Frank Segessenmann

Evaluating various geoinformatic methods of assessing access to water by using Open Data

Case study: Nam Ngum

Helsinki Metropolia University of Applied Sciences

Bachelor of Engineering

Degree Programme in Environmental Engineering

Bachelor Thesis

5 May 2017

Helsinki Metropolia University of Applied Sciences

Author(s) Title Number of Pages Date	Frank Segessenmann Evaluating various geoinformatic methods of assessing access to water by using Open Data. Case study: Nam Ngum 55 pages + 8 appendices 5 May 2017
Degree	Bachelor of Engineering
Degree Programme	Environmental Engineering
Specialisation option	Waste and Water Treatment Technology
Instructor(s)	Marko Kallio, Mentor, Ph.D. Student in Aalto Kaj Lindedahl, Senior Lecturer, Metropolia UAS

This thesis document contains a review of the major researches on different water indicators and indexes.The research on the subject first started in the mid-seventies, and the results are still valuable. The study chosen to be replicated was the Spatial unit non-overlapping WSI model (Sun-WSI) which use the resources found in the surrounding area of an active cell to calculate water stress. The Sun-WSI area defining the surroundings is a circle. Marko Kallio proposed to implement GIS distance analysis methods by using a village point to simulate the access available to water.

One of the biggest difficulties in replicating the methods was the hydrological modelling software. It was not possible to use the same software as in the Sun-WSI study. Marko Kallio had worked for his Bachelor thesis and for other projects in the area of Nam Ngum. He provided the hydrological modelling software from a Finnish company. The main benefit was the access to the output and input time series in the area of the Nam Ngum catchment. The IWRM model generated the discharge layer for the dry and wet season.

The next step was to implement the distance analysis methods. In order to create the surface cost layer, it was necessary to find the appropriate factors to define the cost. In this study slope steepness, road networks and land use was used emphasizing the access from any village. An accumulation layer was necessary in order to create an irregular shape around the starting point, and the starting point, a cost surface layer and a threshold was required to generate the accumulation cost.

Finally, by using Matlab, different algorithms were created to compensate for some limitations found along the way. As much as possible from the Sun-WSI was replicated, but most of its sophisticated techniques were not possible to put in place. On the other hand, the idea proposed by Marko Kallio was met with success. The Village's Accumulation Cost Area (VACA) was generated for each of the villages. The threshold and the accumulation cost layer were generated explicitly for each village. As expected, the water stress was reduced after the resource available inside the VACA was taken into account in the calculation.

Keywords	hydrological modelling, water stress indicator, GIS, distance analysis, hyper resolution

Contents

List of Figures

List of Tables

List of Abbreviations

BOD5	The 5 day biochemical oxygen demand
DEM	Digital Elevation Map
EIA	Environmental Impact Assessment Center of Finland
ET	Evapotranspiration
GDP	Gross Domestic Product
GIS	Geographic Information System
GLDAS	Global Land Data Assimilation System
HDI	Human Development Index
HDII	Handbook of Drought Indicators and Indices
HydroSHEDS	Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales
HWSD	Harmonized World Soil Database
IWMI	International Water Management Institute
IWRM	Integrated Water Resources Management
km^2	Kilometer square
m^3/cap/yr	Cubic meter per capita per year
MODIS	Moderate Resolution Imaging Spectroradiometer
MRC	Mekong River Commission
MS-DTVGM	Hydrological Model-Distributed Time-Variant Gain Model
OSM	OpenStreetMap
PDR	People's Democratic Republic
PDSI	Palmer Drought Severity Index
Sun-WSI	Spatial unit non-overlapping WSI model
TRMM	Tropical Rainfall Measuring Mission
T-test	Statistical test using the Student's t-distribution
USGS	U.S. Geologic Survey
VACA	Village's Accumulation Cost Area
WaterGAP	Water Global Assessment and Prognosis
WSI	Water Stress Indicator
WTA ratio	Water Resources Vulnerability Index

1 Introduction

In an age of excessive consumption and under the pressure of climate change, 193 states met in 2015 to adopt 17 goals in order to achieve a better society (1). Water is essential to humans and there are still more than 1.8 billion people around the world who do not have access to safe drinking water. The consequences are alarming; approximatively 42 children die every hour due to diarrheal diseases that are associated with water and sanitation issues (2). This, among other reasons, has led the United Nations to establish *clean water and sanitation* as one of the 17 sustainable development goals to be accomplished by 2030 (1).

The problem of accessing safe drinking water can be due to different factors such as improper infrastructure or lack of information, but it can also rise due to an increase in demand. Today the United Nation estimated that more than 40 percent of the world's population live in an area where water has become scarce. The scientist's projections that take into account the effect of climate change are showing an increase in water scarcity (3). Despite the climate change's effects, the population will continue to grow; consequently, the scarcity of resources such as water will increase.

The objective of this thesis was to review different water related indicators and indexes applied in the field of geographic information system (GIS). Water stress and scarcity indicator have been researched for more than 40 years, and a large number of research papers has been produced. Despite the great quantity of research, the technology available was, until recently, very limited. However, in recent years, there has been a sharp increase in technologic capabilities, from computer calculation power to investments in space technology. These increases will undoubtedly open new possibilities in the field of both GIS and environmental monitoring in the future.

After reviewing the current state of literature on the subject of water stress and scarcity indicators, the thesis focused in combining the potential of freely available information, such as OpenStreetMap, with the latest geoinformatic methods. The aim was to define water accessibility from a GIS perspective.

2 Theoretical background

2.1 Introduction

Inside this chapter is a summary of the major researches in the field of water stress and scarcity indicators. At first the indexes that define a drought will be introduced, and then a deeper review of water stress indicators will be presented. A special focus will be on a recent study, *Spatial unit non-overlapping WSI model (Sun-WSI)*, from China. The end of this chapter is on the methods required for the practical application of the case study on Nam Ngum.

2.2 Major drought indexes

A drought can be defined as an extended meteorological anomaly which has the properties of an unusual moisture deficiency (4). However, the definition of a drought has been the source of a long debate. There is a good reason for this confusion. The definition of a drought can be seen from different perspectives such as a meteorological, agricultural, hydrologic, geologic or geophysical. To define an index, one must choose which point to study. Below is a chronologic list of the major development in drought indexes. Each of the indexes is categorized for their ease to use as green, yellow or red; green being the easiest to implement. The categorisation is from the Handbook of Drought Indicators and Indices (HDII); it was made by the world meteorological organization (6). The aim of this classification is to help organizations around the world configure their own monitoring system and to provide a better ground for international cooperation.

1965	Palmer drought severity index (PDSI)
1965	Rainfall Anomaly Index (RAI)
1967	Deciles
1968	Crop Moisture Index (CMI)
1982	Surface Water Supply Index (SWSI)
1993	Standardized Precipitation Index (SPI)
1995	Crop-specific Drought Index (CSDI)
1996	Reclamation Drought Index (RDI)

2.2.1 Palmer drought severity index

Palmer based his study on a dataset from Iowa and Kansas. He required precipitation and temperature from the dataset in order to calculate the PDSI. The method is able to calibrate supply and demand for moisture at the top soil layer. In his method, the top soil layer is divided in two levels. From the precipitation and temperature his equation is able to calculate the potential evapotranspiration, potential recharge, potential loss and other sub-variables necessary (4). His study was the first to have such an extent on the subject and is still the most widely used regional index to monitor drought however PDSI has some limitations. For instance the index does not take into account snow and it underestimates runoff as it assumes all surface layers of the soil need to reach saturation before runoff happens. It is considered to be more adapted to agricultural drought than hydrological drought however there are modified versions of the PDSI that answer to other categories, like the Palmer Hydrological Drought Severity Index. The index is standardized so it can be compared across regions and it has stood the test of time (5). The PDSI is classified as yellow because of its complexity and the need for a serially complete dataset (classification from HDII).

2.2.2 Deciles

The Decile Index used precipitation as the only input and has a flexible timescale. The method has a well-defined threshold. The first decile contains the lowest 10 percent of the values and the fifth is the median. The index can be used for agricultural, meteorological and hydrological drought. By using only one input parameter, the index is limited in drought detection. The way to create Deciles is extremely easy and the Australian Bureau of Meteorology provides examples; thus it is categorized as green (HDII) (6).

2.2.3 Crop Moisture Index

Palmer continued his work on his drought index by making the Crop Moisture Index (CMI), which is crop oriented. The input parameters are precipitation and temperature, and the timescale necessary is a week. This allows the CMI to compare moisture levels from the long term average. The necessity for weekly values classifies it as yellow (classification from HDII). The main limitation of the index is that it is oriented to rapidly changing situations and as such does not compute properly when there is a recovery from a long drought (6).

2.2.4 Standardized Precipitation Index

The World Meteorological Organization (WMO) decided in 2009 to endorse the standardized Precipitation Index (SPI) as the principal meteorological drought index. The National Drought Mitigation Centre (NDMC) is proposing guide book and software package to assist with the implementation of the SPI. Those combined efforts are guiding countries around the world to adopt the SPI. The index requires only one input, precipitation, but it is recommended to run the method with monthly data for as long as 30 years, not missing any data. The timescale is flexible, from one month to two years or longer. The negative aspect of the SPI is to consider only precipitation, which makes scenarios difficult to compare. The classification of the SPI is green (6).

2.3 Major water stress indicators

Water stress can be defined at different scales like the river basin or a community, but it is generally analyzed at a bigger scale such as a country or worldwide. Several researches have been carried out. Their aim is to allow decision-makers to plan better policies in order to prevent water related issues to emerge or aggravate. Here is a chronological list of the most reviewed researches:

1989	The Falkenmark water stress indicator (FWSI)
1996	Basic Human Water Requirements (BHWRs)
1997	Water Resources Vulnerability Index (WTA ratio)
1998	Social Water Stress Index (SWSI)
1998	The physical and economical scarcity indicator (IWMI model)
2003	Water Poverty Index (WPI)
2005	Geospatial indicators of water stress
2005	Water Stress Indicator integrating environmental need (WSIEWR)
2007	The Watershed Sustainability Index (WSusI)
2010	The Water Supply Stress Index (WaSSI)

There are different ways to approach water stress. In some cases, the water is considered of same values in every location, but, in reality, water can have various quality levels. Some studies have included the infrastructure such as a water or wastewater treatment plant. The different parameters that each study decided to include depended on their objective and on the information available to them.

2.3.1 Falkenmark water stress indicator

The most commonly referenced method is the Falkenmark water stress indicator (FWSI) made in 1989. The FWSI define water scarcity by looking at the total runoff available per capita per year.

There are 3 different thresholds to classify a country using the FWSI. When a country drops under 1'700 cubic meter per capita per year ($m^3/cap/yr$), it is then classified as *water stress*. *Water scarcity* is considered when the WSI indicator is lower than 1'000$m^3/cap/yr$ and *absolute scarcity* when the indicator falls below 500$m^3/cap/yr$ (7).

Among its advantages, the FWSI can show the difference between man-made and naturally occurring water scarcity, the information required to produce the indicator is widely accessible and easy to comprehend. Those advantages have made it a very popular indicator. The over simplification of a complex issue has of course some disadvantages. Firstly, the water available is taken from the annual national averages, which will omit local and temporal scarcity. Secondly, it does not integrate infrastructures, and, thirdly, the thresholds are defined in a way that implies everybody has the same requirements. Therefore, it is widely accepted that the WSI cannot provide an accurate index for small area (7) (8).

2.3.2 Basic Human Water Requirements

The discussion surrounding the subject of basic human water requierements started already 20 years ago. It was discovered that the basic need for water could be separated in four categories; drinking water for survival, water for hygiene, water for sanitation services and modest household needs for preparing food. The author, Peter Gleick, advised governments to provide a minimum of 50 litres per day per person with no social exclusion. This minimum was set to be independent of external factors such as technology or culture (9). Falkenmark and Gleick went on to develop a standard indicator of 1'000 cubic meters per person per year which was approved by the World Bank.

2.3.3 Water Resources Vulnerability Index

The Water Resources Vulnerability Index (WTA ratio) is a division of the total annual water used by humans by the water available (see Equation 1) (7). The concept comes from the Use-to-Resource Ratio (common WSI) which was first introduced by Falkenmark and Lindh in 1976 (10). This large scale analysis was supported by the work on supply-demand analysis from a group led by Professor Igor Shiklomanov at the State Hydrological Institute in St. Petersburg. In 1993 they published an estimation of the water use by sector of human activities at that moment and future. The thresholds are as follows: when a country uses from 20 percent to 40 percent of its annual supply, it is considered as *water scarce* and when a country uses more than 40 percent it is considered as *severely water scarce* (8). This approach is providing a more precise indicator as it relies on the estimated water use supported by Shiklomanov's early results. However, the water available is not categorized by its quality. The WTA ratio does not provide information on the type of use and it does not integrate the capacity of a society to adapt to stress (7).

$$WTA_i = \sum_j water\ use_{ij} / water\ available_i \qquad (1)$$

Where "i" represent each location (e.g. grid cell, watershed and country) and "j" represent the sector of human activities (e.g. irrigation and domestic).

2.3.4 Social Water Stress Index

Ohlsson made the Social Water Stress Index (SWSI) with the aim to indicate the adaptive capacity of a nation. He used the Human Development Index (HDI) as a multiplier to the indicator made by Falkenmark (FWSI) in 1989 (8). The HDI integrate the life expectancy, education index and the standard of living to categorize a nation's development. The HDI was published by the United Nation Development Programme (UNDP) first in 1990 and is still in use today (11).

2.3.5 Physical and economical scarcity indicator

The physical and economical scarcity indicator was created by the International Water Management Institute (IWMI). The IMWI model takes a more comprehensive approach than the previous mentioned researches. Elements taken into account are the existing infrastructures for supplying water, the consumption which leave as evapotranspiration and the return flow, as well as the capacity of society to adapt via rules, investments and the advancement of techniques of irrigation. The IMWI model has the main disadvantage of being complex and therefore not very accessible. The wide variety of data required makes it also difficult to set up but its results have been widely used. The model was made from 2000 to 2025. If by 2025 a country could not provide the water demand, it would be considerate as *physically water scarce* and if the country had to invest a considerable quantity of money into the infrastructure it would be considered as *economically water scarce* (7).

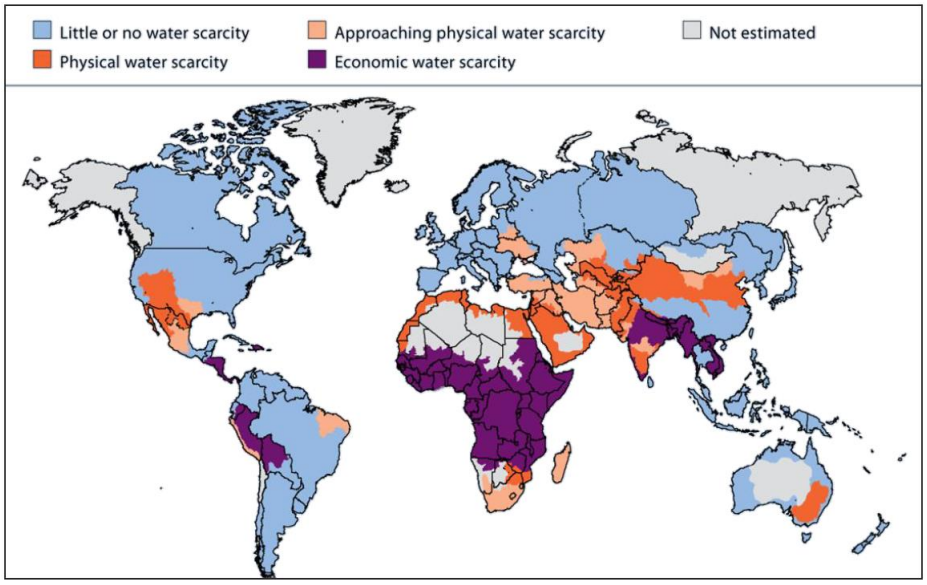

Figure 1. Analysis done on a basin level in 2007 (12).

A more recent use of the IWMI model (see Figure 1) brings an estimation of the situation in 2007. In that more recent study when a basin has more than 25 percent of its water withdrawn for human use, it is considered as *approaching water scarcity*, then with a withdrawal of more of 75 percent categorized it as *physical water scarcity* (12).

2.3.6 Water Poverty Index

The Water Poverty Index (WPI) was developed for a small scale study, for instance a community or a household. The WPI has a holistic approach to water; it takes into account the access, quality, quantity and different use of water as well as the ecological state (7). The method relies on the development of standardized factors to weigh each of the parameters. The largest issue is the assumption that those factors are correct for all locations (8).

2.3.7 Geospatial indicators of water stress

The Water Resources Vulnerability Index (WTA ratio) was used in 2005 with much more recent techniques. The author, C. J. Vörösmarty's aim was to visualize the situation with a homogeneous source of data that is free from political influence in order to do so he diverged from state provided information. He remarks that database such as the Global Runoff Data Centre is increasingly lacking the information necessary to run such analysis. The study was made based on simulated information from a water balance model and indices for the continent of Africa (13). In 2010 Mr. Vörösmarty participated in another article to assess the human water security (HWS) worldwide (14). The results from the 2010 study can be seen in Figure 2.

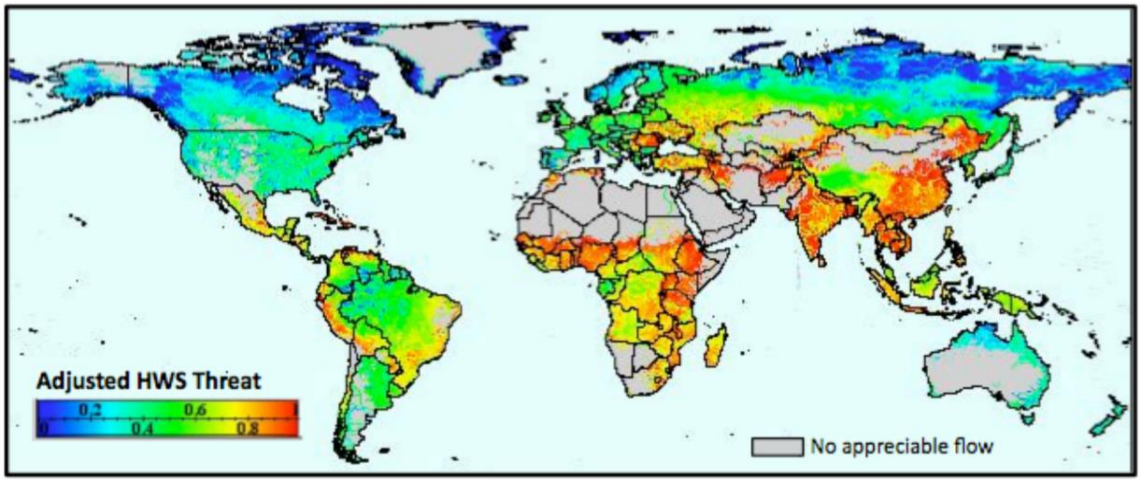

Figure 2. Global Human Water Security Threat Index (14).

2.3.8 Water Stress Indicator integrating environmental need

The indicator WSIEWR use the common WSI method with a slight modification to include the environmental flow required (see Equation 2) by the freshwater ecosystem to maintain life and was applied globally to assess the situation worldwide. It is the first study to try an assessment of the Environmental Water Requirements (EWR) at this scale.

$$WSIEWR = water\ withdrawal/(MAR - EWR) \quad (2)$$

The author named the water supply the Mean Annual Runoff (MAR). The MAR is comprised of the sum of surface runoff and groundwater recharge. The global water model generating the MAR is WaterGAP2. It is calculated on a monthly time frame. The climate inputs are based on a monthly time series from 1961 to 1990. The calibration of the model is based on the measured river discharge of 724 stations worldwide. The calculated long-term MAR and EWR are based on an historical dataset from 1995 to generated conditions as close as possible to the flow as it was before human activities modified it. The EWR was not generated to include seasonal change and is calculated as a monthly flow. The water use was also simulated using the model WaterGAP2 and include sub model for irrigation, domestic, manufacturing industry, livestock and thermoelectric power plant (15).

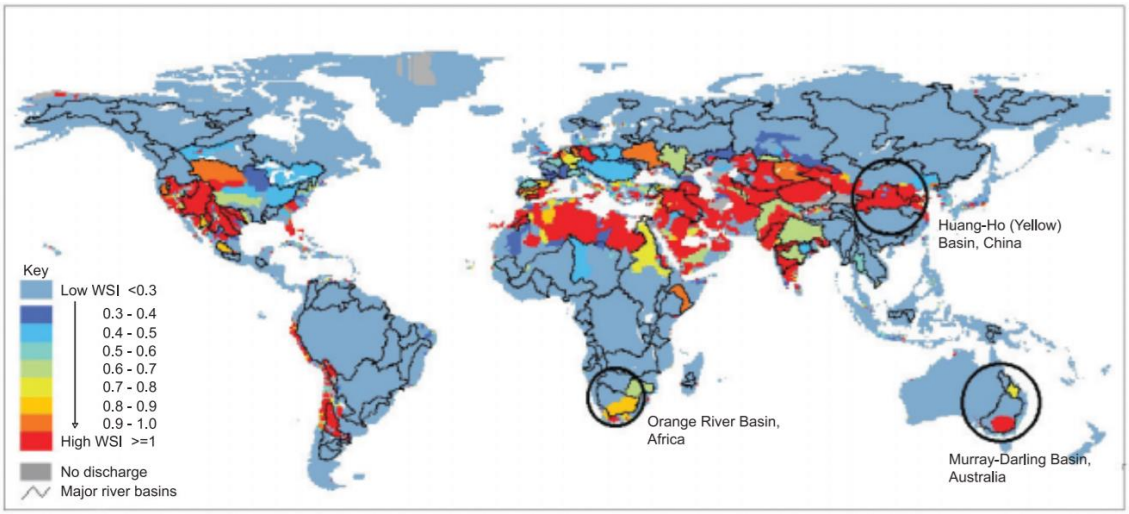

Figure 3. Map of the WSI taking into account EWR (15).

One of the disadvantage of this method as the author pointed out is that the modelled MAR will likely be underestimating the areas with snow covered mountains and overestimating area with arid and semi-arid climate. The study highlights the fact that using the common WSI method underestimates the problem of water stress greatly in some basins; see circled area in Figure 3.

2.3.9 Watershed Sustainability Index

The Watershed Sustainability Index (WSusI) was made to assess a river basin and the upper limit has been defined to be 2'500km^2. The idea behind its conception was to answer a need from UNESCO IHP International Hydrology Programme's initiative called HELP, standing for Hydrology Environment Life Policy. The WSusI took those four parameters (H, E, L and P) together to assess the sustainability of a basin (see Equation 3). Each of the parameters has three properties (pressure, state and response) and all four parameters equal weight (see Table 1) (8).

Table 1. Matrix of parameters and properties for the WSusI (16).

	Pressure	State	Response
Hydrology	- Variation in the basin's per capita water availability in the last 5 years; - Variation in the basin BOD5 (last 5 years)	- Basin per capita water availability - Basin BOD5 (yearly average)	- Improvement in water-use efficiency (last 5 years); - Improvement in sewage treatment/ disposal (last 5 years)
Environment	- Basin's EPI (Rural & urban)	- % of basin area with natural vegetation	- Evolution in basin conservation (Protected areas, BMPs)
Life	- Variation in the basin per capita GDP in the last 5 years	- Basin HDI (weighed by county population)	- Evolution in the basin HDI (last 5 years)
Policy	- Variation in the basin HDI-Ed in the last 5 years	- Basin institutional capacity in WRM	- Evolution in the basin's WRM expenditures in the last 5 years

$$WSusI = (H + E + L + P)/4 \quad (3)$$

Assessing hydrologic pressure is done using two different indicators, one for water quality and one for water quantity. For water quantity, the authors of the model rely on the threshold defined by Falkenmark via the FWSI. For the environmental pressure, the basin's EPI is a modified version of the Anthropic Pressure Index. The life and policy parameters are defined respectively by the income and education indicator from the Human Development Index (HDI) also used in the SWSI. The author created a relationship between the level of education (HDI-Education) and the Water Resource Management (WRM) results. This indicator provides a clear answer to UNESCO demand however the data required to perform the calculation can be challenging to obtain, such as the water quality indicator (BOD5). It does provide a clear relationship between cause and effect and allow the decision makers to assess time related changes (16).

2.3.10 The Water Supply Stress Index

The WaSSI use the common WSI method (see Equation 4). They scaled the research on the U.S. Geologic Survey (USGS) 8-digit Hydrologic Unit Code watershed, which translate into 2'100 watersheds. The water supply is defined by three parameters (total surface water supply, groundwater supply and return flow). The total surface water supply is generated using a water balance model (WBM). The groundwater supply and the return flow for domestic, commercial, industrial, irrigation, livestock, mining and thermoelectric sectors come from annual historical dataset. The water demand integrates the water use of the seven activities, the public use and the losses during the transfer (17). See Figure 4 for the results of the WaSSI. The transition to water stress occurs at 0.2 and from stress to scarce at 0.4.

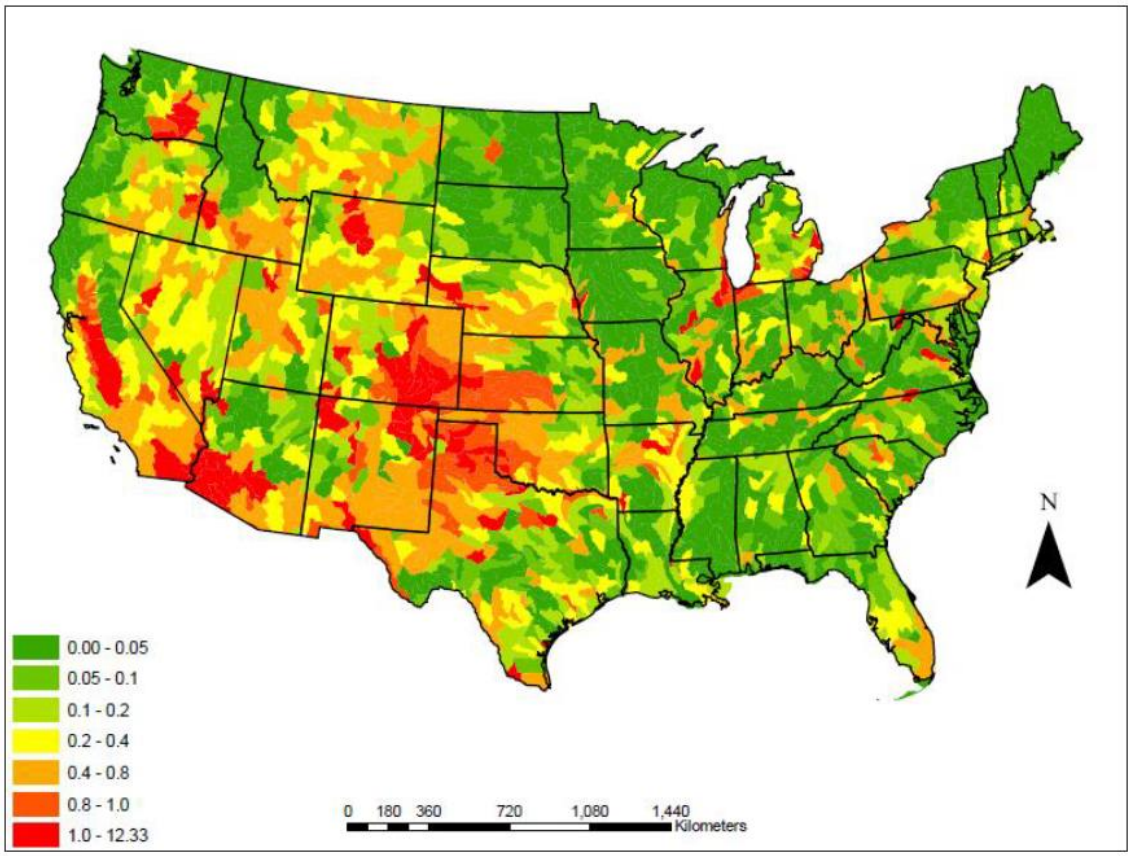

Figure 4. Annual Water Supply Stress Index (WaSSI) (17).

$$WaSSI_x = Water\ Demand_x / Water\ Available_x \quad (4)$$

The study revealed that, for the USA, the sector that would most probably see most of the changes in water withdrawal were irrigation, domestic and thermoelectric power

plants. The WaSSI model was then used in another study to see the relationship between water, carbon and biodiversity. One of their conclusions was that to lower evapotranspiration the optimisation of irrigation was preferable to massive deforestation (18). This model demonstrates that with the use of robust dataset a simple ratio such as the common WSI can provide solid results.

2.4 Life Cycle Assessment and water use

In 2009 a research was done to improve life cycle impact assessment (LCIA) on water use. The research outlined the possible amelioration like for instance the addition of water quality qualification for returning flows. The study case was on the production of cotton worldwide. WTA ratio method was used to calculate the water stress at the watershed level. The WaterGAP2 model was used to calculate the WTA ratio. The database from WaterGAP2 is based on an historical dataset from 1961 to 1990. The WTA ratio was then corrected by using a variation factor to compensate for strongly regulated flows. They relied on Nilsson C. study on “Fragmentation and flow regulation of the world’s large river systems”. To complete the water stress indicator used in the LCIA, they adjusted it using a logistic function (19). See the characterization and damage factors on the watershed level per cubic meter water consumed in Figure 5.

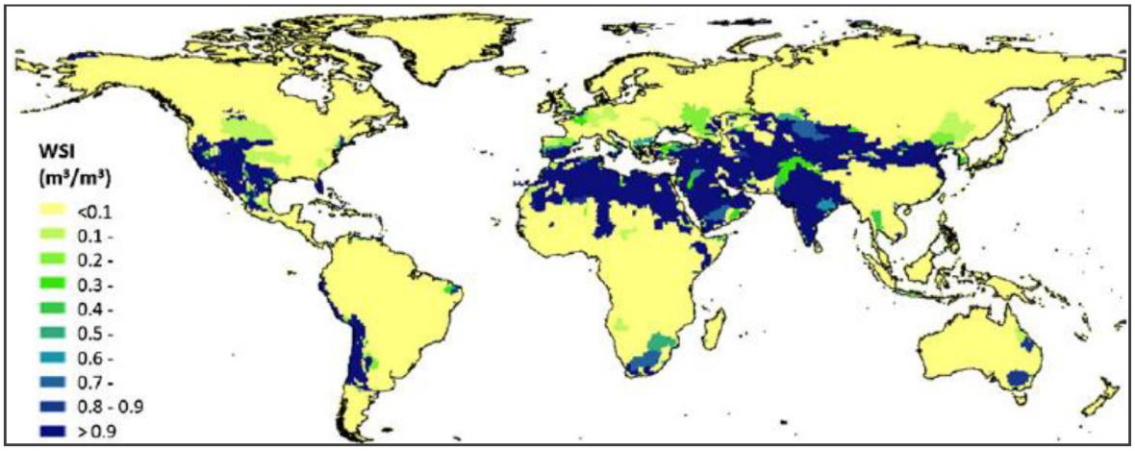

Figure 5. Water Stress Index from the LCIA study (19)

2.5 Water footprint

The concept of water footprint has been mainly developed within an international group called Water Footprint Network (WFN). Their goal has been to find a standard process to define water footprint, provide guideline on accounting and to create a global community. The goal of this process is to account for the type of water being used to produce a product or to provide for a business. A difference is made between three types of water footprints (green, blue and grey; see Figure 6). The green water is the rainfall water that stays on the top soil layer and is available to the vegetation. It will evaporate or transpire through plants. Green water is not part of the run-off or the groundwater. The blue water is the freshwater from rivers, lakes and groundwater. When accounting for blue water, the footprint method does integrate the concept of environmental flow requirement to sustain its ecosystem. The aim is to provide a common base, for instance, to calculate the cost from the cradle to the grave of a product direct and indirect use of water. The grey water footprint plays an important role in this process. It is calculated on the volume of freshwater necessary to dilute and assimilate the pollutants so that the release water has a quality that matches the regulation. The WFN is also working on bringing this methodology into an application to ease the implementation of the methods and create a database on basic industrial processes. The concept is still under development, but the community behind it is growing all the time (20).

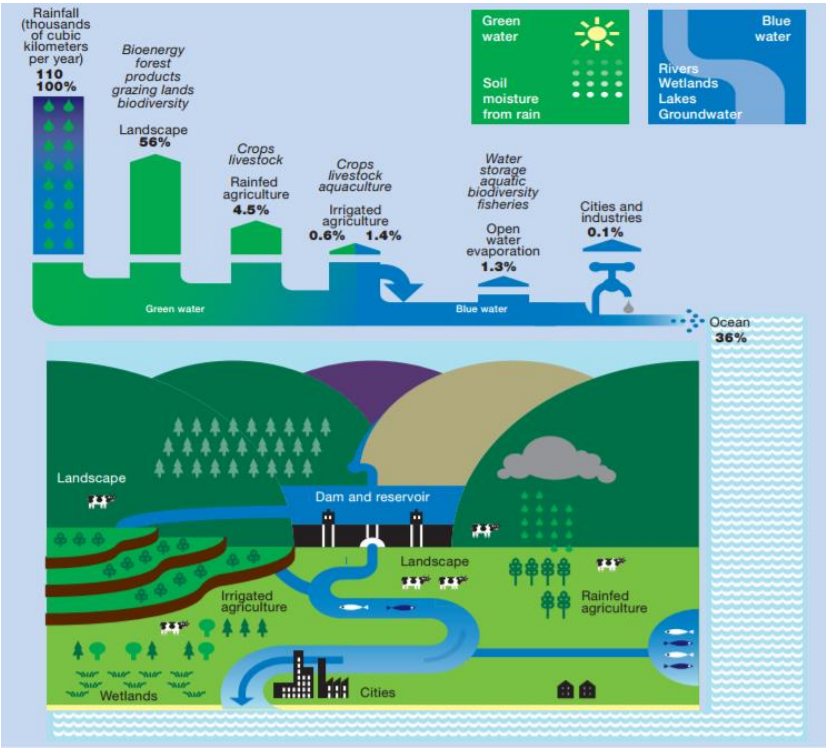

Figure 6. Global water use.

2.6 Spatial unit non-overlapping WSI model (Sun-WSI)

In 2016 a group of researchers from China worked on a study on the Yarlung Tsangpo-Brahmaputra River (TBR) which is flows from the southwest Himalayas of the Tibetan Plateau in China and finish its course in India. They used data from the countries official statistics (GDP, GPP and population density) but in majority from remote sensing sources. To define the indicator, the researchers used the same Use-to-Resource ratio (common WSI method). Normally, this formula (water demand / water available) is used on each grid cell of the map and each cell is treated independently; if there is not enough water to provide for the cell in question, then there is shortage of water (see Figure 7).

Researchers demonstrated this concept does not make sense when pushing the analysis on a map with a higher resolution (for example 1km^2). They, therefore, proposed to use a circle around each cell to test if there is water near the point in question. The radius of the circle is defined by several variables such as the GDP, slope steepness, elevation and distance to the main river. They used a model called Hydrological Model-Distributed Time-Variant Gain Model (MS-DTVGM) (21) in order to obtain the runoff. The inputs of the model are from the following sources:

- Moderate Resolution Imaging Spectroradiometer (MODIS) products
- TRMM rainfall products 3B42 V.6 (daily precipitation)
- Global Land Data Assimilation System (GLDAS) (hourly surface air T.)
- Digital soil map (Harmonized World Soil Database)
- Shuttle Radar Topography Mission (SRTM) digital topographic data

In order to use the GDP, GPP and population density, the authors have used methods developed in several other researches in order to downscale the country indicator into a grid cell mask (22) (23). They named their indicator Sun-WSI. To check if their results made sense, they compared the correlation between their results and a drought index. The authors concluded that their solution provided more accurate results for smaller area.

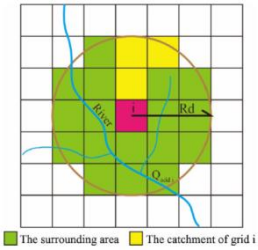

Figure 7. Sun-WSI area.

2.7 Hydrological modelling

A model can simply be a mathematical formula describing a physical phenomenon. In the case of a hydrological model the goal is to predict how water will travel through a geographic region. In general these models are applied in the region of a certain basin. Different models have different parameters but the base consists of rainfall and a river network. The purpose of the model defines the input parameters (24).

A model can be classified using different categories. First the model is deterministic or stochastic; in a stochastic model there is randomness and so with the same input the results will always vary while with a deterministic model with the same input the output is always the same (24).

The model can be lumped or distributed; a lumped model does not take into account space while a distributed model will provide a gridded cell taking into account the spatial dimension of the basin (24).

The model is classified as static or dynamic. A static model will describe the situation at a specific moment in time while a dynamic model is made to answer questions related to continuous timeline (24).

Models can be differentiating between three major groups. First it can be an empirical model which is a type of model that tries to answer the question of what is the relationship between the input and the output. It does not try to base itself on a physical process but rather focus only on the possible mathematical relationship. Examples of empirical models are ANN and unit hydrograph. Secondly are the conceptual models which are based on the generalisation of a process by understanding how the process works and what the biggest contributors are. Conceptual models require a large quantity of meteorological and hydrological data to be properly calibrated. Examples of conceptual models are HBV and TOPMODEL. Thirdly are the mechanistic models, they are based on separated physical processes that describe each component with a mathematical formula. This model will first analyse how the system work and then describe it using independent formulas and coefficients. The mechanistic model is at the opposite thinking of an empirical model. In this model there is not necessarily a need for a large quantity of data but rather information concerning each of the coefficients to be used. Examples of mechanistic models are SHE, MIKESHE and SWAT (24).

2.7.1 Hydrological cycle

Water follows a cycle and each stage of the cycle contributes to the equilibrium of the climate. When modeling the displacement of water in its different physical phases it is important to remember what are the stages of which composed the cycle (see Figure 8).

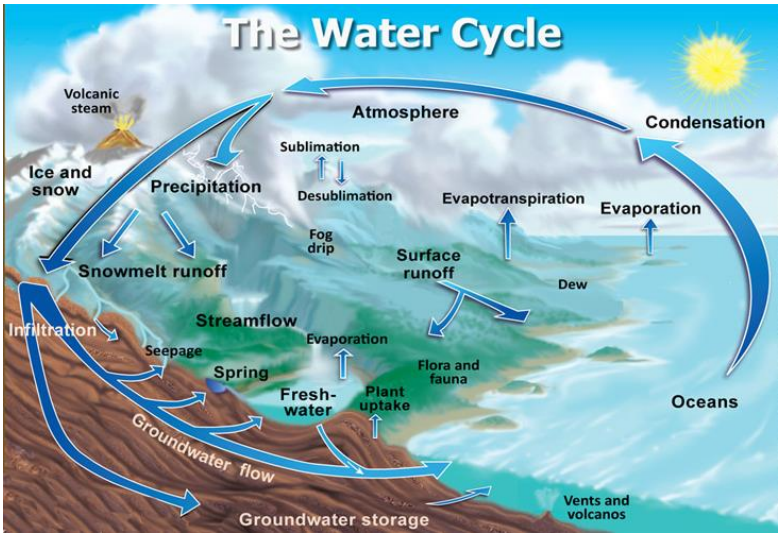

Figure 8. Water cycle.

The major processes are explained below and will most probably be present in the major hydrological models available today. Each of those processes can be simulated using mathematical formulas.

2.7.1.1 Evaporation

The water can evaporate from a drop of rain, the roof of a house or even a leaf of a plant. The change of state from liquid to gas has an energy cost and therefore when water evaporates it cools its surroundings. Sun, temperature, pressure as well as wind are factors affecting the rate of evaporation. With higher temperature the atmosphere can hold more water vapour.

2.7.1.2 Condensation

As water vapour the condensation causes the water to change state from gaseous to liquid, creating clouds or dew. This change of state returns the energy used previously to the environment. Decreases in temperature or saturation of water vapour are the main factors of condensation.

2.7.1.3 Precipitation

Precipitation is when the water drops in the atmosphere reach a considerable size and the gravity entrains it to the ground. It is therefore mostly due to an accumulation of water in a liquid state. When rain is rare then much of the precipitation evaporates back into the atmosphere.

2.7.1.4 Interception

Interception of water is the act of interrupting the process between precipitation and the flow of water in rivers. This can happen for example in a dense forest with trees having large leaves. The water that settles for instance on the leaf is therefore intercepted and can evaporate again (25).

2.7.1.5 Infiltration

Infiltration into the soil is influenced mostly by the type of soil and the soil's conditions at the time water reaches it. When water reaches the ground it tries to enter the upper layer of the ground, if the water does not enter this first layer then it becomes a runoff on the surface of the soil (25).

2.7.1.6 Percolation

Percolation defines the movement of water through the ground. The water moves underground thanks to the force of gravity. The way the water moves depends on the geology of the terrain and the water supply already presents (25).

2.7.1.7 Transpiration

Transpiration occurs when plants react to the sun. Water contained in the leaves is transmitted as a gas into the atmosphere. This allows the leaves to decrease their temperature. The important factors affecting transpiration are related to the biology of plants and the sun (25).

2.7.1.8 Runoff

The flow of rivers is also called runoff. Water that does not infiltrate into the ground is a surface runoff before it reaches the river. The part of the water that seeps into the first layer of the soil is called subsurface runoff and will reach the river. Another flow of water considered as part of the runoff is the groundwater that later joins the river. With direct precipitation into the river all these natural components form the runoff (25).

2.7.2 Integrated Water Resources Management

IWRM stands for Integrated Water Resources Management. The aim of model based on IWRM is to integrate water in the centre of the simulation and connect it to the major factors of the environment. Factors that should be taken into are for instance hydrology, land use, erosion, water quality, agriculture, forestry and many others. The aim is to have as large as possible panel of factors integrated into the same model so that the decisions made from the results are not based on only a specific part of the problem at hand (26).

2.8 Distance analysis

GIS software allows you to define each point on the map in the form of a grid. Each cell in the grid has the same size as the rest of the map. In order to define the access from point A to point B there is a different approach. In general, these methods are called distance analysis.

The most common method is to define a cost surface layer. This layer makes it possible to bring to each cell of the grid a price to cross it. The procedure entails collecting different layers of data that will define the price, for example, a slope of a hill may be one of the parameters used. When it is necessary to define a common price unit between different parameters the process becomes very random. For example, the degree of a slope is not measured in the same unit as the category of a road but each of the parameters will influence the result. It is therefore obvious that defining a common unit of cost it brings uncertainty and therefore is considered as an approximation of the possible result. Once each of the data sources has been classified under the same unit it is necessary to weigh each parameter before adding them together which will lead to the creation of a cost layer (27).

Once this step is completed, the cost of moving from point A to point B can be defined by the cost accumulation method. Those methods are available in the major GIS software. Once the accumulative cost layer has been generated it is possible to see from the point A until where you can travel in each direction for the same price; this creates an area around the starting point.

3 Methodology & Data

The new twist of the Sun-WSI study is, for example, the use of available resources in cells adjacent to the active cell. However, in this study they also mention that the use of a perfect circle to define the area adjacent to the active cell does not respond to a real concept but is rather a simplification of the initial idea. Therefore the aim is to take up this idea and define area of use via the structure of roads combined with other layers to build a cost surface layer to emphasize the accessibility.

The application of Sun-WSI's methods requires a hydrological model to generate a discharge layer which defines the water available in the area of study. In this study they generate the model using Remote Sensing data. The authors establish the need for water in the different sectors of activities such as industrial, agricultural and domestic. This step is done through the use of government statistical data and new innovative methods that allow downscaling country statics to grid cells. Then they evaluate the initial water stress indicator using the resources available in each active grid cells without using adjacent cells. In the end they use the circle around the active cell to calculate a new water stress indicator. The radius of each circle is defined by the following formulas:

$$RRd_{ini,i} = (normalized\ GDP_i) + (1 - normalized\ \ slope\ steepness_i) + (1 - normalized\ elevation_i) + (normalized\ distance\ to\ stream_i) \quad (5.1)$$

$$T_i = normalized\ RRd_{ini,i} \quad (5.2)$$

3.1 Water available and discharge layer

In order to create a set of data in a georeferenced grid concerning the water availability it is necessary to use a method to model the run-off and the river discharge. The use of different models can be seen from various research analyzed previously. The Sun-WSI study used the MS-DTVGM model to generate the hydrological information; unfortunately the software is in Chinese so it could not be used in this thesis project. In the literature review the most often referred model is the WaterGAP2 model. Its implementation does not go through a basic software installation but rather via the addition of several data sources inside a GIS application as modules (28). It has been considered too complicated to implement within the time frame. There is also the possibility to directly download the results obtained via the WaterGAP2 but they are of very low resolution and therefore do not match the objective of this thesis. After a discussion with my mentor, Marko Kallio, it was decided to use the software developed by a Finnish company called Environmental Impact Assessment Center of Finland Ltd. (EIA Ltd). EIA has developed for now more than twenty years "IWRM hydrological model" as a module of the main application. By using this model we can also utilize datasets previously used in other researches done in the area of Lao PDR (29). This also helped to define the area onto which the method would be applied. You can see a print screen directly from the software with a preview of one of the process parameters windows in figure 9.

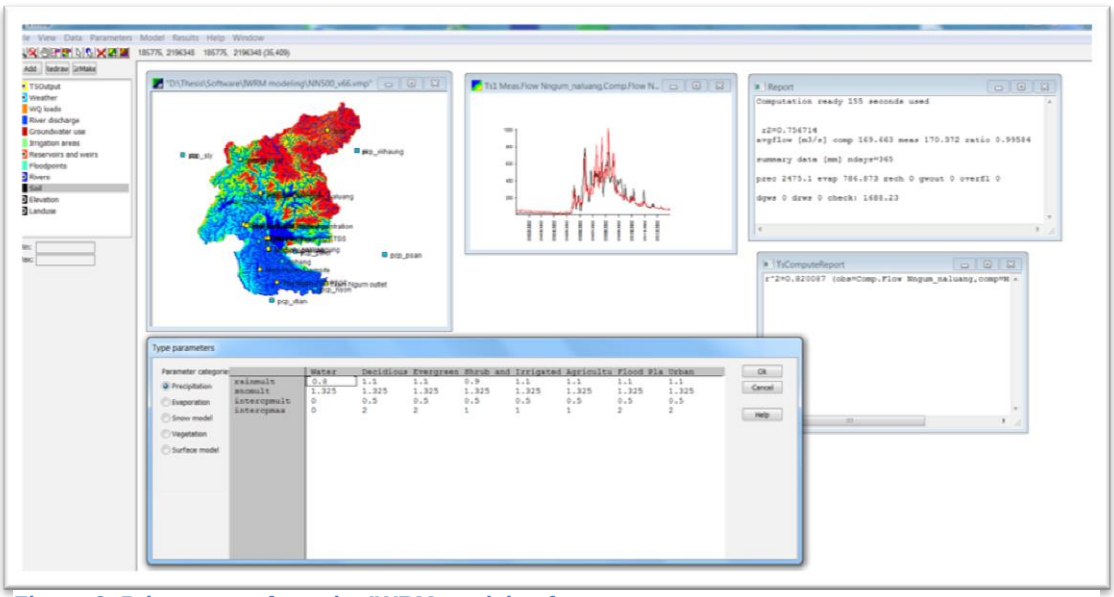

Figure 9. Print screen from the IWRM model software.

3.2 Study area

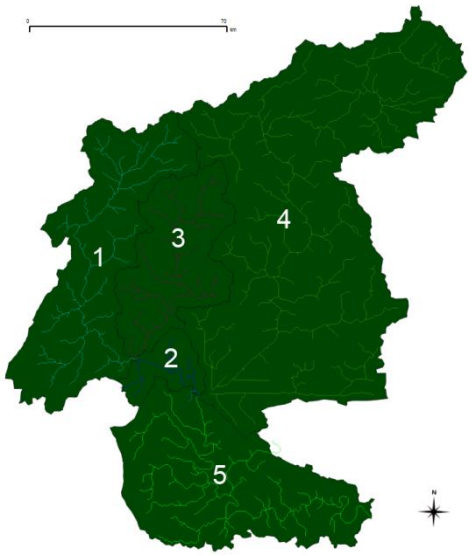

Figure 10. Nam Ngum catchment.

As it was explained in the previous point the data available with the modelling software offers an excellent base to do the work required. As seen in Figure 10 the area in question can be separated in 5 sections. Those 5 parts are as follow:

1. Nam Lik (basin, 3057 km^2)
2. Nam Lik (inter basin, 456 km^2)
3. Nam Song (basin, 1812 km^2)
4. Nam Ngum (basin, 8345 km^2)
5. Nam Ngum (inter basin, 3248 km^2)

The area is also known under the name of the catchment of Nam Nngum. Between the Nam Nngum basin and the inter basin there is the Nam Ngum reservoir. You can see it from satellite imagery on Figure 11. Nam Ngum has a dam which regulates the flow into the inter basin that follow and that point is integrated in the model. The flow of water is regulated by the demand in electricity and as such cannot be predicted by the model. This is the reason why the calibration is done in the Nam Ngum basin area (area 4).

Figure 11. Imagery from ESA satellite Sentinel II of the Nam Ngum reservoir.

3.3 IWRM Hydrological model

The model developed by EIA is an IWRM distributed model. It is partly based on a mechanistic model and on a conceptual model (HBV). The hydrological processes are simplified using physically based formulas. The model is dynamic, deterministic and distributed. Because it is partly based on a conceptual model the time series have a significant effect on the modelling results. The IWRM model has a GIS interface and rely on different types of layers to run:

- Basin boundaries (polygon or also called vector layer)
- River networks (lines, vector)
- Lakes (polygon, vector)
- Structures (point, vector)
- Digital Elevation Model (DEM) (grid of cells or also called raster layer)
- Land use (raster)
- Soil type (raster)

In the case of Nam Ngum's catchment the time series available includes:

- 1 defined reservoir (Nam Ngum reservoir),
- 11 meteorological points (11 with precipitation and 2 with temperature),
- 16 output points with flow measurements

The time series are based on daily measurements from 1995 to 2008.

The following sections of parameters influenced the results for the discharge layer:

- Weather interpolation (precipitation, temperature, and wind)
- Land use type (precipitation, evaporation, vegetation, surface model)
- Soil types (infiltration, soil layer 1 & 2, erosion).

Each of the sub sections contain coefficients for each type of soil and land use. Other sections are present in the software but they are not used and therefore not covered in this document.

An estimate of the quality of the data and results was a necessary step to undertake when the model was received. In order to estimate the model, an exit point was chosen before Nam Ngum dam which is before the Nam Ngum inter basin. The point chosen is an output flow measurements point and it is the point called Nalaung. From the initial parameters and for a one year time period the model is able to explain 68 percent of the total variation (R^2).

3.3.1 Qualitative analysis of the model inputs

The model is based on the application of formulas, which use parameters, onto layers of information. The provided layers for land use looks relatively close to the reality which is not the case for the soil type layer; see Figure 13 below.

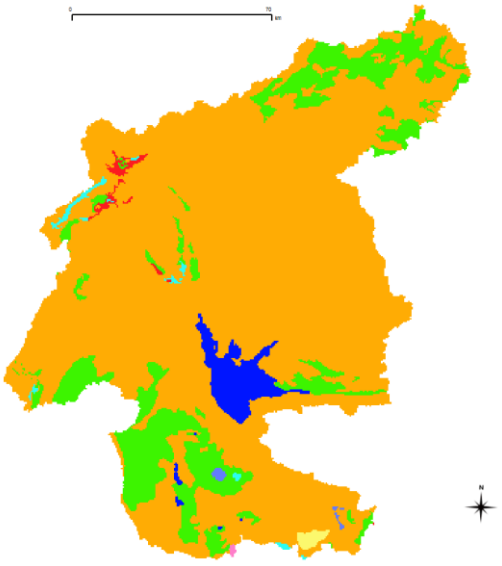

Figure 13. Soil type layer.

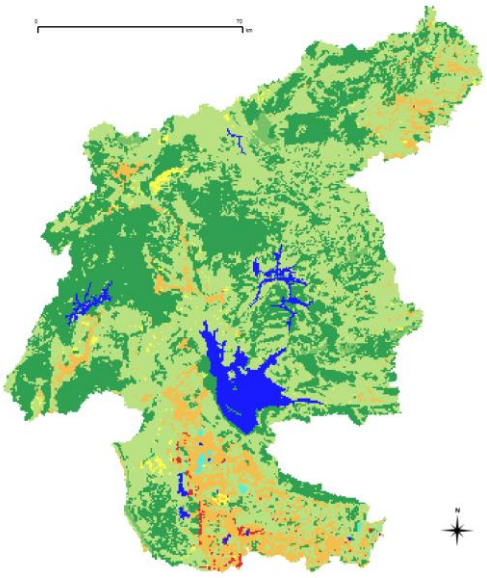

Figure 13. Land use layer.

Soil type categories:

- acrisols(81%)
- lithosols(14%)
- water(3%)
- argic
- histosols
- ferrosols
- alluvial

Land use categories:

- shrub and grassland (49%)
- deciduous forest (36%)
- agriculture (1%)
- water body (4%)
- evergreen forest (1%)
- irrigated agriculture (9%)
- urban (0.4%)
- floodplain (0.2%)

Harmonized World Soil Database offers a soil map from the world. As shown in the Figure 14 the Chinese territory is clearly and precisely mapped while the northern part of Lao PDR has almost no information. For the area of the Nam Ngum catchment proper information for the soil type layer seems to be unavailable and therefore this key element of the model cannot be optimized properly.

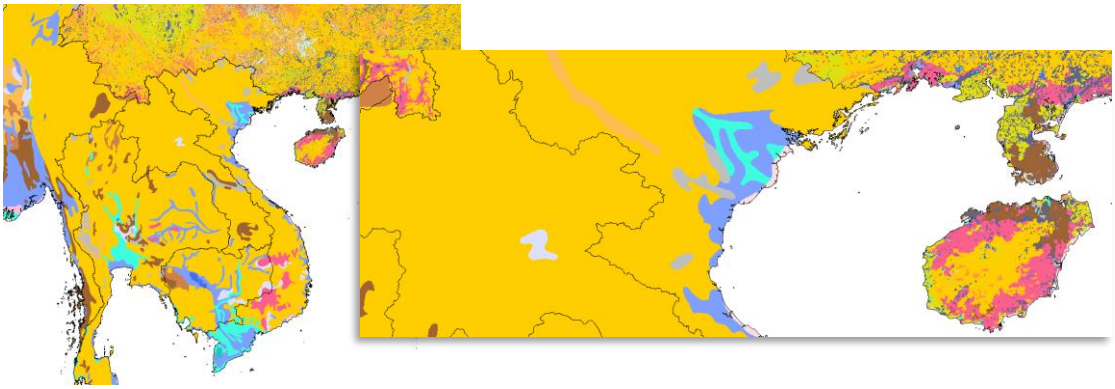

Figure 14. Harmonized World Soil Database view on Loa PDR.

There is another factor playing an important role in the results of the model; the time series. Because the model is partly running on a conceptual base it is important to have enough time series to run the calibration.

The output time series with the flow measurements at different points of the rivers are not widely accessible. One of the websites proposing this type of information in the area is the Mekong River Commission (MRC) but the use of those time series requires special authorization.

The input time series constitute the precipitation and surface air temperature from the 11 weather stations. The temperature is part of the evapotranspiration calculation and the model came with two time series out of the 11 weather station points.

After some research it was found that good resolution datasets on surface air temperature are not easily accessible. For instance the Global Land Data Assimilation System (GLDAS), used in Sun-WSI, provides this type of dataset but require a complex installation of libraries to build remote access software on a Linux operating system.

The study Sun-WSI used the TRMM rainfall products 3B42 from NASA for the daily precipitation so it was decided to retrieve this information and compare it to the dataset obtained via the IWRM model. The spatial resolution of the NASA dataset is 0.25 degrees (approximatively 30km*30km); see the size of square coloured in grey in the Figure 15. The time series from the IWRM model are represented by the weather stations points in the same figure.

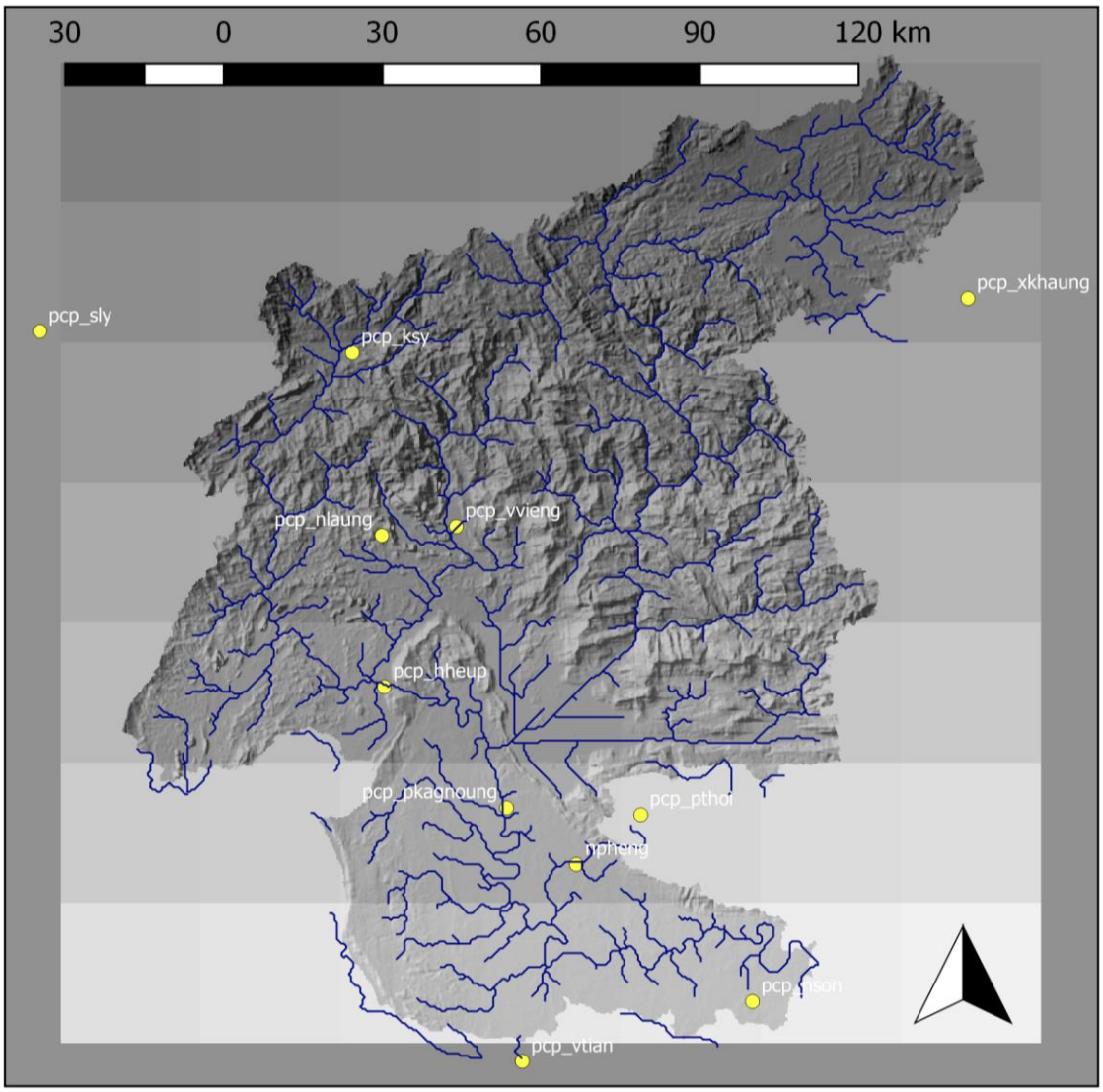

Figure 15. Weather station's position and visible resolution of TRMM products.

There is a possibility that the dataset from IWRM contains errors and the same can be said about the TRMM rainfall products. A quick statistical comparison of each location was done via Matlab. The results can be seen in the Table 2. T-test on the two datasets from the weather station named pcp_xkhaung is not statistically significantly different and the total difference throughout the entire timeframe (daily measurements from 1998 to 2008) is only 75 millimetres however correlation between them is only 35 percent. After zooming on a plot for this specific weather station it is difficult to identify a systematic difference, see Figure 16.

Table 2. Basic statistical comparison of NASA and IWRM rainfall time series.

Name of weather station	Source	Mean	Standard deviation	Range	Total sum of rainfall	Difference in total rainfall	Sum of rainfall likeness	Correlation	t-test with significance level 0.05	t-test p-value
pcp hheup	Nasa	6.5	13.7	143	25'960	5'652	78%	**16.00%**	Statistically different	3.07E-07
	IWRM	5.1	13.1	202	20'308					
pcp ksy	Nasa	5.6	11.4	106	22'499	4'789	79%	30.00%	Statistically different	3.83E-08
	IWRM	4.4	11.8	203	17'710					
pcp nlaung	Nasa	6.2	13.4	185	25'045	-2'880	112%	34.00%	Statistically different	8.20E-03
	IWRM	7.0	16.2	180	**27'925**					
npheng	Nasa	6.7	14.9	173	26'792	2'037	92.40%	17.00%	Not statistically different	9.66E-02
	IWRM	6.2	15.2	151	24'755					
pcp nson	Nasa	6.7	15.0	136	27'022	5'694	78.93%	33.00%	Statistically different	2.40E-08
	IWRM	5.3	12.7	226	21'328					
pcp pkagnoung	Nasa	6.7	14.9	173	26'792	1'816	93.22%	27.00%	Not statistically different	1.32E-01
	IWRM	6.2	16.5	259	24'976					
pcp pthoi	Nasa	7.3	15.5	144	29'321	-4'834	116.49%	34.00%	Statistically different	3.12E-04
	IWRM	8.5	20.6	251	34'155					
pcp sly	Nasa	4.1	8.9	95	16'525	2'211	86.62%	33.00%	Statistically different	1.20E-03
	IWRM	3.6	9.6	122	14'314					
pcp vtian	Nasa	5.1	12.7	241	20'573	810	96.06%	32.00%	Not statistically different	3.94E-01
	IWRM	4.9	13.0	162	19'763					
pcp vvieng	Nasa	6.2	13.4	185	25'045	**-16'769**	166.96%	**40.00%**	Statistically different	**4.75E-36**
	IWRM	10.4	22.3	199	**41'814**					
pcp xkhaung	Nasa	4.4	10.1	91	17'865	-75	**100.42%**	35.00%	Not statistically different	**9.19E-01**
	IWRM	4.5	10.4	123	17'940					

From each comparison a difference arises; sometimes there is a great statistical difference and sometimes there is no statistical difference. In some cases the correlation is low and in other cases a bit higher. With 40 percent of correlation the weather station of vvieng has a difference of more than 166 percent when looking at the total rainfall recorded over the ten year period. It is possible to see the plot of the weather station pcp hheup, pcp nlaung, pcp vvieng and pcp xkhaung in the Appendix 1.

From the Table 2 and the different plots analysed it was concluded that the data does not match. It can be due to a series of different errors included in each of the time series or because of the difference in the resolution. In fact there is a great difference between the tools used to generate the time series. In the case of the TRMM rainfall product 3B42 estimates are calculated from microwave precipitation estimates and infrared precipitation estimates, both from sensors situated on board of NASA satellite (30). Those two estimates are sent through a series of operations to produce the final product which covers three hours of precipitation estimates. Each of the eight parts of

three hours represent together one day. Depending on which latitude the point of interest is, different parts of the eight steps of three hours are added together to form one day; which is not very precise. On the other hand the time series already included in IWRM model came from different weather stations which are using one site sensors. The sensors may change over time and they may also differentiate from one location to another. This explains why the time series might represent the same phenomena but differs.

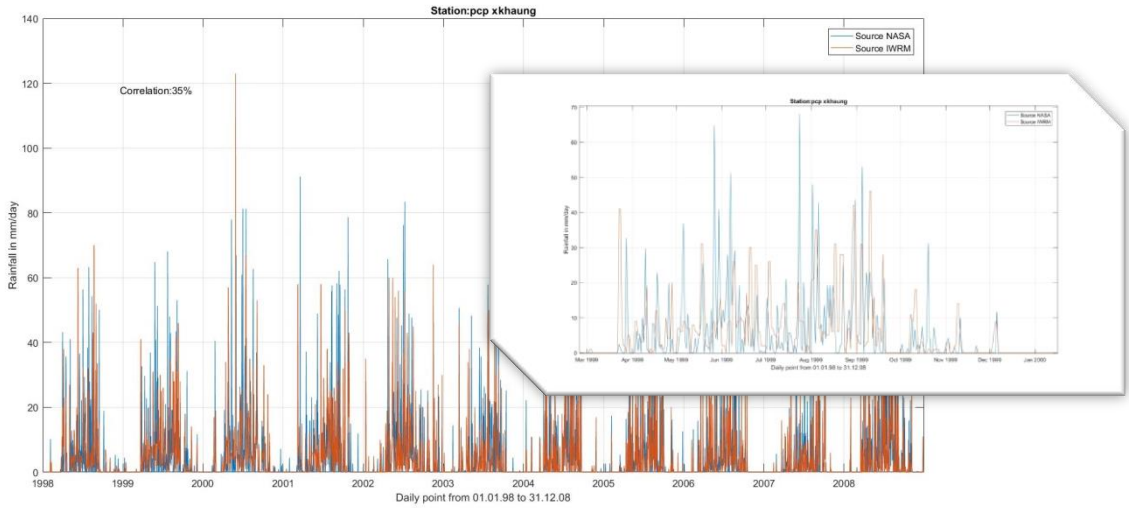

Figure 16. Plot IWRM and NASA rainfall time series, zoom on 1999.

It is also important to note that if the TRMM rainfall product was to be added to prolong the time series from 2008 to 2016 then it would be necessary to do the same for the river flow measurements at the output points. From a statistical point of view it seems that the most important factor is to provide additional time series for the temperature measurements which would have brought more precision on the evapotranspiration process calculation. Additionally the improvements of the soil type layer would have helped choosing more realistic coefficients for the model parameters.

3.3.2 Improvement of the IWRM model's results

From the qualitative analysis of the model's input it was concluded that replacing the time series was not an option and that improving the soil type layer was also not possible therefore the improvement of the model's results could be done through the modification of the coefficients (see Figure 17). After testing the different parameters in IWRM the following parameters have been found to be the most influential for the discharge layer:

- Land use:
  - Precipitation/Rainmult: multiply precipitation values by this number
  - Evaporation/Petcorr: potential evapotranspiration correction coefficient
  - Surface model/Infpratemult: intensity factor of precipitation

- Soil type:
  - Infiltration/infkz: vertical conductivity for the infiltration model
  - Soil layer 1/dz1: layer thickness
  - Soil layer 2/dz2: layer thickness

Each different land use type and soil type has values for each of the parameters above. In this model we have eight different land use and eleven different soil type. The major issue in optimizing the model is that each of the parameters for each of the land use type and soil type are dependent on each other. Therefore it was decided to reduce the number of land use type and soil type to the biggest contributor of the model. Here is the list of the selected type:

- Land use
  - Shrub and grassland (49%)
  - Deciduous forest (36%)
  - Irrigated agriculture (9%)
  - Water (4%)
  - All the others
- Soil type
  - Acrisols (81%)
  - Lithosols (14%)
  - Water (3%)
  - All the others

The strategy to reach for a better model, in this case due to the imprecise layer, is done by searching for the maximum values for both the normal R-squared value and Nash Sutcliffe model efficiency. In other cases this could be achieved by examining the formulas and finding the appropriate values from literature.

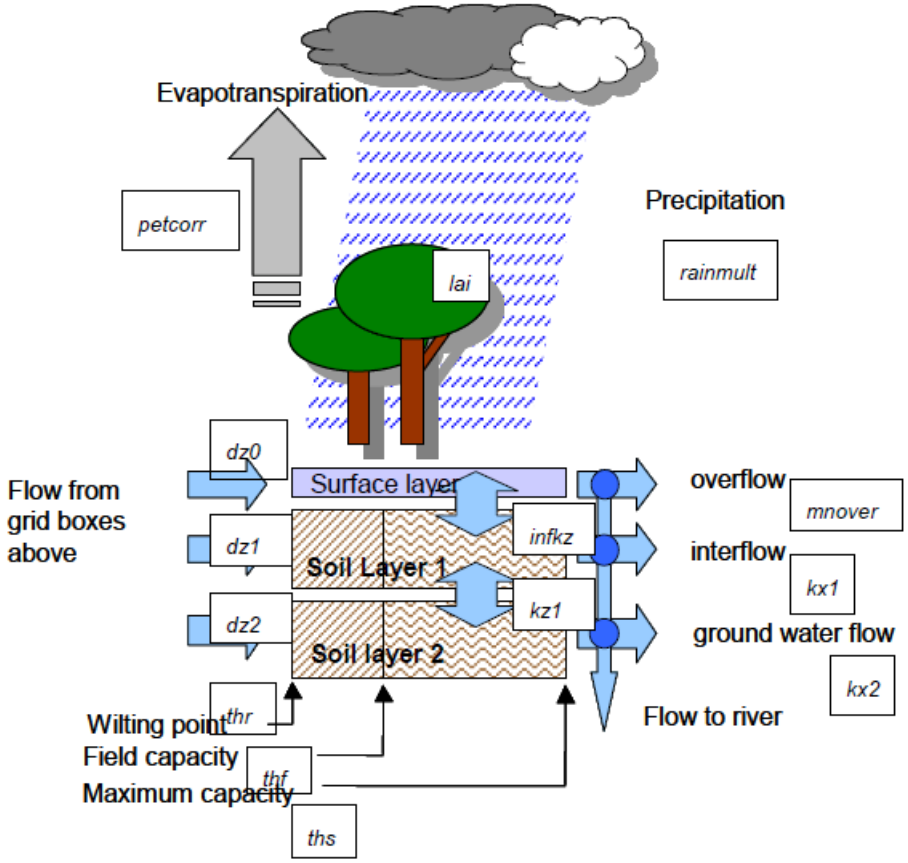

Figure 17. Main model processes and parameters.

Unfortunately the IWRM model is running inside an application that does not allow outside scripting to automatize the process, because of this the optimisation of the parameters was done manually. The year chosen for the optimisation was 2002. Each of the steps will be to run the model with one coefficient changed. This will allow finding the closest maximum values for both indices of efficiency.

The model creates an equation that generated the simulated time series, in this case the flow measured at a specific point called Nalaung. The closer the simulated points are to the measured points, the higher the R-squared. The Nat Sutcliffe model efficiency has been specifically made for hydrological models. It measures the predictive power of the model. In this case it is the most important indicator of the two indices of efficiency.

3.4 Water requirement

The Sun-WSI method proposed to reduce the country's statistics (GDP, GPP, population density) to an estimated grid cells values. However, as they mentioned it in their own research it was too difficult to find the information on other countries than China. The Sun-WSI remains an excellent method but required solid statistical information provided through the government. The study on the indicator WaSSI used an historical database; this method is based on a developed country administration with an advanced access to data from the government. This type of data cannot be used in the case of a country with fewer resources such as Lao PDR. For the purpose of testing the access to water the use of a widely accepted indicator will be enough to fulfill the objective. It would therefore be interesting to use the indicator developed by Falkenmark. For simplification purpose only the first threshold will be used; 1700 cubic meters per capita per year. When the first threshold is passed there is water stress. And to calculate the water stress indicator we will use the formula Use-to-Resource Ratio (formula 1).

3.4.1 Estimation of the initial water stress

The discharge layer use the unit cubic meter per second and the threshold from Falkenmark the unit cubic meters per capita per year. For the purpose of the test it would be interesting to take one day as the common unit. To do that the threshold of 1700 cubic meters was divided by 365 days and multiplied by each village's population. This gives us a table with the position of each village and the requirement for one day from the threshold. All the grid cells of the discharge layer are multiplied by (60 * 60 * 24) so we have cubic meter per day as the flow. The last step is to divide each village's requirements by the corresponding grid cell from the discharge layer. The IWRM model provides two output layers for the discharge, one for the dry season and one for the wet season. The results of the initial water stress were done first for the dry season.

3.5 Defining accessibility

In order to generate the cost surface layer and the accumulation cost layer it was decided to integrate a road network, the slope steepness and the land use layer (seen in Figure 12) as factors of accessibility. The source point will be each village.

3.5.1 Road network from OpenStreetMap

The information available on OpenStreetMap (OSM) is created by volunteers. It is therefore important to consider the limits of this type of collaboration. For example, a dataset generated by a GPS may include a number of inaccuracies due to the quality of the equipment and environmental conditions during the acquisition. Consideration should also be given to the possibility of errors, for instance when a volunteer adds information from an old satellite image. In other cases we can see duplicate and missing information. For instance a research was done within this thesis project to assess the completeness of the data by using different keywords into a search engine related to OSM. This leads to the conclusion that roads are the most commonly mapped items. Buildings and special infrastructure are not commonly mapped in the area of neither Lao PDR nor Cambodia. Although the reliability of the data is not assured the community has taken an important size and the data is reviewed and improved constantly (31). In Figure 18 you can see the power of OpenStreetMap.

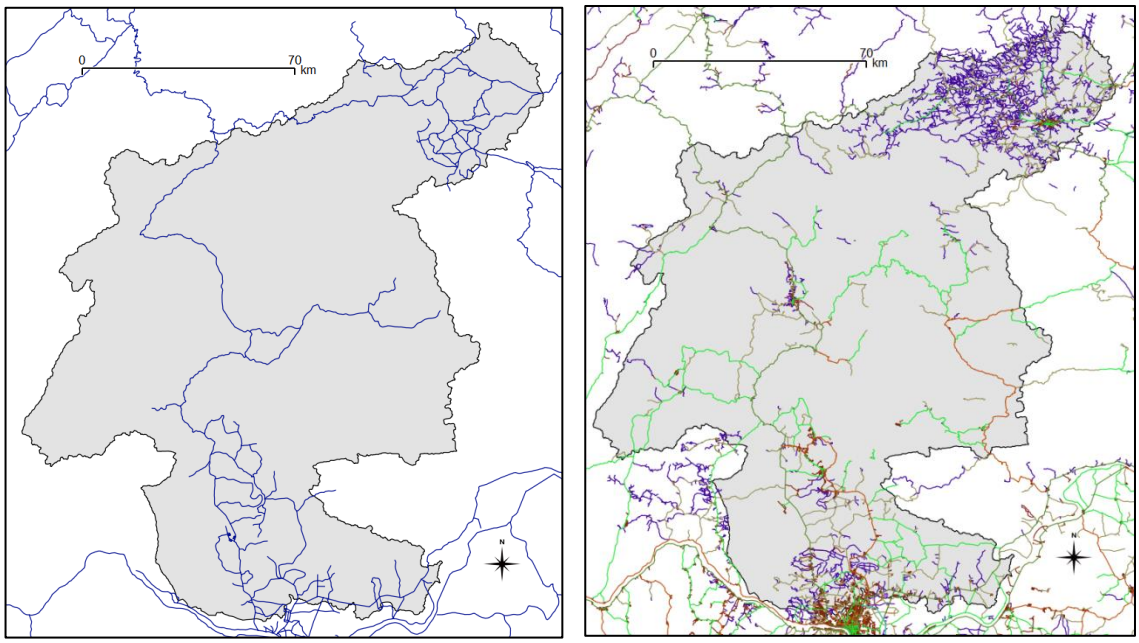

Figure 18. Visual comparison of the road network dataset from MRC and OSM

Different ways exist to extract the OSM information but here is the method used in this thesis. The information was imported from QGIS using Quick OSM plugin. Total of 12 categories were imported and each road type was imported separately. Each vector layer would contain only lines. Using the field calculator a new field is created with a numerical identification number related to its road type. Then each vector layer is transformed into a raster type layer with high resolution. A zoom on the result can be seen in the Figure 19.

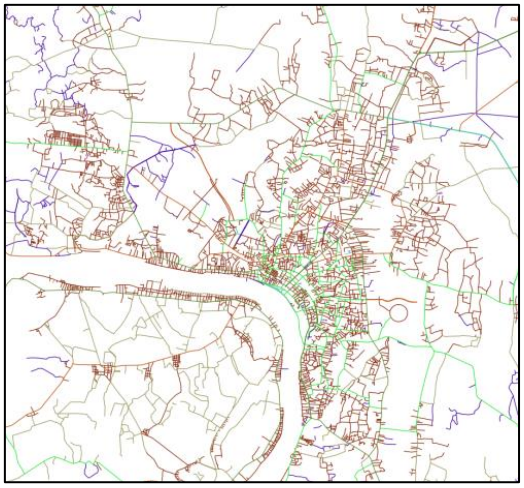

Figure 19. Lao's capital roads (OSM).

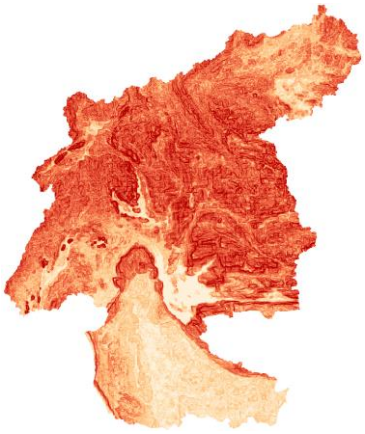

Figure 20. Slope steepness (USGS).

3.5.2 Slope steepness

The first database examined was the Harmonized World Soil Database (HWSD) for its DEM which has a resolution of 30 arc-second. The set in question contains the all world elevation and to clipper the area of Nam Ngum the task would take hours. After more research a set from USGS Earth Explorer was used (GMTED2010N10 E090). The resolution of 7.5 arc-second was much better and the data accessed was already clipper to a very small area. From QGIS the slope steepness layer can be generated using the geoalgorithm from the terrain analysis menu, see results in Figure 21.

3.5.3 Location of the villages

The initial attempt was to create this layer via the combination of two different layers. The first is the global set from NASA called "Gridded Population of the World" and the administrative set of each village (polygons) from the website Lao Decide Info. The polygon set was transformed into points using the centroid geoalgorithm and the number of inhabitants was downscaled to the area of each polygon. The application of the centroid geoalgorithm gives no guarantee as to the real position of the village and this brings an additional possibility for errors. Therefore, Marko Kallio has generously offered access to one of his final layers from his Master's thesis. The layer in question contains not only the precise position of the village but also its population (see Figure 23) and an indicator of poverty. This provides better precision for several calculations such as the distance from the stream and whether there is enough water around the village.

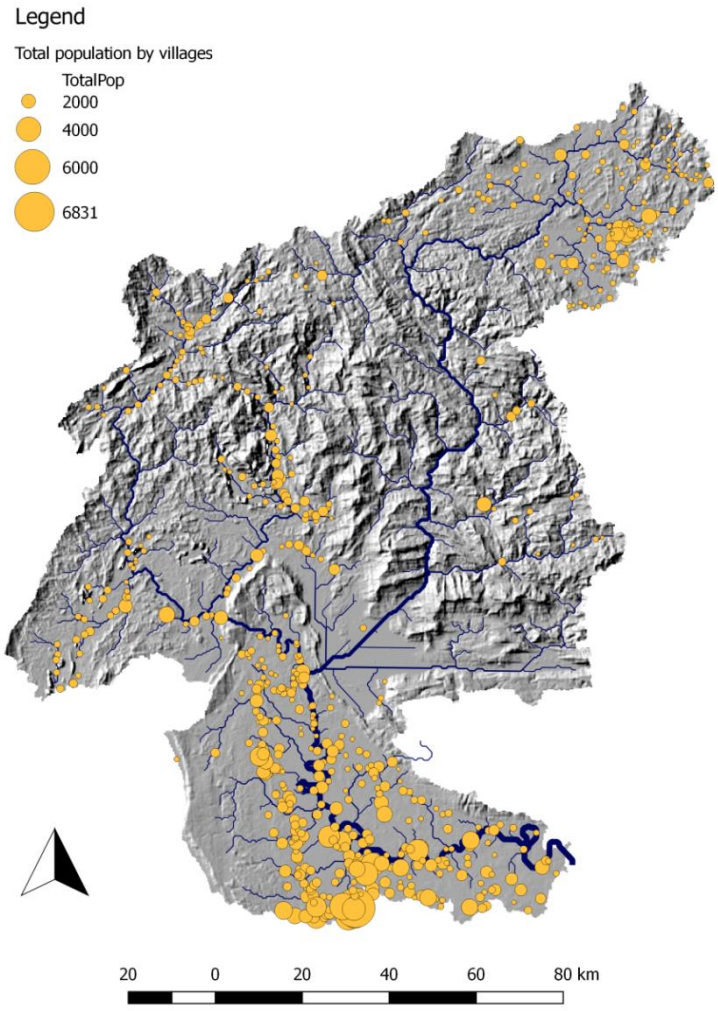

Figure 21. Map of the population in the area of Nam Ngum catchment.

3.5.4 Cost surface layer

There is no defined approach to establish a common unit between parameters that do not share the same unit. In order to define a common unit the problem was approached from a realistic perspective. The assumption taken into consideration was that the person would walk directly to the river and carry some water back home. From this perspective it is possible to use a calorie burning calculator (32). Here is the formula:

$$C = \frac{(W \times MET \times 3.5)}{200} \times t \qquad (6)$$

Where C is the quantity of calories required, W is the weight of the person, MET is the Metabolic Equivalent of Task and t is duration of the activity in minutes. For the purpose of the calculation an estimation of 2.5 km/h was used. We can estimate the time to cross one grid cell as the grid cells are 250 meters by 250 meters. The initial value 5 for the MET can be found in Table 3. This altogether provides the quantity of calories used to carry water on a flat surface across one grid cell; which is 31.5 calories; this calculation provides the normal level of difficulty (level 1 in Table 4).

Table 3. MET values for walking (33).

Activity	MET	SPECIFIC MOTION
walking	5	carrying 7 kilogram load (e.g. suitcase), level ground or downstairs
walking	8	carrying 11 to 22 kilogram load, upstairs

To define a cost surface it is necessary to define a common unit representing the difficulty level to cross each type of layer and to estimate the next level of difficulty it was decided to use the ratio between the MET values in Table 3. The second level of difficulty is ranked as 1.5 times harder than the first level. The same proportion is used to qualify the next level for more difficult slope. The last level (value 4) is made to create the limit in which slope steepness become unpractical, see table 4.

Table 4. Classification of the cost surface parameters.

Slope steepness (degrees)	Land use	Road type	Difficulty level
	Water		0
0 – 1.1	Urban	Basic roads	1
1.1 – 4.6	Agriculture	Small path	1.5
4.6 – 16.7	Forest	None	2
> 16.7			4

The categories of roads were estimated and from the 12 categories only two groups were kept; the basic roads and the small path. The same step was done for the eight categories of land use which were transformed into four groups; water, urban, agriculture and forest. See Figure 22 for the steps to build a cost surface layer.

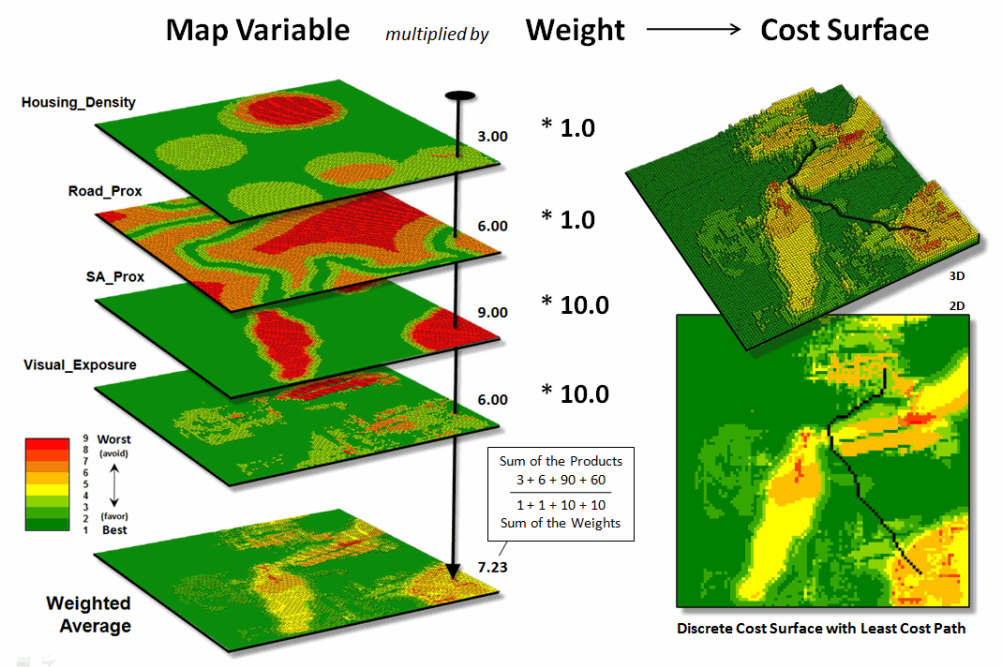

Figure 22. Steps to build of cost surface layer.

In practice generating the cost surface layer should be done via the raster calculator of QGIS however a technical issue was encountered. QGIS interface of the raster calculator is not user friendly additionally the help online is not well developed. Prior calculation each of the layers was scaled down or up to the resolution of 250 meters by 250 meters. During the rescaling operation it was done using a reference raster to align the results. However after using the raster calculator the results would not be aligned, see Figure 23 for a preview. Therefore these steps were undertaken via Matlab.

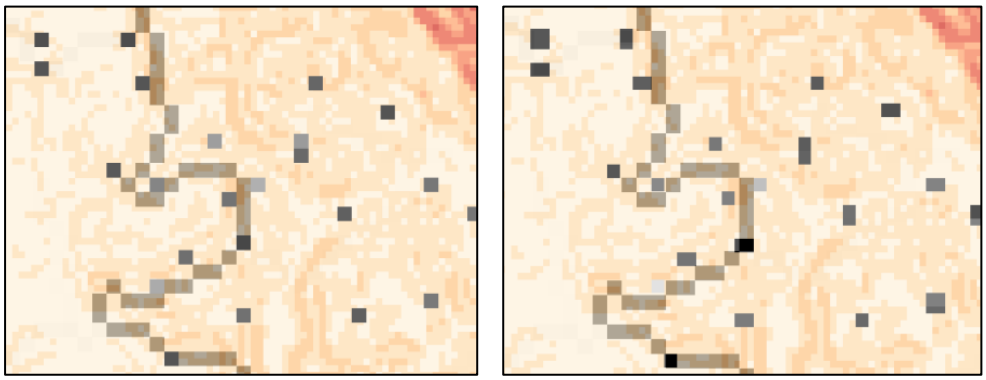

Figure 23. Displacement of the results via the raster calculator of QGIS.

3.5.5 Accumulation cost layer

The accumulation cost is a method in GIS that use the cost surface layer and one or more source points to generate an accumulation cost layer. In order to avoid any errors in applying this geoalgorithm it was necessary to run it through QGIS. To be able to extract each of the village's area it was necessary to run the geoalgorithm with one source point at a time. All villages are initially contained into one layer and to go around this issue it was necessary to first identify each village and create a separate layer with each village. See Figure 24 for a visual representation of an accumulation cost layer.

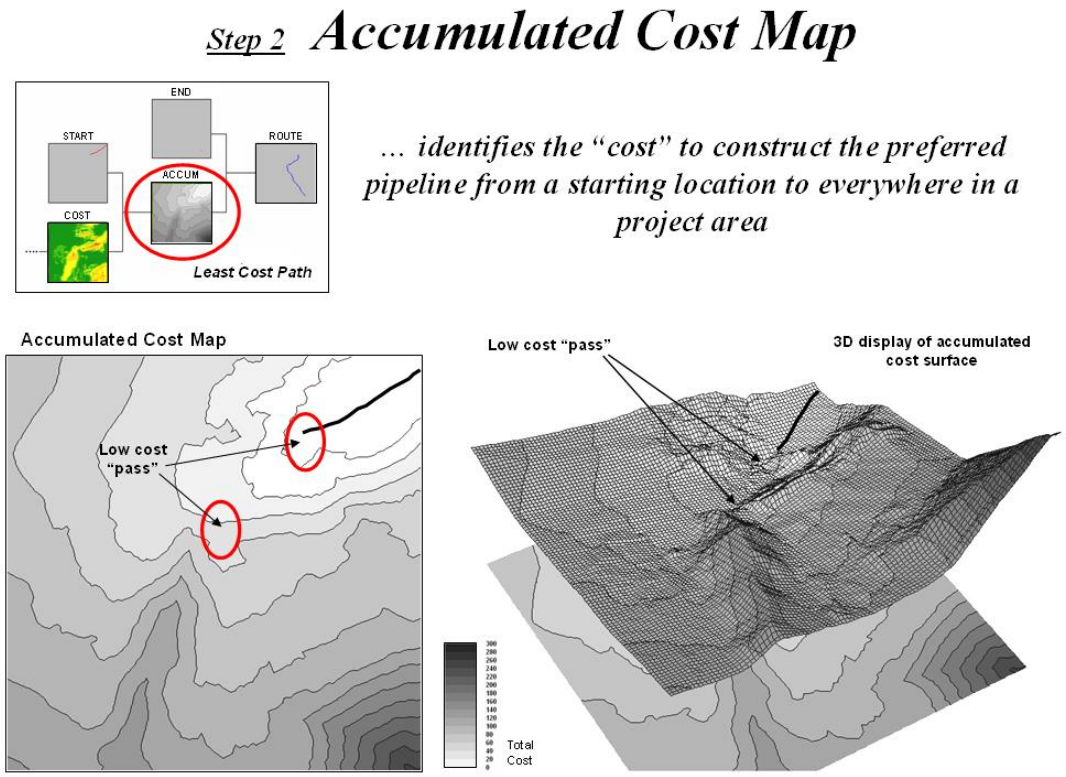

Figure 24. Accumulated cost map.

During the identification process the same order as in Sun-WSI study was used and villages were sorted by elevation. Three criteria were chosen to identify each of the villages separately; first the indicator of poverty, then the distance to the river and finally the population. The proper format was copied from a blank raster created for this purpose and one village was inserted inside each blank copy. This process was done in Matlab but the data from the sorting was found from different raster created through QGIS.

There are different ways to generate the accumulation cost layer but the one chosen for this thesis project was the geoalgorithm SAGA Accumulation Cost (isotropic). In order to apply this geoalgorithm from QGIS the source point, the village in this case, needs to be on a separate layer. All the layers are loaded into QGIS and a Python script is run to apply the method to each of the village's layer. The threshold is not integrated at this point since the aim was to reproduce the methods used in Sun-WSI. In that study the radius of each circle was calculated using the formulas referenced previously as Equation 5.1 and 5.2 in the introduction of the chapter Methodology and Data.

To replace the GDP down scaled method used in the Sun-WSI it was decided to use the indicator of poverty from the village's dataset coming from Marko Kallio Master Thesis. The elevation came from the USGS earth explorer dataset and slope steepness was extracted from the same layer. Finally the shortest distance to the river was calculated using the "v.distance" geoalgorithm from GRASS. To calculate the shortest distance both layers, the village point and the river line, need to be vectors. The river network from the IWRM is a raster and the transformation back into vectors was not successful. Consequently a dataset from HydroSHEDS was used for the river network and its resolution is 30 arc-seconds.

$$P_i = ((2250 \times T_i)/2)/48.46 \quad (7)$$

The average calories intake of Lao PDR (approximatively 2250) (35) multiplied T_i (from Equation 3). To complete the threshold (P_i) calculation the calories available are then divided in two to account for the one way trip. Finally this quantity of calories available is divided by 48.46, cost for the normal level of difficulty for one grid cell (cost surface minimum equal to 0.65), and provides the threshold for the accumulation cost of each of the villages.

3.5.6 Updating dynamically the discharge layer

This step should be processed inside the modelling software. The Sun-WSI integrates the input and output of the model into the same software to retrieve the results. Unfortunately in the IWRM model the input and output parameters need to be recorded manually. Introducing one point where there is a consumption of water will require time. In the area of the Nam Ngum catchment there are approximatively 500 villages that from the initial water stress calculation require water from outside their grid cell. This makes manually adding each point a very long process on itself and does not cover the concept of the accumulation shape generated.

It was thought that this could be processed via a Matlab algorithm. The area of concern contains three basins and two inter basins which make it more difficult. For instance if a village is inside the Nam Ngum basin it will affect also the inter basin of Nam Ngum but will not affect the Nam Lik or Nam Song basin. The only layer of information that was lower resolution than 250 meters was the discharge layer. Consequently the discharge layer was down scaled using SAGA geoalgorithm. The fact that it has been down scaled increased the difficulty for the algorithm. While each grid cell in the original discharge layer had a different value, the downscaled layer contained group of four cells that have the same values. Additionally around the lake areas some irregularities are present.

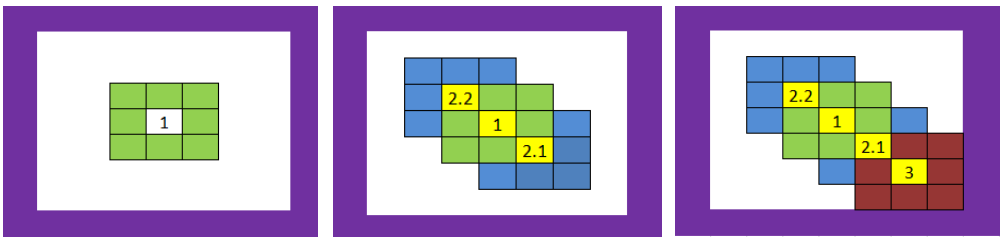

Figure 25. Neighbors indexing in Matlab

The algorithm was built to dynamically find the way the river is flowing by using linear, logical and neighbours indexing practice in Matlab. The yellow cells in Figure 25 show the active cell from which the algorithm makes the neighbourhood calculations. The first step is to find the eight neighbours of the cell 1 (green cells in the Figure 25), then find which of the cells have the maximum values. Very often more than one of the cells has the same values due to the downscaled layer. The next step of the algorithm is to find the most promising next cell. To do that each of the cells containing the maximum values is taken as the next possible cell (2.1, 2.2 in the Figure 25).

Each of these cells is then using the same function to find the eight neighbour's maximum values (blue cells). Then they are compared to each other and the path to the maximum value is found. Sometimes the next possible cell does not contain higher values due to the fact that the layer was downscaled. In that case the process is repeated one more time to find the further best possible next cell (red cells in Figure 25). When there is a lake in the layer the normal algorithm does not compute and a special case needs to be made. In this special case the algorithm was program with manually added exceptions. In total eight exceptions were made for the dry season discharge layer and one exception was added for the wet season discharge layer. For an overview of the path completed by the algorithm it is possible to find a detailed map in the Appendix 5.

At each village the algorithm is first searching for the river. It is searching for a place that does contain enough water. When it has found it will follow the course of the river all the way to the end. The algorithm allows updating the river with what was theoretically needed by each village. The results of this updated layer allow proceeding to the next step, which is the calculation of the final water stress indicator.

3.5.7 Final water stress calculation

To follow the method provided in the Sun-WSI a new water stress indicator was calculated using the resources available in the surroundings of the active cells. The shape used was the Village's Accumulation Cost Area (VACA). The calculation was done in Matlab and was integrated into the algorithm explained in the previous section 3.5.6. The villages were taken in the order established before.

4 Results

4.1 Results for the optimization of the IWRM model

As it was stated before in methodology the optimisation of the model was only possible through the modification of the coefficients inside the IWRM modelling software. Each of the modification for the coefficients was done separately. At each step small value was added or/and subtracted to the initial values. The aim was to see both R-squared and Nash Sutcliffe model efficiency values optimized in the same time. The relationship between each coefficients modification and the change in both R-squared and Nash Sutcliffe model efficiency values (indices of efficiency) is quadratic. This means that it is possible to have more than one maximum for each modification however in this case only one maximum was searched for; the maximum result for both indices of efficiency from the initial values of each of the coefficients. In total 133 trials were done to find the best coefficients (see Figure 26 for the progression).

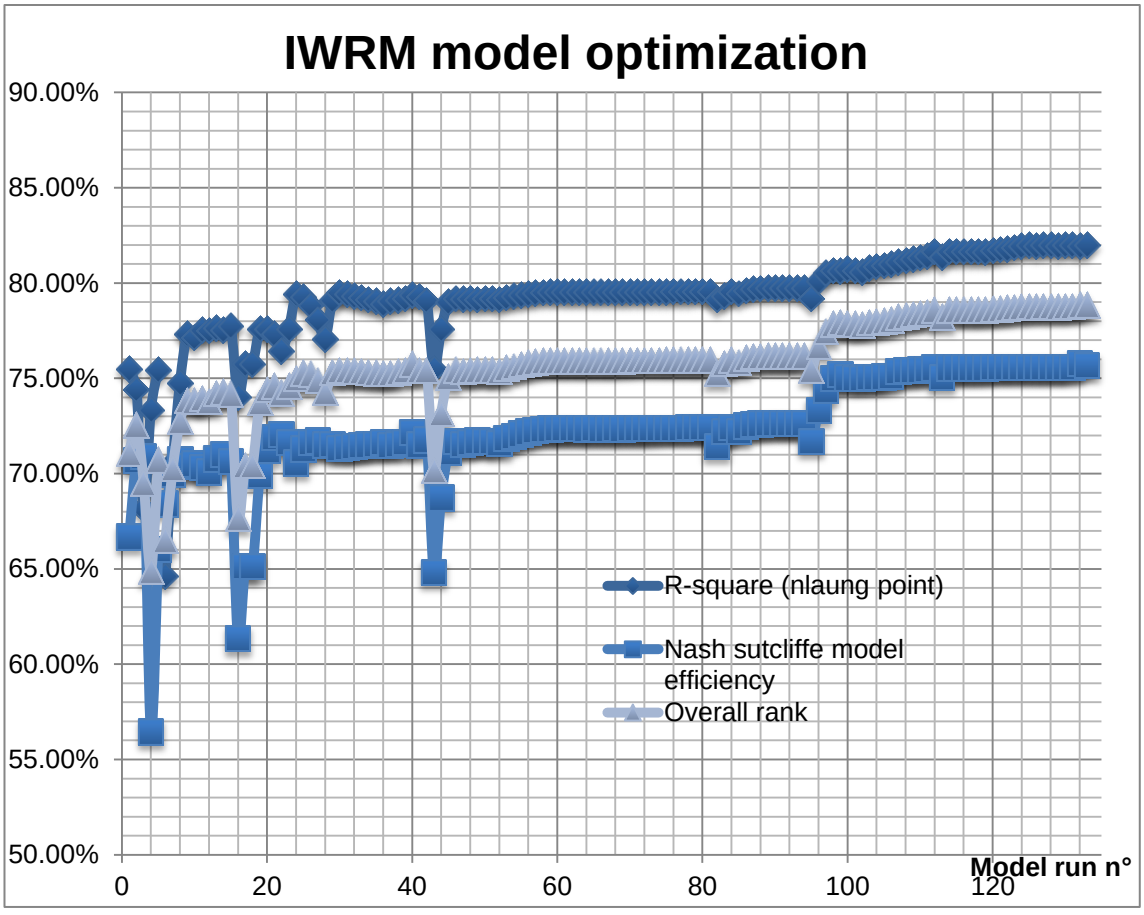

Figure 26. Results for the optimization of the coefficients for the IWRM model.

While the manual optimisation has a good effect on the one year run the effect is not so strong on the overall 13 years of time series, see Table 5. Running one time the model with 13 years of time series requires approximatively 27 minutes which is too long to run an optimization on every parameter selected.

Table 5. IWRM model optimization.

Number of year run	Initial parameters			Final parameters		
	R-squared	Nash & Sutcliffe	Correlation	R-squared	Nash & Sutcliffe	Correlation
1	75%	67%	87%	**82%**	76%	90%
2	72%	56%	85%	78%	65%	88%
3	75%	57%	87%	**80%**	65%	90%
10	62%	7%	84%	66%	20%	86%
13	64%	36%	83%	68%	46%	84%

The optimization of the model was not very successful. The results were not extremely different, the average flow went from 195 to 188 cubic meters per second and the evaporation was a little bit higher. Almost no difference was seen in the total precipitation. It is clear that to really improve the model's results the layers (land use and soil type) should be improved and all the coefficients should be changed to find each of the maximum value. Additional time series for surface air temperature would be a welcome addition since they are part of the calculation of the potential evapotranspiration. The preview is available in the table of Appendix 2. See Figure 27 for a comparison of the results at one specific point (Nalaung) in the river.

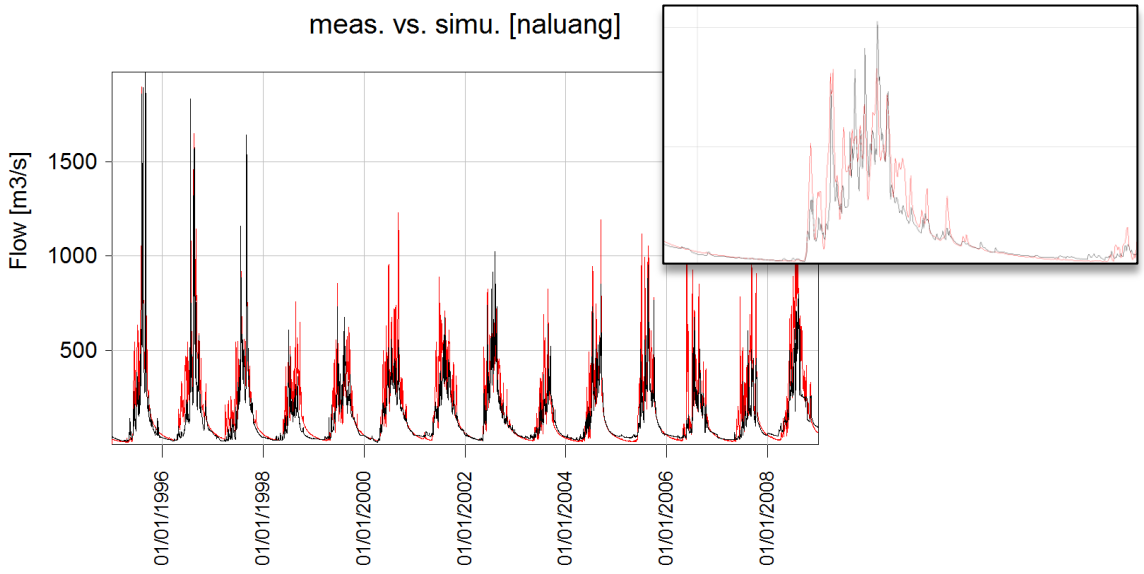

Figure 27. Simulated versus measured flow, zoom on 2002.

4.2 Results for the cost surface and accumulation cost layer

The cost surface clearly shows the road network and the flat terrain as you can see in the Figure 28. The full scale cost surface layer can be found in the Appendix 3. The accumulation cost layer was generated properly and the content was used for each of the villages, accompanied with dynamically calculated threshold. The threshold varies depending on altitude, slope steepness, indicator of poverty and distance to the river (see Equation 7).

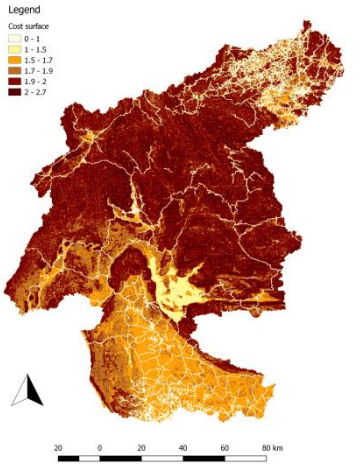

Figure 29. Cost surface layer.

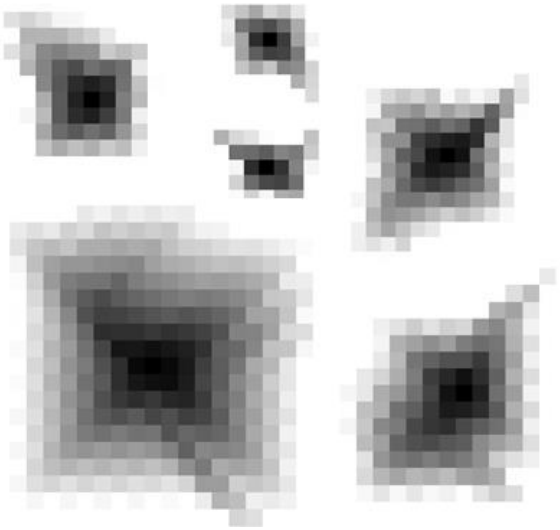

Figure 28. VACA of different size for different village.

See Figure 29 for the VACA from the smallest to the largest; the smallest is found in the mountain area where there is a higher slope on all sides. The middle sized ones are from the north where there are high density roads and flat terrain and the largest ones are found in the south where there is low altitude and flat terrain. They are displayed in Figure 29 proportionally to their actual size so they can be compared (larger size in Appendix 4).

For each season a separate calculation was made. The saved path from the algorithm was also stored separately for both seasons. This, in turn, required a lot of space on the hard-drive. However it was a necessary step because the villages found at high altitude followed a different path depending on the season.

4.3 Results for the algorithm updating the discharge layer

For the dry season, in 81 percent of the cases the script was able to find the point where there is enough water for the village and in 75 percent of the cases the grid cell was within the Village's Accumulation Cost Area. The average cost was 29% of the threshold of each VACA. The average distance to the river calculated via vector layers in QGIS was 1'164 meters. The average of total calories available was 1'020; see the Table 6 for more information.

Table 6. Extraction of some of the values used to verified the algorithm

Average or percent of success	WSI dry seas.	WSI wet seas.
Shortest distance to the river (QGIS, meter)	1'164	1'164
Total calories intake per person	1'020	1'020
Number of cells contained in the VACA	135	135
Number of cells updated (stream affected)	1'448	1'314
When the algorithm found enough water	81%	90%
When the algorithm found water inside VACA	83%	83%
How far water was found from the threshold	29%	21%
Estimated cost of carrying water from the river (calories)	163	107
Difference between available and used calories (calories)	232	316
Number of cells to find the water	5	3
Distance to the cell that contains enough water (meter)	750	450

Figure 30 shows, for a selected number of villages, the difference of water available within the VACA after the algorithm has updated the area. In some cases the villages with a similar elevation are close to each other which affect the water available. In other cases the villages are far from each other and the water available remains the same.

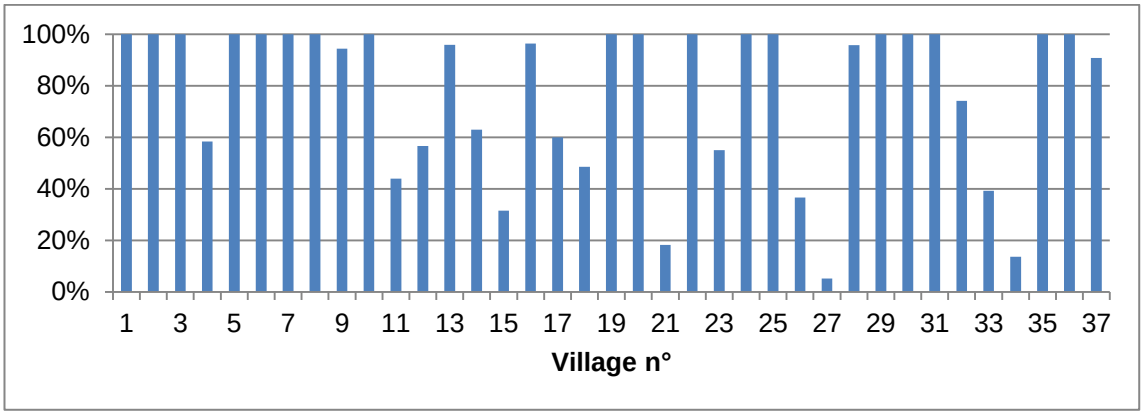

Figure 30. Water available after the algorithm updated the area

4.4 Results for the Water Stress Indicator

The initial WSI calculation is directly proportional to the population as there is almost no water available inside the active grid cell of the discharge layer. You can see a full scale map with both seasons compared in the Appendix 6 and a preview in Figure 31. The initial WSI and the *water stress* are calculated using the same Falkenmark threshold (1700 m^3/cap/yr). Additionally, you can see the place where the water is provided via pipeline.

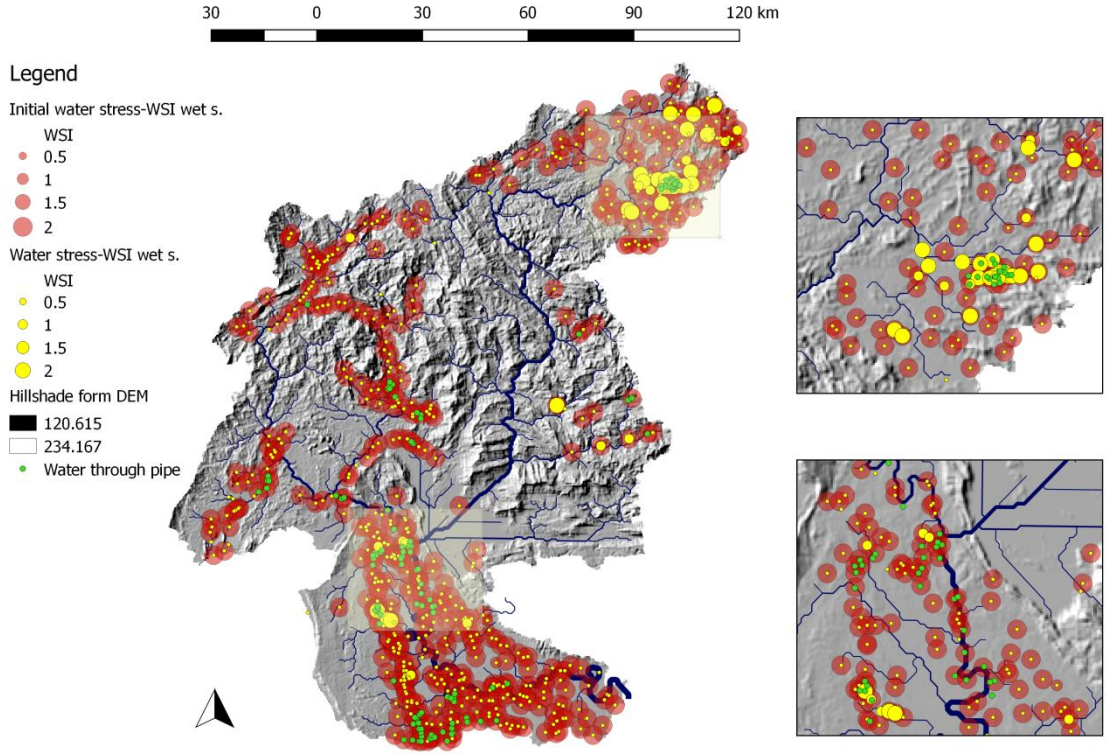

Figure 31. Initial and final WSI for the dry season.

To compare the results from the Final WSI calculation it was decided to use information available from the website Lao Decide Info. Two datasets found on this website provide information on whether people use tab water and whether or not there was a drought in that specific location. As seen in Appendix 6 there have been drought recorded almost everywhere in the Nam Ngum catchment. There is no information concerning the severity or the frequency of each drought recorded. One of the zoom on Figure 31 clearly shows piped water being installed where there is a water stress while the other zoom shows piped water installed but almost no *water stress*. The Sun-WSI compared the drought index results to their Sun-WSI calculation. Generating an indicator of drought

require gridded precipitation records rather than weather station points. This means that it would be necessary to use both in the IWRM model and in the drought index calculation in the same time series.

The final WSI calculation showed only small pockets of *water stress*. They are mostly located in high density habitation and at the starting point of a small river. When looking closely at the data the range of water available within the vicinity of the villages (VACA) is relatively big compared to the theoretical water requirements. The full scale map of the final WSI is in Figure 32. The final WSI for the wet season is available in Appendix 7.

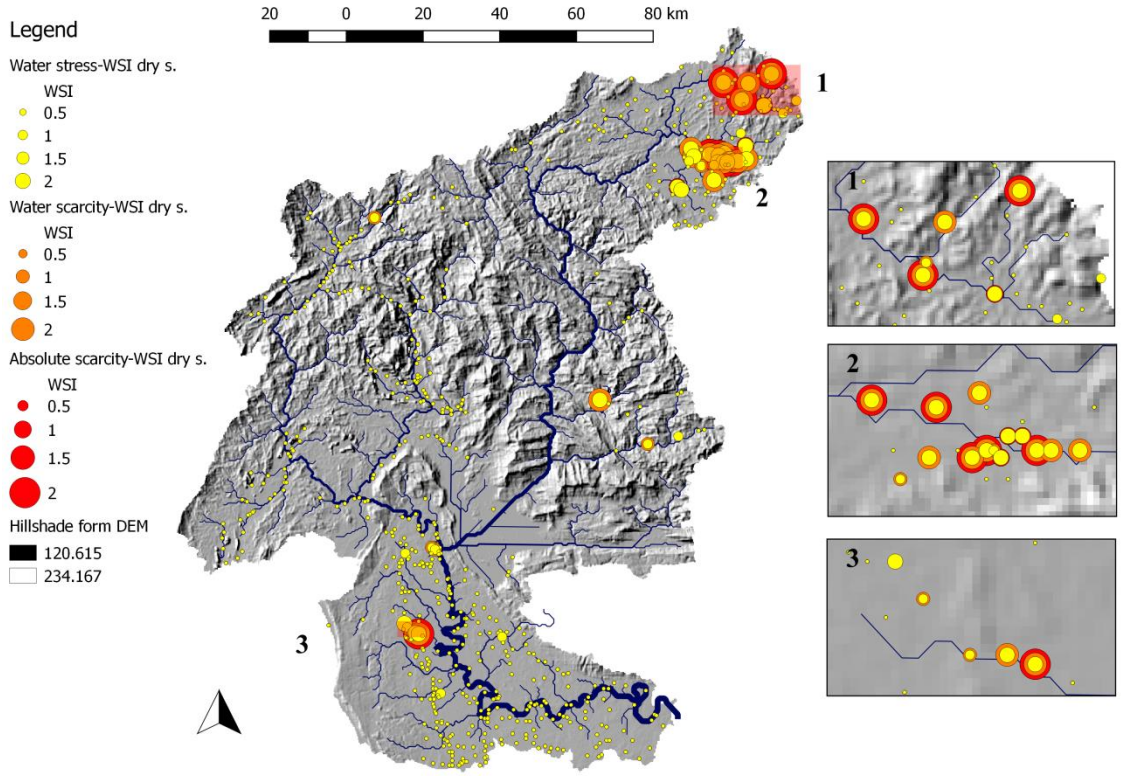

Figure 32. Final WSI for the dry season with the three Falkenmark thresholds.

A map was created to emphasize the quantity of water available within the VACA defined by the dynamic threshold. To be able to compare the water available during the wet and dry season the quantities have been normalized with the maximum value of the wet season and the minimum value of the dry season. The result of this map can be found in the Appendix 8. The dam clearly changes the scale of the flow within the catchment, so much so, that it is not possible to see smaller change if the scale shows all values.

4.5 Discussion & suggestions

The resolution of each dataset used in this type of project influences the overall precision with which a conclusion can be made. In general the actual state of the data freely available has a very coarse resolution. It is difficult to draw a precise conclusion when a dataset comes with a resolution of 30 kilometres by 30 kilometres. As a fact there are variations within a surface of 90 square kilometres and it is difficult to decide what is crucial and which variations may be insignificant. In the case of this thesis project the coarser dataset used was 500 meters by 500 meters and was down scaled using a GIS geoalgorithm therefore the conclusion cannot be more precise than the coarser dataset used. While precision is critical in most cases it is important to limit improvement of the quality and precision by taking into account the time for running the analysis and the space required to store the information. In this project the folder containing all the layers and scripts has reached more than 60 Gigabytes and the complete run of the analysis for accumulation layers and water stress calculations requires more than ten hours to complete. One of the possible optimization of the calculation could be the addition of more factors defining the cost surface layer so it does emphasize more the actual accessibility of a person. This would not necessarily require a better resolution but would improve the shape of the VACA. The distance people seem to settle from the river is much smaller than predicted from different articles read online. This means that people tend to settle as close as possible to a source of water. This also means that to emphasize access it might be better to have a much more detailed dataset, higher resolution than 250 meters, to account for the variations in the surroundings.

The trend is definitively pushing in the direction of Remote Sensing information. Large quantities of world datasets are already available freely online but require complicated and sophisticated techniques. Therefore unless this is changed the research surrounding the use of such powerful tools will remain accessible to only a few. The base of most of the studies and research in the area of water and in direct relation with establishing the state of an area through the use of indicator and indexes are relying heavily on hydrological modelling. During this project the use of the IWRM model through the software developed by EIA was met with some challenging points. The most difficult point was improving the results obtained for the discharge layer. This part of the software could benefit with an automatic overall optimization of the coefficients. After a conversation with the one of the owners of the software it was concluded that the usual technique required in finding optimum values in mathematics (search of an edge on a

surface). It is said to be too time consuming to be applied on so many coefficients at the same time however in the future computer's power will most certainly evolve and it could be a possible addition to the software.

During this project the Falkenmark thresholds of 1700, 1000 and 500 cubic meters per capita per year was used as the theoretical water requirement. This, of course, is a very coarse approximation and does not take into account realistic and location specific water demand for the industry, agriculture and domestic use. It is also not taking into account any return flow from the water withdrawn. There are different ways to make approximations for a specific area but it does require access to sophisticated techniques, which may not be available freely, or rely on the quality and availability of a country's statistical information. My suggestions for further research would be to read the two different gridded downscaled methods (estimating domestic and industrial demand) mentioned in the Sun-WSI study and developed earlier by the same group of people in two separate studies.

The software containing IWRM model was not made to run water stress indicator analysis. It was not designed in that sense and so applying the Sun-WSI methods would not have been possible through that software in any case. However points where the water could be withdrawn with a return flow would have been possible on the form of manually added time series for each location. This is a relevant point also in the optimization of the discharge layer from the model. It is possible that some of the flow fluctuation are due to human's direct influence for instance through a mine wastewater system. Therefore the optimization of the discharge layer is also limited from this side. If one would like to reproduce Sun-WSI study it would be necessary to have the calculation of the discharge layer integrated with the water withdrawn and the return flow. If not, then the discharge information will remain a coarse approximation which in the case of a water stress indicator analysis is not a strong asset. In any case the software running IWRM model was not designed in that purpose so it is normal that it does not provide the integration researched in this thesis project. For instance it would not have been possible to integrate each of the accumulation cost layers within the IWRM model to run the analysis. I suggest finding a good manual guide to the configuration of a model such as WaterGap2 to work on such analysis. From my understanding formed by my research I believe that this type of model would integrate in an environment that allows such analysis.

5 Conclusion

Due to the limitations of the software running the IWRM model the integration of a realistic water demand was not possible as explained in the discussion chapter. The result can be categorized as a coarse approximation of the water stress in the Nam Ngum catchment area. The results show a clear temporal difference between the wet and the dry season. Generating a monthly discharge average would provide a better visualisation of the water stress. The results from the IMWI model, mentioned in chapter 2.3.5, classify the northern part of Lao PDR as economically water scarce. From the results found during this thesis project it is clear that even during the dry the season the catchment has enough water, in average, to provide for each village. If the approximation is correct, there are small pockets of water stress located at the starting point of small rivers; this corroborates the classification as economically water scarce. There is a need for water to be provided via pipeline to compensate for high density settlement.

If I was able to continue the implementation of this project I would try to integrate the water demand inside the modelling software. At this point, however I would have to consider if this would be more productive inside the IWRM modelling software or in a different product. It seems crucial to integrate as many variables as possible inside the same model to be exact inside the same simulation. I do not currently have access to enough knowledge to take a subjective position. I can say, however, that the trend seems to be more in the direction of the use of Remote sensing information within GIS software such as QGIS. What has really amazed me from QGIS is its power to integrate outside software/dataset within its interface which makes it very powerful. I believe that future models should be developed with a good integration to such tools and should enable the use of world dataset such as TRMM rainfall product and other Remote Sensing information freely available to the public.

As expected the use of the surrounding area's resource helped to reduce the initial estimated water stress. Defining the access to water through an irregular shape is an excellent idea by Marko Kallio. Where people do not have access to tab water they will travel to find the resource. I believe that simulating the reach of a person through the cost surface/accumulation cost method is working well. The threshold of the cost is what defines the area which is based on the assumption of the cost to move from one grid cell to another. This approximation is therefore the key component but in the same time very difficult to evaluate. It is based on a set of assumptions that can be incorrect

and as the extra WSI calculation show in figure 31 it does influence the results. This method is really good but it does take a lot of time to generate all the layers of information separately. In case of a bigger area it might become impossible to use it the same way as during this thesis; however to every problem many solutions exist.

To conclude I think there is a promising future in the field of water indicators and as the technology is evolving there will be more possibilities for refined analysis. As the crisis on water is predicted to become increasingly severe I hope that information will become easier to access. Movements such as the OpenStreetMap community are a good example of what we need for the future. Tools such as water stress indicators are key elements for the governments around the world and also indirectly benefit the populations that are affected by the issues.

Figure 33. Water Planet; MODIS on NASA's Terra satellite (Wikimedia).

Bibliography

1. UN. Historic New Sustainable Development Agenda Unanimously Adopted by 193 UN Members. [Online]. New York: DPI; 2015 Sept 25 [cited 2017 February 4]. Available from: http://www.un.org/sustainabledevelopment/blog/2015/09/historic-new-sustainable-development-agenda-unanimously-adopted-by-193-un-members/.
2. UN. Goal 6 Ensure access to water and sanitation for all. [Online]. New York: DPI; 2017 [cited 2017 February 4]. Available from: http://www.un.org/sustainabledevelopment/water-and-sanitation/.
3. UN. Why it matters. [Online]. New York: DPI; 2017 [cited 2017 February 4]. Available from: http://www.un.org/sustainabledevelopment/wp-content/uploads/2016/08/6_Why-it-Matters_Sanitation_2p.pdf.
4. Palmer WC. Meteorological Drought. Research paper NO. 45. Washington: U.S. Weather Bureau, Office of Climatology; 1965.
5. Ashok K. Mishra VPS. A review of drought concepts. Journal of Hydrology. 2010 July;(391): p. 202-216.
6. World Meteorological Organization (WMO) and Global Water Partnership (GWP). Handbook of Drought Indicators and Indices M. Svoboda BAF, editor. Geneva: Integrated Drought Management Programme (IDMP); 2016.
7. Rijsberman FR. Water scarcity: Fact or fiction? Agricultural Water Management. 2006 August;(80).
8. Amber Brown MDM. A Review of Water Scarcity Indices and Methodologies. [Online]. The Sustainability Consortium; 2011 [cited 2017 February 07]. Available from: https://www.sustainabilityconsortium.org/wp-content/themes/sustainability/assets/pdf/whitepapers/2011_Brown_Matlock_Water-Availability-Assessment-Indices-and-Methodologies-Lit-Review.pdf.
9. Gleick P. Basic Water Requirements for Human Activities: Meeting the Basic Needs. [Online]. Oakland: PISDES; 1996 [cited 2017 February 15]. Available from: http://pacinst.org/app/uploads/2012/10/basic_water_requirements-1996.pdf.
10. Paul Raskin PGPKGPKS. Water Futures: Assessment of Long-range Patterns and Prospects. 1st ed. Hultcrantz K, editor. Stockholm: Stockholm Environment Institute; 1997.
11. UNDP. Human Development Index (HDI). [Online]. New York: UNDP; 2016 [cited 2017 February 18]. Available from: http://hdr.undp.org/en/content/human-development-index-hdi.
12. Molden D. Water for Food, Water for Life: A Comprehensive Assessment of Water Management in Agriculture. [Online]. London/Colombo: Earthscan/IWMI; 2007 [cited 2017 February 15]. Available from: http://www.fao.org/nr/water/docs/summary_synthesisbook.pdf.
13. Charles J. Vörösmarty EMDPAGaCR. Geospatial Indicators of Emerging Water Stress: An Application to Africa. Royal Sweedish Academy of Sciences. 2005 May 03; 34(3): p. 230-2306.
14. Charles J. Vörösmarty PBM. Global threats to human water security and river. Nature. 2010 September;(467): p. 555–561.
15. Vladimir Smakhtin CRPD. Taking into Account Environmental Water Requirements in Global-scale Water Resources Assessments. [Online]. Colombo: IWMI; 2004 [cited 2017 February 22] [Comprehensive Assessment of Water Management in Agriculture Research Report 002]. Available from: http://www.iwmi.cgiar.org/assessment/files/pdf/publications/ResearchReports/CARR2.pdf?galog=no&redir=yes.
16. Henrique M.L. Chaves SA. An Integrated Indicator Based on Basin Hydrology. Water Resource Management. 2007;(21): p. 883-895.
17. Sun Ge SGMJAMMECC. Impacts of Climate Change, Population Growth, Land Use Change, and Groundwater Availability on Water Supply and Demand Across the Conterminous U.S. Watershed Update. 2008 August; 6(2).
18. Steven McNulty GSJMMECaPC. Robbing Peter to Pay Paul: Tradeoffs Between Ecosystem Carbon Sequestration and Water Yield. Proceeding of the Environmental Water Resources Institute Meeting. 2010 August;(12).
19. Pfister Stephan AKSH. Assessing the Environmental Impacts of Freshwater Consumption in LCA. Environmental Science & Technology. 2009 April 23; 43(11): p. 4098–4104.

20. Arjen Y. Hoekstra AKCMMAMMM. The Water Footprint Assessment Manual : Setting the Global Standard London: Earthscan; 2011.
21. Qiuwen Zhou SYCZMCHLLLH. Development and implementation of a spatial unit non-overlappingwater stress index for water scarcity evaluation with a moderatespatial resolution. Ecological Indicators. 2016 October;(69): p. 422-433.
22. Yaohuan Huang DJDZYZJF. An improved approach for modeling spatial distribution of water useprofit—A case study in Tuhai Majia Basin, China. Ecological Indicators. 2014 July;(36): p. 94-99.
23. Yang Xiaohuan YHPDHL. An updating system for thegridded population database of China based on remote sensing, GIS and spatialdatabase technologies. Sensors. 2009 February 20;(9): p. 1128-1140.
24. Gayathri K Devia GBPDGS. A Review on Hydrological Models. Aquatic Procedia. 2015 March 17; 4: p. 1001-1007.
25. Wikipedia. Water cycle. [Online]. Wikipedia; 2017 [cited 2017 March 7]. Available from: https://en.wikipedia.org/wiki/Water_cycle.
26. United Nation. Integrated Water Resources Management (IWRM). [Online]. United Nation; 2014 [cited 2017 March 27]. Available from: http://www.un.org/waterforlifedecade/iwrm.shtml.
27. Wiki.GIS.com. Cost surface. [Online]. Wiki.GIS.com; 2016 December 29 [cited 2017 March 22]. Available from: http://wiki.gis.com/wiki/index.php/Cost_surface.
28. Petra Döll JATHFKBLRSS. THE GLOBAL INTEGRATED WATER MODEL WATERGAP 2.1. Technical report. Germany: University of Kassel, Center for Environmental Systems Research; 2001.
29. Kallio M. Effects of Mining and Hydropower on Metals in Surface Waters. Bachelor Thesis. Helsinki: Helsinki Metropolia University of Applied Sciences, Environmental Engineering; 2014.
30. NASA. Algorithm 3B42: TRMM Merged HQ/Infrared Precipitation. [Online]. GODDARD Space Flight Center; 2013 [cited 2017 April 5]. Available from: https://trmm.gsfc.nasa.gov/3b42.html.
31. OpenStreetMap. OpenStreetMap - Accuracy. [Online]. openstreetmap.org: OpenStreetMap; 2016 June 26 [cited 2017 February 05]. Available from: http://wiki.openstreetmap.org/wiki/Accuracy.
32. Seneca Labs Inc.. Calories burned walking calculator. [Online]. Captain Calculator; 2017 [cited 2017 April 2]. Available from: https://captaincalculator.com/health/calorie/calories-burned-walking-calculator/.
33. ProCon.Org. MET Values for 800+ Activities. [Online]. ProCon.Org; 2012 [cited 2017 April 2]. Available from: http://golf.procon.org/view.resource.php?resourceID=004786.
34. innovativegis.com. Topic 19: Routing and Optimal Paths. [Online]. innovativegis.com; 2003 [cited 2017 April 28]. Available from: http://www.innovativegis.com/basis/mapanalysis/topic19/topic19.htm.
35. ChartsBin statistics collector team. Daily Calorie Intake Per Capita. [Online]. ChartsBin; 2011 [cited 2017 April 2]. Available from: http://chartsbin.com/view/1150.
36. UN. Water scarcity. [Online]. New York: DPI; 2014 Nov 24 [cited 2017 February 7]. Available from: http://www.un.org/waterforlifedecade/scarcity.shtml.

Appendix 1: plot of NASA and IWRM rainfall time series

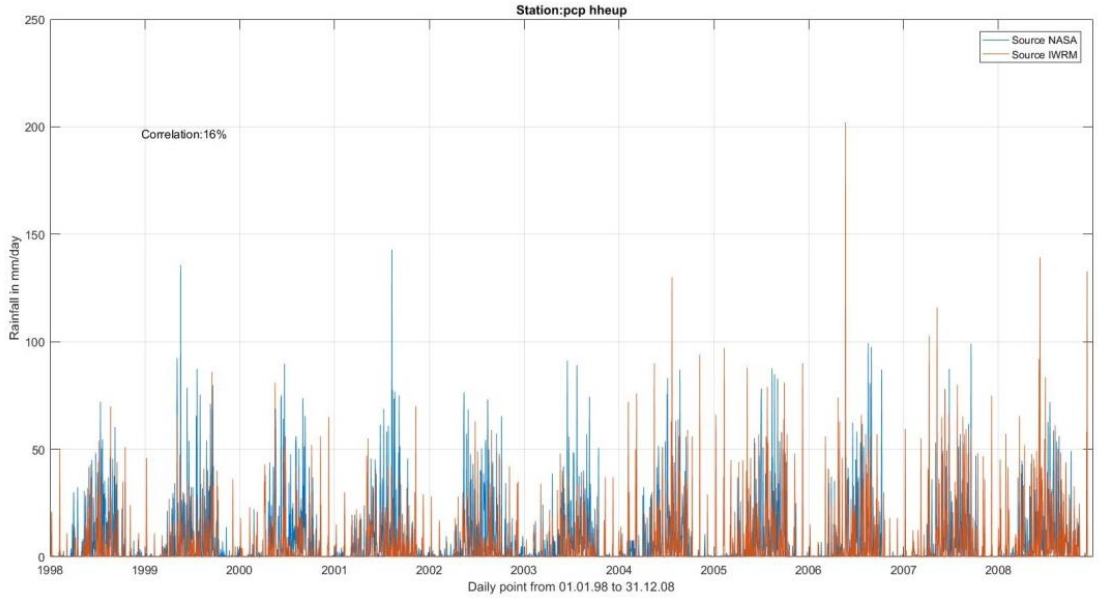

Figure ii. Weather station hheup (1), plot of the time series from NASA and IWRM

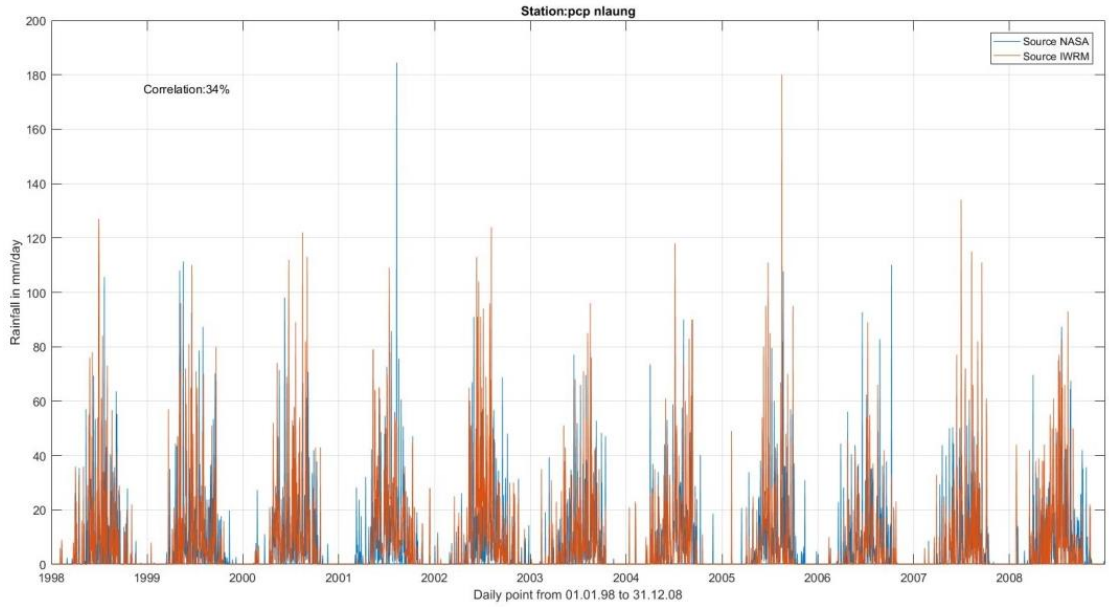

Figure hh. Weather station nlaung (3), plot of the time series from NASA and IWRM

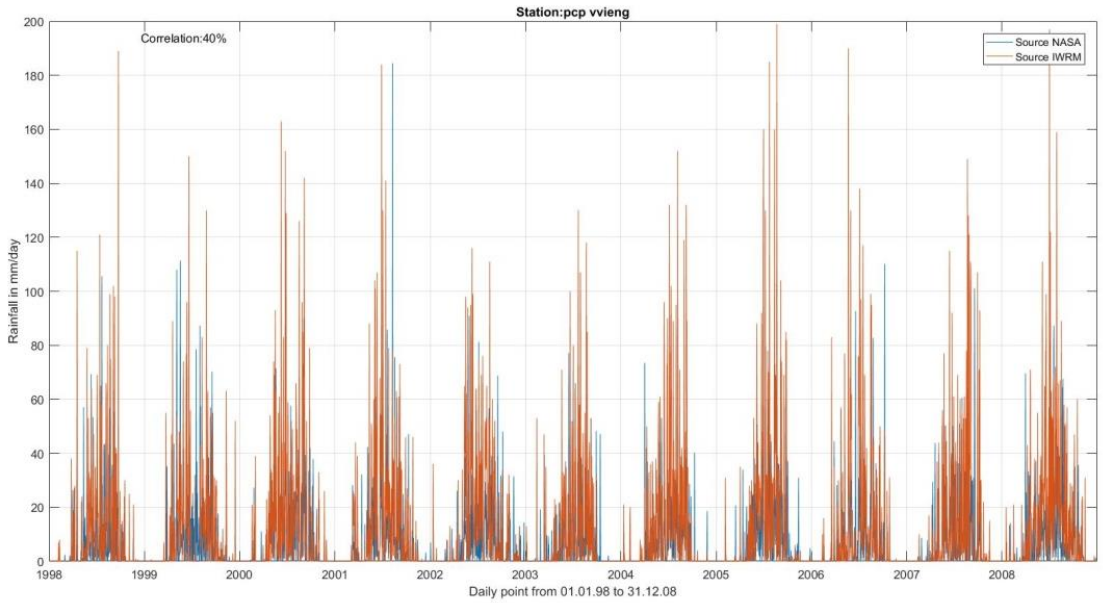

Figure kk. Weather station vvieng (10), plot of the time series from NASA and IWRM

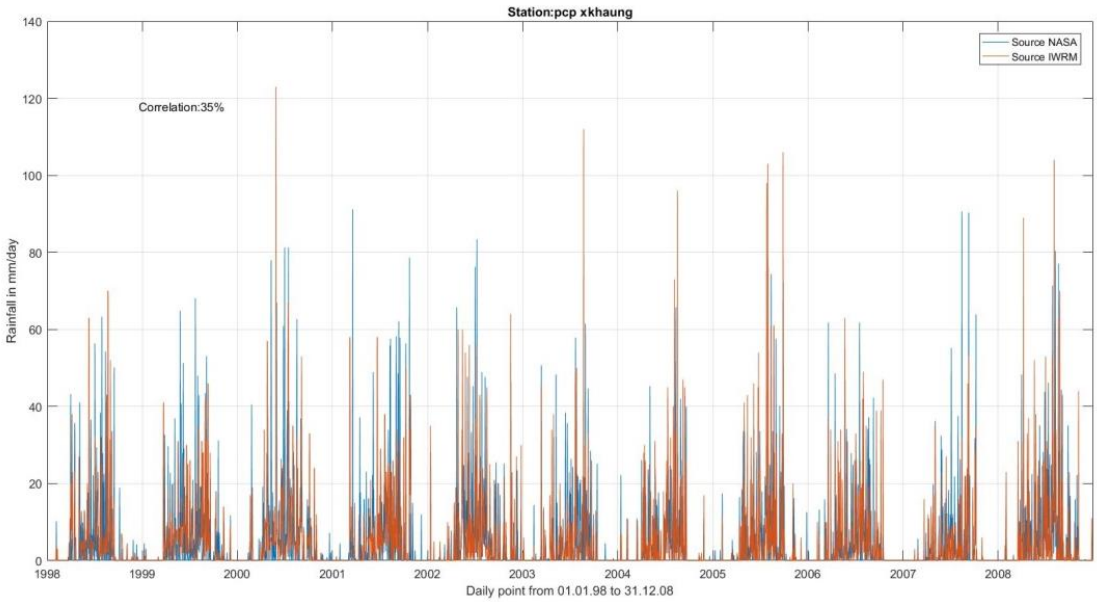

Figure jj. Weather station xkhaung (11), plot of the time series from NASA and IWRM

Appendix 2: Detailed of the IWRM model optimization

Table vii. Initial and final values for the parameters optimized in the IWRM model

Iteration number (1 year) for the model run				1	133
LU	PRE	rainmult	Shrub and grassland	**0.99**	0.9
LU	PRE	rainmult	Deciduous forest	0.99	1.1
LU	PRE	rainmult	Irrigated agriculture	0.99	1.1
LU	PRE	rainmult	Water	0.99	0.8
LU	PRE	rainmult	Other	0.99	1.1
LU	EVA	petcorr	Shrub and grassland	0.55	0.65
LU	EVA	petcorr	Deciduous forest	0.55	0.35
LU	EVA	petcorr	Irrigated agriculture	0.55	0.55
LU	EVA	petcorr	Water	0.55	10
LU	EVA	petcorr	Other	0.55	0.55
LU	SML	infpratmult	Shrub and grassland	12	11.5
LU	SML	infpratmult	Deciduous forest	12	20
LU	SML	infpratmult	Irrigated agriculture	12	6
LU	SML	infpratmult	Water	12	12
LU	SML	infpratmult	Other	12	28
ST	INF	Infkz	Acrisols	0.5	0.5
ST	INF	Infkz	Lithosols	0.5	1
ST	INF	Infkz	Water	0.5	0
ST	INF	Infkz	Other	0.5	0
ST	SL1	dz1	Acrisols	1.5	2.2
ST	SL1	dz1	Lithosols	1.5	0.2
ST	SL1	dz1	Water	1.5	0.1
ST	SL1	dz1	Other	1.5	0.1
ST	SL2	dz2	Acrisols	12	18
ST	SL2	dz2	Lithosols	12	12
ST	SL2	dz2	Water	12	0.1
ST	SL2	dz2	Other	12	0.1
R-squared (nlaung point)				75.50%	82.01%
Nash Sutcliffe model efficiency				66.67%	75.67%
Avgflow (m3/s) comp				179	170
Avgflow (m3/s) meas				170	170
Precipitation (mm/year)				2472	2475
Evaporation (mm/year)				743	787
Difference in indicator				**9%**	**6%**
Global rank				**71%**	**79%**

Appendix 3: Cost surface layer

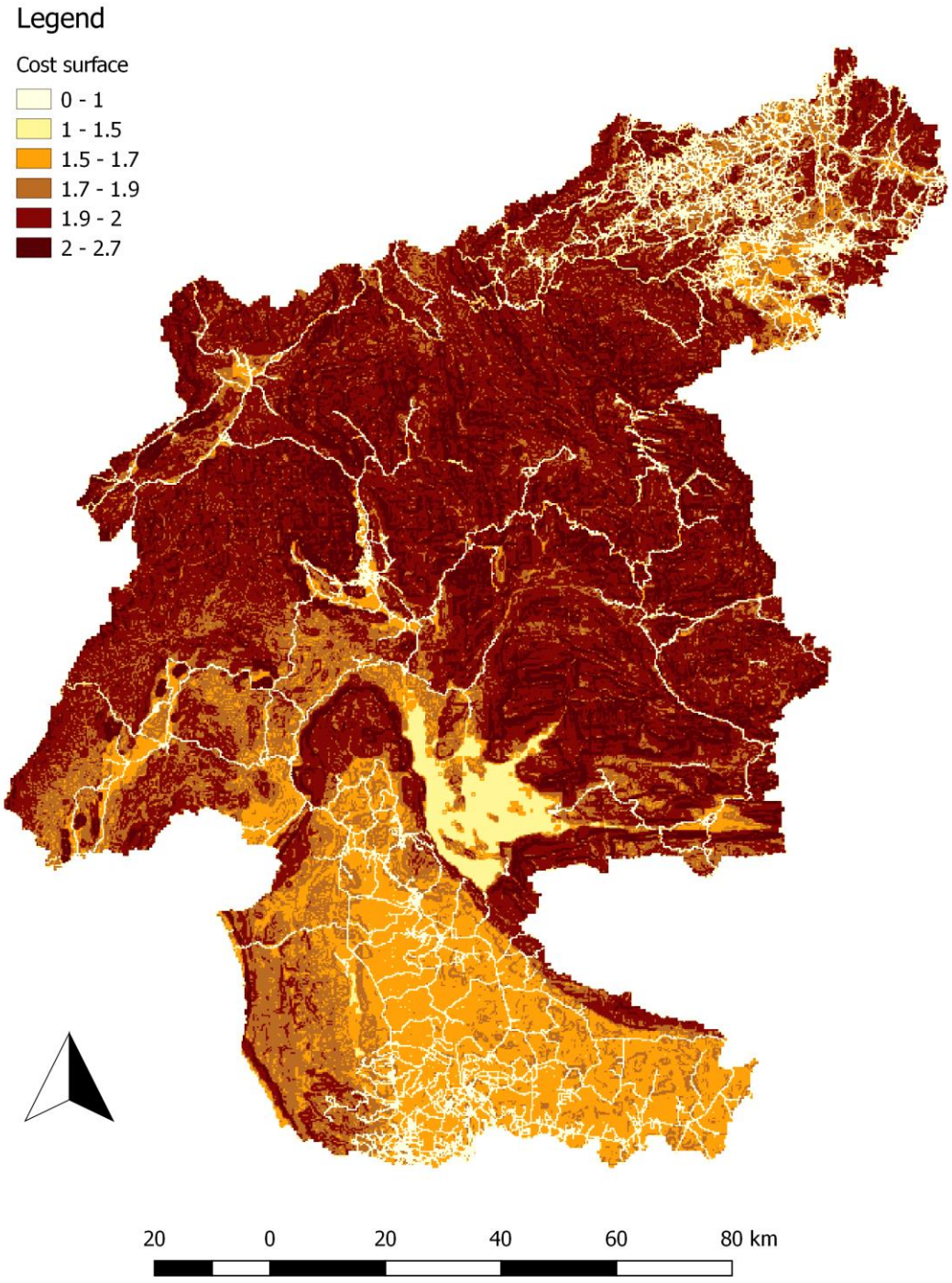

Figure II Cost surface layer

Appendix 4: Examples of the accumulation cost shape (VACA)

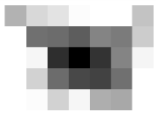

Figure nn. VACA for village 222 threshold 5

Figure mm. VACA for village 159 threshold 7

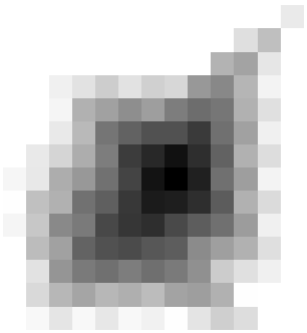

Figure oo. VACA for village 191 threshold 11 (average threshold)

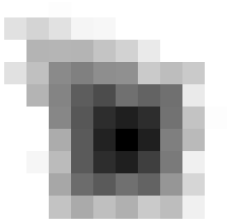

Figure pp. VACA for village 158 threshold 8

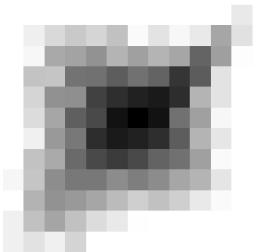

Figure rr. VACA for village 258 threshold 9

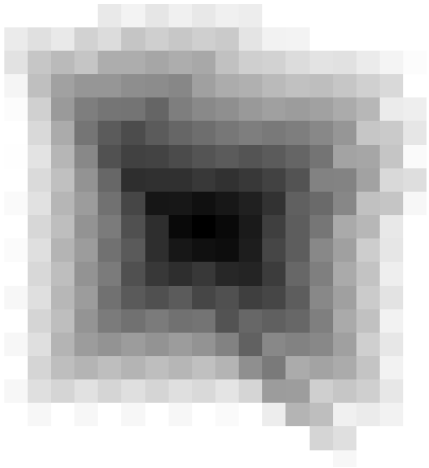

Figure qq. VACA for village 609 threshold 18

Appendix 5: Path found by the algorithm

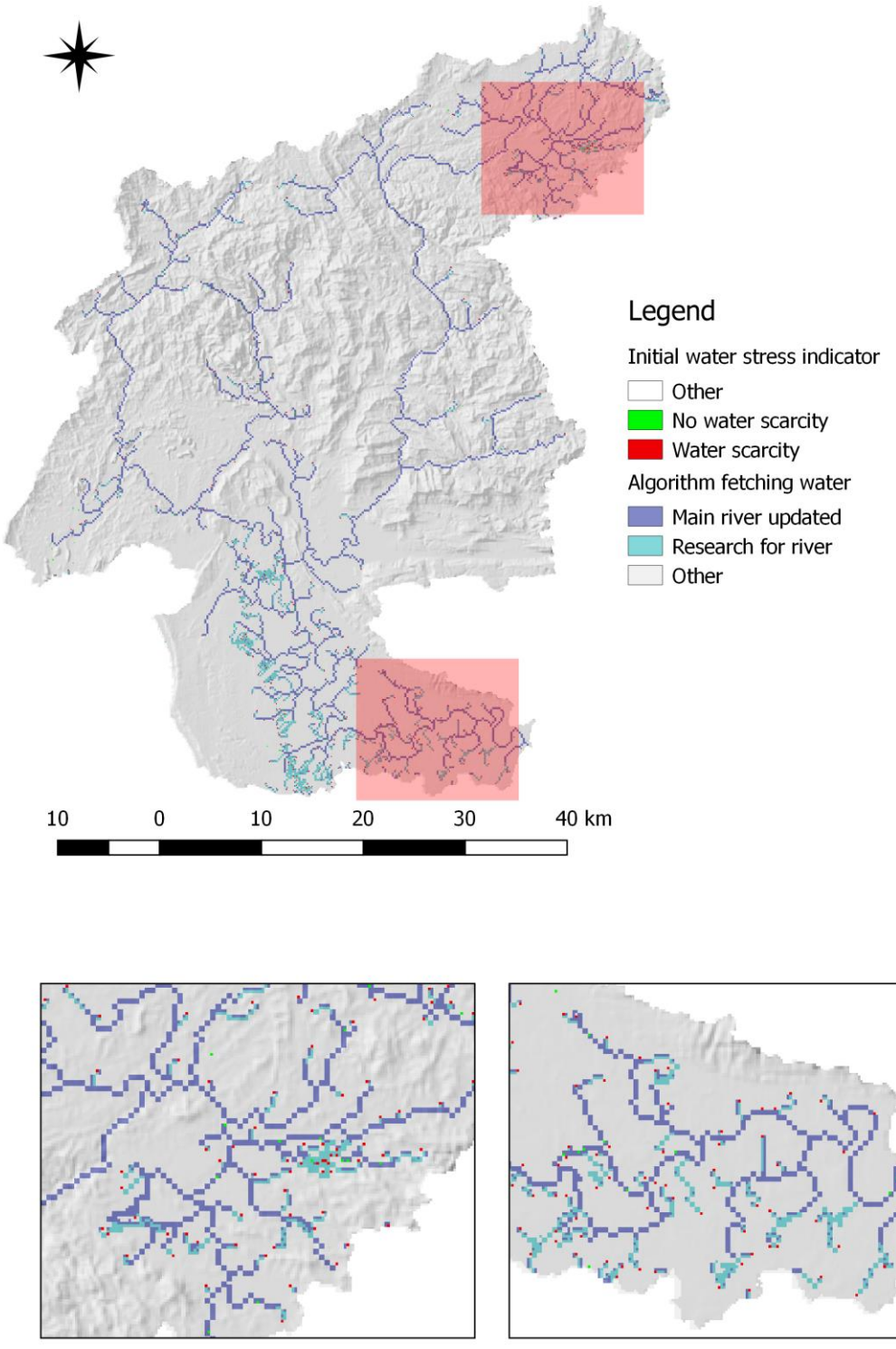

Figure ss. Map from the algorithm updating the discharge layer

Appendix 6: Initial and final Water Stress Index

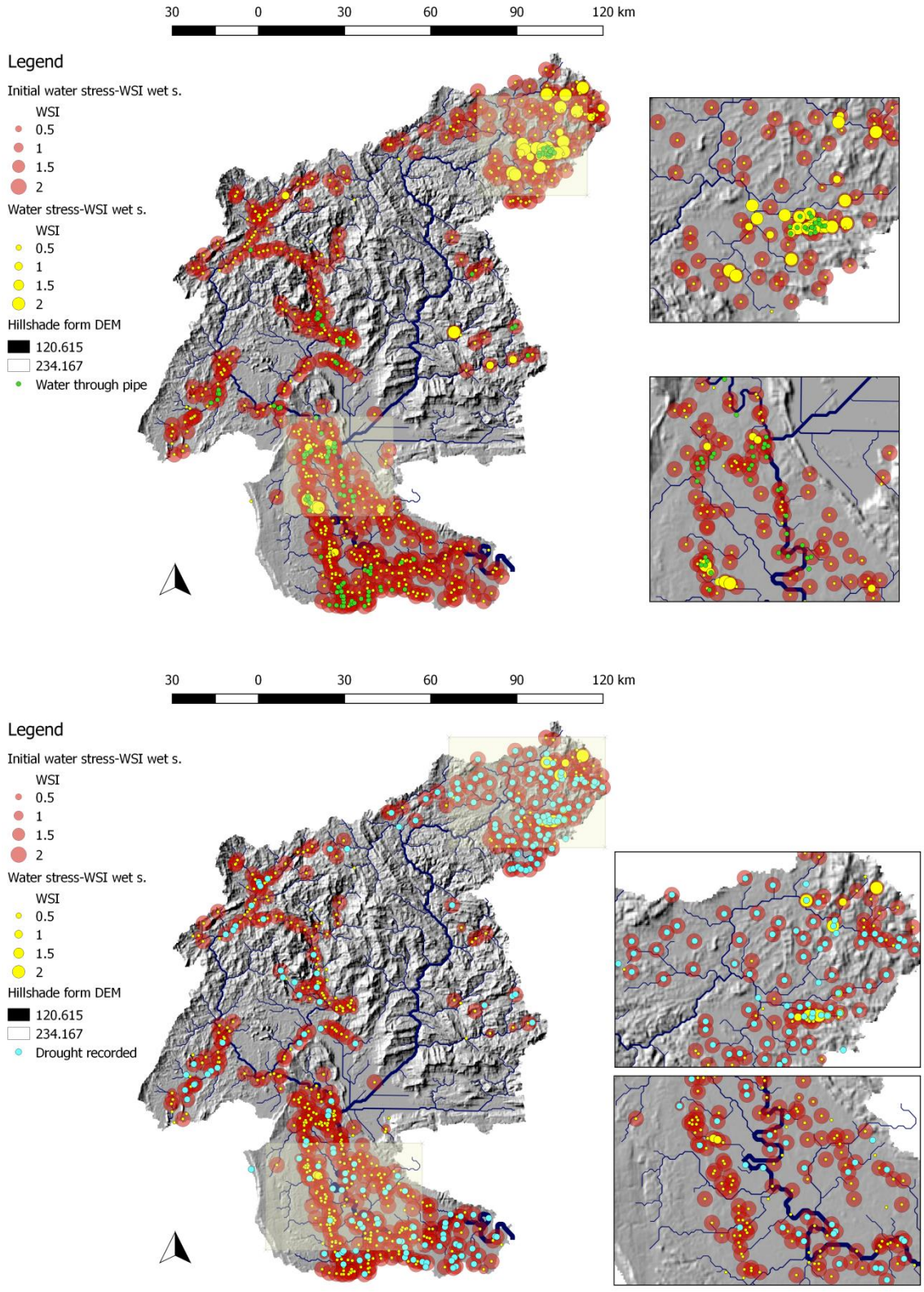

Figure tt. WSI during dry and wet season with additional information.

Appendix 7: Final WSI with the three thresholds from Falkenmark

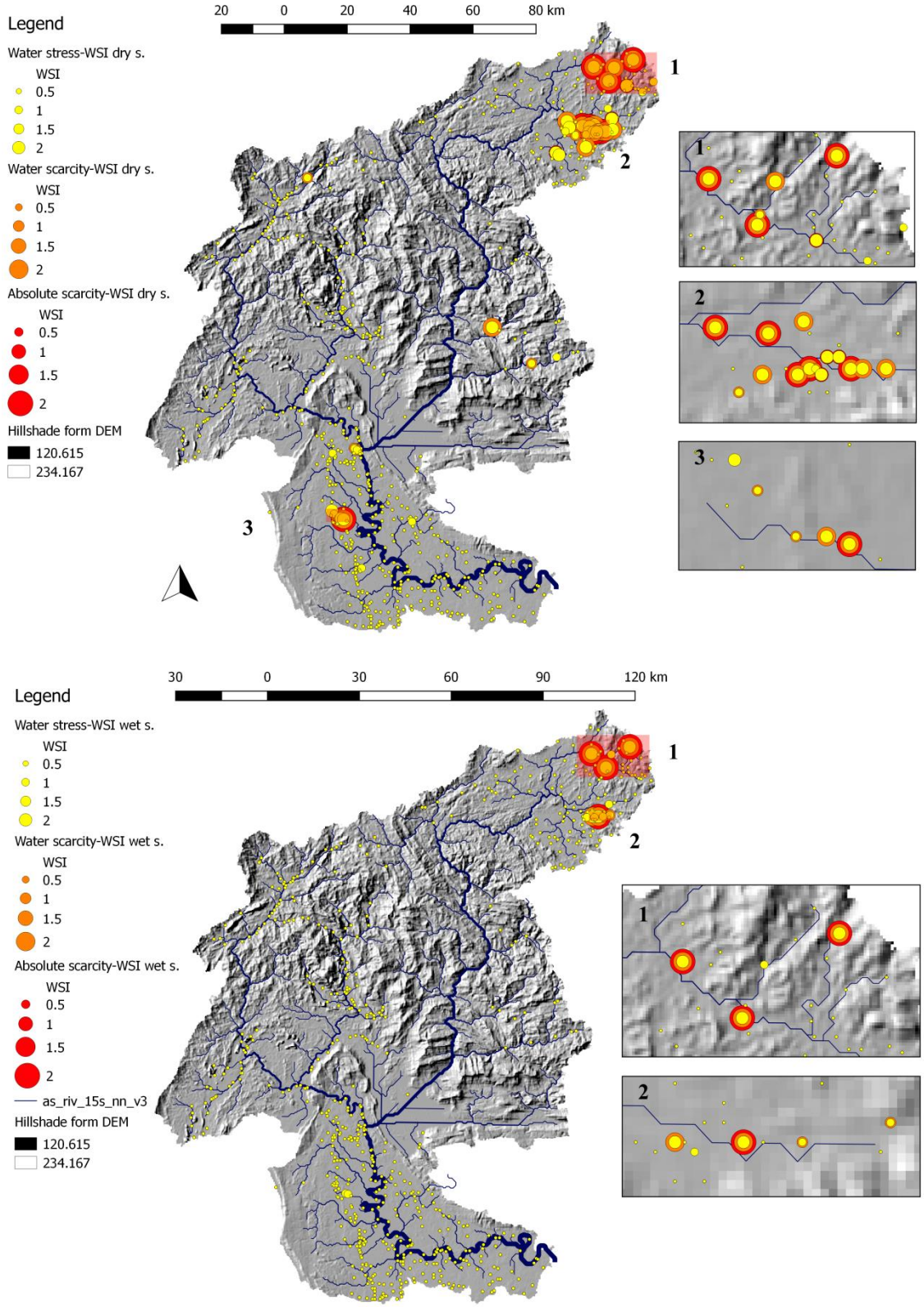

Figure uu. Final WSI with the three thresholds from Falkenmark

Appendix 8: Water available within the VACA

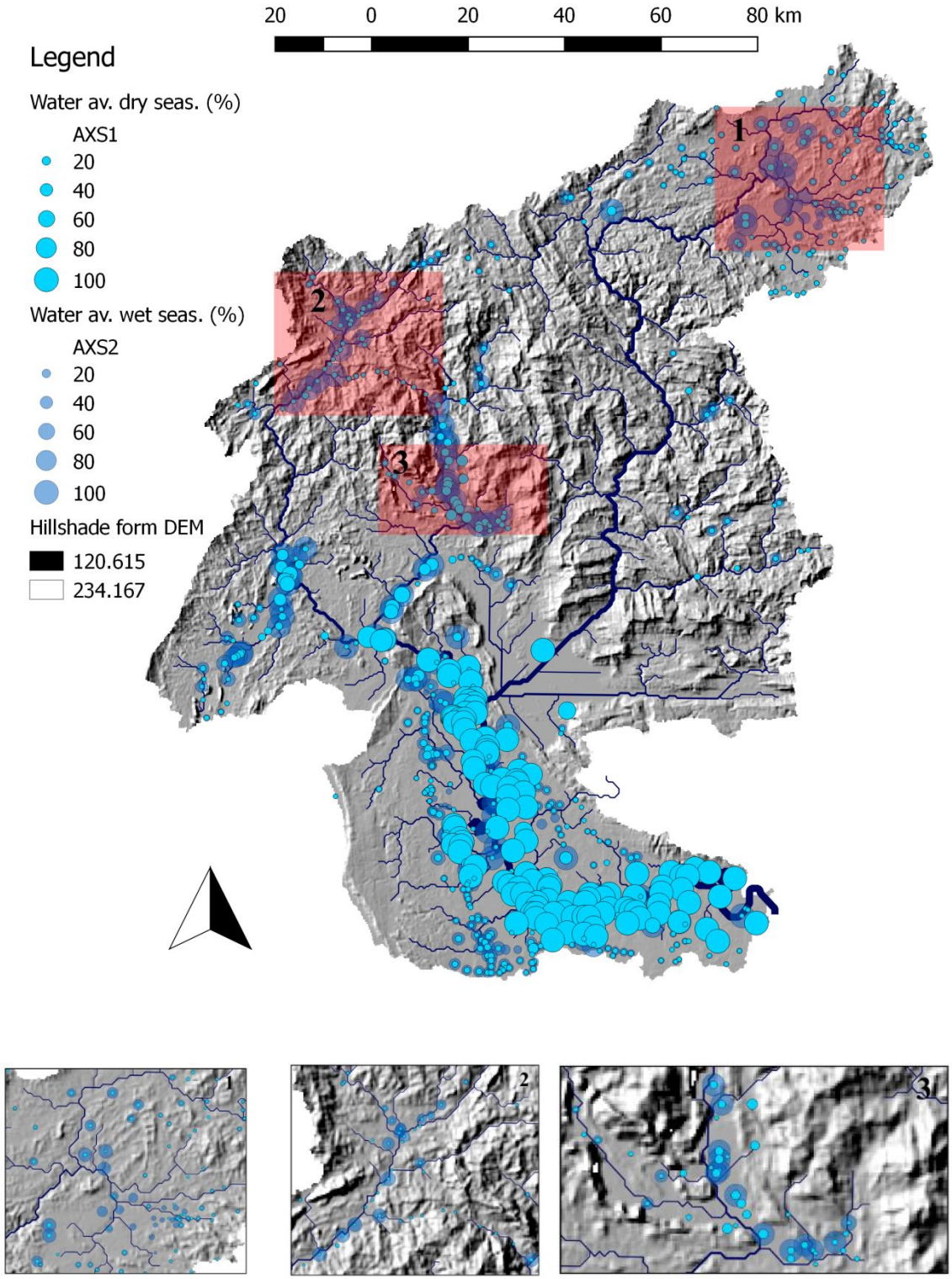

Figure vv. Normalized quantity of water available within VACA