usace.army.mil/usace-docs/eng-manuals/em1110-2-2300/entire.pdf.

Wan, L. et al., 2015. Decadal climate variability and vulnerability of water resources in arid regions of Northwest China. *Environmental Earth Sciences*, 73(10), pp.6539–6552.

Wikipedia, 2016. Bortala Mongol Autonomous Prefecture. Available at: https://en.wikipedia.org/wiki/Bortala_Mongol_Autonomous_Prefecture.

Wolock, D.M., 1995. Effects of subbasin size on topographic characteristics and simulated flow paths in Sleepers River watershed, Vermont. *Water Resources Research*, 31(8), pp.1989–1997.

Wu, S., Li, J. & Huang, G.H., 2008. A study on DEM-derived primary topographic attributes for hydrologic applications: Sensitivity to elevation data resolution. *Applied Geography*, 28(3), pp.210–223.

Wyllie, D.C., 2003. *“Foundation of gravity and embankment dams”, in Foundations on rock: engineering practice*, CRC Press.

Xie, Z., Wang, Y. & Li, F., 2005. Effect of plastic mulching on soil water use and spring wheat yield in arid region of northwest China. *Agricultural Water Management*, 75(1), pp.71–83.

Yasser, M. et al., 2015. An integrated decision support system for dam site selection. *Scientia Iranica*, 22, p.2015.

Yasser, M., Jahangir, K. & Mohmmad, A., 2013. Earth dam site selection using the analytic hierarchy process (AHP): a case study in the west of Iran. *Arab J Geosci*, pp.3417–3426.

Yu, X.I.N., Guangxing, Z. & Xin, Z., 2006. Climate change in Bortala Mongol Autonomous Prefecture, Xinjiang since recent 46 years. *Arid Land Geography*, 29(2), pp.100–105.

Zavadskas, E.K., Turskis, Z. & Kildienė, S., 2014. State of art surveys of overviews on MCDM/MADM methods. *Technological and Economic Development of Economy*, 20(1), pp.165–179.

Zhu, F., Wang, S. & Zhou, P., 2003. Flavobacterium xinjiangense sp. nov. and Flavobacterium omnivorum sp. nov., novel psychrophiles from the China No. 1 glacier. *International journal of systematic and evolutionary microbiology*, 53(Pt 3), pp.853–7.

Zionts, S., 1979. MCDM—If Not a Roman Numeral, Then What? *Interfaces*, 9(4), pp.94–101.

Appendix

Appendix 1. Structure of ISRIC-WISE soil data

Name	Type	Description
CLAF	Text	FAO-Unesco (1974) Legend code
PRID	Text	Profile ID (as documented in table WISEunitC*omposition*)
Drain	Text	FAO soil drainage class
Layer	Text	Code for depth layer (from D1 to D5; e.g. D1 is from 0 to 20 cm)
TopDep	Integer	Depth of top of layer (cm)
BotDep	Integer	Depth of bottom of layer (cm)
CFRAG	Integer	Coarse fragments (vol% > 2 mm)
SDTO	Integer	Sand (mass %)
STPC	Integer	Silt (mass %)
CLPC	Integer	Clay (mass %)
PSCL	Text	FAO texture class
BULK	Single	Bulk density (kg dm^{-3})
TAWC	Integer	Available water capacity (cm m^{-1}, -33 to -1500 kPa conform to USDA standards)
CECs	Single	Cation exchange capacity ($cmol_c$ kg^{-1}) for fine earth fraction
BSAT	Integer	Base saturation as percentage of CEC_{soil}
CECc	Single	CEC_{clay}, corrected for contribution of organic matter ($cmol_c$ kg^{-1})
PHAQ	Single	pH measured in water
TCEQ	Single	Total carbonate equivalent (g C kg^{-1})
GYPS	Single	Gypsum content (g kg^{-1})
ELCO	Single	Electrical conductivity (dS m^{-1})
TOTC	Single	Organic carbon content (g C kg^{-1})
TOTN	Single	Total nitrogen (g kg^{-1})
CNrt	Single	C/N ratio
ECEC	Single	Effective CEC ($cmol_c$ kg^{-1})

Appendix 2. Parameters of original soil classes

SUID	Map unit	SDTO	STPC	CLPC	SUID	Map unit	SDTO	STPC	CLPC
3619	WD3619	45.00	31.00	24.00	3722	WD3722	34.00	38.00	28.00
3619	WD3619	44.00	30.00	26.00	3722	WD3722	31.00	39.00	30.00
3619	WD3619	44.00	30.00	26.00	3722	WD3722	30.00	40.00	30.00
3619	WD3619	43.00	31.00	26.00	3722	WD3722	31.00	40.00	29.00
3619	WD3619	43.00	29.00	28.00	3722	WD3722	31.00	42.00	27.00
4090	WD4090	63.00	19.00	18.00	3967	WD3967	40.00	30.00	30.00
4090	WD4090	62.00	20.00	18.00	3967	WD3967	40.00	30.00	30.00
4090	WD4090	57.00	24.00	19.00	3967	WD3967	41.00	29.00	30.00
4090	WD4090	49.00	27.00	24.00	3967	WD3967	43.00	28.00	29.00
4090	WD4090	49.00	27.00	24.00	3967	WD3967	44.00	28.00	28.00
3314	WD3314	67.00	15.00	18.00	3094	WD3094	40.00	30.00	30.00
3314	WD3314	62.00	15.00	23.00	3094	WD3094	40.00	30.00	30.00
3314	WD3314	60.00	16.00	24.00	3094	WD3094	41.00	29.00	30.00
3314	WD3314	60.00	16.00	24.00	3094	WD3094	43.00	28.00	29.00
3314	WD3314	59.00	16.00	25.00	3094	WD3094	44.00	28.00	28.00
6997	WD6997	-1.00	-1.00	-1.00	6998	WD6998	-2.00	-2.00	-2.00
6997	WD6997	-1.00	-1.00	-1.00	6998	WD6998	-2.00	-2.00	-2.00
6997	WD6997	-1.00	-1.00	-1.00	6998	WD6998	-2.00	-2.00	-2.00
6997	WD6997	-1.00	-1.00	-1.00	6998	WD6998	-2.00	-2.00	-2.00
6997	WD6997	-1.00	-1.00	-1.00	6998	WD6998	-2.00	-2.00	-2.00
3043	WD3043	98.00	1.00	1.00	6693	WD6693	37.00	34.00	29.00
3043	WD3043	98.00	1.00	1.00	6693	WD6693	36.00	31.00	33.00
3043	WD3043	98.00	1.00	1.00	6693	WD6693	36.00	30.00	34.00
3043	WD3043	98.00	1.00	1.00	6693	WD6693	37.00	29.00	34.00
3043	WD3043	98.00	1.00	1.00	6693	WD6693	36.00	29.00	35.00
3732	WD3732	67.00	15.00	18.00					
3732	WD3732	62.00	15.00	23.00					
3732	WD3732	60.00	16.00	24.00					
3732	WD3732	60.00	16.00	24.00					
3732	WD3732	59.00	16.00	25.00					

Department of Physical Geography and Ecosystem Science, Lund University

Lund University GEM thesis series are master theses written by students of the international master program on Geo-information Science and Earth Observation for Environmental Modelling and Management (GEM). The program is a cooperation of EU universities in Iceland, the Netherlands, Poland, Sweden and UK, as well a partner university in Australia. In this series only master thesis are included of students that performed their project at Lund University. Other theses of this program are available from the ITC, the Netherlands (www.gem-msc.org or www.itc.nl).

The student thesis reports are available at the Geo-Library, Department of Physical Geography and Ecosystem Science, University of Lund, Sölvegatan 12, S-223 62 Lund, Sweden. Report series started 2013. The complete list and electronic versions are also electronic available at the LUP student papers (https://lup.lub.lu.se/student-papers/search/) and through the Geo-library (www.geobib.lu.se).

1 Soheila Youneszadeh Jalili (2013) The effect of land use on land surface temperature in the Netherlands
2 Oskar Löfgren (2013) Using Worldview-2 satellite imagery to detect indicators of high species diversity in grasslands
3 Yang Zhou (2013) Inter-annual memory effects between Soil Moisture and NDVI in the Sahel
4 Efren Lopez Blanco (2014) Assessing the potential of embedding vegetation dynamics into a fire behaviour model: LPJ-GUESS-FARSITE
5 Anna Movsisyan (2014) Climate change impact on water and temperature conditions of forest soils: A case study related to the Swedish forestry sector
6 Liliana Carolina Castillo Villamor (2015) Technical assessment of GeoSUR and comparison with INSPIRE experience in the context of an environmental vulnerability analysis using GeoSUR data
7 Hossein Maazallahi (2015) Switching to the "Golden Age of Natural Gas" with a Focus on Shale Gas Exploitation: A Possible Bridge to Mitigate Climate Change?
8 Mohan Dev Joshi (2015) Impacts of Climate Change on *Abies spectabilis*: An approach integrating Maxent Model (MAXent) and Dynamic Vegetation Model (LPJ-GUESS)
9 Altaaf Mechiche-Alami (2015) Modelling future wheat yields in Spain with LPJ-GUESS and assessing the impacts of earlier planting dates
10 Koffi Unwana Saturday (2015) Petroleum activities, wetland utilization and livelihood changes in Southern Akwa Ibom State, Nigeria: 2003-2015
11 José Ignacio Díaz González (2016) Multi-objective optimisation algorithms for GIS-based multi-criteria decision analysis: an application for evacuation planning
12 Gunjan Sharma (2016) Land surface phenology as an indicator of performance of conservation policies like Natura2000
13 Chao Yang (2016) A Comparison of Four Methods of Diseases Mapping
14 Xinyi Dai (2016) Dam site selection using an integrated method of AHP and GIS for decision making support in Bortala, Northwest China