TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Environmental Sciences Institute for Cartography

Master Thesis

Digital Morphometry Applied to Geo-Hazard Risk Assessment: A Case Study from Germany

submitted by **Karen Byrne**

born on 29.07.1988 in Tipperary, Ireland

submitted for the academic degree of
Master of Science (M.Sc.)

Submission on 10.10.2018

Supervisors
Dr.rer.nat Nikolas Prechtel
Technical University of Dresden

Dr.rer.nat Ourania Kounadi
University of Twente

Statement of Authorship

Herewith I declare that I am the sole author of the thesis named

“Digital Morphometry Applied to Geo-Hazard Risk Assessment:
A Case Study from Germany”

which has been submitted to the study commission of geosciences today.
I have fully referenced the ideas and work of others, whether published or unpublished. Literal or analogous citations are clearly marked as such.

Dresden, 10/10/2018 Signature:

Acknowledgement

I would like to take this opportunity to thank Dr. Nikolas Prechtel and Andreas Knobloch for their continuous support and guidance throughout this thesis.

Also I would like to extend my gratitude to the team at Beak Consultants Gmbh, I thoroughly enjoyed my time working at the company and I am forever grateful for all of your advice and help throughout my time there.

Thank you to my family for their unfailing support and belief in me, and for all the pep talks. To my husband Michell, for his never ending patience and positive attitude, for you I am eternally grateful.

Lastly but definitely not least, I would like to extend a special thank you to my friends Natalie and Satarupa for their never ending support, advice, guidance, I don’t think I could have done it without you and for everything I am eternally grateful.

Thank you.

Table of Contents

List of Figures

List of Tables

Appendix

List of abbreviations

ANN – Artificial Neural Network

ATKIS - Authoritative Topographic Cartographic Information System

AUC – area under the curve

BS – Prediction Model (Basic parameters plus soil)

BKG - German Federal Agency for Cartography and Geodesy

BP – Back propagation Algorithm

DD – Drainage Density

DEM – Digital Terrain Model

DI – Dissection Index

DTF – Distance to faults

FOC – Factor of safety

FP – Prediction Model (basic parameters plus new test parameters)

FPS - Prediction Model (basic parameters plus new test parameters & soil)

FR – Frequency response

FVRM - Prediction Model (basic parameters plus optimised parameters & soil)

GIS – Geographic Information system

HI – Hypsometric Integral

KL – Geological category: fault susceptibility

LfULG - Saxon State Office for Environment, Agriculture and Geology

LS - Landslide

LSS – Landslide Susceptibility

LR – Logistic regression

MLP – Multi Layered Perceptron

NN – Neural Network

RDE - Relation Declivity Extension

RMSE – Root Mean Square Error

ROC – Receiver operator curve

TF – Geological category: prone to lineation's and separating surfaces/contacts.

TRI – Topographic Roughness Index

TPI – Topographic Position Index

TWI – Topographic Wetness Index

SPI – Stream Power Index

SVM – Support vector machine

VRM – Vector Roughness Measure

WoE – Weight of evidence

ZM – Zwickauer Mulde

Abstract

A landslide susceptibility analysis is performed utilising an artificial neural network (ANN), in order to model the nonlinear relationships between landslide occurrence and a set of newly developed parameters. The main research aim was to create and test new morphometric parameters datasets in an ANN to determine whether higher accuracies in landslide susceptibility predictions can be obtained. The ANN was applied to a test area near the city of Zwickau in the south west of Saxony, Germany. Seven new parameters were developed in ArcGIS model builder, and used as input for a new higher resolution models. The new parameters included were, stream power index, topographic wetness index, dissection index, vector ruggedness measure, topographic ruggedness measure, hypsometric integral, and knickpoints generated from a relative extension index (RDE). Other new datasets were also tested such as geological lineaments and soil. All datasets were modelled alongside traditional parameters such as slope, flow accumulation, curvature and lithology. The data were pre-processed, normalised and and imported from ArcGIS into the ANN. Final evaluation of both the training phase and application phase results showed that when the newly developed parameters were included in the modelling process, notable improvements were made in the prediction classification accuracy.

Keywords: *Landslide susceptibility, ANN, GIS, morphometric parameters.*

1. Introduction

1.1 Purpose and Scope:

The evolving geomorphology of our earth's surface, has led to the creation of a diverse patchwork of complex landforms. Variations in landscape occur due to ongoing erosional and deformational processes, however as we continue to expand into new areas while dealing with the consequences of climate change, there is a need for improved modelling practices to deal with arising issues. Infrastructure construction and intensive land use practices lead to increased pressure on land resources and modification of the natural surface process of the earth. By studying surface morphology, we can gain a wealth of knowledge about the landscape around us, which could potentially help mitigate negative impacts to our way of life.

One such example of this is landslide susceptibility mapping which uses morphological geological and environmental factors to predict potential locations susceptible to landslides. This can be utilized for the prevention of infrastructure damage and impacts to livelihood. Studies over the last two decades have identified the impact of human interaction with our environment through both construction and environmental means which can lead to the initiation or re-initiation of landslides (Bruschi VM, 2013; Meusburger K, 2008; Van Den Eeckhaut M, 2009; Vanacker V, 2003). The German government has identified geohazards such as landslides as a potential future threat to the countries road and rail network, under the changing climate conditions (Klose, Auerbach, Herrmann, Kumerics, & Gratzki, 2017). Landslides by nature occur as local phenomena but are a widespread issue for road safety and operations, especially when coupled with storm events with excessive rainfall and extreme conditions (Krauter, Kumerics, Feuerbach, & Lauterbach, 2012). The creation of higher accuracy models can facilitate the identification of vulnerable areas, which, while currently stable, may be liable to fail in the future under changing climate conditions.

Traditionally morphological parameters were obtained from aerial photographic and topographic map interpretations. Technological improvements in remote sensing, Lidar, plus DEM technologies have opened a pathway for new types of analysis methods. The area of Digital Terrain Analysis (DTA) has seen a flurry of activity in the recent decades due to these advancements. Notably, the higher resolution DEM's currently available, are frequently being used to calculate parameters for direct use in the DTA process (Jordan, 2004).

A number of toolboxes and programs have been developed to automatically extract morphological parameters from DEM's for terrain analysis. Examples of these include TecDEM, TecLines, SAGA, Topotoolbox and various others (Pérez-Peña, Azañón, & Azor, 2009; Schwanghart & Scherler, 2014; Shahzad & Gloaguen, 2011a, 2011b). To model these parameters for the purposes of landslide susceptibility prediction many different methods have been used including weight of evidence, logistic regression, frequency ratio, and Artificial Neural Networks.

In the past computational solutions for analysing multiple complex datasets for the purpose of landslide modelling were quite restricted. Recently we can profit from the creation of sophisticated statistical systems such as Artificial Neural Networks (ANN) which, with the addition of training data, can act as a powerful method of identifying landslide susceptibility zones (Paraskevas Tsangaratos & Benardos, 2014).

This aim of this study is thus, to explore the inclusion of new morphological datasets in the landslide susceptibility modelling process, to evaluate their performance in an artificial neural network and to assess whether the overall prediction accuracy for the test area can be improved compared to previous lower resolution models. Morphological parameters will be reviewed from previous tectonic and soil studies to determine their applicability within the study. From these suitable parameters will be selected and the methodologies adapted to GIS. The original modelling methodology put in place by Beak Consultants Gmbh, will be used as the foundation upon which the new models shall be built. The work will be conducted using their Advangeo® modelling software and the research will be undertaken using higher DEM resolution of 10m, compared to the previously used 20m resolution.

1.2 Objectives

The main objective of this research is to explore whether the addition of new DEM-derived morphological parameters can enhance landslide susceptibility prediction. These parameters will be tested in a GIS-based Artificial Neural Network.

Any gain in the accuracy to a landslide susceptibility prediction can be utilised to further prevent damage to property and life and for future infrastructure planning decisions.

For this study the following objectives were formulated:

- Review and selection of suitable morphological parameters, with a specific focus on tectonic morphological parameters, for the use in Landslide susceptibility modelling.
- Investigate the relationship between tectonic morphology and landslide prediction and occurrence.
- Assessment of the available morphological analysis toolboxes for DEM's, i.e. TecLines, TecDEM, SAGA, Topotoolbox.
- Adaption of techniques and datasets derived from selected toolboxes, to a GIS-based artificial neural network (ANN).
- Collection of training data for the ANN, through site-based fieldwork.
- Determine the effect of the chosen datasets when incorporated into the ANN modelling process.
- Asses various model validation methods to determine if the newly introduced parameters increase the accuracy of the model results.
- Evaluation and selection of the best methods for visualising the results of the study.

1.3 Research questions:

- What are the tectonic morphological parameters that can be derived from a DEM?
- Of these, which are the greatest importance according to previous literature?
- Many toolboxes for morphological analysis exist, can their methodology and produced datasets be adapted for the purposes of the project, and what is the best approach to take?
- Does functionality already exist in ArcGIS which can be implemented in Advangeo®?
- What methods can be used on the model output for validation and assessment?
- Which dataset can be said to add the most benefit to the prediction model?
- Can an overall improvement in the susceptibility prediction be obtained?

1.4 Thesis Structure

This thesis consists of seven chapters. Chapters one through four describe the background and theory behind the project, and set up the foundation upon which the study is based. They include an introduction to the thesis, followed by an extensive review of previous literature in Chapters three and four, and in Chapter 4 the study area is discussed. Chapters five through seven, then detail the body of the work of the project, describing the methodology used, and finally presenting the results, conclusions and further recommendations.

2. Background

2.1 Landslides

A landslide is a complex process defined as the downslope movement of rock, debris and soil material most commonly under the force of gravity. The complexity of this process has led to many definitions in the research community (L. M. Highland & Bobrowsky, 2008). Landslide has become the widely used term for mass movement but in essence, this is not correct, as mass movement does not necessarily involve sliding. Rock falls, for example, have no sliding mechanism hence they would not classify under the title of Landslide. Despite this, the wider academic community has come to commonly use the term landslide and are aware of its overarching meaning (Varnes, 1978).

Many attempts have been made to classify mass movement, however, Varnes (1978) classification (Table 2.1) is considered the standard landslide schema. He classifies by two elements, the material, and the movement mechanism. A category is also assigned for complex movement which incorporates a combination of types.

TYPE OF MOVEMENT			TYPE OF MATERIAL		
			Bedrock	ENGINEERING SOILS	
				Predominantly Coarse	Predominantly fine
Falls			Rock fall	Debris fall	Earth fall
Topples			Rock topple	Debris topple	Earth topple
Slides	Rotational	Few Units	Rock Slump	Debris slump	Earth slump
	Translational	Many Units	Rock block slide/ Rock Slide	Debris block slide/ Debris slide	Earth block / Earth slide
Lateral spreads			Rock Spread	Debris spread	Earth spread
Flows			Rock Flow (deep creep)	Debris flow	Earth flow
				Soil creep	
Complex			Combination of two or more principal types of movement		

Table 2.1 Adapted from Varne's classification of mass movement (Varnes, 1978)

The term "slide" in the stricter sense, applies to mass movements which move along distinctive zones of weakness, this plane of weakness separates the material which makes up the slide from the stable ground beneath. The two most prominent forms are rotational and translational (Figure 2.1). These vary in the type of plane along which a rupture occurs. Translational landslides occur along planar planes and are associated with slopes ranging from 20 to 40 degrees. Rotational on the other hand occur along curved planes, are generally shallower than translational slides and can range from very small to large regional failures. Translational slides are one of the most common types of landslides and occur worldwide.

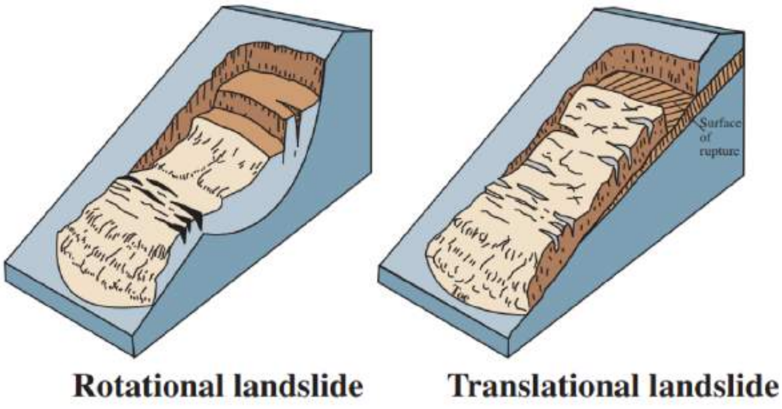


Figure 2.1 Rotational and translational landslides (adapted from USGS).

Falls and topples can be interpreted as a detachment of the material from the source area with little to no shearing influence involved, falling downwards under gravity (Figure 2.2). Falls and topples are considered as a more abrupt displacement, and the material general breaks on impact forming scree or rubble piles below the rupture site.

Flows are a rapid form of mass movement commonly seen worldwide after intense rainfall or snowmelt events. They are defined by a mixing of the material as it moves downslope. Flows are most often associated with a large water content however, this is not always the case, as rock and sand flows most often contain low amounts of moisture (L. Highland, 2004).

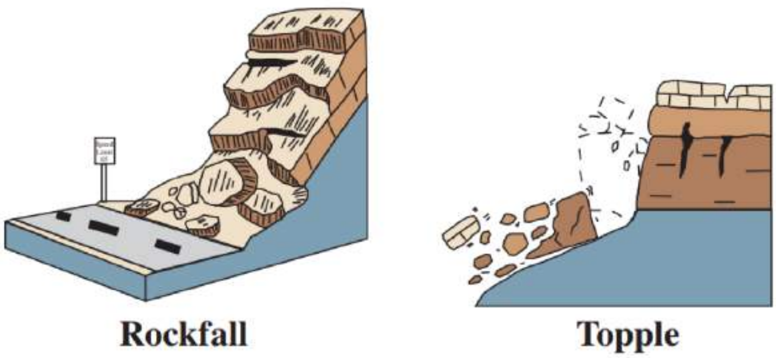


Figure 2.2 Rockfall and Topple failures (adapted from USGS)

As stated previously gravity is not necessarily always a factor in the mass movement of material. Despite the association of landslides with higher elevation areas, they can also occur in areas of lower elevation. The collapse of mining dumps and in particular coal mining dumps are a common occurrence in un-managed dumps. Quarry's and open pit mining create artificial areas of higher slopes with the potential to fail. Spreading is a form of displacement which occurs on relatively low angle slopes or flat areas. Spreading occurs in the process of liquefaction. During this process waterlogged, loose and non-binding sediments namely of sand and silt change from a solid state to a liquefied state. When a stable material layer such as bedrock or soil, sits on top of a layer of liquefiable material, an event such as oversaturation, earthquakes, weight-loading etc. can cause fracturing and extension in the top layer, which may be followed by subsidence, rotation or flow (L. M. Highland & Bobrowsky, 2008).

Overall a mass movement from any of the categories can be regarded as a movement of material from its source point to more stable conditions. Mass movements are a natural process with the aim of reaching an equilibrium between contributing forces. Landslides are characterised by internal and external conditioning factors and triggering factors. Triggers can change previously stable conditions resulting in destabilisation, many factors exist (Figure 2.3) however, the three main ones to be considered are water (in the form of intense precipitation, runoff erosion, over-saturation, and flooding), seismic activity and volcanic activity (L. Highland, 2004). Triggers can also be anthropogenic in nature, land use practices lead to soil erosion and greater runoff, deforestation plays a big role in this in some countries. (Crozier, 1984).

Geological Causes	Morphological Causes	Human Causes
Weak or sensitive material	Tectonic or volcanic uplift	Slope or toe excavation
Weathered materials	Glacial rebound	Slope or cret loading
Sheared, jointed or fissured materials	Subterranean erosion	Drawdown (reservoirs)
Contrast in permeability and/or stiffness of materials	Slope or crest loaded with deposition	Deforestation
Adversely orientated discontinuity (bedding, faults, unconformity,	Vegetation removal (natural causes)	Irrigation
	Thawing of frozen ground	mining
	Freeze-thaw weathering	Artificial Vibration
	Shrink-swell weathering	Water leakage from utilities
	Fluvial, wave or glacial erosion of slope toe or	

Figure 2.3 Categorised triggering factors in mass movement (adapted from USGS).

In contrast to triggers, internal and external conditions are generally pre-existing but can change over time, these include internal conditions such as geological lithology, structures, and extent of near-surface weathering. External conditions include relief, soil coverage, and morphology. The morphology of the landscape greatly affects slope stability, hence parameters derived from the morphology such as slope, aspect, curvature, and roughness, can subsequently be used as indicators to highlight (predict) susceptible areas (R. J. Pike, Evans, & Hengl, 2009). This is done on the premise that landslides of similar scale and environmental background will occur under a similar set of conditions as past landslides. Morphological parameters will be discussed further in Chapter three. Another element that deserves mentioning is the scale of the landslide. Landslides can range from small-scale shallow landslides, where the material movement is more superficial than catastrophic, to regional scale landslides. Small landslides are common and widespread throughout the world, whereas large failures occur less commonly and generally in areas with certain predisposing conditions (Glade & Crozier, 2005).

2.2 Landslide susceptibility mapping

Landslide susceptibility mapping is the use of the past to predict the future. Conditions where LS's previously occurred are said to exhibit conditions which can be used to predict areas vulnerable to future failures (Fausto Guzzetti, Carrara, Carrara, Cardinali, & Reichenbach, 1999). LS assessment is based on the conditioning factors discussed previously, it allows for the depiction of the spatial distribution of the susceptibility, without the necessity to determine magnitude or temporal aspects of the phenomena. In simpler terms, it estimates where a landslide is most likely to occur as a result of geo-environmental conditions, this does not take into account triggering events. Varnes (1978), describes susceptibility as "the probability that a landslide of a given type may occur in a given area ". Landslides due to their complexity are regarded as more difficult to model and assess than other phenomena such as flooding and earthquakes. The reason for this is down to the wide range of process which must be assessed (Glade & Crozier, 2005). Landslides susceptibility can be on various scales from, small (<1:100,000), medium (1:100,000 to 1:25,000) to large (1:25,000 to 1:5,000) (Fausto Guzzetti et al., 1999).

In previous literature, much confusion occurs between the terms "susceptibility" and "hazard" (Reichenbach, Rossi, Malamud, Mihir, & Guzzetti, 2018). Landslide "Hazard" mapping factors the probability of a landslide of a certain magnitude occurring within a specified time period. Landslide "susceptibility", solely deals with presenting spatially the relationship between mass movements and conditioning factors, while assuming that future events will occur under the same conditions (Fausto Guzzetti et al., 1999).

2.3 Methods of Susceptibility Mapping

Many techniques have developed over time for the production of LS maps. These techniques can be broadly grouped into quantitative and qualitative methodologies (Figure 2.4). Early methodologies focused mainly on qualitative methods, however better processing capabilities have meant that larger more detailed datasets can now be produced as part of quantitative predictions (S. Ali, P. Biermanns, R. Haider, & K. Reicherter, 2018).

Despite which methodology is used to derive the landslide susceptibility, a set of assumptions exists, these have been touched upon in this chapter but can be summarised in the following points according to Fausto Guzzetti et al. (1999). Firstly, landslides display recognisable imprints and/or indicators on the landscape which can be identified and mapped using aerial and satellite imagery, and direct mapping in the field (Alessandro Mondini, 2008; Varnes, 1978). Secondly, both natural and man-made conditions control landslide occurrence, these can be assessed using qualitative or quantitative methodologies. Factors related to slope failure can be gathered and utilized for prediction models (Costanzo, Rotigliano, Irigaray, Jiménez-Perálvarez, & Chacón, 2012; Dou et al., 2015). Thirdly, as stated previously, past landslide conditions are key to predicting future failures, by assuming the conditions of past and future landslides will the same (Varnes, 1978).

Qualitative or knowledge-driven heuristic methods rely on expert user input to determine susceptibility categories and as a result are inherently subjective. These categories tend to be described in terms of "very low" to "very-high" susceptibility (Dragićevića, Laia, & Balram, 2015). The main disadvantage to this method is that the prediction accuracy depends on the level of experience of the user. On the other hand, quantitative methods create predictions by assessing the relationship between landslide events and causative factors by statistical or deterministic analysis. This is an objective approach which reduces bias when weighting causative factors (Reichenbach et al., 2018). Deterministic methodologies assess failures by factor of safety (FOS) (Gorsevski, Gessler, Boll, Elliot, & Foltz, 2006). Statistical predictions are more popular in recent times. Statistical classification methods include, bivariate and multivariate, logistic regression (LR), weight of evidence (WoE), frequency response (FR), support vector machines (SVM) and Artificial Neural Networks (ANN) (S. Ali, P. Biermanns, R. Haider, & K Reicherter, 2018; Reichenbach et al., 2018). Statistical approaches employ indirect methodologies to objectively determine the relationships between dependent and independent factors or parameters by feeding the system training data and are subsequently verified using validation data. All possible parameters are entered and then compared with training and validation sets derived from the inventory data. to determine their effect (F. Guzzetti, 2006; Fausto Guzzetti, Reichenbach, Ardizzone, Cardinali, & Galli, 2006). GIS can be integrated into the pre-processing phase to help manage and analyse the heavy datasets. One disadvantage of quantitative methods is that they require high accuracy data spread out over the study area, this can be a problem in areas with little or no inventory and poor satellite data (Demoulin & Chung, 2007). Despite the numerous papers available comparing methodologies, the overall best approach is still a matter of debate (Carrara & Pike, 2008).

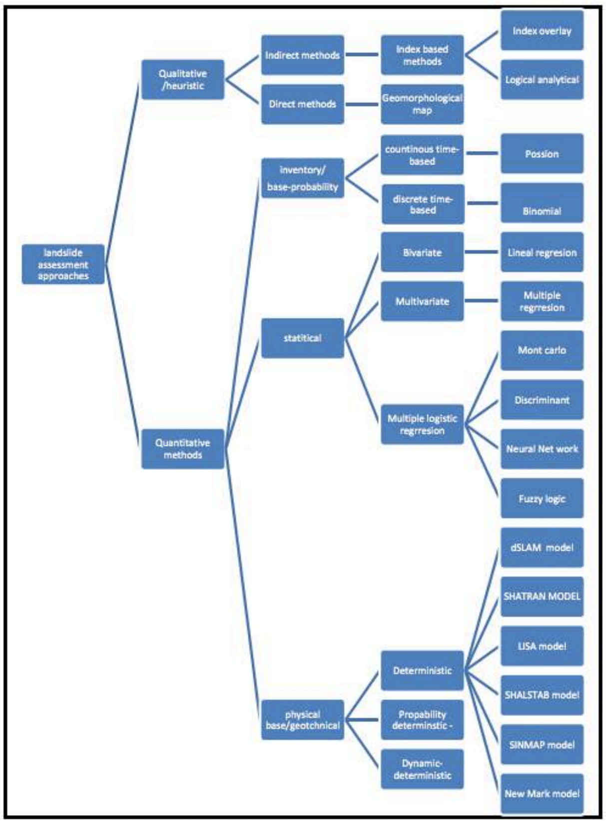


Figure 2.4 Landslide methodology classification from Carrara & Pike, 2008.

S. Lee and Talib (2005) noted that the selection of optimised factors for landslide susceptibility modelling can improve the prediction accuracy. They described how this can reduce discrepancies in the model. Pradhan and Lee (2010) also adopted this approach and removed parameters with smaller weight values, reducing the overall amount of final input parameters. Reichenbach et al. (2018) discusses the numbers of parameters used based on an extensive literature review. According to his research, he determined that the parameters used ranged from two to twenty-two for a single model, with the average being nine.

2.3.1 Landslide inventory

A landslide inventory, in theory, is a comprehensive record of all failures which have occurred in an area both past and present. An up to date and accurately recorded landslide inventory is necessary for landslide susceptibility, however, these rarely exist (Reichenbach et al., 2018). It is difficult to conduct studies with poor data quality, statistical predictions, in particular, require greater amounts of accurate data for modelling susceptible areas. A number of methods exist for the collection of this data including the direct collection of data in the field, historical record keeping of events, and identification through aerial and satellite imagery. Linking triggering events to landslide occurrence requires accurate recording of the factors surrounding the failure (extreme weather events, seismic events etc.), the type and scale of the event and its location (Fausto Guzzetti et al., 2012; Van Westen, Castellanos, & Kuriakose, 2008). The problem in this is that the conditions that existed to initiate the landslide in the first place often change dramatically. This can mean that a slope is now more susceptible to other another type of failure, as is the case with a landslide to debris flow scenarios, or the slope no longer exhibits the conditions which caused the event to begin with. Hence, great care and consideration is needed whilst capturing data.

The standardisation of landslide event records has yet to be agreed upon and the literature discusses the problems that this poses when conducting susceptibility predictions (Fausto Guzzetti et al., 1999).

The thesis research has been conducted using an Artificial Neural Network (ANN) called Advangeo©, developed by Beak Consultants. Hence, this methodology will be the focus of further discussion in this chapter.

2.4 Artificial Neural Networks in Landslide susceptibility

Artificial Neural Networks (ANN) in geosciences are used for multivariate, statistical based predictions of spatial phenomena. ANN's are "generic non-linear function approximators that were developed for pattern recognition and classification" (McCulloch & Pitts, 1943). "An ANN classifier defines a potentially complicated decision boundary in feature space" (Woods & Bowyer, 1994). They are capable of handling large quantities of data and learning complex model functions by "training" the system, similar to how our brains function. The brain consists of billions of interconnected neurons that form a neural network, we learn/process based on complex connections between these neurons (Ermini, Catani, & Casagli, 2005). ANN's strive to replicate this type of learning architecture by a combination of artificial intelligence and statistical analysis. It is an indirect or quantitative technique, which combines input parameters with training sets of the phenomena locations, and the system then "learns" the relationships from the data it is fed (Van Weston, Rengers, & Soeters, 2003). Traditional statistical modelling, on the other hand, requires previous knowledge of data relationships (Farrokhzada, Bararib, Choobbastia, & Ibsenb, 2011). In comparison to other methods, with ANNs the user selects the input parameters and criteria, and the system determines the susceptibility, hence it is inherently objective. ANN's are considered "black box models" because of the level of difficulty that exists in interpreting and manipulating the inner workings of the system (Saro Lee, Ryu, Min, & Won, 2003).

Various ANN architectures exist, one of the more common approaches is using a Multi-Layered Perceptron (MLP) with a back propagation learning algorithm (BPN), and a sigmoid activation function (Bishop, 1995). MLP's are a multi-layered feed-forward architecture, where the information is fed forward through multiple layers. The basic set up consists of three layers, an input layer with neurons for the selected input parameters, a hidden layer of neurons through which the information propagates and an output layer with generally one to two neurons (in the case of landslide susceptibility). Numerous hidden layers can be used, however previous studies noted that it was difficult to assess how many layers provide the most benefit as each layer added increases the complexity of the model and it's processing time and power, in the end, most studies opted for one hidden layer (Pradhan & Lee, 2010). The decision boundary in the feature space of the hidden layer is formed by a non-linear combination of a set of hyperplanes. Each node of the hidden layer is defined by a hyperplane (Woods & Bowyer, 1994).

There are two phases to the ANN modelling, the training and application phases. During the training phase weights of influence of each parameter, input are determined by comparing the relationship between the parameters and known occurrence and converted into spatial probabilities. The hidden and output layers multiply each input by a corresponding weight, summing the product of this, before processing the total using a non-linear transfer function (sigmoid) (Pradhan & Lee, 2010). The network "learns" by use of the back-propagation algorithm (BP), which adjusts the weights between the nodes ("neurons") in the input layer and the hidden layer and the hidden layer and the output layer, in response to errors between the initial actual output values and target output values (Figure 2.5) (Saro Lee et al., 2003; P. Tsangaratos & Bernardos, 2013). It does this over an assigned set of "epochs" or cycles of the system (Zhou, 1999). As the training phase progresses the error naturally drops (Ermini et al., 2005). At the end of the training phase, the ANN provides a model which can predict a target value from a given input value (Pradhan & Lee, 2010).Due to the use of a training data set to develop the relationships, ANN is considered a supervised classification method (Atkinson & Tatnall, 1997). The second phase is the application phase where the derived weights are used to apply the neural net on the rest of the area.

ANN's have many advantages, they are good at detecting patterns that are not always apparent to our human perception and interpretation. New complex and non-linear relationships (weights) can be analysed as they are independent of the statistical distribution of the datasets (Farrokhzada et al., 2011). Of significant importance is the ability of ANN's to view problems differently which cannot be solved by statistical methods due to theoretical relationships. ANN models are considered adaptive and capable of generalisation. They can also handle imperfect or incomplete data (Saro Lee et al., 2003).

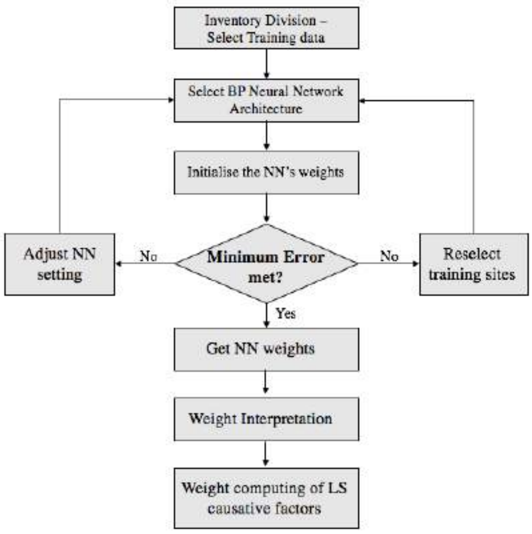


Figure 2.5 Flow chart for weight determination using an ANN model (After Pradhan, 2010)

There are some disadvantages associated with ANN's, for example, the BP algorithm can involve long execution times with a heavy computing load (Pradhan & Lee, 2010). Another issue with the BP is the local minima problem. This occurs during gradient descent following the slope of the RMS error value down along with changes in all weight values. The weights are continuously adjusted until the error value is no longer decreasing. This final position should be on the global minimum however, due to the complex nature of the RMS error value with many parameter weights, the network may instead converge into a local minimum Ultimately this means that more complex models require multiple re-runs, to asses the models and determine the best model produced.

2.5 Validation

Few authors have truly discussed methodologies in detail for the evaluation of model prediction performances (Reichenbach et al., 2018). That is not to say that studies have not conducted using validation methods, but that the purpose has mostly been for model comparison to assess which produced model is the best (Kalantar, Pradhan, Naghibi, Motevalli, & Mansor, 2018; Reichenbach et al., 2018). Prediction result values for LSS can be assessed by analysing the network error, the statistic distribution of the prediction results, cross-validation and field work (Kalantar et al., 2018). Other popular methods include receiver operator curve (ROC), area under the curve (AUC) and frequency ratio (Woods & Bowyer, 1994).

Reichenbach et al. (2018) strongly emphasises that while methodologies such as receiver operator curve (ROC), area under the curve (AUC) and success/prediction rates measure the overall performance, they cannot capture local conditions of an area or relevant geomorphological conditions. Models with high AUC values display a better statistical performance than lower values. However, the lower AUC value model may be more reliable and useful from a geomorphological standpoint.

ROC is a common method of evaluating classification performance and has been used in many studies (Choi, Oh, Won, & Lee, 2009; Dou et al., 2015; Fabbri, 2003; Pradhan & Lee, 2010; Shahabi & Hashim, 2015; P. Tsangaratos & Bernardos, 2013; Woods & Bowyer, 1994). ROC involves the calculation and plotting of true positives against false positives at multiple threshold setting. The further the line inflects towards the upper left of the plot, the better the model is interpreted as. The AUC determines the model accuracy figure of the model prediction. The values range from 0.5 to 1; 1 indicating perfect performance and 0.5 is achieved in the case of weak models (P. Tsangaratos & Bernardos, 2013). In chapter six, the results of a ROC, AUC validation alongside other evaluation techniques will be presented.

3. Geomorphometry and morphometric parameters

3.1 Introduction

"Geomorphometry is the science of quantitative land-surface analysis" (R. J. Pike et al., 2009). The primary objective of Geomorphometry is to characterise discrete surface features or landforms through the analysis of the earth's geomorphology (Ivan Marchesini1 & Mondini1 2014). Classical morphometry as a domain previously focused on the areas of surface form, calculating averages for elevation and slope, relative relief, contour maps, and drainage density. Modern Geomorphometry on the other hand now encompasses GIS data extraction and analysis of detailed continuous surfaces and the study of distinctive landforms such as watersheds (R. J. Pike et al., 2009). Landscapes are molded by tectonics, lithology, and rivers. Complex process interact between, tectonics, erosion, and sedimentation to create water gaps, knick points and meanders, as well as other tectonic and geomorphic features (Pirasteh, Pradhan, & Rizvi, 2009). Morphometric parameters are powerful indicators of these processes which shape our landscape.

Geo-hazard susceptibility mapping relies greatly on these parameters to make accurate assessments. Traditionally landslide susceptibility studies were limited by the ability to process large amounts of data and were based on qualitative classifications and interpretation. In recent times, however, this has changed thanks to the onset of advanced techniques in GIS analysis, data extraction and improved processing power using more sophisticated statistical modelling methodologies (R. J. Pike et al., 2009). An advantage to morphometric parameters is that they tend to be less specific than geo-environmental variables such as geology, climatology, soil, and land-use. Despite dependency on the kernel size and the DEM resolution used to derive these parameters, they are still more simplified than the geo-environmental variables which can be detailed and area specific, hence morphometric parameters can be applied to different areas with less difficulty (Reichenbach et al., 2018).

Morphometric parameters which have been used in the study of tectonics include isobase, drainage dissection & incision, surface roughness, hypsometric integral, bifurcation ratio, mountain front sinuosity, stream length index, knick-points and basin asymmetry (Gloaguen & Mahmood, 2011; Keller & Pinter, 2002; Kirby & Whipple, 2001; Kirby & Whipple, 2012; Mahmood & Gloaguen, 2012; Strahler, 1957). In soil and hydrological studies another set of

parameters are traditionally used, some of which include stream power index, topographic wetness index, stream order, stream length, elongation ratio, and drainage density.
According to R. J. Pike et al. (2009), there are five steps to geomorphometric analysis, first the sampling or generation of a surface such as a DEM, the correction of the surface model, the calculation of surface parameters or objects and the application of the results to the research problem. The basis behind surface parameters is the DEM, a representation of the surface of the earth. Layers derived from the DEM carry specific information that can be interpreted as features (R. J. Pike et al., 2009). Geomorphometric data can be classed into three types of data; basic, hydrological and climatological. Hengl and MacMillan (2009) reported that more than 100 basic and complex surface parameters exist for characterising a landscape.
Complex analysis such as landslide susceptibility incorporates a basic set of these parameters, such as slope, aspect, curvature, and flow accumulation, together with predisposing factors like geology, soil and land-use (Fausto Guzzetti et al., 1999; Reichenbach et al., 2018). Newly developed free software and packages such as TecDEM[1], SAGA[2], ILWIS[3], GRASS[4], Landserf[5], MicroDEM[6] and TauDEM[7], have been developed to take advantage of the increased quality and availability of modern DEM's. These toolboxes were not developed with landslide susceptibility, but rather for tectonic interpretation, soil analysis and general geomorphometry in mind. Some of these parameters are applicable to landslide studies and can be used cross-domain and incorporated into the modelling process. Detailed below are selected parameters which have been reviewed for this study.

3.2 Core parameters in LSS

Nowadays, while many new Landslide prediction techniques exist, the core assessment parameters used in these predictions have remained for the most part constant. Studies have focused on what are regarded as the core indicative parameters such as slope, aspect, geology,

---

[1] https://tecdem.soft112.com/
[2] https://saga-gis.org
[3] https://www.ilwis.org/open_source_gis_ilwis_download.htm
[4] https://grass.itc.it
[5] https://landserf.org
[6] https://www.usna.edu/Users/oceano/pguth/webiste/microdem/microdemdown.htm
[7] http://hydrology.usu.edu/taudem/taudem5/index.html

elevation, curvature, land-use, soil, drainage density and distance to faults (Dou et al., 2015; R. J. Pike et al., 2009).

3.2.1 Slope

It is known that slope is one of the most influential parameters on landslide occurrence and is one of the most commonly used parameters in the past and modern landslide susceptibility predictions (Costanzo et al., 2012; Kalantar et al., 2018; Reichenbach et al., 2018). Slope refers to the rate of change in height over the distance between two points. Low slope values represent flatter terrain while higher values represent steeper near vertical terrain. The slope is the foundation for two other important parameters which describe the slope, aspect, and shape (curvature) (Fausto Guzzetti et al., 1999)

3.2.2 Aspect

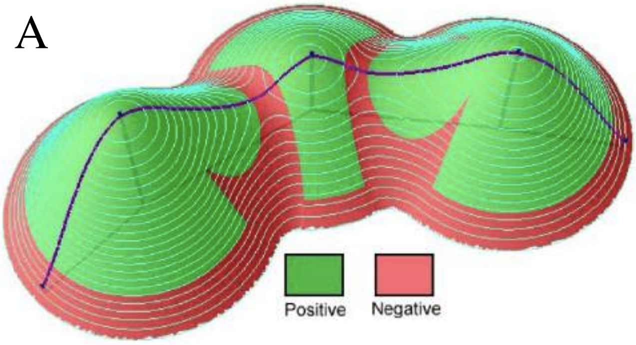
Aspect describes the orientation of a slope or the direction to which it faces, this parameter has been used worldwide for LS studies at different scales (Costanzo et al., 2012; Qiqing, Wenping, Wei, & Hanying, 2015; Roşca et al., 2015). Values can be represented with the cardinal directions, North, South, East, West, or 0-360°. A slope can be subjected to different climatological conditions depending on the direction to which it is orientated. This can create stress on natural processeses related to soil erosion and weathering of the underlying lithology and lineaments, and affect the overall moisture retention (Yalcin & Bulut, 2007). Slope aspect is also regarded as having an impact on vegetation cover, affecting landslide occurrence (Othman, Gloaguen, Andreani, & Rahnama, 2018)

3.2.3 Elevation

Elevation as anan indicator has also been linked to landslide susceptibility, with the most landslides occurring at intermediate elevations, as these elevations are normally characterised by steep slopes since they lie in the higher to lower elevation transition zone, hence these slopes tend to be covered by thin layers of colluvium which is prone to failure. Landslides commonly occur at very high elevations differences in shear strength. Lower elevations, on the other hand, are not commonly considered landslide-prone areas unless there are flooding or general water table changes (Dragićevića et al., 2015). Changes in local climate conditions at different elevations also affect slope stability leading to failures (Othman et al., 2018)

3.2.4 Curvature

The natural shape of a landscape can be described by the change of slope angles or curvature; this is a second derivative of elevation (or the slope of the slope). Curvature plays an important role on the erosional and run-off processes which influence the land surface (R. J. Pike et al., 2009). Overall curvature can be divided into, plan or profile and negative and then sub-divided into negative and positive (Figure 3.1).



Source: http://www.et-st.com/et_surface/userguide/Raster/ETG_RasterCurvature.htm

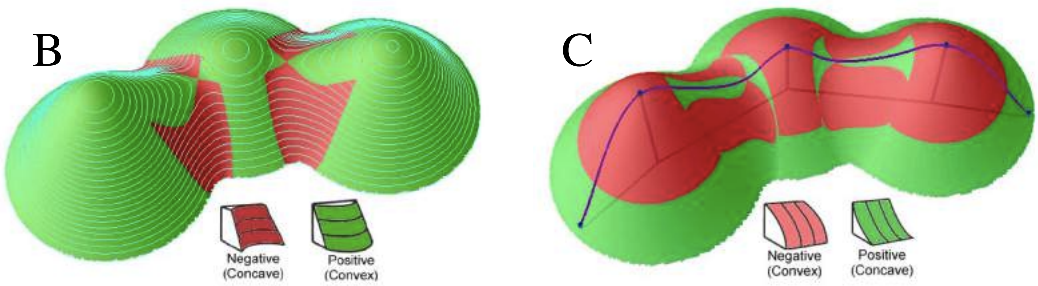


Figure 3.1 General (a), profile (b) and plan (c) slope curvature, with positive and negative areas highlighted.

(Source: http://www.et-st.com/et_surface/userguide/Raster/ETG_RasterCurvature.htm)

3.2.5 Drainage Density

Drainage systems have an adverse effect on slope instability in the form of surface water runoff and it's density and intensity (Yalcin & Bulut, 2007). Runoff patterns affect the undercutting and general erosion of slopes and areas with poor runoff lead to over-saturation. Dou et al. (2015) noted that in "increasing density of the drainage network causes increasing occurrences of landslide frequencies".

Drainage Density (DD) can be defined as the "total stream length per unit area"(Horton, 1932). First used by Horton (1932), it has since been used in many hydrological studies. The pattern and configuration of stream channels denote the efficiency of the drainage system. The drainage density itself is the result of interacting factors controlling surface runoff, however, on the other hand, it to influences the runoff and sediment/water output from a system (Gregory & Walling, 1968). DD has been known to be influenced by climate, vegetation, soil and rock types and relief (Moglen, Eltahir, & Bras, 1998).

3.2.6 Distance to faults

Distance to faults (DTF) is a commonly used variable in susceptibility mapping (Costanzo et al., 2012; Fausto Guzzetti et al., 1999). This zone around the fault refers to the area of influence of the fault, surrounding which there can be altered rock mechanics and hydrological properties. Reichenbach et al. (2018) discusses how these zones are more conditioned for landslide occurrence and how the distance of influence of each fault zone is variable. These fractured zones tend to be more susceptible to failure when triggered by earthquakes, however, the response is completely variable between different zones. Water runoff and permeation can also be affected greatly by fractured ground.

3.3 New Parameters in LSS

As technology advances, the processing power of the systems behind predictive models opens up new possibilities in the realm of DEM-derived parameters for landslide modelling.
Potential exists for ANN to explore and define new relationships in a non-linear way by including parameters which could possibly have an effect on the modelling process. Outlined below are selected parameters which have been determined as most applicable for the study out of a variety of available parameters.

3.3.1 Stream power index

Stream power index (SPI) is a compound topographic attribute which has been previously used for landslide studies (Costanzo et al., 2012; Dou et al., 2015; Kalantar et al., 2018; Roşca et al., 2015; Yilmaz, 2009). The index describes the erosive power of surface flowing water and has been traditionally used for the study of erosion, sediment transport, and geomorphology. It is based on the assumption that surface runoff is proportional to the upslope contributing area (Moore, Grayson, & Ladson, 1991). It predicts net erosion in profile and convexity or high flow acceleration, convergence zones, and net deposition in profile concavity areas which display decreasing velocity (Pourghasemi, Pradhan, Gokceoglu, & Moezzi, 2013).

3.3.2 Topographic Wetness Index

Also known as the compound topographic index (CTI), the topographic wetness index (TWI) is used widely in conjunction with SPI in the study of surface runoff and the development of ephemeral gullies. TWI was developed by Beven and Kirkby (1979) to study special scale effects on hydrological processes. The index represents the tendency of water to accumulate at any point of the catchment, and the natural process of gravity moving water downslope (Poudyal, Chang, Oh, & Lee, 2010). TWI can be considered an indicator of soil moisture spatial patterns (Dragićevića et al., 2015).

TWI has been used various landslide susceptibility models (Costanzo et al., 2012; Dou et al., 2015; Wilson, 2012). Yilmaz (2009) spoke of their findings, high TWI values distributed in higher elevations alludes to the infiltration of surface water into slope-forming materials, and decreasing shear strength occurs alongside an increase in pore pressure. Overall it was found that landslides were less common at higher elevations with high TWI values. Moore et al. (1991) concluded that the thresholds which give meaning to the TWI and SPI index values will vary from area to area.

3.3.3 Roughness

Topographic roughness can be broadly defined as the variability or irregularity of the terrain; however, the definition varies depending on the calculation used. Scale is an important factor in any roughness calculation, for example, is the surface roughness characterising a localised or regional scale landscape? Surface also varies depending on the landscape, urban landscapes have a different set of surface roughness calculations (Jhaldiyal, Gupta, Gupta, Reddy, & Kumar, 2016). According to Pawley, Hartman, and Chao (2017), topographic roughness is useful in the characterisation of landslide morphology. Roughness measures have been used for various landslide susceptibility studies in the literature (Alkhasawneh, Ngah, Isa, & Al-batch, 2013; Costanzo et al., 2012). Mumipour and Nejad (2011) used roughness alongside basin analysis to interpolate the tectonics of the Zagros Mountains region in Iran. Gosh (2015)

discusses earlier uses and definitions and suggests that roughness is slightly more advanced than slope and dissection index in that it gives an overall view of the evolutionary rhythmic process acting on the landscape. It is found that DEM grid spacing greatly affects topographic indexes (Mukherjee, Mukherjee, Garg, Bhardwaj, & Raju, 2013). The whole concept of surface roughness and what is considered rough or not can be ambiguous and there are many variations of the roughness measure exist in the literature and have been summarised[8] and also reviewed and analysed against each other[9]. Many studies have used the basic "standard deviation of the elevation" as the roughness measure (C. T. Lee, Huang, Lee, Pan, & Lin, 2008). For this study, two more advanced parameters have been selected which claim to consider more factors and used at multiple scales, terrain ruggedness index (TRI) and vector ruggedness measure (VRM).
Topographic Ruggedness Index (TRI): was proposed by Riley, DeGloria, and Elliot (1999) who states that it provides an "objective quantitative measure of topographic heterogeneity and that the algorithm calculation can be used at any scale for the purposes of a study. The higher the value is the more rugged the terrain is. TRI may be influenced by rock and soil characteristics and thus is frequently used to model landslide distribution (Conoscenti, Rotigliano, Cama, & Lombardo, 2016). It was originally developed for the use in habitat studies but has also been used in LSS models.
Vector Ruggedness Measure (VRM): was originally proposed by Hobson (1972) and further developed by Sappington, Longshore, and Thompson (2007) for the purpose of habitat studies. It has also been adapted for landslides studies in recent times, however, these studies focus more on landslide detection than susceptibility mapping. Pawley et al. (2017) previously used VRM to measure the slope and aspect variation together. The measure takes into account both aspect and slope and attempts to solve a problem with ruggedness measures. Many other measures focus on slope for the calculation, but just because a slope is steep does not necessarily mean it is rough. This method calculates vector dispersion and is less related to the slope. Sappington et al. (2007) found that locally that VRM quantifies the ruggedness independently of slope than the other measures tested, one of which is the aforementioned TRI.

[8] http://gis4geomorphology.com/roughness-topographic-position/

[9] http://www.let-group.com/lecture/l4061ar-dem-based-terrain-roughness-analysis-for-landsl4061e-characterization-4061.html

3.3.4 Dissection Index (DI)

Dissection Index can be defined as the ratio between relative relief and absolute relief. It can also be considered a roughness of sort but in regards to roughness caused by river incision and the extent of this incision. DI is a useful indicator in the study of landscape morphology evolution. Dissection can refer to both basin and overall landscape morphology and gives insight into the age and processes of a system. Lithology, relief, slope and drainage density play an impact on the overall dissection (Deolia & Pande, 2014). In the literature, both Gosh (2015) and Pandey, Sharma, and Bandooni (2018) included dissection index in the LSS modelling process and the latter study concluded that drainage density and dissection index were important in that particular study.

3.3.5 Hypsometric Integral (HI)

Is a derivative of the hypsometric curve which is used to describe the maturity and evolution of a basin and eludes to the "cycle of erosion" (Strahler, 1957). Both R.J. Pike and Wilson (1971) and (Strahler, 1957) discuss the similarities of HI to a parameter called elevation-relief ratio. HI is closely related to the degree of dissection by a drainage network and hence can be used to discriminate between landscape types. The HI is considered the area below the hypsometric curve (Figure 3.2), hence it corresponds to the shape of the curve. This is based on the thought that "a mountain is rapidly uplifted without serious denudation and then increases in dissection with a lowering in mean elevation (Davisian scheme)" (Pérez-Peña, Azañón, Booth-Rea, Azor, & Delgado, 2009).

A value greater than 0.6 indicated an elevated landscape with a significantly entrenched network. HI values from 0.35 to 0.6 correspond to noticeably eroded areas with well developed V-valley shaped valley systems. Below 0.35 is considered a relatively flat terrain with little incision (Strahler, 1957). Othman et al. (2018) used HI in the context of landslide susceptibility mapping and found HI to be a stronger indicator than curvature, improving AUC values by 2%.

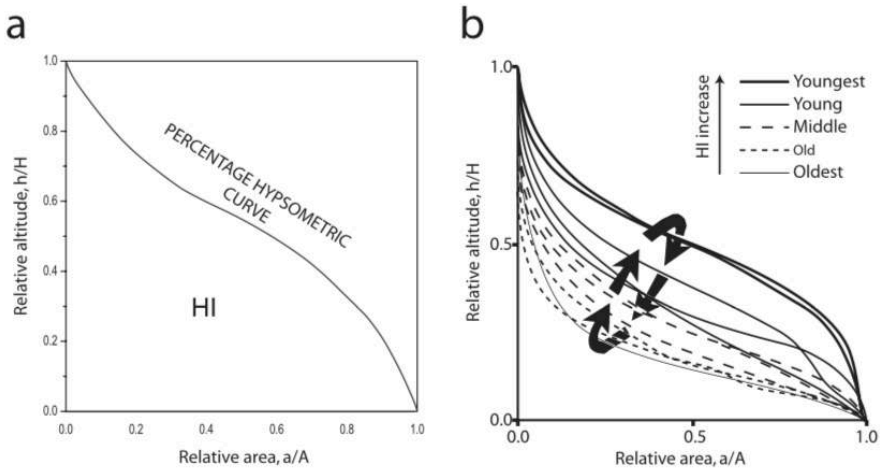


Figure 3.2 (a) Hypsometric curve from Strahler (1952). The HI is the area of the region under the curve. (b) hypsometric curves demonstrating the evolution of basin shape, from (Ohmori, 1993).

3.3.6 Knickpoints

A knickpoint is a location on a river where a notable inflection or sharp change occurs in the natural channel slope profile. Knickpoints are believed to form as a result of river incision into a bedrock that has experienced uplift due to various influences. Multiple factors may cause this including tectonics (active and dormant fault-lines), lithology and mass movement debris altering the channel profile. Human interference in the natural drainage pattern can also change the slope of a channel, resulting in knickpoints. Examples of this are dams, re-routing and funneling of channels. Attempts have been made to calculate and link these points to tectonic process (Hayakawa & Oguchi, 2006; Kirby & Whipple, 2001; Kirby & Whipple, 2012; Lopes Queiroz, Salamuni, & Do Nascimento, 2015; Zahra, Paudel, Hayakawa, & Oguchi, 2017).

Discovering knickpoints can be done in a variety of ways such as analysing stream cross-section profile using Hack (1973)'s stream length-gradient index (SLI) or Etchebehere, Saad, Perinotto, and Fulfaro (2004)'s SLI derivative, Relation Declivity Extension (RDE). SLI is the ratio of slope to length. RDE gives an idea of the current energy in a particular drainage segment and varies with the slope and discharge (Lopes Queiroz et al., 2015).

Troiani, Galve, Piacentini, Della Seta, and Guerrero (2014) used ordinary kriging interpolation on SLI values and were able to make correlations and interpretations from the data generated to benefit the study. Moussi, Rebaï, Chaieb, and Saâdi (2018) used RDE to detect river channel anomalies for neo-tectonic studies and found that the derived knickpoints were correlated well to large fault lines. Andreani, Stanek, Gloaguen, Krentz, and Domínguez-González (2014) applied knickpoint extraction to the Erzgebirge Mountain area using the TecDEM and reported overall positive correlations, this covers some of the area from this research but on a much smaller scale study related to the tectonic interpretation.

4. Study Area

4.1 Location

The study area is located in the East German state of Saxony, covering an area of 531km^2 surrounding the urban centre of Zwickau (Figure 4.1). The district of Zwickau constitutes the majority of the overall land coverage. In the South West, the study falls partially within the district of Vogtland. The area corresponds to four 1:25,000 topographic map quadrangles focused on the urban centre of Zwickau. The city is located at the foothill of the Western Erzgebirge Mountains (Ore Mountains). It has developed along the widened valley of the Zwickauer Mulde river. Once a prominent coal mining area, there is now a strong focus on agriculture with the landscape dominated by pasture and arable land. Due to geological and geomorphological conditions, the area is prone to flooding and earthquakes.

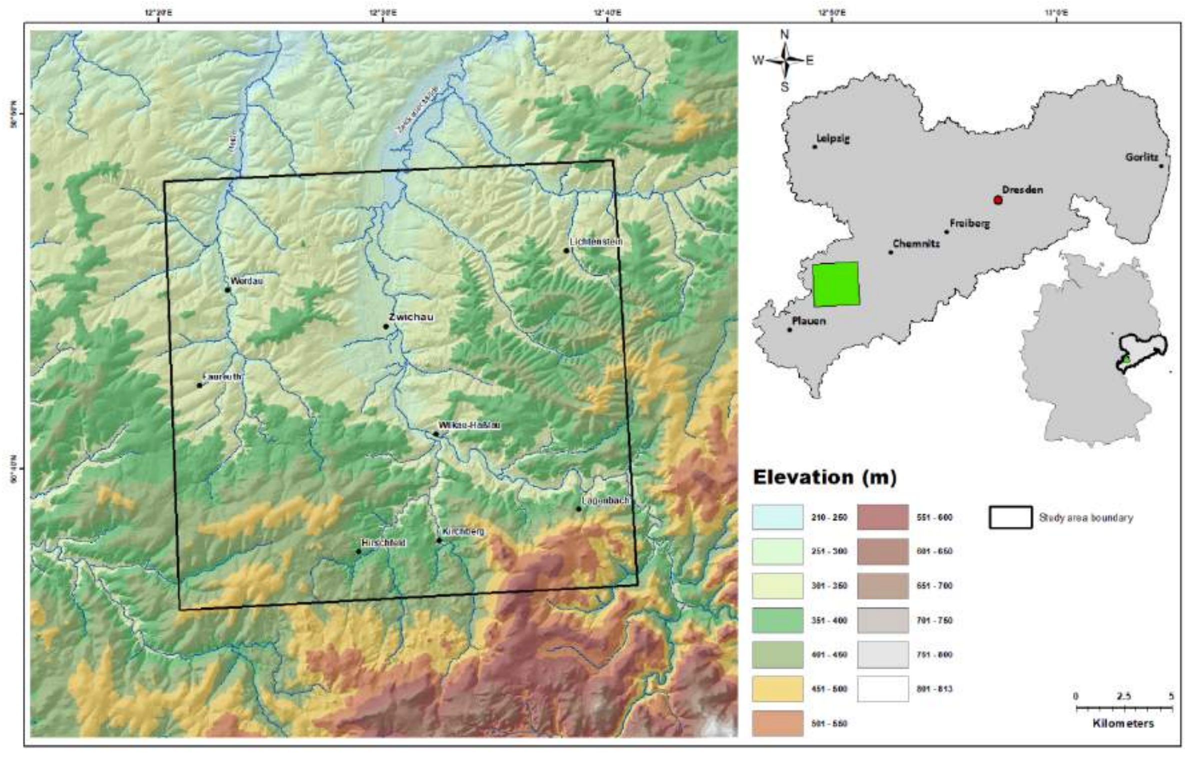


Figure 4.1 Location of study area

4.2 Climate

Lower altitudes of the study area are classified as warm and temperature whereas the higher sections in the SE are classified as cold and temperature. The temperature in the southern lowlands averages 8.2°C, varying between -1.3°C in January to 17.3°C in July. Precipitation here ranges from 32- 70mm during February and July respectively. Annual average precipitation is 573mm[10]. In comparison temperatures in higher areas in the south and south east, annually average 7.2°C, varying between 16.3°C in July to -2.2°C in January. These are known for higher snowfall rates. Precipitation averages 639mm annually ranging from 37mm in February to 78mm in July[11].

The state of Saxony has suffered from severe storm events and intense long lasting rainfall in the past, most notably the rain events of June 2013, August 2002 and July 1954 (Krauter et al., 2012). These events are defined by specific low-pressure conditions caused by the interaction between the continental and oceanic climatic systems which influence the area (Horlacher et al., 2007). Events like these can trigger landslides in vulnerable areas prone to failure due to oversaturation of the soil, hydrological pressure, and slope destabilisation.

4.3 Geomorphology and Geology

The geomorphology of the study area is characterised by the higher relief of the western Ore Mountains (Erzgebirge) in the south and by the lowland hills and valleys of the Erzgebirge foreland in the north. The elevation ranges from 210-813m respectively (Figure 4.2).

The urban area of Zwickau has developed along the flatter valley floor of the Zwickauer Mulde river (ZM). Here the river departs the higher energy environment of the upper southern elevations, upon reaching flatter relief the river slows down and widens. Frequent flooding events have widened the valley and a significant floodplain has developed.

The geomorphology of the study area is closely related to the underlying geology, hence any discussion about geomorphology must be done in reference to the geology. The higher elevations in the south are dominated by the more durable Kirschberg Plutonic Granite which is less susceptible to erosion. This Variscan age granite body was intruded into the surrounding

---

[10] https://de.climate-data.org/location/22790/
[11] https://de.climate-data.org/location/23109/

sedimentary rock creating a progressive contact zone of metamorphosed to semi-metamorphosed meta-sediments of Ordovician age, namely phyllite, shale, and mica-schist. Faulted blocks of Devonian, Silurian and Carboniferous age sediments mark the boundary between the North-East edge of the Erzgebirge with the rolling lowland hills of the Erzgebirge Basin. The flatter landscape of the Erzgebirge foreland is geologically composed of late Palaeozoic age reddish sandstones and, conglomerates from the Permian and Upper Carboniferous. (LfULG, 1875-1900). During the Carboniferous period, forested bogs developed in the basin during interruptions in sediment deposition, forming the coal deposits which were historically exploited in the region. Evidence of past coal extraction can be seen in the numerous mining dumps and a tailings pond located in and around the urban landscape of the city (Schneider et al., 2005).

The slopes around Zwickau's urban centre are characterised by sedimentary rocks of Permian age (Rotliegendes) conglomerates, siltstones and sandstones, and a Tertiary age capping of sands and gravels. These sediments are juxtaposed by numerous faults SE-NW trending faults. Incision patterns take advantage of these natural avenues of least resistance, which ultimately plays an important role in the overall geomorphological presentation of the area. The Quaternary deposits in the area are characterised by river sediments of various composition (Syrbe et al., 2014). These loose sediments line the rivers and valleys of the Zwickau area.

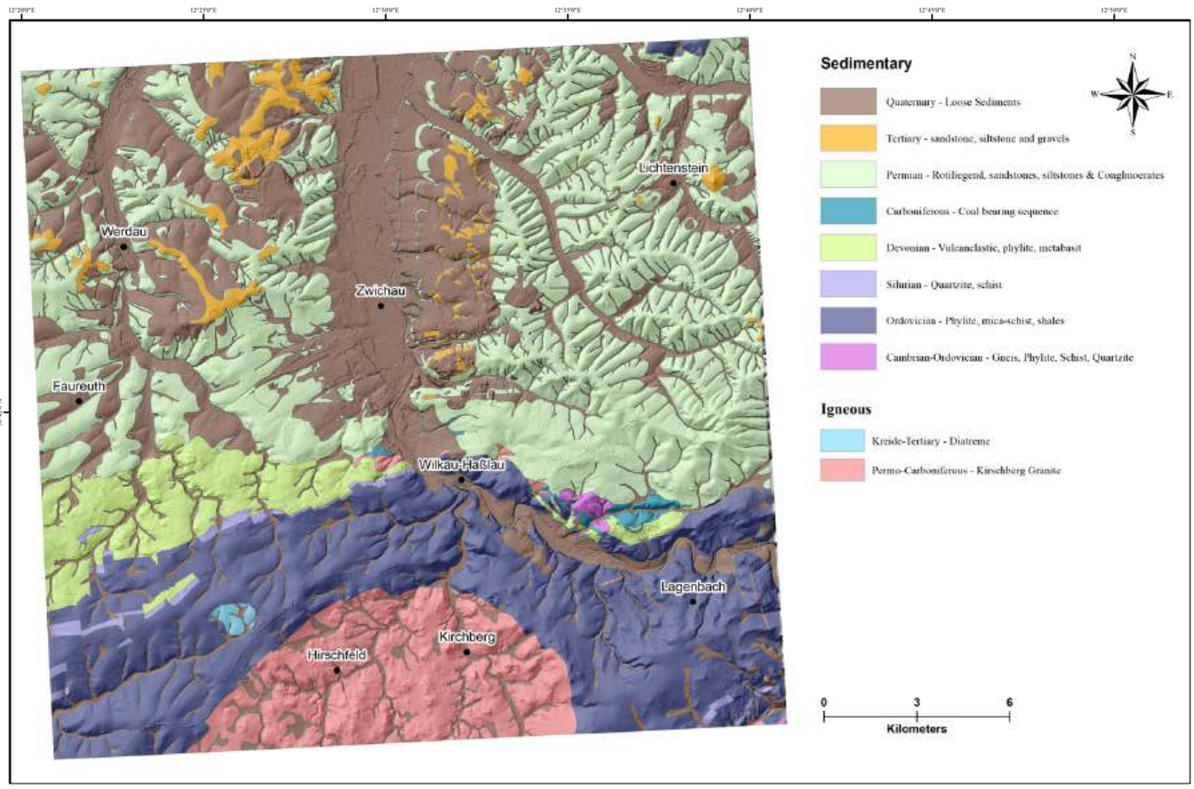


Figure 4.2 Geological Units of the study area, by age

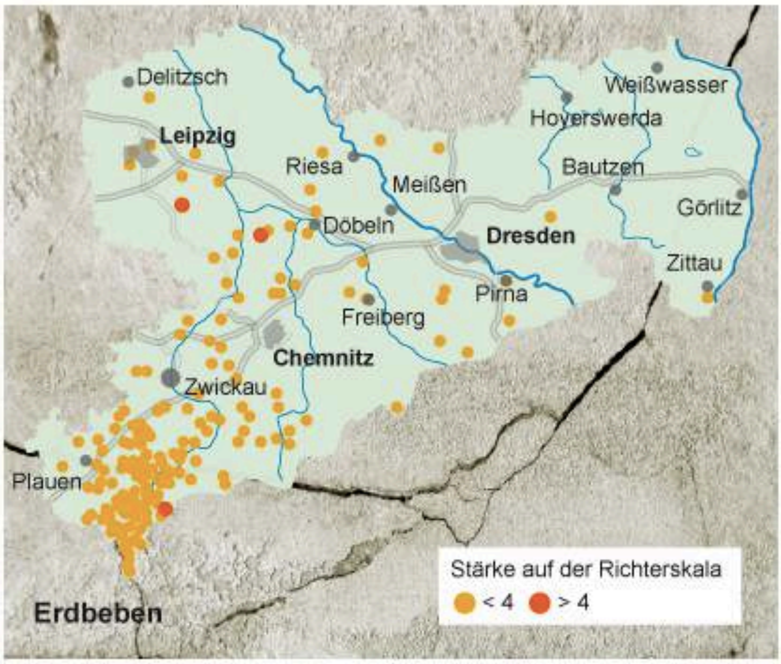


Figure 4.3 Earthquake occurrence in the State of Saxony (http://www.naturgefahren.sachsen.de/erdbeben-erdrutsch.htm)

Tectonically, the area is heavily faulted, the most notable feature being the Gera-Jáchymov fault zone, trending from SE- NW. Significant earthquake swarm events have occurred in the region due to the tectonic interplay along the western extent of the Erzgebirge Mountains and to a lesser extent induced by previous mining activities (Korn, Funke, & Wendt, 2008). (See Figure 4.3)

Evidence of past localised block tilting can be seen in the asymmetrical pattern along river sources.

The drainage basins of the Zwickauer Mulde (ZM) and the Weiße Elster crosscut the foreland basin in a N-S direction (Figure 4.4). The ZM and it's tributaries contribute most to the geomorphology of the study area in comparison to the Pleiße River, a tributary of the Weiße Elster, which plays only a minor role to the west of Zwickau. The ZM valley widens significantly as it leaves the higher elevations of the Erzgebirge mountains and enters the foreland landscape of the Erzgebirge basin. Upriver it is orientated along the Gera-Jáchymov fault zone however upon reaching the foreland basin it re-orientates to the north. The overall drainage network creates a varied pattern of incised river valleys and steep valley slopes in the southern mountainous region, which upstream concentrates the flow into the ZM valley and its floodplain. This rapid runoff from higher elevations fosters conditions which attribute to the intense flooding events that have been experienced in the region.

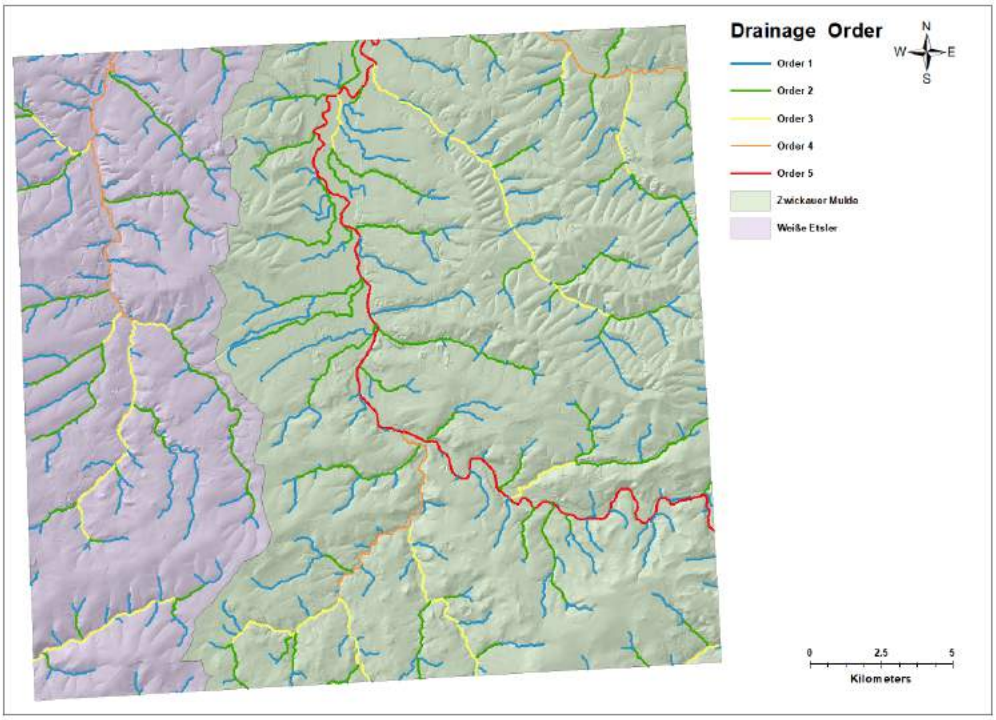


Figure 4.4 Drainage basins and network by Strahler order

4.4 Soil

Three main soil type zones are present in the area. The rolling landscape of the north consists predominantly of silty brown earth Loess of peri-glacial origin, that are derived from mostly sandstone, conglomerates clay or shale, which are interposed with clays, brown pseudo-Gleyes and minor brown Podzols (Figure 4.5) (LfULG, 2011). In flat areas alongside and between rivers, waterlogged pseudo-Gleyes and alluvial clays/silts have developed (LfULG, 2018b).
Soils in urban settlements are for the most part classified under the title of anthropogenic origin. These include Hortisoil, mixed fill construction material and various dumps of Regosol/Pseudo Gleye mixed with waste associated with coal and ore mining activities. The southern zone represents soils with a high proportion of acidic to intermediate magmatites and metamorphites. These zones represent the transition in the underlying geology, from the sedimentary rocks of the Erzgebirge foreland into the granite and metamorphic rocks of the Erzgebirge mountains. In the incised valleys of the south, Vega Gleyes, silts and sands have

developed over fluvial gravels. Podzols and peat soils are more common on flat terrain of the higher elevations in the south.

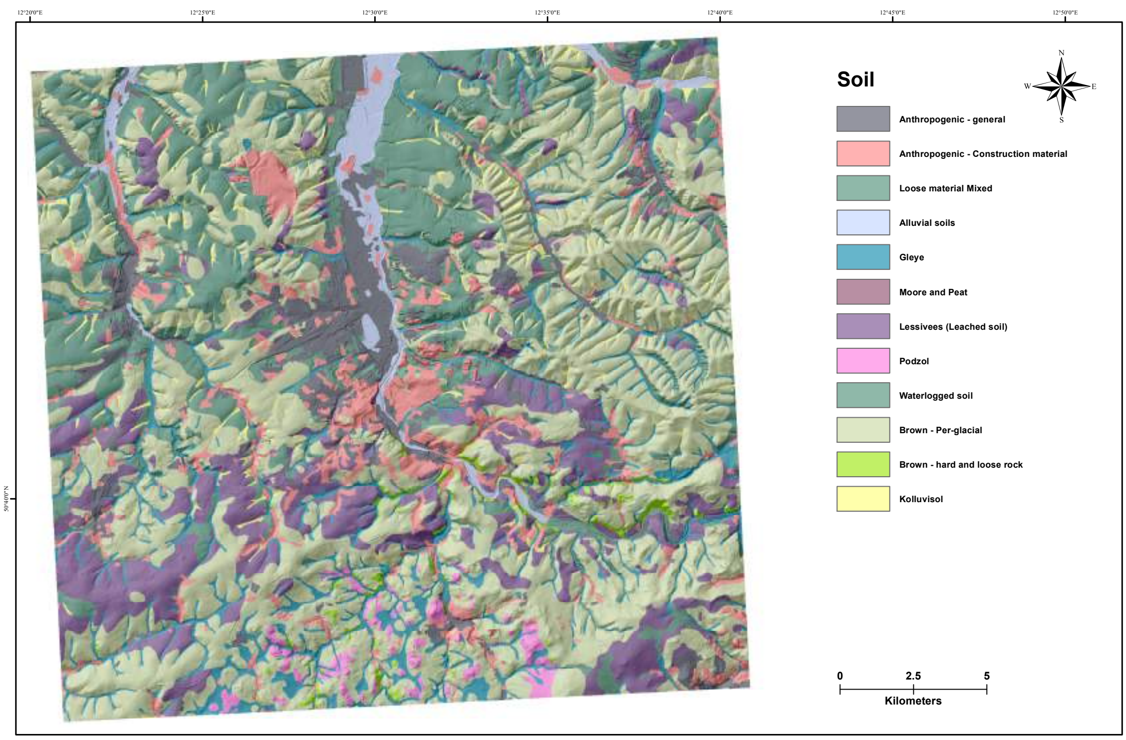


Figure 4.5 Soil map of the area (LfULG, 1875-1900).

4.5 Landslide occurrence

In comparison to other areas of Saxony, the study area is not very susceptible to landslides (see Figure 4.6). Landslides in this area can be classified as small failure shallow failures according to Varne's classification in Chapter Two (Varnes, 1978). Interaction and modification of the landscape by humans and reoccurring flood events can be identified as the main cause of landslides. From a geological perspective, it can be assumed that destabilisation of vulnerable slopes due to earthquakes may also occur however, there is no evidence of this in the area at present.

Though all the landslides have not been recorded, during the fieldwork many locations were identified as potentially susceptible. In most instances, these are marked along deeply incised river valleys, steep slopes, along floodplains or human interaction. All but two locations were recorded as small-scale shallow translational landslides. The remaining two locations are classified as Rockfalls according to Varne's classification. The Rockfalls recorded occur at locations which still exhibit conditions similar to shallow landslides, so for the purpose of this study, they were included for modelling.

It was noted that many steep valleys have been previously vegetated with dense forestry as part of good land management practices to prevent soil erosion, water funneling during heavy rain events and landslides. In many steep sections, evidence of preventative engineering measures were visible. These good practices may be the reason for low landslide occurrence in the area, but that is not to say that in the future under changing climatic and land use scenarios these measures will be sufficient to stop potential failures.

Figure 4.6 Landslide occurrence in the State of Saxony

(Source: http://www.naturgefahren.sachsen.de/erdbeben-erdrutsch.htm).

5. Methodology

5.1 Data Sources

For this study, all data was projected into ETRS_UTM 32N with a D_ETRS_1989 datum. All sources are provided originally in German (see
Table 5.1). The following sources were all pre-processed in preparation for modelling, however, on further evaluation land-use was not included in the model.

5.1.1 DEM

The Digital Elevation Model “**D**igitales **G**elände**m**odell DGM10 – Gitterweite 10 m“ which was used to derive the majority of the parameters for this study, was supplied by the German Federal Agency for Cartography and Geodesy (GeoBasis-De, BKG), in line with the ATKIS project (Authoritative Topographic Cartographic Information System) (AdV, 2015b). The DGM10 version being used dates from 2017 and has a locational resolution of 10m and height resolution of 0.01m. The accuracy of the DEM is terrain type dependent and varies from 0.5-2m for both location and height. The production involved using height data taken directly by the Land Survey Administration using various methods including, laser scanning, photogrammetry, and contour line digitisation. The digital terrain model is provided in the ETRS89_UTM position reference system and the DHHN2016 height reference system (AdV, 2015a).

5.1.2 Geological Map

A 1:25,000 scale Geological map “**G**eologische Karte des Freistaates Sachsen GK25" was employed for this study, supplied by the Saxon State Office for Environment, Agriculture and Geology (LfULG). Detailed descriptions of the rock types and lithological groups and ages were provided with the map. The study area is based on four sheets digitised from this map, 5240 Zwickau (2008), 5241 Zwickau East (2008), 5340 Planitz – Ebersbrunn(1884), 5341 Kirchberg-Wildenfels (1900), which range from 50°36’ to 50°48’N and 12°20’ to 12°40’E.

5.1.3 Soil Map

A 1:50,000 map "Bodenkarte des Freistaates Sachsen BK50" was the source for soil data. This was also provided by the Saxon State Office for Environment, Agriculture, and Geology (LfULG). Detailed descriptions of the soil zones, type and substrate are provided. The map is based on evaluations from existing databases, incorporating current data collected especially for the project. The soil map for Saxony is still being updated since extensive mapping finished in 2011. A free online digital version of the soil map exists since 2012 (LfULG, 2018a). All other map derived data for this study is on a 1:25,000 scale, with the soil map being the exception at 1: 50,000, however, the detail is of a high quality for the study.

5.1.4 Land-cover map

Land cover data was derived from the 2016 "Digitales Landschaftsmodell Basis-DLM (AAA)" Landscape model was supplied by the German BKG. The datasets used are from the ATKIS Basis-DLM of the German Federal state. The model is a digital, object-structured vector dataset that is updated continuously (AdV, 2016). Land cover was given in percentage of vegetation, with 100% corresponding to forested areas and 0% to urban areas.

Classification	Sub-Classification	GIS Data Type	Scale	Source
Inventory	Landslide	Point	Saxony wide	LfULG
Surface Model	DEM	Grid	10m	AdV
	Geological	Polygon	1:25,000	LfULG
Map	Land-use	Grid	1:25,000	AdV
	Soil	Grid	1:50,000	LfULG

Table 5.1 Source data and information

5.1.5 Landslide Inventory and Field Work

It can be said that past and present landslide locations are key in the prediction and prevention of future events (Dou et al., 2015). Logically from this, the compilation of a landslide inventory is the first step in undertaking Landslide Susceptibility (LSS) Modelling. The current inventory was obtained from the Saxon State Office for Environment, Agriculture, and Geology. Only four event points exist within the extent of the study area. Out of these two were not positioned incorrectly in regards to the original landslide location. One was situated directly on train tracks and the other was placed adjacent to the failure area. Not much information is supplied about the landslides from the original inventory. Such few data points are insufficient for modelling purposes. Hence fieldwork was undertaken, the purpose of which was to collect more data points for the modelling process and validate results from a previous lower resolution model. Twenty-nine locations in total were recorded during the course of the field work, data and photographs for each location were collected for each data point. Overall from the fieldwork and previous inventory thirty-three points in total were available for modelling.

5.2 Software

The data extracted from the DEM and various source data was digitised/extracted and pre-processed using ArcGIS 10.2. The Landslide susceptibility modelling was carried out using an Artificial Neural Network prediction software developed by Beak Consultants Gmbh. Other packages such as SAGA 5.0 and MatLab R2018a, were used during the background review to assess various package methodologies for deriving surface parameters from DEM's.

5.2.1 ArcGIS 10.2

Is an ESRI software, marketed as a complete Geographic Information System (GIS) with strong mapping and analytics capabilities (Esri, 2013). ArcGIS is composed of toolboxes grouped by function type, including Spatial Analyst, 3D Analyst, Data Management tools. External toolbox packages developed for ArcGIS are also available, of these CalHypso, Vector Topographic Roughness (VRM), Basin Asymmetry, and TopoToolbox were reviewed as part of the study.

5.2.2 Advangeo© Prediction Software.

Advangeo© offers software solutions from data capture to prediction developed by Beak Consultants Gmbh. The package includes an Artificial Neural Network (ANN) which can be used for the prediction of spatial events and phenomena such as geo-hazard susceptibility modelling. It is fully integrated within the ArcGIS platform and as such GIS layers can be accessed directly from ArcGIS. The ANN is a Multi-layered Perceptron (MLP) and with a Back Propagation learning algorithm.

5.3 Pre-processing of the source data

Data sources in their original state required pre-processing to make them GIS and modelling compatible for the purposes of the research. Some tools require prior processing to ready the surface for parameter calculation.

The geological data was first digitised from the original map sheets into polygon format within ArcGIS. These sheets were then mosaicked together to join the separate layers and subsequently re-projected. The sediment, hard rock and Geological lineaments (fault lines) maps were supplied separately. Since all of the landslides recorded were shallow in nature, the decision was made to join the sediment and hard rock maps into one layer to represent more accurately the geology of the shallow subsurface.

The geological map units were then divided into different classes according to rock type (GK), whether the rock is liable to faulting (KL) and by the structure of the rock (TF). The rock types were classed as hard rock, loose sediments or anthropogenic. The loose sediments were then further sub-dived into fine, mixed or coarse grain (see

Table 5.2). The classifications for each category were based on a geological assessment of individual properties for each rock type in regards to strength, porosity, grain size and tendency to deform. Only the susceptible (2) categories from KL and TF parameters were used.

Around the fault lines from the geological lineaments map, a linearly graduated buffer of 80 m was created. These buffered polygons represent the extent of influence of the fault zone. For this study, it was decided that 80m was a sufficient representation for the fault zones in the study area for the LSS model.

GK - Rock Class		**KL - Fault Susceptibility**		**TF - Plane susceptibility (Cleavage, jointing, foliation)**	
Class	Description	Class	Description	Class	Description
1	Hard Rock	2	Susceptible	2	Susceptible
2	Loose Rock - fine grain	0	unknown	0	unknown
3	Loose Rock - mixed grain	1	not susceptible	1	not susceptible
4	Loose Rock - coarse grain				
5	Anthropogenic				
6	Loose to hard rock - fine grain				
7	Loose to hard rock - mixed grain				
8	Loose to hard rock - coarse grain				

Table 5.2 Geological Data classes

The land-use and soil maps provided by the relevant German authorities were already in raster format and as such, minimal processing was required. The datasets were mosaicked and re-projected. The resultant soil raster was broken down into soil types applicable to the study (Table 5.3). Not all soil types correspond to the training points and hence were not relevant to the modelling process.

Class	Soil Type
1	Anthropogenic
2	Loose material - Mixed
3	Alluvial soils
4	Gleye: wetland soils
5	Moore & Peat Soils
9	Water-logged soil
10	Brown peri glacial earth
12	Brown - mixed rock
14	Fluvial hummus soil

Table 5.3 Final list of soil types used in the modelling process

The nationwide Geo-basis DEM from the Adv was supplied in raster tiles which were mosaicked and re-projected. All GIS parameters were initially calculated from a DEM larger than the study area and clipped to the study boundary polygon during the ArcGIS to Advangeo© import phase. The calculation of the various parameters from the DEM utilized kernels of varying size, so by using a larger DEM surface continuity of the data around the study boundary could be maintained.

5.4 Data Creation and tool development

The 10m resolution DEM acts as the base for all extracted parameters used in the LS modelling process. After careful review of the literature, parameters were selected for development and assessment based on suitability for the study purpose and scale. These parameters traditionally come from predominantly tectonic and soil/hydrology backgrounds.

All extraction and processing of the data was done in ArcGIS using Model Builder and various ArcGIS compatible external scripts and toolboxes. ArcPy was used in conjunction with model builder to construct custom toolboxes for extracting parametric data. A combination of different methodologies from previous literature were used in the creation of these custom toolboxes. Various neighbourhood sizes were tested and final sizes were selected based on the suitability and resolution of each parameter. Upon generation of the parameter data, the data output was checked for any outliers and/or holes in the data and to ascertain whether the parameters were generated successfully.

5.4.1 Data from available tools

The GIS layers for slope, aspect, curvature, drainage density, and flow accumulation were created using functions already available as part of the Spatial Analyst toolbox in ArcGIS. These functions automatically extract the data from the DEM surface values. All these parameter surfaces were calculated on a cell-by-cell basis, as fitted through that cell and its eight surrounding neighbours.

The *slope* raster can be easily generated from the DEM using spatial analyst. It is the first deviation of the DEM raster. The values for the slope raster can be set to degrees or radians, and for the purpose of this study, degrees were used which gives the inclination angle of the slope in degrees. Values range from 0-90°, 0° being flat and 90° vertical.

In ArcGIS, *aspect* is calculated by the direction of maximum rate of change in value from each cell to its neighbours (Esri, 2001). Aspect raster values range between 0-360 degrees, 0° is true North and 180° South and so on. For this study raster values from 0-360° were reclassified into five divisions based on the orientation, 0.0001- 45 and 315-360 for North, 45-135 for East, 135 – 225 for South, and 225 – 315 for West. Flat areas are represented with the values -1 to -0.0001, referring to a lack of aspect.

The *curvature* function produces rasters for the overall curvature, plan and profile curvature. The plan and profile rasters were then further subdivided into positive and negative curvature, resulting in four curvature layers overall. The curvature is the second derivative of the DEM raster. A 0 value indicates a flat surface and this needs to be addressed in some calculations as it can cause issues in "raster calculator" from ArcGIS, this also applies to zero slope raster values.

Flow accumulation is calculated in ArcGIS using the D8 algorithm, which calculates the outflow number for each cell based on the surrounding 3x3 matrix (Esri, 2013). The resulting raster represents the accumulated flow to each cell, this is determined by the accumulated weight for all cells in the neighbourhood that run downslope into the cell. Zero flow locations are characterised as topographic highs and can be used to identify ridges. For this study, the flow accumulation raster was created as part of the drainage extraction process (Figure 5.1).

Drainage extraction can be done using inbuilt functions in the Spatial Analyst toolbox. First, any pits in the DEM are filled to maintain flow-lines. The flow direction is then calculated for each cell, followed by the flow accumulation. Depending on the size and processing capability of the system used the flow accumulation can take some time to calculate. Flow accumulation values are spread over a very large range and not all of these values are relevant to the main drainage pattern. Hence, the values were logarithmically scaled to better represent the data. To separate the drainage network itself from all surface flow, the raster calculator is used to assign values less than 5000 to zero so the areas of concentrated flow can be identified. From this, the raster stream segments of the created network are assigned order numbers according to the Strahler methodology. In ArcGIS 10.2 this is the standard methodology used for stream ordering. The stream to feature function is subsequently used to convert the raster drainage network to polylines. Another step that was undertaken for the knickpoint generation but not shown in Figure 5.1, was the conversion of the stream network to a 3D network, this was done using the 3D Analyst toolbox.

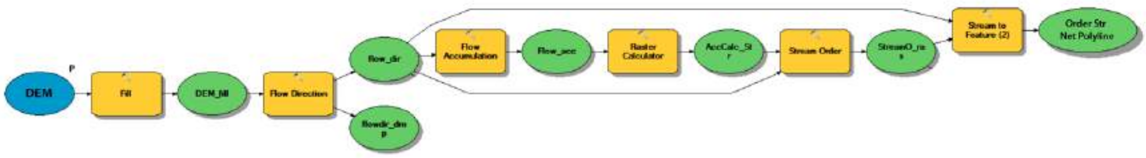


Figure 5.1 Drainage Extraction workflow in model builder

Drainage Density (DD) was calculated using the line density function from the spatial analyst toolbox in ArcGIS. The drainage network discussed previously is used as the function input. Line density calculates the density of a line features for each output raster cell. The calculation is based on the length of each linear segment within a circle of chosen radius around a pixel. The length is multiplied by the population field and the resulting figures summed, before being divided by the area of the circle. The measurement is given in units of length per unit area and sq km was set as the area unit. The drainage feature layer generated from the DEM is used as input. A circle radius of 100m and 500m was used to generate the drainage density for this study.

5.4.2 Data from developed tools

After an extensive review of the pre-existing literature, the following parameters were extracted based on formulas and methodologies described in the literature. Rectangular and circular neighbourhoods were both tested and rectangular neighbourhoods were chosen and used for all raster calculation.

5.4.2.1 Stream Power Index (SPI)

In GIS terminology, SPI is a function of the erosive power of runoff acting on each cell and is defined by Moore et al. (1991) as:

$$SPI = A_s tan(\beta)$$

A_s is the upslope contributing area or flow accumulation and β is the slope in degrees. When both A_s and β increase, so too does the amount of water supplied from the upslope contributing area and the velocity of the water, this results in an increase in SPI and the risk of slope erosion. The formula was translated into model builder to create the final raster used in the modelling process (Figure 5.2).

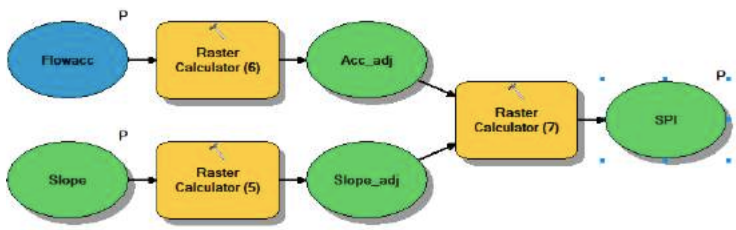


Figure 5.2 SPI workflow in model builder

The basic formula requires some manipulation to work in model builder. The slope in degrees cannot be used in raster calculator and must first be converted to radians. Various methodologies were tested but did not produce satisfactory results. After some trial and error, the methodology chosen was developed by Danielson (2013) as it produces a full and complete raster for interpretation. The final formula used in model builder requires the slope to be in percentage rise format. A variable of 0.001 was added to each raster to avoid zero calculation discrepancies. The flow accumulation was multiplied by the cell resolution and the natural log was used to stretch the values.

5.4.2.2 *Topographic Wetness Index (TWI)*

Topographic wetness index examines the relationship between slope and flow accumulation and as such is regarded as a compound parameter, as is SPI. TWI as defined by Beven and Kirkby (1979) as:

$$TWI = \ln(A_s/\tan\beta)$$

Where A_s is the cumulative upslope area draining through a cell (flow Accumulation) and β is the slope angle at the point. TWI values in high accumulation areas will be greater.

A model was constructed in model builder to calculate TWI (Figure 5.3), using slope and flow accumulation as input. The slope was first converted from degrees to radians, as the raster calculator function does not accept rasters in degree format. The tan of the slope was then calculated using the raster calculator, a conditional equation was used to calculate values greater than zero only, as to avoid any discrepancies. The final equation was constructed in raster calculator to create the TWI raster. For both SPI and TWI, they are classified as indexes and hence do not have any measurement value and are rather described as high, medium and low values.

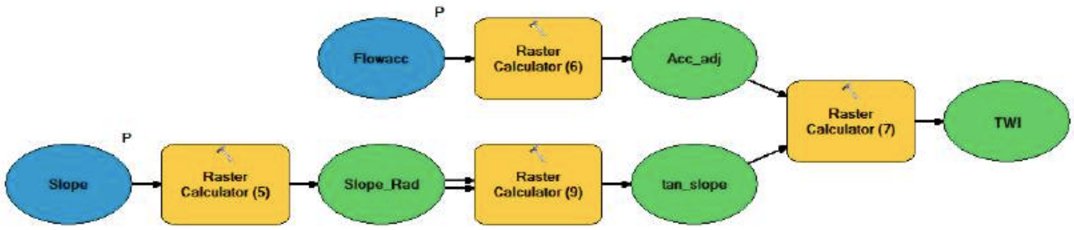


Figure 5.3 TWI workflow in model builder.

5.4.2.3 *Roughness*

As discussed in Chapter three, many methodologies exist for measuring roughness. For this study tools for both Riley's Terrain Ruggedness Index (TRI) and the more recent Vector Ruggedness Measure (VRM) were developed.

Terrain Roughness Index (TRI): calculates the sum change in elevation between a grid cell and its neighborhood, following the method developed by Riley et al. (1999). It was developed by Riley et al. (1999) to quantify the elevation difference between adjacent cells of a DEM. The equation for the calculation is:

$$TRI = \sqrt{(maxDEM)^2 - (minDEM)^2}$$

A model builder toolbox was created to calculate the parameter surface in ArcGIS based on the formula (Figure 5.4). It calculates the difference in elevation value from the center cell and its neighbors, which is dependent on the kernel size selected. The elevation differences are then squared to create positive averaged values before taking the square root of the average.
The created model calculates the TRI by first using the focal statistics function to create the maximum and minimum rasters. Following this, the raster calculator was used to square each raster and then get the difference of the two rasters using a subtraction. The result of this was squared using raster calculator and then the absolute values were calculated to avoid any discrepancies in the resultant raster. Finally, a mean focal statistics function using a 3x3 neighborhood was run to smooth out any outliers from the final resultant raster. Inputs for the model include the DEM and neighborhood sizes. Three output TRI rasters were created using different neighborhood kernel sizes, a 3x3, 5x5 and 9x9.

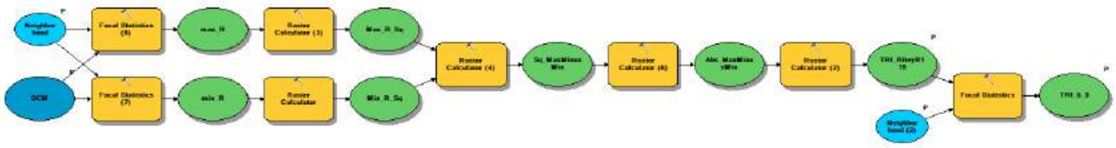


Figure 5.4 TRI workflow in model builder

Vector ruggedness measure (VRM): The raster for VRM was created using an available Arcpy script[12] developed by Sappington et al. (2007) for use in ArcGIS as part of their habitat study. VRM was first proposed by Hobson (1972) before being adapted by Sappington et al. (2007), who notes that it appears to decouple slope dependency from the ruggedness calculation better than other popular ruggedness indexes such as TRI. The calculation combines the analyses of both slope and aspect into a single measure. "Vector analysis is used to calculate the dispersion of vectors normal (orthogonal) to grid calls within the specified neighbourhood" (Sappington et al., 2007). The script calculated the VRM by first creating x,y and z rasters and calculating the sum of these for the selected neighbourhood size (rectangular) (Figure 5.5). It then calculates the resultant vector raster before finally creating the final VRM raster.

---

[12] https://www.arcgis.com/home/item.html?id=9e4210b3ee7b413bbb1f98fb9c5b22d4

The resulting values vary from 0 (no variation in the terrain) to 1 (full terrain variation). Sappington et al. (2007) states that typical values for natural terrains tend to fall between 0 and 0.4. Three resultant rasters were created using the aforementioned script in ArcGIS for using 15x15, 9x9 and 5x5 cell neighbourhoods.

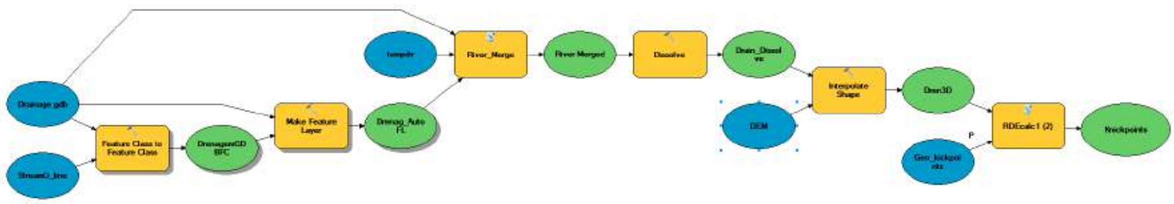


Figure 5.5 VRM workflow in model builder

5.4.2.4 Dissection Index (DI)

As calculated by Nir (1957) and described in chapter three, DI is the ratio between relative and absolute relief of a particular area. Values generally range from 0-1, however, Farhan, Anbar, Enaba, and Al-Shaikh (2015) noted that in some cases, these values can be exceeded.

$$DI = relative\ relief / absolute\ relief$$

A toolbox was created in model builder to calculate the index raster (Figure 5.6). The inputs for the model are the 10m DEM and the desired neighbourhood size. This involved using the focal statistic function in ArcGIS to first calculate the maximum and then the minimum rasters using a user designated neighbourhood, and deriving the relative relief from their difference. The minimum raster value was extracted from the DEM and used to calculate the absolute relief. The DI index was then calculated using raster calculator, as the ratio of relative and absolute relief. Any values larger than 1 are considered discrepancies and are set to null to create the final raster. Rasters were generated using 5x5, 9x9 and 15x15 cell neighbourhoods.

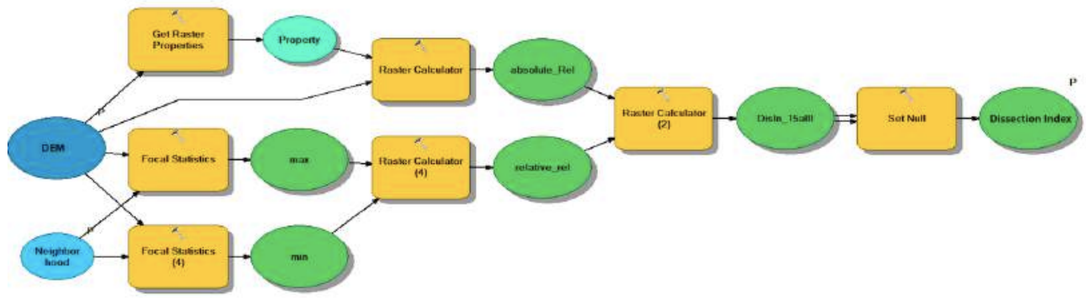


Figure 5.6 DI workflow in model builder

5.4.2.5 Hypsometric Integral (HI)

As discussed in chapter 3, Hypsometric Integral is an adaption of the Hypsometric curve which compares the incision of the surface by the relief. It is also known as elevation/relief ratio and Hypsometric Index. A model was built in ArcGIS to calculate the HI (Figure 5.7). Inputs for the model are the DEM and the neighbourhood size required. The minimum, mean and maximum rasters are calculated and from these, the maximum, minimum difference and the mean/minimum difference is calculated. The HI is then derived from the ratio of the two according to Strahler (1957) who found that it is inversely correlated with the steepness of a slope, the DD, the channel gradient and the total relief. It is expressed as a percentage and indication of the erosion and tectonic process of a basin. It is defined as:

$$HI = Elev_{mean} - Elev_{min}/Elev_{max} - Elev_{min}$$

A 9x9 kernel was implemented on the HI using focal statistics (mean), to smooth out outliers in the raster. HI rasters were created using 25x25, 15x15, 9x9 cell neighbourhoods.

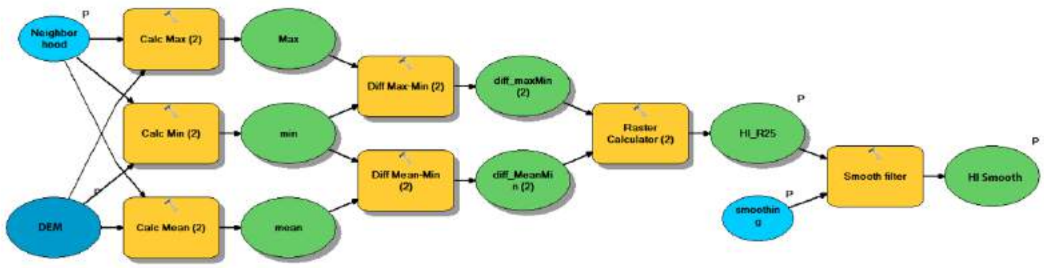


Figure 5.7 HI workflow in model builder

5.4.2.6 *Knickpoints*

The most common methodology for the calculation of knickpoints is by analysing the change in the run and rise variations along drainage networks, to look for anomalies in the stream profiles. This can be done manually through by inspecting longitudinal river profiles, however, this can be time-consuming. For this study, an automated knickpoint extraction workflow was developed by the adaption of a python script created by the Universidade Federal do Paraná in Brazil[13] (Lopes Queiroz et al., 2015). This script implements the RDE calculation proposed by Etchebehere et al. (2004); (Lopes Queiroz et al., 2015) and is calculated based on the relationship between RDEs (stretch index) and RDEt (total index).

$$RDEs = (\Delta H/\Delta L).L$$

The change in elevation (ΔH) and change in length (ΔL) refer to the difference in height and length between the extremities of a particular segment being examined. L indicates the distance between the lower point of the segment and the source of the river (Figure 5.8).

$$RDEt = (\Delta H/\Delta L).\ln(L)$$

The total RDE (RDEt) calculation is the similar to RDEs except it refers to the total length of a river and accounts for the slope ($\Delta H/\Delta L$) between the source and mouth points of the river and the natural logarithm of its entire length (Lopes Queiroz et al., 2015).

When the RDEs/RDEt rato is greater than two, the segment is considered to be anomalous. A value between two and ten is designated as a 2[nd] order anomaly and values greater than ten as a 1[st] order. The code first calculates the RDEt for the whole river and then calculates the RDEs for each segment. When a segment drop value (from the previous) exceeds the elevation equidistance entered by the user, the code then calculates the relationship between the RDEs and RDEt and assigns 1[st] order or 2[nd] order designation based on the resultant value.

---

[13] https://github.com/silverlq/KnickpointFinder/blob/master/README.md

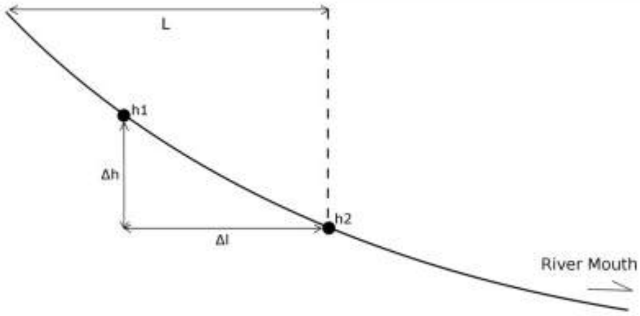


Figure 5.8 Visual representation of the RDE calculation (from Lopes Queiroz et al. (2015))

The code above was adapted for use in the study. The code was broken down and modularised in model builder to make it easier for the user to save various outputs from the process (Figure 5.9). Some elements of the pre-calculation of flow direction and flow accumulation were removed as this slows down the process considerably as is not required. Part of the code merges river segments based on their “to” and “from” nodes, these segments are then merged and interpolated to 3D using the DEM as input. There is no function in ArcGIS for merging rivers in this way. By modularising this part of the code, the user can run multiple models faster, as this step is only required once, further models with different user input can run the last script of the workflow to calculate the RDE relationships.

Three values for elevation equidistance were given, 20m, 50m, and 100m producing a set of point data representing the knickpoints. These points were then interpolated into a surface using the point density (hotspot) function in ArcGIS. To produce a smoother result, a large search area radius and lower resolution pixel format were chosen.

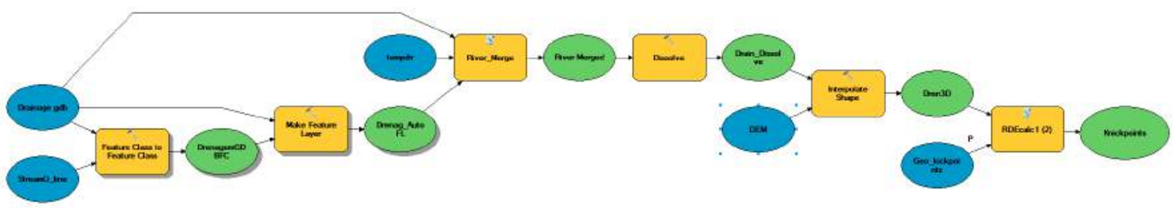


Figure 5.9 RDE workflow in model builder

5.4.3 Collection and processing of Inventory data

The basis of any good study is on the input data that it uses. This is especially true for LS inventories. They provide a window into the past and so that the conditions which lead to the failure can be analysed. As stated in section 5.1.5 of this chapter, the result of the fieldwork conducted produced thirty-three points in total, in the end only thirty-one of these points were used as one of the original inventory points was incorrectly located and another point collected during the fieldwork was deemed as not applicable due to it's close proximity to another point. During the fieldwork, careful attention was given to the locational position of each point. Areas where failures have occurred tend to no longer exhibit the conditions which once led to it in the first place. The points from this study were placed where possible as close to the scarp on the uphill part of the slide. Parameter maps were used as a guide when searching for potential locations during the field work. Every location was recorded using GPS and later imported to ArcGIS (Appendix 2).

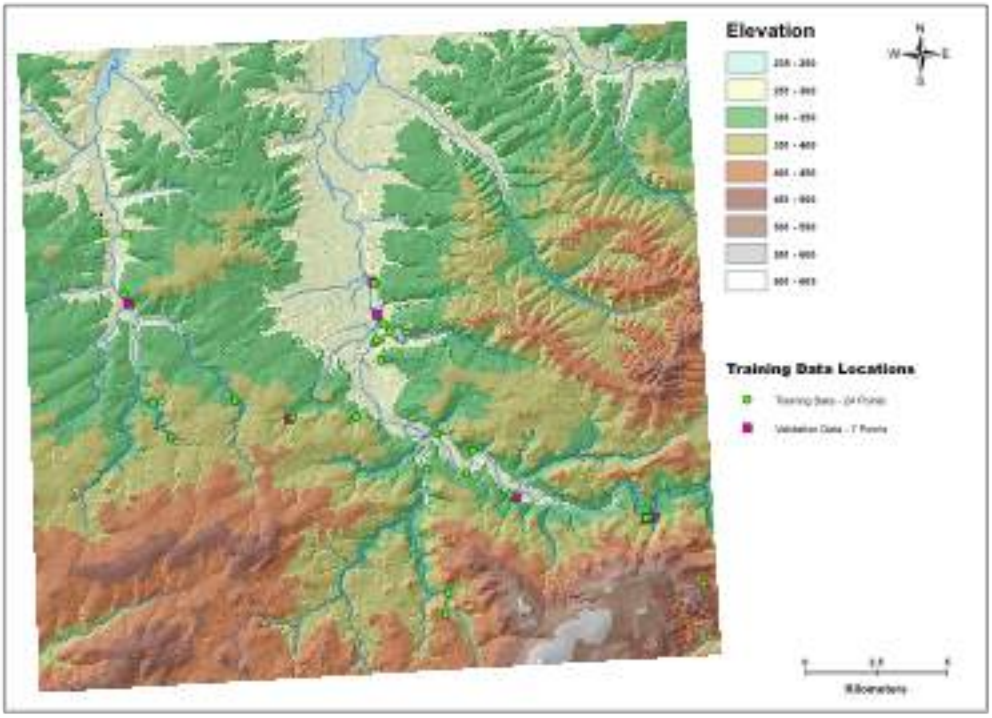


Figure 5.10 Landslide Inventory locations

The final points were randomly divided on a roughly 80/20 split into training and validation sets respectively (Figure 5.10). This equated to twenty-four validation points and seven training points, which is a relatively small inventory, however, as discussed in chapter two, ANN has been proven to work well with limited but good quality datasets. The validation and training point layers were converted to pixels in two binary rasters using the project extent, each inventory points received the value one. It is worth noting that even though polygon extents can be used for the LS extent, due to the high resolution of the data and small scale of the landslide, it was deemed more accurate to represent each LS by one pixel. Further processing of the training raster was then required. The ANN model needs both positive (LS present: 1) and negative (LS absent: 0) cells to train the network. A 50% random raster was created and combined with the training raster points, this reduced the processing load on the system, as the no data cells are ignored.

5.5 ANN set up and modelling

Advangeo© software from Beak Consultants Gmbh was used to model the parameters outlined in the previous section. The software is fully integrated with ArcGIS and the interface presents similar to the ArcGIS user interface (Figure 5.11), hence it reduces the complexity of transferring and viewing the data. Despite this, pre-processing of the produced rasters was necessary. Firstly, a file data structure was created to hold the rasters once imported. A base raster and polygon extent are required for the processing of the data in the software. All further imported data is clipped to the extent polygon and snapped to the base raster to maintain a consistent grid system. A total of forty-six data layers were created from pre-existing and created functions in ArcGIS and subsequently imported into Advangeo© for modelling (Appendix 1). All layers were converted to either continuous or binary rasters in Advangeo© using bi-linear interpolation. The software can only process these two formats of data.
Each category for the soil, geology, and aspect rasters had to be converted to separate binary layers. For example, soil type two, present (1) or not present (0). Parameters such as slope, curvature, SPI and TWI etc are continuous in nature.

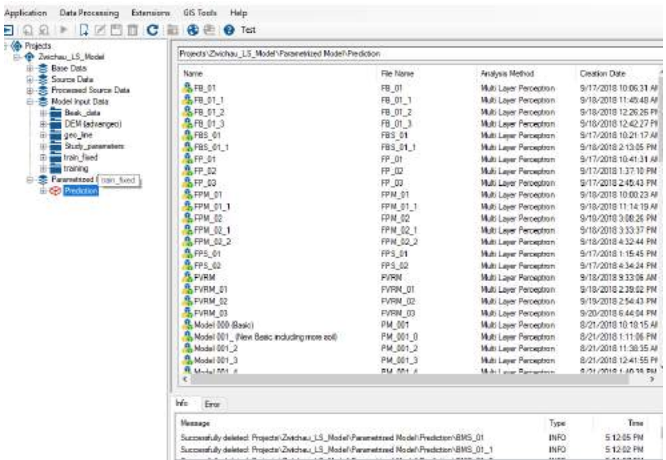


Figure 5.11 Advangeo interface and file structure

Once all data was imported and processed in Advangeo© into the file system, the modelling phase could begin. The model used is an MLP with a BP algorithm and a sigmoid activation function. The MLP contains three layers, an input, hidden and output layer (Figure 5.12). The hidden layer contained double the number of input nodes (neurons). The output layer contains one neuron which classifies the result between 0 and 1. One being most susceptible and 0 being the least susceptible.

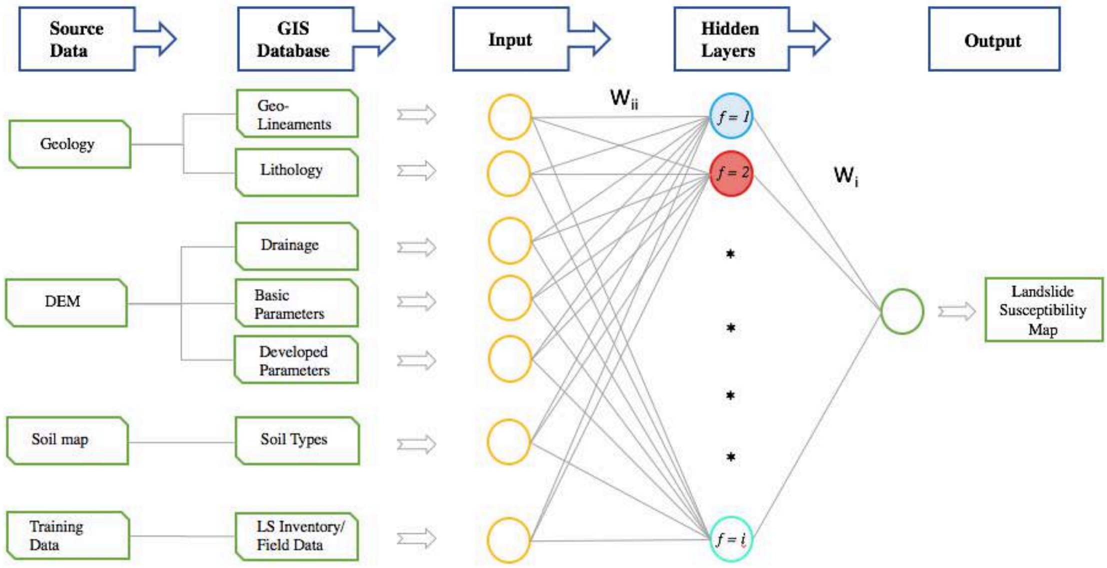


Figure 5.12 ANN workflow and setup

For each model, a set of parameters and the training data layer were chosen. Due to processing limitations of the system, the number of parameters could not exceed twenty-one. The model settings were kept constant for all produced models.

For the training phase, the epoch number was set to 100, with the initial min and max weights for the BP algorithm at -0.1 and 1 respectively. The error function used was tanh and a learning rate of 0.7 was assigned. On completion of the training phase, the derived weights and error curves were evaluated (this will be discussed in more detail under the Validation section). Many parameters combinations were run and assessed, changing the number of hidden layers, the learning rate, using 100% of inventory points instead of an 80/20%, and using various combinations of parameter rasters. Once an acceptable final model was achieved the training weights were applied to the whole study area to create the final susceptibility raster.

Due to the limitation on the number of input nodes (21), a preliminary assessment was carried out to reduce the number of possible parameter combinations included for modelling. Parameter data for each training point raster cell was extracted and the values compared

(Appendix 3). Using a combination of graph and mapped visual aids an early assessment of each parameter was made. This was especially applicable for parameters calculated with different kernel sizes. Parameters were eventually selected based on the distribution and frequency of each dataset.

In the end, five prediction models were selected. To negate the effect of the local minimum versus global minimum issue outlined in chapter two, each of the five models were run multiple times. These rasters were then averaged to harmonise the resultant data.

To series of maps and graphs were created to validate and display the output from both the models and the parameters rasters generated. Difference rasters were calculated to compare the resultant model predictions and comparison maps to highlight the differences between kernel sizes for TRI, VRM, DI, DD, and HI.

5.6 Validation

Upon completion of the modelling, a variety of measures were used to assess the quality of the produced models. Validation of ANN can be complex due to the black box nature of the system, however, on a basic level, certain benchmarks must be achieved for the model to be considered accurate.

Once a model training phase is completed, the root mean square error (RMSE) should fall below 0.01 for the model to be considered (Figure 5.13). The classification of the training data and connection weights must also be analysed. All training pixels are input with a value of 1, after system training, these values are recalculated based on the weights derived by the BP. The ideal output is that all of these ones remain ones. The quality of the calculated weights can be seen in the assigned values. A more realistic requirement to be met is that the new pixel values fall within the values 0.9-1 (Figure 5.14). This means the model weight correlation from the BP was strong and adheres strongly to the training value of 1 assigned. The same applies to zero values, which are expected to fall between 0 and 0.1.

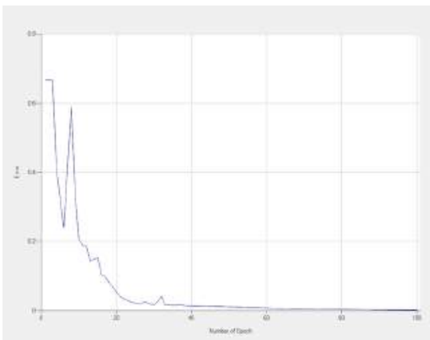


Figure 5.13 Sample error curve

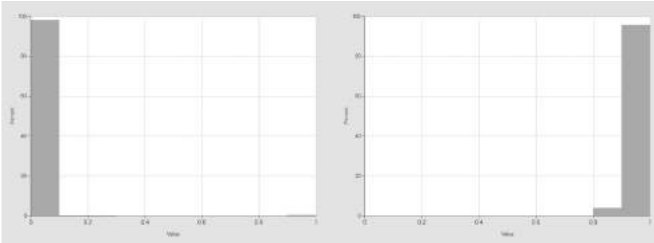


Figure 5.14 Sample histograms showing the recalculated training data (right) and the training area (left).

The final assessment on the result of the training phase is to scrutinise the resultant parameter model weights. The weights produced should increase steadily from 0.1 upwards with no abnormally high weight values. When a connection weight value is becoming too polarised, the system assigns a random -15,000 value, which indicates to the user that the training phase was not successful at correctly assigning the weights.

Once the weights from the training phase are applied over the whole raster and a model is produced the results can be assessed using other methods. First, each of the raster models were reclassified by intervals of 0.1 from 0.6-1. The pixel counts for each threshold were compared for each model as these show the distribution and amounts of the values assigned. Ideally, the pixels should be evenly distributed to indicate a more precise classification.

Finally, the receiver operator curve (ROC) was plotted and the area under the curve (AUC) was calculated for each model result. This was done using the ROCR package[14] in R_studio[15]. ROC is a plot of the true positive against the false positive rates. The more inclined the curve is to the upper left corner the better the model result is meant to be. The same applies to AUC, the greater the value the better the result.

$$TPR = TP/(TP + FN)$$
$$FPR = FP/(FP + TN)$$

Maps and graphs were also used for visual assessment of both the parameter and prediction data produced. Difference, comparison, and normal maps were created to better display the complexities of the data.

---

[14] https://cran.r-project.org/web/packages/ROCR/ROCR.pdf
[15] https://www.rstudio.com/

6. Results & Discussion

A total of seven new parameters were developed in ArcGIS as part of this study. These data sets were then tested in an ANN and the resulting models evaluated using various techniques. All models were trained using an 80/20% split of the data, resulting in 24 training points and 7 validation points. From this five LSS prediction models were created for further evaluation. The results of the body of work will be presented in the following chapter.

Beak Consultants Gmbh previously modelled LS susceptibility for the whole of Germany with a 20m resolution. The parameters used for this model are the foundation for the modelling approach on which this research was built. Soil data, fault-line and drainage density are newly included data sets for this research, however, they were not used as part of the original model from Beak. The soil and fault datasets were pre-processed from the source soil & geological maps but were not developed in ArcGIS as with the other parameters. Drainage density was derived from the DEM using an existing ArcGIS function and was not developed.

6.1 Parameters & Landslide Occurrence

The parameters can be dived into three sets as seen in Figure 6.1:

1 – Standard parameters previously used by Beak Consultants Gmbh to create a nationwide German susceptibility model, from source maps and ArcGIS functions.

2 - New parameters developed to be tested in an ANN.

3 - Soil datasets were also included in the new model for analysis, however, this data was not used as part of the original model from Beak. Soil, fault-lines, and DD will be included here for discussion.

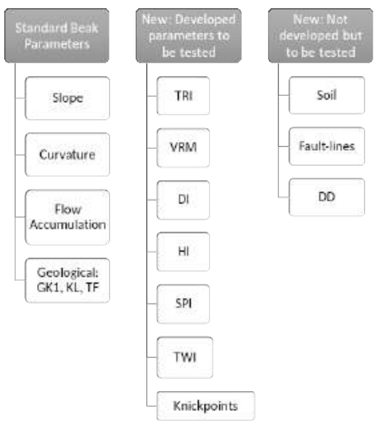


Figure 6.1 List of parameters modelled

6.1.1 Standard Parameters

The standard parameters are datasets previously used by Beak Consultants Gmbh to create a nationwide German susceptibility model. These include slope, curvature, flow accumulation, and three geological layers (TF, KL, GK1). All percentage distributions graphs for the inventory points can be viewed in Appendix 4 for the standard parameters.

Elevation of the study area ranges from 236m in the North up to 606m in the SE. The lowest elevations exist in the Zwickauer Mulde (ZM) lower valley and floodplains. 84% of landslides from the inventory occurred between 266-350m.

The *flow accumulation* raster was created during the drainage extraction process and is a continuous raster (Figure 6.2). High accumulation values exist in areas of concentrated flow such as rivers and streams, with the highest values present in the Zwickauer Mulde river. Ridges and flat areas can be observed as areas of low flow accumulation, as there is no flow at these points. Geomorphological and geological features can clearly be seen in the distribution of flow accumulation. The Kirschberg Granite body in the south of the study, is demarcated by irregular ridge patterns. In the NE section of the map, east of the ZM very distinctive flow patterns exist in the form of numerous small valleys. The majority of landslides occur in the lower flow accumulation range (90% <1.6). This is understandable because steeper slope areas are not associated with extreme levels of flow accumulation. Higher accumulation is linked to river and streams directly and the majority of inventory points were collected close to these areas but not close enough to yield extreme values. However, there are some outliers. The value range for flow accumulation is high, but most pixels fall into lower value categories making the data difficult to visualise, hence log10 was applied to distribute the range better.

The *curvature* is a very important parameter in LS susceptibility. Four rasters were created that divide the overall curvature into plan positive and negative, and profile positive and negative. Landscape form can be easily distinguished by observing patterns of curvature. The curvature map in Figure 6.2, displays the distribution of the general curvature values throughout the study area, positive values refer to convex slopes and negative to concave slopes. It can clearly be seen that distinctive valley prone areas in the NE display higher concentrations of positive curvatures. In general, positive convex slope shape tends to occur on upper slopes in contrast to the concave slope shape seen as negative values on the map, these occur at slope bases near to the valley floor. Intermediate values refer to flatter slopes or terrain with little curvature. Once again features such as the granite body in the south can be seen quite clearly in the curvature patterns. For plan curvatures, landslide occurrence is focused mainly in the lower value ranges. The locations are relatively evenly distributed between negative (55%) and positive (45%). Of the 55% negative values, 35% occur between -1 and 0. Of the 45% positive values 27% occur from 0-1. This shows that the landslide locations for this area, are not characterised by extremes in curvature but more gentle plan curvatures, meaning the convergence/divergence nature of plan curvature is not prominent. In contrast occurrences of landslides for profile curvature are skewed towards more positive values amounting to 73%. Of this 73%, 55% occur between 0 and 3 or in other words, the lower half of the positive value range. This means that while that more landslides occur on upwardly concave slopes.

Slope and *aspect* maps display the steepness and orientation of a slope. Eight orientations (N, NE, E, SE, S, SW, W, NW) were defined in the aspect map in, from this the nature of the geomorphology can be seen (Figure 6.3). The slope values for the study area range from 0 to 61°, higher slope values occur in the east, with the highest values occurring along the eastern edge of the Zwickauer Mulde floodplain near Zwickau and in the SE higher elevation areas where the drainage has incised the landscape. 70% of the landslides were recorded on slopes between 20-40°, with the highest cluster being from 30-40° at 36%. The distribution of LS in regards to aspect is more evenly spread, no landslides occurred on east-facing slopes, whereas the highest concentration was on SW facing slopes (25%).

Geologically, 61% of landslides were recorded in the hard rock category (GK1). TK & KL distributions were less conclusive. Out of the 3 categories present for each of these parameters only 36% (TK) and 27% (KL) of the landslides, occurred in what would be considered conditions more susceptible to failures, namely category 2. For KL this is the fault susceptible lithological units and for TF this is the lithological units with a prevalence to separation planes such as jointing, bedding, foliation etc.

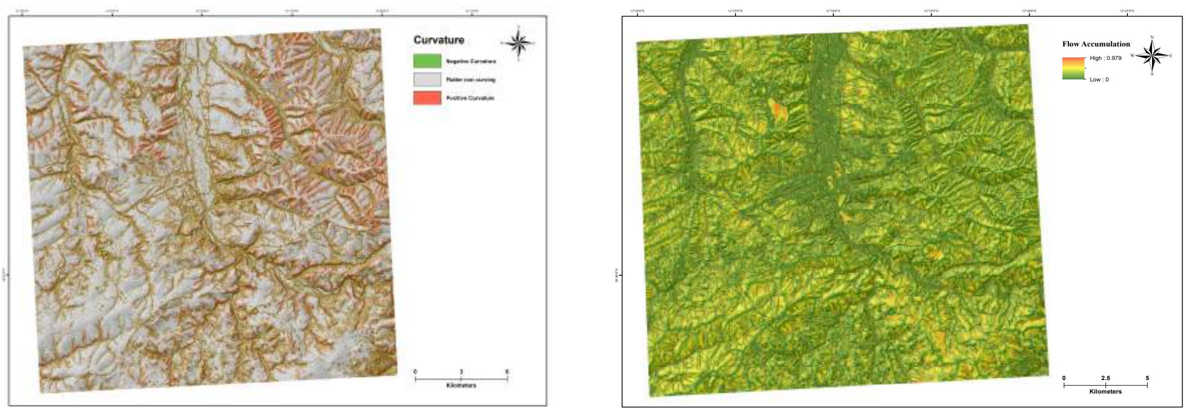


Figure 6.2 General Curvature & Flow Accumulation maps

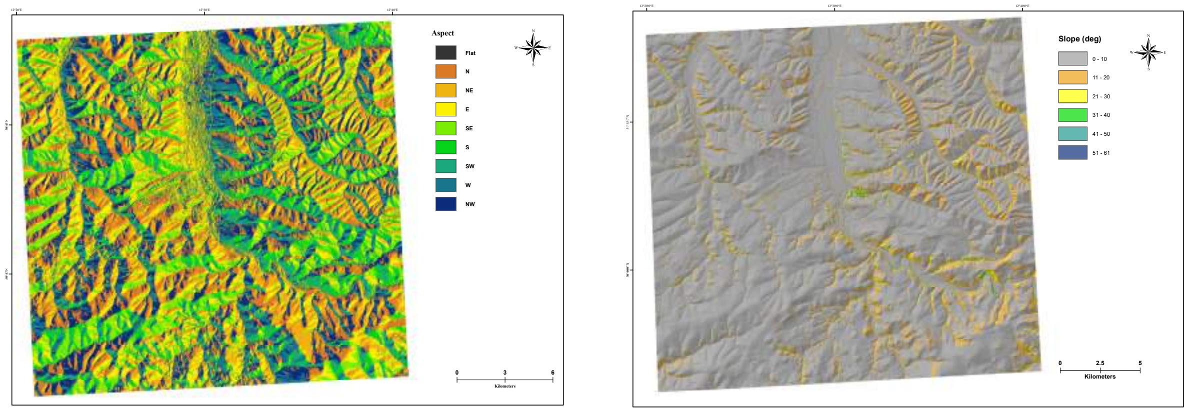


Figure 6.3 Slope & aspect maps

6.1.2 New Parameters

A total of seven new parameters were created to be tested in the ANN alongside the previously used standard parameters. As mentioned in the previous section soil and fault results will be grouped in this section with the newly developed parameters as they were not included in the original Beak Gmbh model. All landslide occurrence data graphs for each parameter can be found in Appendix 5.

The *distance to faults* raster produced buffered areas around each fault by 80m. An intense network, of strongly SE-NW and SE-NW trending faults, with clusters in the SW and SE, can

be seen in

Figure 6.4. A strong negative correlation exists in relation to the fault lines. 52% of the landslides occurred in areas with zero correlation to fault zones. However, this may not negatively affect the overall result as the other 48% occur within 80m of a fault line, so this may provide some information for the training process.

Figure 6.4 Faultline map with 80m buffers.

Soil types in the study area are well distributed in the study area and have been discussed in chapter 4. Most landslides were observed in loose material areas (type 2 at 24%). The nearest values closest to this were type 10 (Brown peri-glacial earth) and type 1 (anthropogenic material) at 15% each. Overall the soil per landslide occurrence is well distributed across six of the nine categories on the graph.

As is to be expected, the highest intensity of *drainage density (DD)* values can be seen along the major rivers and streams. Two radius sizes were chosen for the DD raster calculation. The density radiuses of 100m and 500m yield comparatively different results. Comparing them is not directly possible due to the dependency of the calculation on the radius supplied, however, they can be described in terms of the data distribution. Landslide occurrence for the 100m radius raster groups mainly around mid-range values and 12% of the dataset has zero value, which is not ideal. On the other hand, the 500m radius tends to cluster around lower value ranges but does not have any points which have zero value (Figure 6.5).

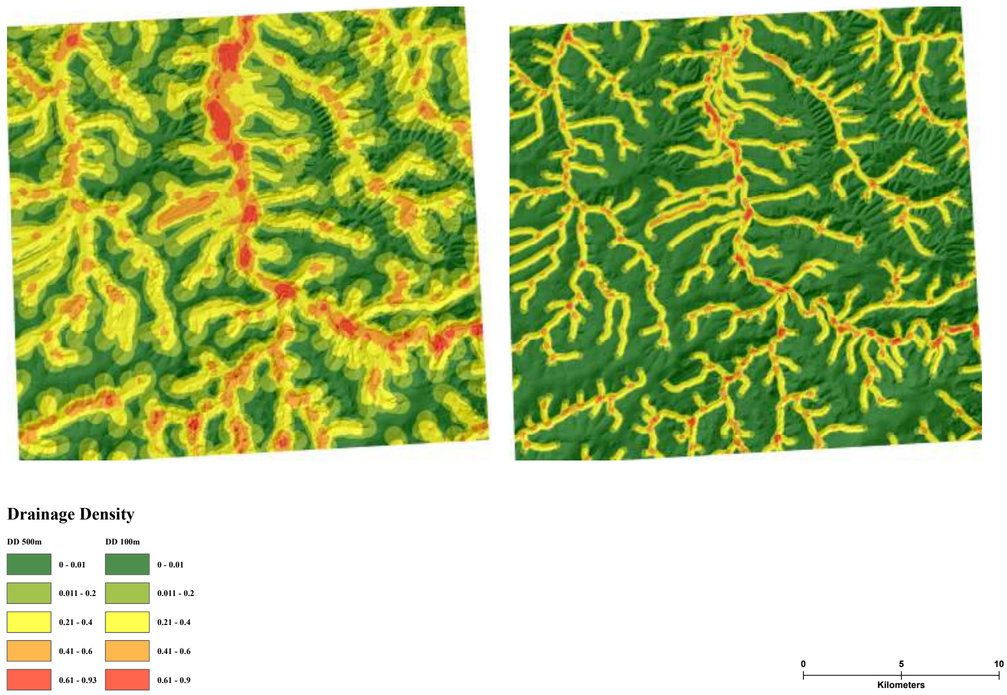


Figure 6.5 Drainage Density map, 500 & 100m.

Stream power index (SPI) values were standardised form 0-1, as index parameter values tends to vary between different study areas. Lowest SPI values can be seen along ridge lines and valley floors. In contrast to the low values on the valley floors, active stream, and river areas display the highest SPI values. Higher values can also be seen more frequently in the south, especially in the SE, this is more than likely due to the difference in elevation and incision process between the north and south. Landslide occurrence in relation to SPI, is distributed in the mid- to high ranges, which can be linked to steep slope areas with higher accumulation values (Fig?).

The *topographic wetness index or compound topographic index (TWI/CTI)* values were also standardised between 0-1. The results of the TWI calculation display value distribution as expected. Areas prone to saturation are represented by higher values whereas, higher relief areas generally exhibit lower values. This is very obvious along the ZM floodplain, which is relatively flat and receives a lot of water into the area. An area of high TWI in the north-west of the ZM valley can be identified as a water dammed location where mining waste is stored (fig?). 85% of the landslide occurrence in relation to TWI occurs in mid-range values in a standard distribution (Figure 6.6).

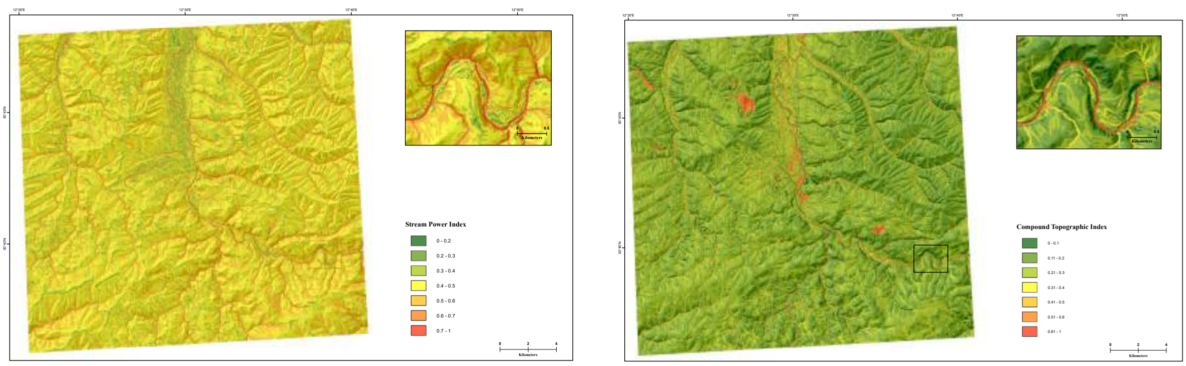


Figure 6.6 SPI & TWI (CTI) maps

Riley's *terrain ruggedness measure (TRI),* was calculated with three kernel sizes (3,5,9) and these were compared to evaluate which kernel size would be more suitable for modelling. The data is standardised from 0-1. The highest values present on steep slope sections and overall it can be seen that the eastern section of the map and in particular the south-east, generally has higher values than the rest of the study area. Flat valley areas like the ZM valley exhibit the lowest values. Comparing the different kernel sizes, it is obvious that the distribution of values varies greatly depending on which size is chosen. Upon examination of the TRI landslide data, it is obvious that each kernel has strong correlations between data peaks and occurrence. However, they do vary in the value ranges and a shift in the occurrence peaks per kernel. The 3x3 cell kernel groups LS occurrence in the lower range of values, this is easily observed in Figure 6.7 as the 3x3 kernel map shows a very low distribution of higher values. The percentage of higher values in the 3x3 cell kernel is much lower than the 5x5 and 9x9.

The *vector ruggedness measure (VRM)* follows similar patterns to the TRI as they are both roughness measures, however notably differences are present. In comparison with TRI, the VRM values are less focused on slope sections due to the differences in the calculations, and VRM can be seen to characterise the slope in more detail (Figure 6.7). Three kernel sizes were also used for the VRM calculation (5x5, 9x9, 15x9). After observing the distributions of TRI data, it was decided to try a larger 15x15 kernel for VRM and leave out the 3x3 kernel. Also, in this case, the 5x5 kernel value distribution was too limited and in the case of the 15x15 kernel, too generalised. LS occurrence percentage distributions are more varied for VRM. This can be attributed to the methodology, which places less bias on steeper slopes, resulting in a less homogeneous distribution of values. For both the 15x15 and the 9x9 kernels, roughly 30-35% of LS's occur at mid-range values.

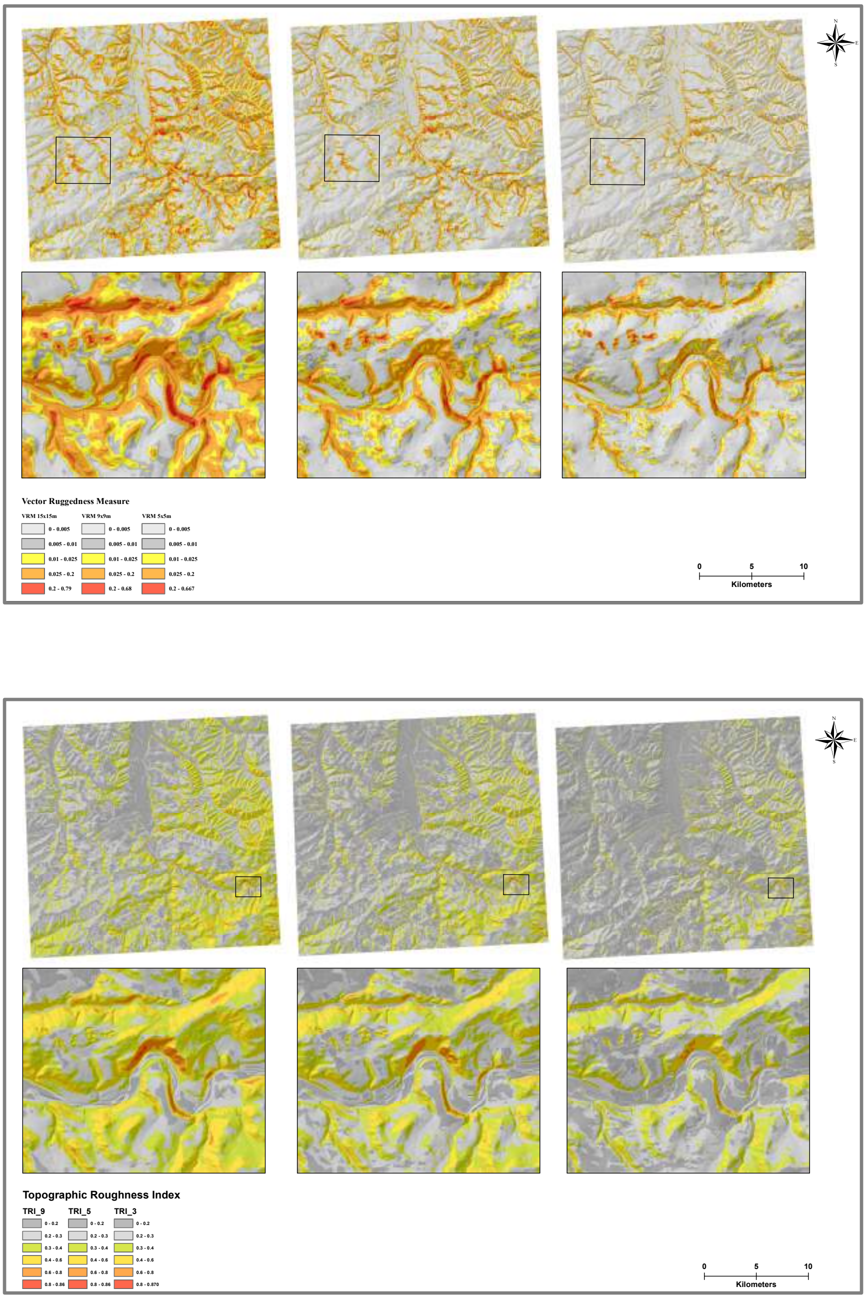


Figure 6.7 VRM & TRI kernel comparison maps

Three kernel sizes were also used for the *dissection index (DI)* generation, 15x15, 9x9 and 5x5. Notable differences in value distribution can be seen between the three sizes. The 5x5 cell calculation resolution is too small to pick up the dissection effectively and can be discounted. Both the 9x9 and 15x15 raster show more evenly distributed data and pick up the dissection more successfully (Figure 6.8). High DI values correspond to steeper slopes adjacent to incised valleys on the map. Landslide occurrence is comparable for both 9x9 and 15x15 kernels. However, no clear correlation exists between LS occurrence and the DI value as the distribution is over the full range of values.

For the *Hypsometric Index (HI),* three larger kernel sizes were trialled (25x25, 15x15 & 9x9). All three rasters show good distributions of values across the landscape, however, the 9x9 distribution tends to be too limited to the lower range values and doesn't pick up certain features. The 15x15 raster, over generalises the landscape features, whereas the 15x15 seems to represent the data best. High HI values are present on topographic highs; lower values are represented by lower flatter areas. The NE section of the study area displays slighter higher values overall; this may have something to do with the interplay of many small dissecting valleys. Landslide occurrence per HI value shows a normal distribution around the mid-value range. For the 9x9, 15x15 & 25x25 kernels, 97%, 94% & 97% respectively, occur between values of 0.3 and 0.6. In this case, deciding on the optimum kernel is best done by visual inspection of the map (Figure 6.8).

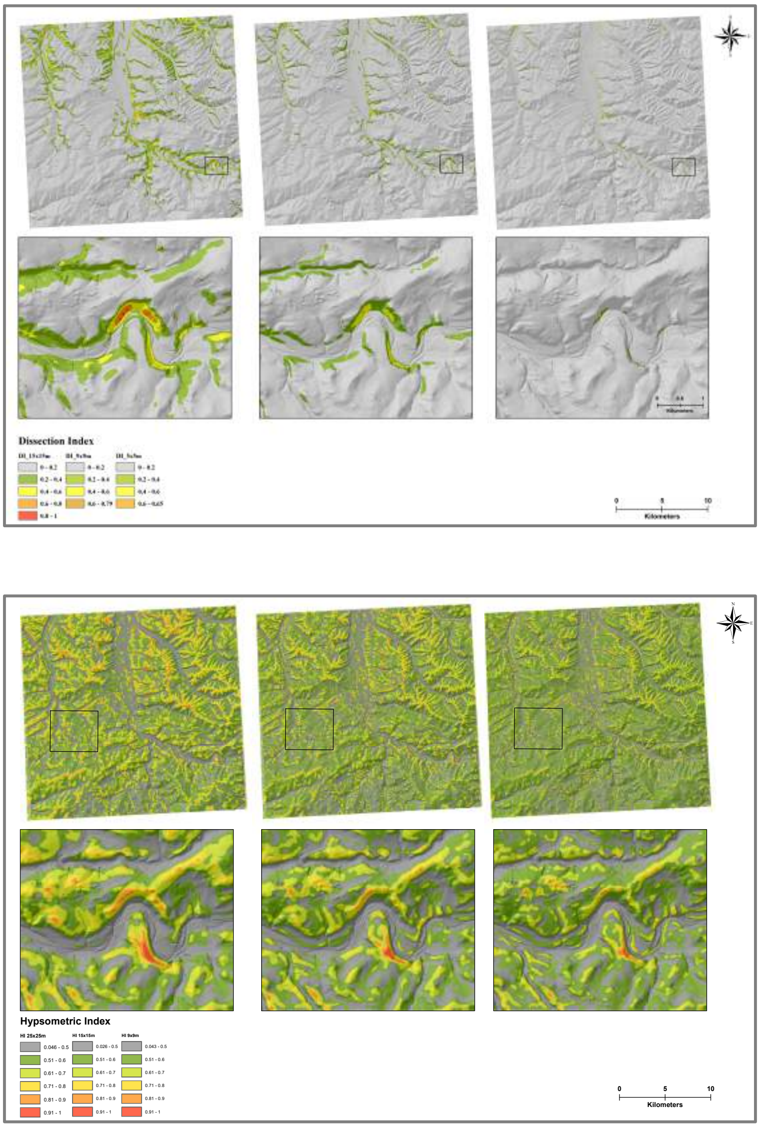


Figure 6.8 DI & HI kernel comparison maps

Three knickpoint data sets were produced using the user input of 20, 50 and 100 meters. On examination of the data produced, the 50m threshold was selected as being optimal for this area (Figure 6.9). Using 20m as the threshold resulted in an oversensitive result that placed knickpoints at every small variation in the river course. On the other hand, 100m proved not to be sensitive enough. Cross sections plots were examined to confirm or not the presence of the knickpoints in the stream section, and satellite images were carefully examined to rule out whether the points were due to man-made interference with the river path. The main reasons for the knickpoint occurrence were found to be faults, lithological boundaries and man-made alterations to the natural river path. In more tectonically active areas, the aim of such a parameter would be to detect majorly fault lines which due to active movement would exhibit more defined knickpoints. LS's per knickpoint density shows slightly higher occurrences in the lower value ranges, 64% between 0 and 0.2.

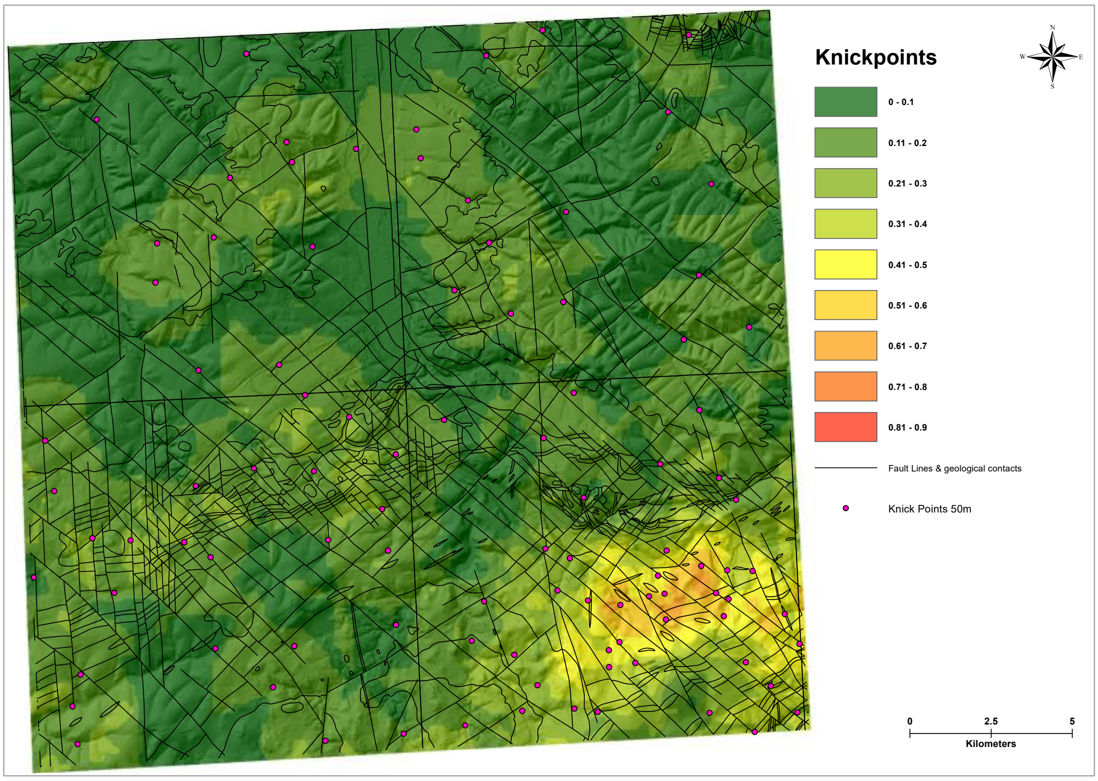


Figure 6.9 Knickpoint density map, with fault superimposed

6.2 Parameter selection

Due to the limitation on the number of input nodes in the ANN model, parameter outputs were analysed based on visual distribution from maps and the landslide occurrence inventory data. From this, a final list of parameters could be decided upon.

Initially, the land-cover was to be included as part of the modelling process, however, through observations in the field, it was noted that the majority of LS's occur in highly vegetated areas (90%). Hence, the model would see this as being a strong predictor of LS occurrence and would place a high weight amount. Open un-vegetated slopes are more commonly thought of having a higher likelihood of failure because they lack the deep roots to stabilise the ground. From this, it was decided to exclude land-cover from the final parameters.

In the last section, a review of the parameters results showed variability in the many layers produced. From this, certain kernel sizes and parameters were chosen and tested. For example, for TRI the 9x9 kernel was selected due to better generalisation and distribution of the values. For DI both 15x15 and 9x9 kernels were assessed to be of potential for modelling. Only soil types with the highest landslide occurrence were chosen. For the lithological types only hard rock was used (GK1), this was to keep consistency with the original Beak Consultants model but also due to the high occurrence of landslides in this category.

6.3 Models produced using an Artificial Neural Network.

As part of the modelling procedure, numerous parameter combinations were tested to obtain the best prediction result and ascertain what parameters were most effective for landslide susceptibility. Any parameters consistently scoring low weights during the training phase were excluded and the models re-run. Once a stable acceptable model was produced, it was re-run multiple times to obtain an average result for the weights and accuracies produced.

All the difference maps used in this section for model comparison have been categorised into 1 (>>), 0.5 (>), 0, -0.5 (<) and -1 (<<), with 1 (>>) and -1 (<<) being the greatest differences between the maps. All model maps are located in the appendix from 6 to 12.

Parameter weights vary between each model re-run and also between the different models due to the use of different parameter combinations, hence, the most effective method of comparing their importance is through ranking and not by the values themselves.

6.3.1 Basic Model (Basic)

The basic model tested as part of this body of work included the same nine parameter rasters used in the original Beak Consultants Gmbh model; slope, flow accumulation, GK1, TK, FL, and the four curvature rasters. The data resolution used to produce the basic model is higher than the original, using a 10m and a 1:25,000 geological which includes fault-lines. The original model was also applied to buffered areas around the main road networks, however, for this study a prediction for the complete area was generated. The training data sets are also very different, the original model was trained based on a German-wide inventory of hundreds of inventory points, and for this particular area, only the four original inventory points were available. In ways, this is both an advantage and a disadvantage. The more training points that are available for the model to learn from means there is less bias, however, for this study, there is an advantage in having area-specific data but the disadvantage of not having many data points overall for the system to learn from.

Overall system configuration for the MLP is 9-19-1 neurons, which equates to 9 input parameters, 19 hidden neurons, one output neuron, leading to 190 total connections. The new basic model produced an accuracy of 0.0080 during the training phase, averaged from multiple model runs. The result classification of 1's and 0's during the training phase was reasonably good, with on average 86% of the training points being classified between 0.9 and 1 an example is shown in Figure 6.10.

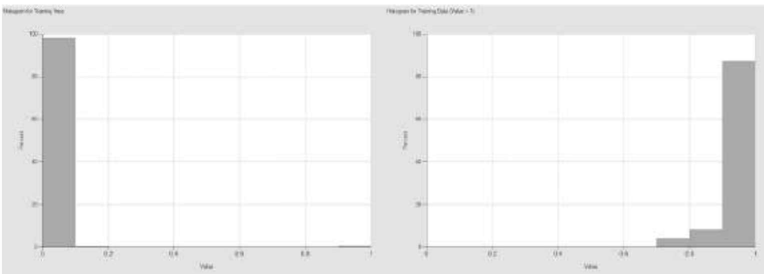


Figure 6.10 Sample result histogram from the basic model

On visual inspection of the new higher resolution model, some differences are apparent. The two models were compared using a difference map (Appendix 1), as the prediction resolution is so detailed it is hard to observe the difference fully by comparing only the two resultant maps.

The results of this comparison are variable, on one hand, the prediction locations are more precise and less generalised as in the original model, however, the new model does tend to over predict in many areas, this can be seen by pixel outliers throughout the map. In a very noticeable area of high LS prediction from the new model can be seen, this area was not previously modelled. Running a low pass filter over the basic model reduces the number of outliers seen in the prediction and brings the prediction more in line with the Beak model. The purpose of this basic comparison was done as a foundation on which to assess the basic model and assess in general the effects of higher resolution data.

Parameter rankings for the basic model can be seen in Table 6.1. From this, it is obvious that all four curvature layers play the biggest role in the prediction followed by slope, with the lowest contributor being the hard rock (GK1) classification.

Rank	Parameter
1	Curvature - Profile Neg
2	Curvature - Profile Pos
3	Curvature - Plan Neg
4	Curvature - Plan Pos
5	Slope
6	Flow Acc
7	TF
8	KL
9	GK1

Table 6.1 Basic model parameter weights rank

6.3.2 Basic model with soil (BS)

This model was run specifically to test the effect of the soil data alone with the basic model before the other parameters are added for testing. Soil type layers 1, 2, 10 and 12 were added to the basic parameters for training the system, as these are the layers in which the highest percentages of landslides occurred according to the inventory data. The system configuration was 13-27-1 with a total of 378 connections. The average accuracy achieved was 0.006, an improvement on the previous model. Classification of the training data after training resulted in 91% of the points on average being classified between 0.9 and 1, also an improvement on the previous model.

In Table 2.1 curvature also ranks highly out of the 13 parameters. Soil types 1 and 2 scored highest out of the four layers included. Overall the inclusion of soil data in the model has seen

a separation of the original parameters. KL, for example, has been pushed lower in the rank and GK1 still remains lowest on the weight ranking.
Visual evaluation of the BS against the basic models shows an overall reduction in outlier predictions in certain areas.
Soil types 1 and 2 scored highest out of the four layers included. Overall the inclusion of soil data in the model has seen a separation of the original parameters. KL, for example, has been pushed lower in the rank and GK1 still remains lowest on the weight ranking.
Visual evaluation of the BS against the basic models shows an overall reduction in outlier predictions in certain areas.

Rank	Parameter
1	Curvature - Profile Neg
2	Slope
3	Curvature - Profile Pos
4	Curvature - Plan Pos
5	TF
6	Flow Acc
7	Curvature - Plan Neg
8	Soil Type 1
9	Soil Type 2
10	KL
11	Soil type 12
12	Soil type 10
13	GK1

Table 6.2 BS parameter weight ranking

6.3.3 Basic model with new parameters included (FP)

The third model produced included the selected new parameters with the basic parameters to compare their effect on the model without the soil. The advantage of this is that more nodes are available for the inclusion of more developed parameters. By running the models in this way, possible dependent relationships can be identified, i.e. some parameters react differently when included together rather than separately.
The system configuration was 19-31-1 with a total of 780 connections. The average accuracy achieved was 0.0046, a notable improvement from the last model. Classification of the training data after training resulted in 93% of the points on average being classified between 0.9 and 1.

Visually comparing the basic, BS and FP maps, clear changes in the accuracy of the prediction can be seen, high prediction areas show progressively stronger, more detailed classifications between the thresholds and further reductions in outliers can be seen. The prediction for this model now looks much cleaner.

The parameter rankings for this model show that the weight ranking given to the VRM 9x9 parameter now sits higher than plan positive curvature and flow accumulation (Figure 6.9). It is important to remember that these rankings are based on averages of the same model run multiple times to negate the effects of the local/global minimum problem. Drainage density for both radius sizes ranked lowest on the weight influence for the model.

Rank	Parameter
1	Curvature - Profile Neg
2	Curvature - Plan Neg
3	Curvature - Profile Pos
4	Slope
5	VRM 9
6	Curvature - Plan Pos
7	Flow Acc
8	TWI
9	TF
10	DI 9
11	HI 15
12	Faults
13	SPI
14	TRI 9
15	GK1
16	KL
17	Knick
18	DD 500
19	DD 100

Table 6.3 FP model parameter weight rankings

6.3.4 Basic model with parameters and soil included (FPS)

For this model, both the new parameters and the soil were tested together in the model with the basic set of parameters. The system configuration was 21-47-1 with a total of 1034 connections. The average accuracy achieved was 0.0036, once again a further improvement from the last model. Classification of the training data after training resulted in 96% of the points on average being classified between 0.9 and 1.

Comparing the FPS prediction map with the previous prediction map (FP) shows that not only do the parameters have an effect but combined they further improve the prediction and reduce outliers. The highest-ranking parameters from this model in order are VRM 9x9, DI 9x9, and SPI. Soil type 12 ranks much higher than the other soil types and sits alongside VRM 9x9 as high in the rankings (Table 6.4).

Rank	Parameter
1	Curvature - Plan Neg
2	Curvature - Profile Neg
3	Curvature - Profile Pos
4	Slope
5	Curvature Plan Pos
6	Flow Acc
7	VRM 9
8	Soil 12
9	TF
10	Faults
11	DI 9
12	SPI
13	TWI
14	Soil 2
15	KL
16	Knick
17	HI 15
18	Soil 1
19	TRI 9
20	GK1
21	DD 500

Rank	Parameter
1	Curvature - Profile Pos
2	Curvature - Profile neg
3	Curvature - Plan Neg
4	Slope
5	Curvature - Plan Pos
6	VRM 9
7	Flow Acc
8	Soil 2
9	KL
10	Soil 12
11	Faults
12	TF
13	DI 9
14	HI 15
15	SPI

Table 6.4 Average FPS(left) and FVRM (right) model parameter rankings

6.3.5 Basic model with optimised parameters and soil (FVRM)

This model and the last are for the most part the same, however, for this model the slected parameters were modelled (Table 6.4). To determine if the accuracy could be further improved, by optimising the parameters used. The system configuration was 21-47-1 with a total of 1034 connections. The average accuracy achieved was 0.0034, once again a further improvement from the last model. Classification of the training data after training resulted in 97% of the points on average being classified between 0.9 and 1. One of the models achieved 100% classification of the training pixels between 0.9-1

6.3.6 Summary

Overall an increase in the accuracy of the training sets for each model can be seen both in the calculated figure and visually. Interpreting the maps directly can be difficult due to the low susceptibility of landslides in the area and the high resolution of the prediction. The results are best reviewed and analysed in GIS, however, the use of difference maps can also aid in visual interpretation of the whole area. The FPS and FVRM models show the most promising results, as a notable reduction in outliers can be seen.

Due to the variance in the connection weights for each model, it can be hard to assess which parameters play the biggest role and show the most promise for further studies. However, by assessing their rank in each model patterns begin to emerge. Curvature, slope, flow accumulation ranks the highest across all models. The most surprising result was the low ranking of the hard rock lithological class (GK1) across all models because 61% of LS from the inventory occurred in this class. The reason for this is not easily explained. The interaction between different combinations of parameters in artificial neural networks is complex and not fully understood. Including new parameters has shown a marked increase in the model prediction, of these the most influential and highest ranking without a doubt is the VRM 9x9. Presented here are three averaged models in which the VRM 9x9 was included and for which the weight ranking was 5, 7, & 7 respectively. Numerous models not presented here were trialled and analysed as part of the research, and throughout these, the VRM 9x9 was consistently high ranking. On the other hand, DD was repeatedly assigned low weights in the training phase and shows the least promise for future modelling. Distance to fault lines also showed some correlation in the model along with soil type 2 (Mixed loose material). Loose material is more inclined towards small failures as observed during fieldwork.

6.4 Model Performance & Validation

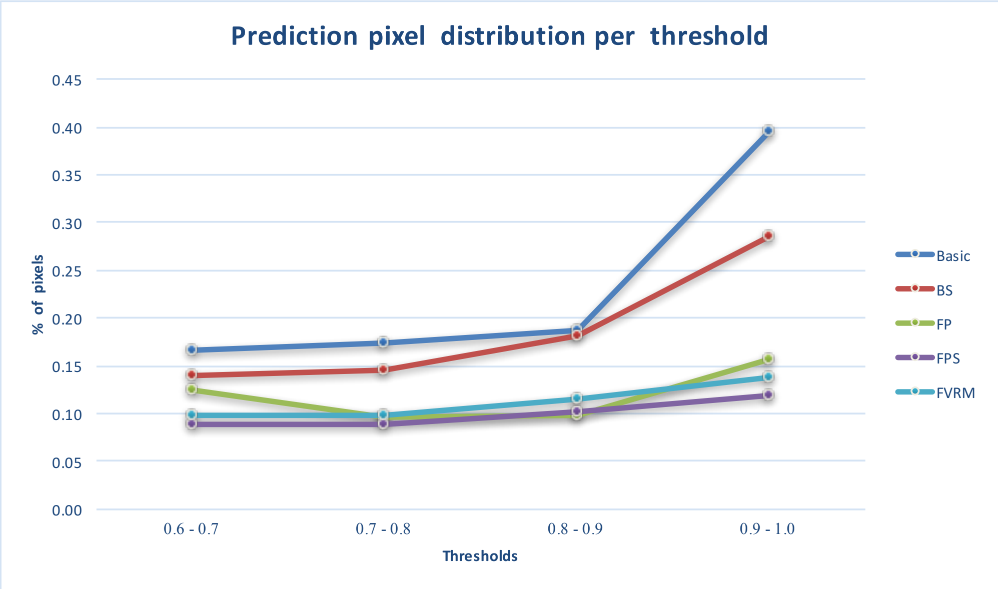
The results of the training phase and the validation sample were assessed for each model. No one method of evaluation or validation of the models is sufficient, however, a combination of techniques can build a bigger overall picture of the models and how they differ from each other.

As detailed previously the training phase results can be evaluated based on training pixel classification, and the RMSE error output, in all models a steady improvement in these factors are seen (Table 6.5). The differences between the final FPS and FVRM models are minimal but still an improvement. The error decreased by 0.0002 and the pixel classification increased by 1%.

Model Code	Description	Input #	Avg Error	Training pixel classification 0.9-1
Basic	High resolution 10m model using basic set of parameter inputs that were previously run by Beak Consultants Gmbh at a lower resolution of 20m .	6	0.00804566	86%
BS	The basic including soil data.	13	0.00637834	91%
FP	Basic model plus tested parameters	19	0.00459187	93%
FPS	Basic model plus tested parameters and soil	21	0.00359103	96%
FVRM	Basic model plus optimised/selcted parameters	15	0.00337845	97%

Table 6.5 Model summary table of training phase results

So far only the quality of the training phase has been assessed. To determine the accuracy of the final susceptibility prediction two methods were used. Firstly, by comparing the pixel distribution per threshold for the prediction rasters. Each raster was classified into thresholds and the pixel count for each plotted (). Only thresholds from 0.6 to 1 were compared as 0.6 was determined as the threshold above which a slope is liable to failure and also represents a natural break in the prediction data. All models have 99%+ of the pixel prediction classification from 0-0.6. This is a true representation of real life, as the study area is not overly susceptible to landslides, so a prediction of 99% as not LS susceptible is a good representation of reality. In the higher ranges, the pixel percentage values diverge, this is particularly seen above the 0.9 threshold (Figure 6.11). The basic and BS models classify 0.39% and 0.29% of the pixels into this category, and this is reflected visually in the maps, many higher pixel values are scattered as outliers over the model. In contrast, the percentage of pixel classification for the FPS and FVRM models has dropped to 0.12% and 0.14%, and the overall distribution of the prediction pixels are more evenly spread.



Threshold	Basic	BS	FP	FPS	FVRM
0 - 0.6	99.08	99.25	99.52	99.60	99.55
0.6 - 0.7	0.17	0.14	0.12	0.09	0.10
0.7 - 0.8	0.17	0.15	0.10	0.09	0.10
0.8 - 0.9	0.19	0.18	0.10	0.10	0.12
0.9 - 1.0	0.39	0.29	0.16	0.12	0.14

Figure 6.11 Comparison of prediction pixel classifications per model (% of total pixels)

In order to further verify the results of the LSS prediction model the ROC curves for each model were plotted and the AUC value calculated with the ROCR package in R. The resulting plots and values are presented in Figure 6.12. The ROC and AUC are derived from the comparison of the validation set (20% - a binary classifier) with the prediction raster results. The ROC plots the true positives rates against false positives rates. It provides information about the degree of reliability of the model, while a larger area under the curve meaning higher accuracy achieved, and, therefore, the AUC values provide a quantitative evaluation. The five distribution plots for the ROC are so similar that the plots overlap each other and the result AUC values confirm this as the results vary by only 0.001. The validation set of seven points is not optimum for plotting of ROC, this results in very contrast in values between the true positive rate and the false positive rate.

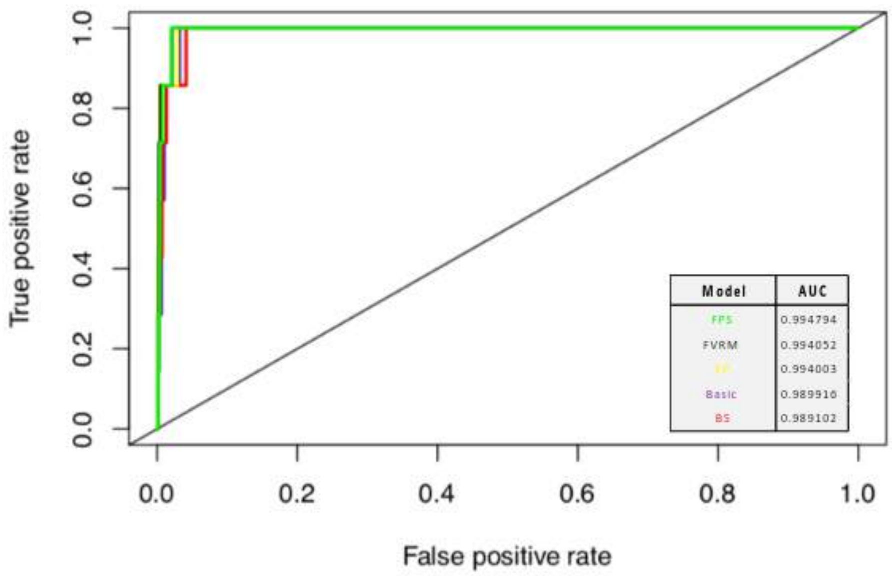


Figure 6.12 ROC plot and AUC values for all prediction models

To develop on the ROC plot and try to show more clearly and break down the differences between the models, the TP's against each threshold were plotted for the basic, BS and FP models, to show differences between the validation points and the prediction models. Once again it can be seen that in Figure 6.13 including the new parameters in the modelling process increases the prediction. At the 0.5 threshold, 100% of the FP predictions classified as positive matches with the validation points. The BS model also showed a higher classification rate of 71% in comparison to the basic model which only classified 57%.

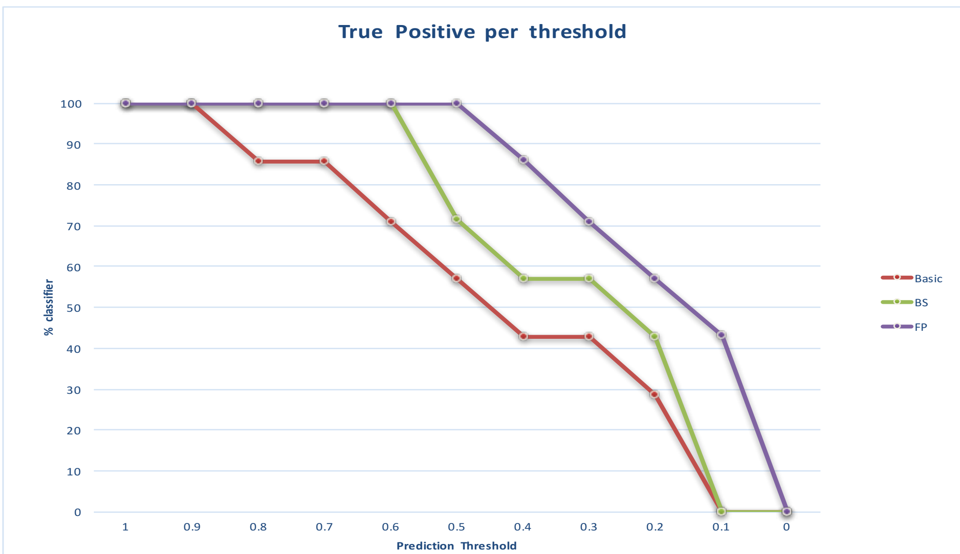


Figure 6.13 True positive rate plotted per threshold

7. Conclusion and Recommendations

7.1 Conclusion

In this study, it was found that new parameters developed in a GIS and modelled in an ANN can improve landslide susceptibility predictions. The ANN was trained from a created spatial database of the study area which encompasses the city of Zwickau in SW Saxony, Germany. A landslide inventory was generated from in the field recordings of landslide locations, combined with four points available from the Saxony state LS inventory. For modelling and validation purposes, the dataset was divided in an 80/20% split, 80% used to train the ANN and 20% kept for validation.

After an extensive review of the literature and available toolboxes, seven parameters in total were implemented using model builder and adapted scripts. After performance assessment using weight rankings, instead of the variable weight values, it was determined that the parameter *vector ruggedness measure* (VRM) calculated using a 9x9 cell kernel the parameter with most influences on the model.

From the training phase, the error was reduced by 0.00466721 between the basic and FVRM models and the classification of training pixels improved by 11%. The final susceptibility predictions were assessed using ROC and AUC but were found to not be sensitive enough to determine the accuracy for such a small data set (seven validation cells from 5,309,739). A more effective assessment was made by making a direct comparison of true positives per threshold. From this, the model differences were more clearly seen. Introducing soil and new parameters into different models (BS & FP) increased the prediction accuracy in comparison to the basic model.

7.2 Recommendations

The landslide occurrence in the study area is relatively low, there is potential for newly developed parameters such as VRM to be tested in other areas, particularly in high slope areas. VRM is not strongly polarised by slope, so this could potentially lead to more accurate assessments in higher elevation terrain. Kernel cells sizes would need to be re-examined in this case.

Due to the low performance of the drainage density, a possible suggestion for further work may be to swap this parameter for “distance to drainage”, to see if this offers improvements.

Attempts were made to model and assess knickpoint distributions using kriging and stream profile analysis, this work could be elaborated on in future research.

Running a low pass filter over the final prediction raster reduced the number of outliers while still maintaining good prediction resolution. Due to the variabilities of high-resolution DEMs, more outliers inevitably occur, a low pass filter to smooth the result could be a solution.

7.3 Objectives & research questions

7.3.1 Objectives

Review and selection of suitable morphological parameters, with a specific focus on tectonic morphological parameters, for the use in Landslide susceptibility modelling.

- *After extensive research, nine parameters from both tectonic, soil and classic morphological studies were evaluated and developed, however, due to further evaluation of suitability in relation to the study area, two parameters, namely Valley Asymmetry and Isobase were deemed to be non-applicable due to scale and model dependencies. Hence, only the seven remaining parameters will be discussed during this thesis.*

Investigate the relationship between tectonic morphology and landslide prediction and occurrence.

- *Due to the limited spatial extent of the high-resolution source data for this study, it was concluded that the relationship between tectonics and landslides is impossible to detect at this scale, deeper geological studies would be needed to confirm any relationship at this scale, which is not possible. Interpretations can be made from parameters that work on more regional settings however, any data produced from these parameters would produce homogenous results that are impossible for the ANN to interpret.*

Assessment of the available morphological analysis toolboxes for DEM's, i.e. TecLines, TecDEM, SAGA, Topotoolbox.

- *Various toolboxes were reviewed during the research review phase and the core methodologies used by these toolboxes were adapted and, in some cases, improved upon (i.e. VRM – Vector Ruggedness Measure instead of more basic methodologies).*

Adaption of techniques and datasets derived from selected toolboxes, to a GIS-based artificial neural network (ANN).

- *As mentioned in the previous point, methodologies and formulas specifically were adapted successfully, through the use of model builder and adapted python scripts in ArcGIS. The derived parameter datasets were then pre-processed for use in the ANN.*

Collection of training data for the ANN, through site-based fieldwork.

- *Fieldwork was undertaken over the whole study area for the acquisition of new landslide inventory data needed to train the ANN. Geological and morphological observations were recorded as part of the process. A total of twenty-eight locations were mapped, which when added to the original inventory brought the total inventory to thirty-two points. One of the original four inventory points that were supplied by the Sachsen state was then removed for modelling due to inaccurate user input, bringing the total number of inventory points to thirty-one.*

Determine the effect of the chosen datasets when incorporated into the ANN modelling process.

- *By comparing the average weighting ranks for each model, it was possible to determine the effect of each parameter, in regards to both the basic and newly included parameters. VRM proved the be the best performer from the the new parameters.*

Asses various model validation methods to determine if the newly introduced parameters increase the accuracy of the model results.

- *Receiver Operator Curve (ROC), Area under the curve (AUC), TP plots, pixel classification comparison and visual interpretations were made to assess the quality of the models produced.*

Evaluation and selection of the best methods for visualising the results of the study.

- *Due to the very precise prediction locations of the results, it is hard to visualize the results on paper and to see the changes clearly, however, a range of difference maps and comparison graphs have been created to better display the data.*

7.3.2 Research questions:

<u>What are the tectonic morphological parameters that can be derived from a DEM?</u>

- *Asymmetry, Isobase, Roughness, mountain front sinuosity, Stream Length Index (used in knick point calculations), drainage basin shape, Hypsometrical Integral, and Dissection Index.*

<u>Of these, which are the greatest importance according to previous literature?</u>

- *These parameters have not been used in this particular context so one cannot gauge the importance of these parameters. However, in the context of LSS, various forms of a roughness calculation have been used in the modelling process throughout the literature. Hypsometrical Integral has also been included in some studies. The other parameters have been used in the context of tectonic interpretation in relation to landslides but not using modelling techniques.*

<u>Many toolboxes for morphological analysis exist, can their methodology and produced datasets be adapted for the purposes of the project, and what is the best approach to take?</u>

- *The parameters can be adapted with various levels of difficulty. Many toolboxes use similar methodologies but each one tends to focus on specific outcomes in terms of soil analysis, tectonic interpretation, landscape classification etc. The best approach was to combine the use of adapted scripts with in-built ArcGIS functions in model builder to calculate the parameter datasets.*

<u>Does functionality already exist in ArcGIS which can be implemented in Advangeo®?</u>

- *Some functionality exists in ArcGIS, but not as stand-alone functions, they must be combined within model builder or some cases with python scripts to achieve the desired outcome.*

<u>What methods can be used on the model output for validation and assessment?</u>

- *ROC, AUC, TP plots, pixel classifications and distributions, visual validation, difference rasters and running different divisions of the training data. The final models presented are based on a 80/20% split of the inventory data in 80% training and 20% validation, however, 100% of the inventory points were also used as a test and this did not show massive variation in the model predictions.*

<u>Which dataset can be said to add the most benefit to the prediction model?</u>

- *The Vector Ruggedness Measure (VRM) has been shown to consistently rank highest out of all newly tested parameters.*

Can an overall improvement in the susceptibility prediction be obtained?

- *Yes, the classification of prediction values in higher thresholds becomes much more precise when the new parameters are added to the model. Training phase error was reduced and the training data classification improved.*

8. References

AdV. (2015a). *Digitales Geländemodell Gitterweite 10 m DGM10*. Retrieved from

AdV. (2015b). *Documentation on the Modelling of Geoinformation of Official Surveying and Mapping (GeoInfo Doc)*. Retrieved from http://www.adv-online.de/Publications/AFIS-ALKIS-ATKIS-Project/binarywriterservlet?imgUid=8e1708ee-765f-8551-2357-c133072e13d6&uBasVariant=11111111-1111-1111-1111-111111111111

AdV. (2016). *Digitales Basis-Landschaftsmodell (AAA-Modellierung) Basis-DLM (AAA)*. Retrieved from

Alessandro Mondini, F. G., Fausto Guzzetti, Reichenbach Paola, Reichenbach Paola, Francesca Ardizzone. (2008). Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images. *Remote Sensing of Environment, 115*(7), 1743-1757. doi:10.1016/j.rse.2011.03.006

Ali, S., Biermanns, P., Haider, R., & Reicherter, K. (2018). Landslide susceptibility mapping by using GIS along the China Pakistan economic corridor (Karakoram Highway), Pakistan *Nat Hazards Earth Syst Sci*. doi:https://doi.org/10.5194/nhess-2018-39

Ali, S., Biermanns, P., Haider, R., & Reicherter, K. (2018). Landslide susceptibility mapping by using GIS along the China–Pakistan economic corridor (Karakoram Highway), Pakistan. *Natural Hazards and Earth System Sciences Discussions*, 1-28. doi:10.5194/nhess-2018-39

Alkhasawneh, M., Ngah, U., Isa, N. A., & Al-batch, M. (2013). Determination of Important Topographic Factors for Landslide Mapping Analysis Using MLP Network. *Scientific World Journal*. doi:10.1155/2013/415023

Andreani, L., Stanek, K., Gloaguen, R., Krentz, O., & Domínguez-González, L. (2014). DEM-Based Analysis of Interactions between Tectonics and Landscapes in the Ore Mountains and Eger Rift (East Germany and NW Czech Republic). *Remote Sensing, 6*(9), 7971-8001. doi:10.3390/rs6097971

Atkinson, P. M., & Tatnall, A. R. L. (1997). Introduction Neural networks in remote sensing. *International Journal of Remote Sensing, 18*(4), 699-709. doi:10.1080/014311697218700

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin, 24*(1), 43-69.

Bishop, C. (1995). *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.

Bruschi VM, B. J., Remondo J, Gómez-Arozamena J, Rivas V, Barbieri M, Capocchi S, Soldati M, Cendrero A. (2013). Land management versus natural factors in land instability: some examples in northern Spain. *Environ Manage, 52*(2), 398-416.

Carrara, P., & Pike, R. (2008). GIS technology and models for assessing landslide hazard and risk. *Geomorphology, 94*, 257-260.

Choi, J., Oh, H.-J., Won, J.-S., & Lee, S. (2009). Validation of an artificial neural network model for landslide susceptibility mapping. *Environmental Earth Sciences, 60*(3), 473-483. doi:10.1007/s12665-009-0188-0

Conoscenti, C., Rotigliano, E., Cama, M., & Lombardo, L. (2016). Exploring the effect of absences selection on landslide susceptibility models: A case study of Sicily,Italy. *Geomorphology, 261*, 222-235.

Costanzo, D., Rotigliano, E., Irigaray, C., Jiménez-Perálvarez, J. D., & Chacón, J. (2012). Factors selection in landslide susceptibility modelling on large scale following the gis matrix method: application to the river Beiro basin (Spain). *Natural Hazards and Earth System Sciences, 12*(2), 327-340. doi:10.5194/nhess-12-327-2012

Crozier, M. J. (1984). Field assessment of slope instability. In B. D. a. P. D. (Eds.) (Ed.), *Slope Instability* (pp. 103-142): London,

John Wiley & Sons Ltd.

Danielson, T. (2013). *Utilizing a High Resolution Digital Elevation Model (DEM) to Develop a Stream Power Index (SPI) for the Gilmore Creek Watershed in Winona County, Minnesota*. Retrieved from Central Services Press. Winona, MN:

Demoulin, A., & Chung, C. J. F. (2007). Mapping landslide susceptibility from small datasets: a case study in the Pays de Herve (E Belgium). *Geomorphology, 89*, 391-404.

Deolia, R., & Pande, A. (2014). Spatial Distribution of Dissection Index (Erosion Intensity) versus Geomorphological Environment in Parkha Watershed, Central Himalaya. *Indian Journal of Geography & Environment, 13*, 10-22.

Dou, J., Tien Bui, D., Yunus, A. P., Jia, K., Song, X., Revhaug, I., . . . Zhu, Z. (2015). Optimization of Causative Factors for Landslide Susceptibility Evaluation Using Remote Sensing and GIS Data in Parts of Niigata, Japan. *PLoS One, 10*(7), e0133262. doi:10.1371/journal.pone.0133262

Dragićevića, S., Laia, T., & Balram, S. (2015). GIS-based multicriteria evaluation with multiscale analysis to characterize urban landslide susceptibility in data-scarce environments. *45*(2), 114-125.

Ermini, L., Catani, F., & Casagli, N. (2005). Artificial Neural Networks applied to landslide susceptibility assessment. *Geomorphology, 66*(1-4), 327-343. doi:10.1016/j.geomorph.2004.09.025

Esri. (2001). ArcGISTM Spatial Analyst: Advanced GIS Spatial Analysis Using Raster and Vector Data. In E. USA (Ed.), (An Esri white paper ed., Vol. J-8747). 380 New York Street, Redlands, California, USA.

Esri. (2013). ArcGIS for desktop (Vol. 136353). 380 New York Street, Redlands, California, USA.

Etchebehere, M. L. C., Saad, A. R., Perinotto, J. A. J., & Fulfaro, V. J. (2004). Aplicação do Índice BRelação Declividade-Extensão - RDE^ na Bacia do Rio do Peixe (SP) para detecção de deformações neotectônicas. *Revista do Instituto de Geociências USP, 4*(2), 43-56.

Fabbri, C.-J. C. a. A. G. (2003). Validation of Spatial Prediction Models for Landslide Hazard Mapping. *Natural Hazards, 30*, 451-472.

Farhan, Y., Anbar, A., Enaba, O., & Al-Shaikh, N. (2015). Quantitative Analysis of Geomorphometric Parameters of Wadi Kerak, Jordan, Using Remote Sensing and GIS. *Journal of Water Resource and Protection, 7*, 456-475.

Farrokhzada, F., Bararib, A., Choobbastia, A. J., & Ibsenb, L. B. (2011). Neural network-based model for landslide susceptibility and soil longitudinal profile analyses: Two case studies. *Journal of African Earth Sciences, 61*(5), 349-357. doi:https://doi.org/10.1016/j.jafrearsci.2011.09.004

Glade, T., & Crozier, M. J. (2005). The nature of landslide hazard impact *Landslide Hazard and Risk*.

Gloaguen, R., & Mahmood, S. A. (2011). Analyzing Spatial Autocorrelation for the Hypsometric Integral to Discriminate Neotectonics and Lithologies Using DEMs and GIS. *GIScience & Remote Sensing, 48*(4), 541-565. doi:10.2747/1548-1603.48.4.541

Gorsevski, P. V., Gessler, P. E., Boll, J., Elliot, W. J., & Foltz, R. B. (2006). Spatially and temporally distributed modeling of landslide susceptibility. *Geomorphology, 80*, 178-198.

Gosh, D. (2015). Landslide susceptibility analysis from morphometric parameter analysis of Riyong Khola Basin, West Sikkim, India: a geospatial approach. *International Journal of Geology, Earth & Environmental Sciences, 5*(1), 54-65.

Gregory, K. J., & Walling, D. E. (1968). The variation of drainage density within a catchment. *International Association of Scientific Hydrology, 13*(2). doi:10.1080/02626666809493583

Guzzetti, F. (2006). *Landslide Hazard and Risk Assessment.* (P.h.D.), University of Bonn. Retrieved from https://geomorphology.irpi.cnr.it/Members/fausto/PhD-dissertation

Guzzetti, F., Carrara, A., Carrara, A., Cardinali, M., & Reichenbach, P. (1999). Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology, 31*, 181-216.

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K.-T. (2012). Landslide inventory maps: New tools for an old problem. *Earth-Science Reviews, 112*(1-2), 42-66. doi:10.1016/j.earscirev.2012.02.001

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., & Galli, M. (2006). Estimating the quality of landslide susceptibility models. *Geomorphology, 81*(1-2), 166-184. doi:10.1016/j.geomorph.2006.04.007

Hack, J. T. (1973). Stream profile analysis and streamgradient index. *Research U.S. Geol Survey, 1*(4), 421-429.

Hayakawa, Y. S., & Oguchi, T. (2006). DEM-based identification of fluvial knickzones and its application to Japanese mountain rivers. *Geomorphology, 78*(1-2), 90-106. doi:10.1016/j.geomorph.2006.01.018

Hengl, T., & MacMillan, R. A. (2009). Geomorphometry : a key to landscape mapping and modelling (1st ed., Vol. 33). Geomorphometry : concepts, software, applications: Elsevier.

Highland, L. (2004). *Landslide types and processes.* Retrieved from http://pubs.usgs.gov/fs/2004/3072/pdf/fs2004-3072.pdf

Highland, L. M., & Bobrowsky, P. (2008). *The Landslide Handbook - A guide to understanding landslides* (Vol. Circular 1325): USGS - US Geological Survey.

Hobson, R. D. (1972). Surface roughness in topography: quantitative approach. In R. J. Chorley (Ed.), *Spatial analysis in geomorphology* (pp. 221–245). New York, USA: Harper and Row.

Horlacher, H.-B., Heyer, T., Carstensen, D., Bielagk, U., Bielitz, E., & Müller, U. (2007). *Analysis of dyke breaks during the 2002 flood in Saxony:Germany*. Retrieved from

Horton, R. E. (1932). Drainage-basin characteristics *Transactions American Geophysical Union, 13*(1). doi:https://doi.org/10.1029/TR013i001p00350

Ivan Marchesini1 , M. R., 2 , Alessandro, & Mondini1 , M. S., Francesco Bucci. (2014, June 10-13). *Landslides morphometric signatures.* Paper presented at the OGRS Open Source Geospatial Research and Education Symposium Helsinki, Espoo, Finland.

Jhaldiyal, A., Gupta, K., Gupta, P. K., Reddy, S., & Kumar, P. (2016). Review of Methods for Estimating Urban Surface Roughness. *Vayu Mandal, 42*(2).

Jordan, G. (2004). *Terrain modelling with GIS for tectonic geomorphology. Numerical Methods and Applications.* Acta Universitatis Upsaliensis, Comprehensive summaries of Uppsala Dissertations. . (1031)

Kalantar, B., Pradhan, B., Naghibi, S. A., Motevalli, A., & Mansor, S. (2018). Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). *Geomatics, Natural Hazards and Risk, 9*(1), 49-69. doi:10.1080/19475705.2017.1407368

Keller, E. A., & Pinter, N. (2002). *Active Tectonics: Earthquakes, Uplift, and Landscape* (2nd ed.). New Jersey.

Kirby, E., & Whipple, K. X. (2001). Quantifying differential rock-uplift rates via stream profile analysis. *Geology, 29*(5), 415-418.

Kirby, E., & Whipple, K. X. (2012). Expression of active tectonics in erosional landscapes. *Journaal of Structural Geology, 44*, 54-75.

Klose, M., Auerbach, M., Herrmann, C., Kumerics, C., & Gratzki, A. (2017). Landslide Hazards and Climate Change Adaptation of Transport Infrastructures in Germany *Advancing Culture of Living with Landslides* (pp. 535-541).

Korn, M., Funke, S., & Wendt, S. (2008). Seismicity and seismotectonics of West Saxony - new insights from recent seismicity observed with the Saxonian seismic network. *Stud Geophys Geod, 52*(479). doi:https://doi.org/10.1007/s11200-008-0033-z

Krauter, E., Kumerics, C., Feuerbach, J., & Lauterbach, M. (2012). *Abschätzung der Risiken von Hang- und Böschungsrutschungen durch die Zunahme von Extremwetterereignissen.* Retrieved from

Lee, C. T., Huang, C. C., Lee, J. F., Pan, K. L., & Lin, M. L. (2008). Statistical approach to earthquake-induced landslide susceptibility. *Eng. Geology, 100*, 43-58.

Lee, S., Ryu, J.-H., Min, K., & Won, J.-S. (2003). Landslide susceptibility analysis using GIS and artificial neural network. *Earth Surface Processes and Landforms, 28*(12), 1361-1376. doi:10.1002/esp.593

Lee, S., & Talib, J. (2005). Probabilistic landslide susceptibility and factor effect analysis. *Environmental Geology, 47*(7), 982-990.

LfULG (Cartographer). (1875-1900). Geologische Karte des Freistaates Sachsen GK25. Retrieved from http://www.geologie.sachsen.de/geologische-karte-gk25-14093.html

LfULG (Cartographer). (2011). Bodenkarte des Freistaates Sachsen BK50 [Soil Map]. Retrieved from https://www.umwelt.sachsen.de/umwelt/boden/11639.htm

LfULG. (2018a). Bodenkarte (BK 50). Retrieved from https://www.umwelt.sachsen.de/umwelt/boden/11639.htm

LfULG. (2018b). *Erzgebirgsbecken (EGB).* Retrieved from https://www.umwelt.sachsen.de/umwelt/natur/22498.htm:

Lopes Queiroz, G., Salamuni, E., & Do Nascimento, E. R. (2015). Knickpoint finder: A software tool that improves neotectonic analysis. *Computers & Geosciences, 76*, 80-87. doi:10.1016/j.cageo.2014.11.004

Mahmood, S. A., & Gloaguen, R. (2012). Appraisal of active tectonics in Hindu Kush: Insights from DEM derived geomorphic indices and drainage analysis. *Geoscience Frontiers, 3*(4), 407-428. doi:10.1016/j.gsf.2011.12.002

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activities. *Biophys, 5*, 115-133.

Meusburger K, A. C. (2008). Impacts of anthropogenic and environmental factors on the occurrence of shallow landslides in an alpine catchment (Urseren Valley, Switzerland). *Nat Hazards Earth Syst Sci, 8*, 509-520. doi:10.5194/nhess-8-509-2008

Moglen, G. E., Eltahir, E. A. B., & Bras, R. L. (1998). On the sensitivity of drainage density to climate

change. *Water Resources Research, 34*(4), 855-862.

Moore, I. D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes, 5*, 3-30.

Moussi, A., Rebaï, N., Chaieb, A., & Saâdi, A. (2018). GIS-based analysis of the Stream Length-Gradient Index for evaluating effects of active tectonics: a case study of Enfidha (North-East of Tunisia). *Arabian Journal of Geosciences, 11*(6). doi:10.1007/s12517-018-3466-x

Mukherjee, S., Mukherjee, S., Garg, R. D., Bhardwaj, A., & Raju, P. L. N. (2013). Evaluation of topographic index in relation to terrain roughness and DEM grid spacing. *J. Earth Syst. Sci., 122*(3), 869-886.

Mumipour, M., & Nejad, H. T. (2011). Tectonics Geomorphology Setting of Khayiz Anticline Derived from GIS Processing, Zagros Mountains, Iran. *Asian Journal of Earth Sciences, 4*, 171-182. doi:10.3923/ajes.2011.171.182

Nir, D. (1957). The Ratio of Relative and Absolute Altitude of Mt. Carmel. *Geographical Review, 27*, 564-569.

Othman, A. A., Gloaguen, R., Andreani, L., & Rahnama, M. (2018). Improving landslide susceptibility mapping using morphometric features in the Mawat area, Kurdistan Region, NE Iraq: Comparison of different statistical models. *Geomorphology, 319*, 147-160. doi:10.1016/j.geomorph.2018.07.018

Pandey, V., Sharma, K., & Bandooni, S. (2018). Characteristics of Large Landslides and Application of Frequency Ratio Model for Susceptibility Assessment, Lower Jalal Catchment (Himachal Pradesh). *International Journal of Research in Geography (IJRG), 4*(1), 17-26. doi:10.20431/2454-8685.0401004

Pawley, S. M., Hartman, G. M. D., & Chao, D. K. (2017). *Examples of Landslides and the Geological, Topographic, and Climatic Factors That Contribute to Landslide Susceptibilities in Alberta: A Cross-Reference to the Explanatory Notes for AGS Map 605*. Retrieved from

Pérez-Peña, J. V., Azañón, J. M., & Azor, A. (2009). CalHypso: An ArcGIS extension to calculate hypsometric curves and their statistical moments. Applications to drainage basin analysis in SE Spain. *Computers & Geosciences, 35*(6), 1214-1223. doi:10.1016/j.cageo.2008.06.006

Pérez-Peña, J. V., Azañón, J. M., Booth-Rea, G., Azor, A., & Delgado, J. (2009). Differentiating geology and tectonics using a spatial autocorrelation technique for the hypsometric integral. *Journal of Geophysical Research, 114*(F2). doi:10.1029/2008jf001092

Pike, R. J., Evans, I. S., & Hengl, T. (2009). Chapter 1 Geomorphometry: A Brief Guide *Geomorphometry - Concepts, Software, Applications* (pp. 3-30).

Pike, R. J., & Wilson, S. E. (1971). Elevation-relief ratio, hypsometric integral, and geomorphic area-altitude analysis, Bull. *Geol. Soc, 82*, 1079-1084.

Pirasteh, S., Pradhan, B., & Rizvi, S. M. (2009). Tectonic process analysis in Zagros Mountain with the aid of drainage networks and topography maps dated 1950–2001 in GIS. *Arabian Journal of Geosciences, 4*(1-2), 171-180. doi:10.1007/s12517-009-0100-y

Poudyal, C. P., Chang, C., Oh, H. J., & Lee, S. (2010). Landslide susceptibility maps comparing frequency ratio and artificial neural networks: A case study from the Nepal Himalaya. *Environmental Earth Sciences, 61*(5), 1049-1064. doi:10.1007/s12665-009-0426-5

Pourghasemi, H. R., Pradhan, B., Gokceoglu, C., & Moezzi, K. D. (2013). A comparative assessment of prediction capabilities of Dempster–Shafer and Weightsof-evidence models in landslide susceptibility mapping using GIS. *Natural*

Hazards and Risk,, 4(2), 93-228. doi: 10.1080/19475705.2012.662915

Pradhan, B., & Lee, S. (2010). Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environmental Modelling and software, 25*(6), 747-759.

Qiqing, W., Wenping, L., Wei, C., & Hanying, B. (2015). GIS-based assessment of landslide susceptibility using certainty factor and index of entropy models for the Qianyang County of Baoji city, China. *J. Earth Syst. Sci., 124*(7), 1399-1415.

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. *Earth-Science Reviews, 180*, 60-91. doi:10.1016/j.earscirev.2018.03.001

Riley, S., DeGloria, S. D., & Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogenity. *Intermountain Journal of Sciences, 5*, 1-4.

Roşca, S., Bilaşco, S., Petrea, D., Fodorean, I., Vescan, I., & Filip, S. (2015). Large scale landslide susceptibility assessment using the statistical methods of logistic regression and BSA – study case: the sub-basin of the small Niraj (Transylvania Depression, Romania). *Nat. Hazards Earth Syst. Sci., 3*, 7171-7201. doi:10.5194/nhessd-3-7171-2015

Sappington, J. M., Longshore, K. M., & Thompson, D. B. (2007). Quantifying Landscape Ruggedness for Animal Habitat Analysis: A Case Study Using Bighorn Sheep in the Mojave Desert. *Journal of Wildlife Management, 71*(5), 1419-1426. doi:10.2193/2005-723

Schneider, J. W., Hoth, K., Gaitzsch, B. G., Berger, H. J., Steinborn, H., Walter, H., & Zeidler, M. K. (2005). Carboniferous stratigraphy and development of the Erzgebirge Basin, East Germany. *Zeitschrift der Deutschen Gesellschaft für Geowissenschaften, 156*(3), 431-466. doi:10.1127/1860-1804/2005/0156-0431

Schwanghart, W., & Scherler, D. (2014). Short Communication: TopoToolbox 2 – MATLAB-based software for topographic analysis and modeling in Earth surface sciences. *Earth Surface Dynamics, 2*(1), 1-7. doi:10.5194/esurf-2-1-2014

Shahabi, H., & Hashim, M. (2015). Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment. *Sci Rep, 5*, 9899. doi:10.1038/srep09899

Shahzad, F., & Gloaguen, R. (2011a). TecDEM: A MATLAB based toolbox for tectonic geomorphology, Part 1: Drainage network preprocessing and stream profile analysis. *Computers & Geosciences, 37*(2), 250-260. doi:10.1016/j.cageo.2010.06.008

Shahzad, F., & Gloaguen, R. (2011b). TecDEM: A MATLAB based toolbox for tectonic geomorphology, Part 2: Surface dynamics and basin analysis. *Computers & Geosciences, 37*(2), 261-271. doi:10.1016/j.cageo.2010.06.009

Strahler, A. N. (1957). Quantitative analysis of watershed geomorphology. *Transactions, American Geophysical Unit, 38*(6).

Syrbe, R.-U., Grunewald, K., Mannsfeld, K., Slobodda, S., Bastian, O., & Klenke, F. (2014). *Naturräumliche Potenziale, Empfindlichkeiten und Landnutzung im Freistaat Sachsen*. Retrieved from http://www.umwelt.sachsen.de/umwelt/natur/22499.htm

Troiani, F., Galve, J. P., Piacentini, D., Della Seta, M., & Guerrero, J. (2014). Spatial analysis of stream length-gradient (SL) index for detecting hillslope processes: A case of the Gállego River headwaters (Central Pyrenees, Spain). *Geomorphology, 214*, 183-197. doi:10.1016/j.geomorph.2014.02.004

Tsangaratos, P., & Benardos, A. (2014). Estimating landslide susceptibility through a artificial neural network classifier. *Natural Hazards, 74*(3), 1489-1516. doi:10.1007/s11069-014-1245-x

Tsangaratos, P., & Bernardos, A. (2013). *APPLYING ARTIFICIAL NEURAL NETWORKS IN SLOPE STABILITY RELATED PHENOMENA*. Chania, Crete, Greece.

Van Den Eeckhaut M, P. J., Van Gils M, Van Rompaey A, Vandekerckhove L. (2009). How do humans interact with their environment in residential areas prone to landsliding? A case-study from the Flemish Ardennes. In: Proceedings of the international conference on "landslide processes: from geomorphologic mapping to dynamic modelling [Press release]

Van Westen, C. J., Castellanos, E., & Kuriakose, S. (2008). Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Eng. Geology, 102*(3-4), 112-131.

Van Weston, C. J., Rengers, N., & Soeters, R. (2003). Use of Geomorphological Information in Indirect Landslide Susceptibility Assessment. *Natural Hazards, 30*(3), 399-419. doi:10.1023/B:NHAZ.0000007097.42735.9e

Vanacker V, V. M., Govers G, Willems E, Poesen J, Deckers J, De Bievre B. (2003). Linking hydrological, infinite slope stability and land-use change models through GIS for assessing the impact of deforestation on slope stability in high Andean watersheds. *Geomorphology, 52*, 299-315. doi:10.1016/S0169-555X(02)00263-5

Varnes, D. J. (1978). *Slope Movement Types and Processes*. Retrieved from National Academy of Sciences:

Wilson, J. P. (2012). Digital terrain modeling. *Geomorphology, 137*(1), 107-121. doi:10.1016/j.geomorph.2011.03.012

Woods, K. S., & Bowyer, K. W. (1994). *Generating ROC curves for artificial neural networks*. Winston-Salem, NC, USA.

Yalcin, A., & Bulut, F. (2007). Landslide susceptibility mapping using GIS and digital photogrammetric techniques: A case study from Ardesen (NE Turkey). *Nat Hazards, 41*(1), 201-226.

Yilmaz, I. (2009). The effect of the sampling strategies on the landslide susceptibility mapping by conditional probability and artificial neural networks. *Environmental Earth Sciences, 60*(3), 505-519. doi:10.1007/s12665-009-0191-5

Zahra, T., Paudel, U., Hayakawa, Y. S., & Oguchi, T. (2017). Knickzone Extraction Tool (KET) – A new ArcGIS toolset for automatic extraction of knickzones from a DEM based on multi-scale stream gradients. *Open Geosciences, 9*(1). doi:10.1515/geo-2017-0006

Zhou, W. (1999). Verification of the nonparametric characteristics of back propagation neural networks for image classification. *IEEE Transactions on Geoscience and Remote Sensing, 37*, 771-779.

9. Appendix

Source	Factor	GIS Layer	Modelled?	Layer Type	Comments
Geology Map 1:25,000	Geology - Hard rock and sediments	GK 1 - Hard Rock	Y	Binary raster	All hard rock types
		GK 2 - Sediment	N		Unconsolidated loose sediments - fine grained
		GK 3 - Sediment	N		Unconsolidated loose sediments - mixed grained
		GK 4 - Sediment	N		Unconsolidated loose sediments - coarse grained
		KL - Fault Susceptibility	Y		Based on hard rock classifications
		TF - Joint, cleavage prone	Y		Based on hard rock classifications
	Fault Lines	Crossing buffer	N	Continuous raster	Zones where geological lineaments cross
		80m buffer	Y		Area of influence of an individual fault-line
Soil Map 1:25,000	Soil	Type 1	Y	Binary raster	Anthropogenic : Construction materials
		Type 2	Y		Mixed loose & hard rock
		Type 3	N		Alluvial soils
		Type 4	N		Gleye: wetland soils
		Type 5	N		Moore & Peat soils
		Type 9	N		Black Earth
		Type 10	Y		Water logged soil
		Type 12	Y		Brown peri glacial earth
		Type 14	N		Fluvial hummus soil
Landscape Map 1:25,000	Land-cover	Percentage of vegetation	N	Continuous raster	Not used in the modelling, the majority of the LS from the inventory occur in 90-100% vegetated areas, this would have created too much positive weighting towards forested areas in the model.
		Percentage of sealed ground	N		
DEM – Geo-morphological	Curvature	Profile Positive	Y	Continuous raster	Curvature describes the convex/concave shape of the slope and is an important landslide parameter.
		Profile Negative	Y		
		Plan Positive	Y		
		Plan Negative	Y		
	Slope	Slope	Y		Slope throughout the literature has always been considered an important parameter in LSS.
	Aspect	Aspect N,NE,E,SE,S,SW,W,NW.	N	Binary raster	No correlation found and excluded from the final model.
	Flow Accumulation	Flow Accumulation (Log)	Y	Continuous raster	Drainage patterns are important in LS modelling
	Elevation	Elevation	N		Traditionally elevation is seen as an important parameter, however the elevation range within the study are is relatively low and no correlation was made during modelling.
	Roughness	VRM 5m	N		The VRM 9x9m kernel was assessed as the most stable value range and was selected.
		VRM 9m	Y		
		VRM 25m	N		
		TRI 9m	Y		The TRI 9x9m kernel was assessed as the most stable value range and was selected.
		TRI 5m	N		
		TRI 3m	N		
DEM – Hydrological	Hpysommetric Index	HI 25m	N		The HI 15m kernel was assessed as the most stable value range and was selected.
		HI 15m	Y		
		HI 5m	N		
	Dissection Index	DI 15m	N		The DI 9m kernel was assessed as the most stable value range and was selected.
		DI 9m	Y		
		DI 5m	N		
	Drainage Density	DD 500m	Y		The DD 500m buffer was selected as the most stable value range.
		DD 100m	N		
	Stream Power Index	SPI	Y		Closely linked to erosion processes
	Compound topographic Index	TWI/CTI	Y		Closely linked to accumulated water
	Stream Length Index (RDE)	Knickpoints 20m	N		50m knickpoint segments were selected and the density calculated using 250m buffers. The 100m segments were not sensitive enough to possible knickpoints and 20m segments proved to be too sensitive and selected very localised knickpoints
		Knickpoints 50m	Y		
		Knickpoints 100m	N		

Appendix 1. GIS raster layers used as part of the study.

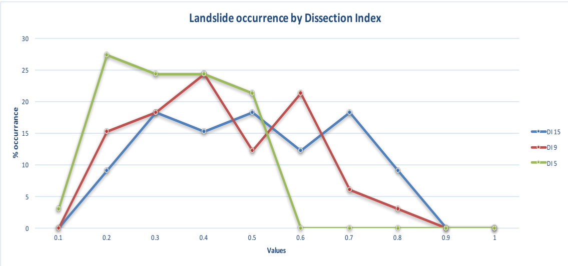
INFORMATION									
ID	X	Y	Z	Source	Train/ Val	Landform	Landuse	Comments	Failure Class
1	747,347.49	5,622,564.78	278.1743	Field Points	Training	slope	dump	Old mining dump in Zwickau above federal highway, small walls of slag for terracing, heap in Bewegu	Landslide
2	747,874.29	5,622,973.92	269.098	Field Points	Training	slope	dump	old heap in the city area Zwickau, creeping processes prevailing, erosion gullies from time without vegetation	Landslide
3	747,414.46	5,622,676.29	285.9836	Inventory	Training	Slope	dump	Failure on mining waste dump	Landslide
4	752,086.58	5,617,242.72	306.3311	Field Points	Training	slope	forest	steep slope under castle, loose debris on slope, evidence of slope creep and bending of trees.	Landslide
5	757,145.04	5,616,455.95	317.513	Field Points	Validation	slope	forest	beside stream gully, evidence of movement due to high flowing water, fresh debris imovement	Landslide
6	756,989.85	5,616,463.43	313.642	Field Points	Training	slope	road	Rocky cliff on road just before Langenbach, Phyllite pending, rocks shattered and fall prone	Rockfall
7	757,160.32	5,616,453.17	314.6923	Field Points	Training	valley	forest	Side valley to the Zwickauer Mulde, intensive erosion by water	Landslide
8	756,886.93	5,616,432.99	322.0667	Field Points	Validation	slope	forest	In forest near Langenbach, large landslide	Landslide
9	747,301.81	5,624,670.77	276.7542	Field Points	Training	slope	forest	wooded slope above Zwickauer Mulde in Zwickau	Landslide
10	752,303.80	5,617,150.95	291.9533	Field Points	Validation	slope	forest	As above, Steilhaunterhalb Castle	Rockfall
11	738,519.64	5,624,263.11	292.2506	Field Points	Training	slope	forest	Steep slope near Werdau, intensive channeling and mass transportation	Landslide
12	738,597.96	5,626,380.03	274.9482	Field Points	Training	slope	forest	Steep hillside road in Werdau, entire slope in motion	Landslide
13	749,525.44	5,619,377.38	282.2066	Field Points	Training	slope	road	Rockfall and debris at base of slope underneath the autobahn bridge, Wilkau-Hasslau,	Rock Fall
14	747,268.09	5,624,715.21	265.8589	Field Points	Validation	slope	forest	soil and trees have fanned out from over step area, clay rich soil	Landslide
15	744,328.49	5,619,908.76	347.1886	Field Points	Validation	slope	forest	outcropping rock, change in slope texture, fall to sw, possible movement on slope	Landslide
16	738,650.89	5,623,971.09	299.2048	Field Points	Validation	valley	forest	massively cut side valley with lots of water	Washout
17	756,868.98	5,616,431.53	323.348	Field Points	Training	slope	forest	large slide, rock and soil, fan deposited into terraces, steep scarp face with insitu rock.	Landslide
18	746,692.02	5,619,971.26	345.7381	Field Points	Training	slope	village	Steep slope in the center of Zwickau, no endangerment, but the slope is clear, but it is not shown here	Landslide
19	747,624.31	5,623,238.11	266.4738	Field Points	Training	slope	grassland	Slope above roadway	Landslide
20	747,425.74	5,623,565.77	271.0089	Field Points	Validation	slope	shrubs	Steep slope above Mulden-Aue, above slope development	Landslide
21	740,165.52	5,619,278.59	364.3956	Field Points	Training	hill	forest	Steep slope under Schönfels Castle, Metasediment, many large loose rocks	Rockfall
22	739,525.21	5,620,465.01	325.1195	Field Points	Training	slope	forest	slope failure on steep slope, rock outcrop upslope, looks like progressive sliding, matches vrm data	Landslide
23	744,454.77	5,619,948.12	362.8653	Field Points	Training	hill	grassland	long slope, break in vegetation, apex originating from top of slope, Fault Line?	Landslide
24	737,814.91	5,626,433.06	300.0803	Inventory	Training	hill	grassland	slope on corner of road intersection	Landslide
25	742,419.61	5,620,531.56	345.36	Inventory	Training	slope	Quarry	Man made cause, over-steepened from quarrying, engineering methods used, house at base of slope	Rockfall
26	737,727.64	5,627,060.88	287.3519	Inventory	Validation	hill	Infrastructure	train track failure	Landslide
27	749,920.57	5,613,787.81	367.9903	Field Points	Training	slope	forest	steep slope behind houses, evidence of engineering methods at preserving the slope, some houses abandoned	forest
28	749,828.22	5,613,028.29	394.9484	Field Points	Training	slope	forest	very steep slope section, forested, man made feature from quarrying	Rock fall
29	750,566.06	5,617,991.90	304.911	Field Points	Training	slope	forest	exposed rock outcrop slope beside train track, slope around is forested.	Rock Fall
30	750,790.17	5,618,799.80	296.068	Field Points	Training	slope	forest	At corner of bend in river, steep section down to river, housing estate at top of slope, forested	Landslide
31	749,174.13	5,618,156.53	309.5271	Field Points	Training	slope	forest	River at base of slope, steep forested slope, flat pastureland at base	Landslide
32	747,568.26	5,621,958.85	291.0931	Field Points	Training	slope	forest	On steep slope above housing estate, underneath school	Landslide
33	758,917.34	5,614,124.42	456.1094	Field Points	Training	hill	grassland	Partly vegetated slope above forest. curved	Landslide

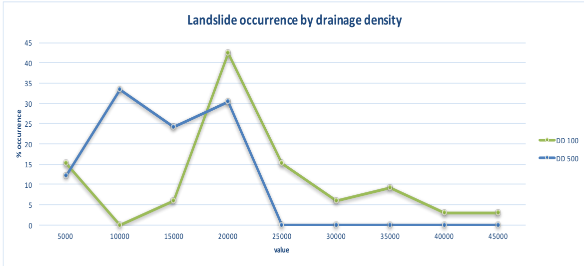
Appendix 2 Field work data per inventory point

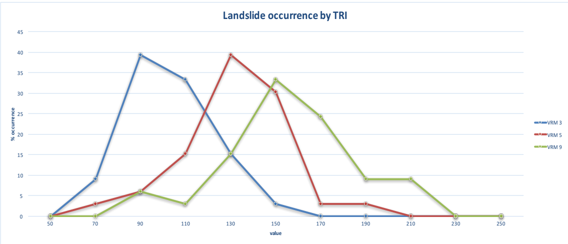
	GIVEN PARAMETERS						EXTRACTED FROM DEM																					
ID	GK	KL	TF	GEOLIN	SOIL	VEG	ID	SLOPE	FLACC	CPLAN	CPRO	VRM 15m	VRM 9m	VRM 5m	TRI 9m	TRI 5m	TRI 3m	HI 25m	HI 15m	HI 9m	KNICKDEN	SPI	CTI	DD 500m	DD 100m	DI 15m	DI 9m	DI 5m
1	5	0	0	0.530	1	100	1	27.886	1.445	1.281	1.394	0.047	0.039	0.021	146.858	118.174	85.845	0.312	0.342	0.383	0.080	1.704	4.119	11747.100	17308.801	0.631	0.580	0.386
2	4	0	0	0.937	4	100	2	25.867	1.413	-0.145	2.789	0.076	0.043	0.027	141.186	103.646	72.548	0.354	0.328	0.335	0.080	2.608	4.980	9864.510	15461.800	0.733	0.634	0.352
3	5	0	0	0.684	1	100	3	35.213	1.547	-0.563	-0.589	0.030	0.020	0.005	153.793	133.043	99.712	0.409	0.498	0.549	0.080	3.197	4.870	13325.900	17580.000	0.582	0.564	0.444
4	1	1	2	0.379	12	100	4	41.579	1.619	0.876	1.543	0.055	0.065	0.035	150.193	137.211	110.134	0.303	0.345	0.382	0.398	0.044	3.208	16249.700	37595.602	0.434	0.391	0.340
5	1	1	2	0.000	4	100	5	23.659	1.374	0.679	0.694	0.030	0.035	0.030	128.506	106.167	81.635	0.499	0.500	0.502	0.482	2.658	5.554	18227.000	20972.301	0.300	0.245	0.176
6	1	1	2	0.662	10	100	6	26.572	1.424	1.680	0.960	0.029	0.019	0.014	145.673	114.196	85.285	0.406	0.368	0.367	0.566	1.804	4.329	18344.801	41901.500	0.395	0.330	0.212
7	4	0	0	0.000	4	100	7	11.068	1.044	-0.609	2.680	0.028	0.033	0.032	117.853	88.477	64.591	0.501	0.511	0.502	0.488	3.570	9.221	18321.500	20312.699	0.293	0.217	0.119
8	1	1	2	0.117	2	100	8	34.288	1.535	0.690	-1.240	0.038	0.038	0.024	156.335	125.579	92.603	0.438	0.415	0.426	0.625	1.513	3.652	17764.301	30553.600	0.403	0.355	0.236
9	1	2	2	0.000	3	100	9	38.077	1.581	0.080	1.213	0.029	0.028	0.015	139.525	123.741	96.164	0.472	0.501	0.509	0.080	2.461	4.130	11911.200	16890.900	0.612	0.527	0.465
10	3	0	0	0.000	5	21	10	27.677	1.442	-0.755	5.666	0.038	0.035	0.026	158.066	123.418	89.442	0.371	0.319	0.309	0.410	1.630	4.755	17236.100	18780.500	0.658	0.504	0.333
11	4	0	0	0.000	2	100	11	22.890	1.360	-2.033	3.356	0.051	0.066	0.066	146.444	111.431	85.411	0.547	0.534	0.492	0.080	1.923	6.196	7508.700	14730.600	0.604	0.427	0.245
12	1	2	2	0.000	1	82	12	33.418	1.524	-1.610	2.498	0.023	0.021	0.019	129.617	111.299	88.345	0.408	0.397	0.412	0.159	2.519	4.415	8898.070	15515.100	0.597	0.473	0.363
13	3	0	0	0.000	10	100	13	16.934	1.229	0.545	4.258	0.055	0.059	0.052	178.200	128.478	79.561	0.325	0.320	0.317	0.000	2.045	5.690	15670.000	26737.000	0.797	0.759	0.406
14	3	0	0	0.122	3	100	14	28.151	1.449	0.308	3.028	0.038	0.036	0.023	134.954	116.438	84.037	0.403	0.405	0.400	0.080	1.784	4.204	13053.600	18088.801	0.653	0.599	0.494
15	4	0	0	0.591	4	100	15	23.529	1.372	0.666	2.497	0.028	0.031	0.021	143.586	118.661	86.832	0.399	0.363	0.370	0.176	2.079	4.697	7090.830	16173.400	0.233	0.222	0.161
16	1	2	2	0.000	2	100	16	30.016	1.477	-3.042	1.352	0.026	0.023	0.012	167.104	128.340	94.670	0.483	0.479	0.473	0.080	4.403	6.268	10199.300	15832.700	0.666	0.513	0.321
17	1	1	2	0.000	2	100	17	36.706	1.565	-0.781	0.359	0.038	0.035	0.021	161.845	132.752	98.297	0.447	0.450	0.467	0.633	2.122	3.906	17794.100	31218.600	0.409	0.374	0.264
18	1	2	1	0.303	10	100	18	29.421	1.469	-0.217	-0.897	0.015	0.009	0.006	170.216	133.460	99.746	0.548	0.551	0.556	0.243	3.087	5.080	4402.380	15113.500	0.419	0.320	0.201
19	1	1	2	0.640	3	99	19	23.491	1.371	0.538	3.178	0.026	0.028	0.022	121.032	106.601	77.518	0.409	0.412	0.407	0.000	2.503	5.057	16137.200	27023.699	0.518	0.483	0.409
20	3	0	0	0.065	3	100	20	32.283	1.509	-0.140	0.251	0.019	0.024	0.021	104.772	99.338	83.393	0.450	0.492	0.502	0.000	1.961	4.054	16614.801	22612.801	0.378	0.348	0.312
21	1	2	1	0.442	2	100	21	35.516	1.550	-0.575	0.578	0.030	0.015	0.007	171.240	142.743	110.361	0.363	0.393	0.434	0.140	1.492	3.604	13165.600	24126.100	0.327	0.264	0.188
22	1	2	1	0.323	2	100	22	36.143	1.558	-0.266	1.273	0.038	0.043	0.021	152.976	130.127	99.947	0.418	0.378	0.369	0.080	2.315	4.079	7385.030	16637.199	0.339	0.318	0.239
23	1	2	1	0.000	2	100	23	16.705	1.223	2.018	-0.368	0.010	0.013	0.012	122.061	93.460	72.147	0.469	0.536	0.533	0.159	-0.434	4.133	7536.890	13674.200	0.229	0.136	0.084
24	2	0	0	0.000	9	100	24	10.945	0.000	1.346	-1.204	0.011	0.007	0.004	82.382	71.072	54.194	0.557	0.553	0.534	0.159	-3.372	4.982	8293.940	4873.030	0.167	0.128	0.106
25	1	2	1	0.000	2	50	25	42.818	1.632	5.916	0.751	0.036	0.046	0.066	135.418	127.577	108.541	0.375	0.411	0.443	0.104	-1.439	2.997	9239.270	23996.301	0.237	0.200	0.182
26	4	0	0	0.000	14	80	26	10.858	1.036	-0.863	-0.230	0.010	0.010	0.010	74.244	64.972	51.033	0.585	0.656	0.625	0.080	-1.940	5.041	1315.640	0.000	0.194	0.129	0.113
27	1	1	1	0.000	10	78	27	38.861	1.590	-0.149	0.185	0.035	0.034	0.027	147.824	136.463	113.644	0.357	0.368	0.408	0.389	1.689	3.548	12109.400	30102.100	0.224	0.192	0.166
28	1	1	1	0.000	1	100	28	49.138	1.691	-0.387	-1.067	0.058	0.065	0.028	198.333	185.634	146.795	0.509	0.526	0.536	0.398	-0.342	3.477	6147.980	0.000	0.290	0.274	0.251
29	1	1	2	0.000	12	100	29	40.881	1.612	0.014	-0.343	0.038	0.029	0.010	169.822	145.516	110.588	0.498	0.511	0.510	0.147	3.730	5.063	11220.800	17070.000	0.568	0.504	0.396
30	1	1	2	0.743	12	100	30	33.852	1.530	1.329	0.965	0.029	0.037	0.032	131.845	120.935	97.775	0.498	0.532	0.542	0.221	4.450	6.069	9226.520	15317.700	0.409	0.351	0.296
31	1	1	2	0.000	12	100	31	42.540	1.629	-1.715	5.740	0.039	0.027	0.011	202.526	163.564	118.806	0.409	0.407	0.416	0.000	3.594	4.812	8429.370	17743.000	0.748	0.649	0.455
32	1	2	1	0.509	10	50	32	33.563	1.526	-1.091	1.434	0.032	0.031	0.011	145.407	132.144	101.370	0.511	0.486	0.489	0.099	2.497	4.380	2299.450	0.000	0.499	0.455	0.401
33	5	0	0	0.468	1	100	33	26.005	1.415	-1.142	-0.041	0.021	0.015	0.009	190.224	147.636	108.353	0.667	0.601	0.547	0.591	3.429	5.639	3924.760	0.000	0.199	0.166	0.106

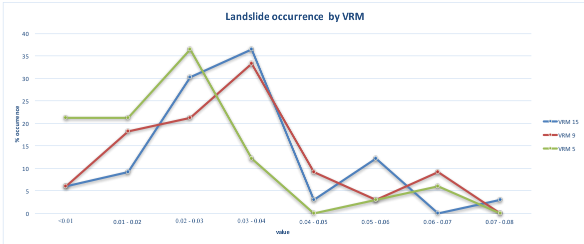
Appendix 3 Parameter data per inventory point

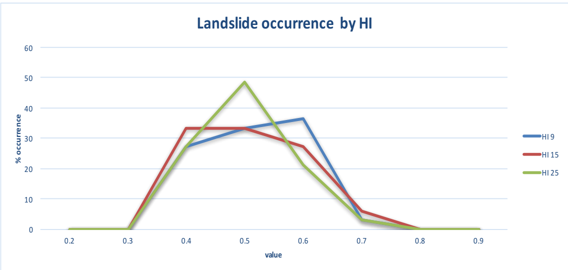
Appendix 4 Landslide occurrence plots per standard parameter from inventory

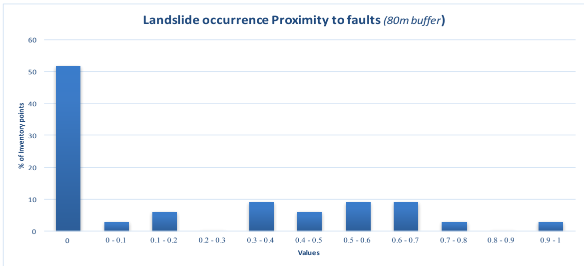


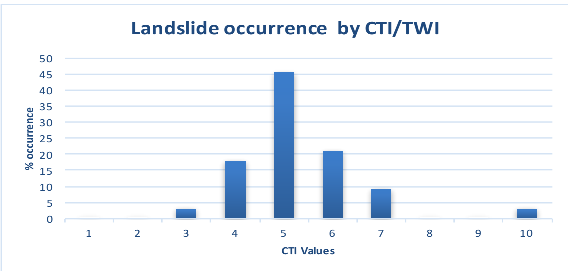


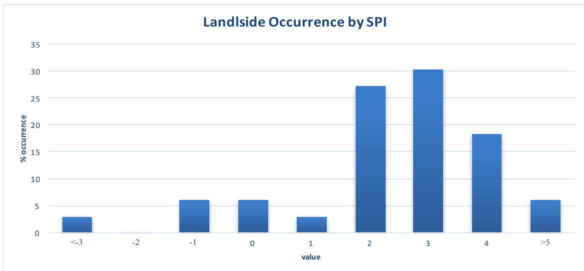




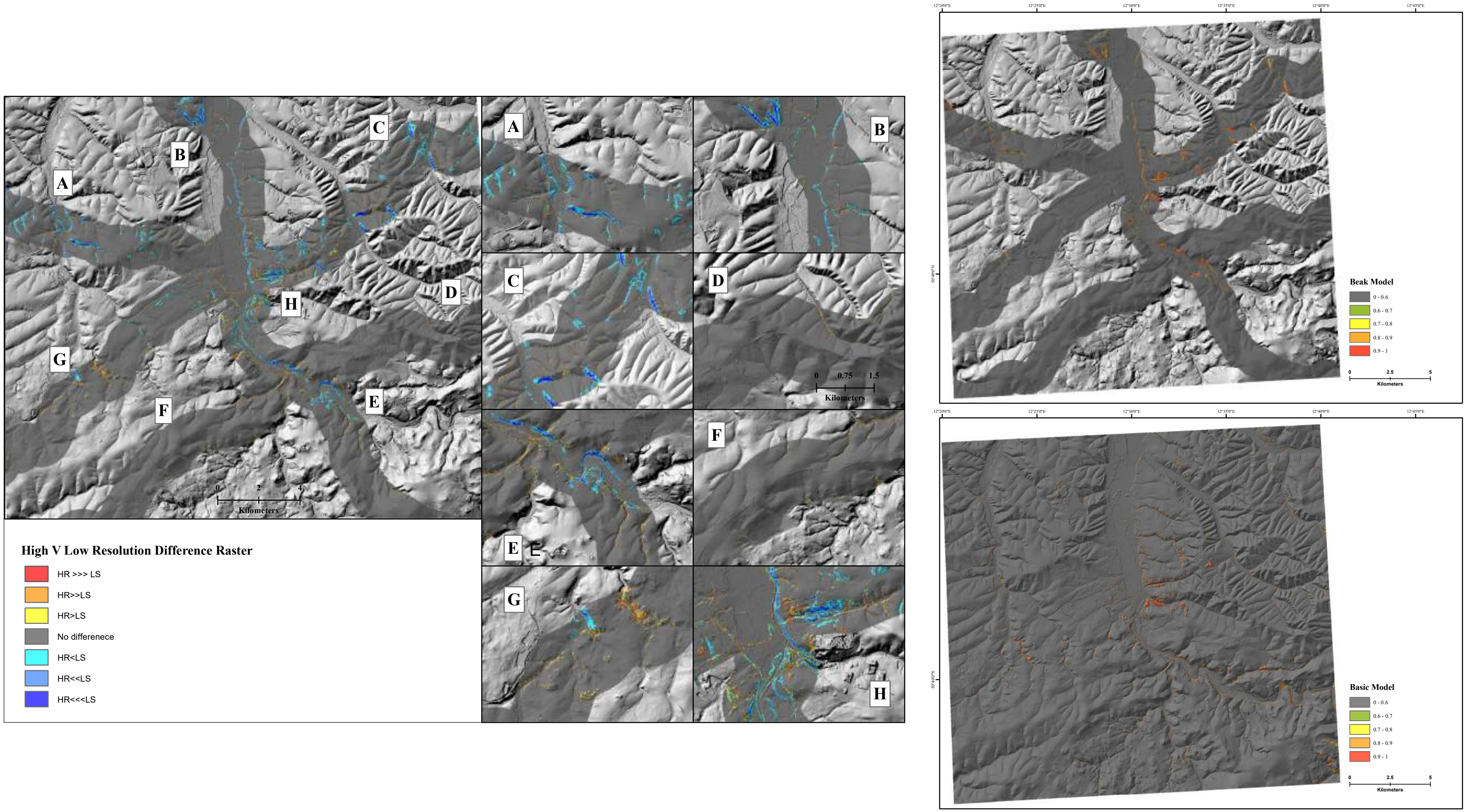




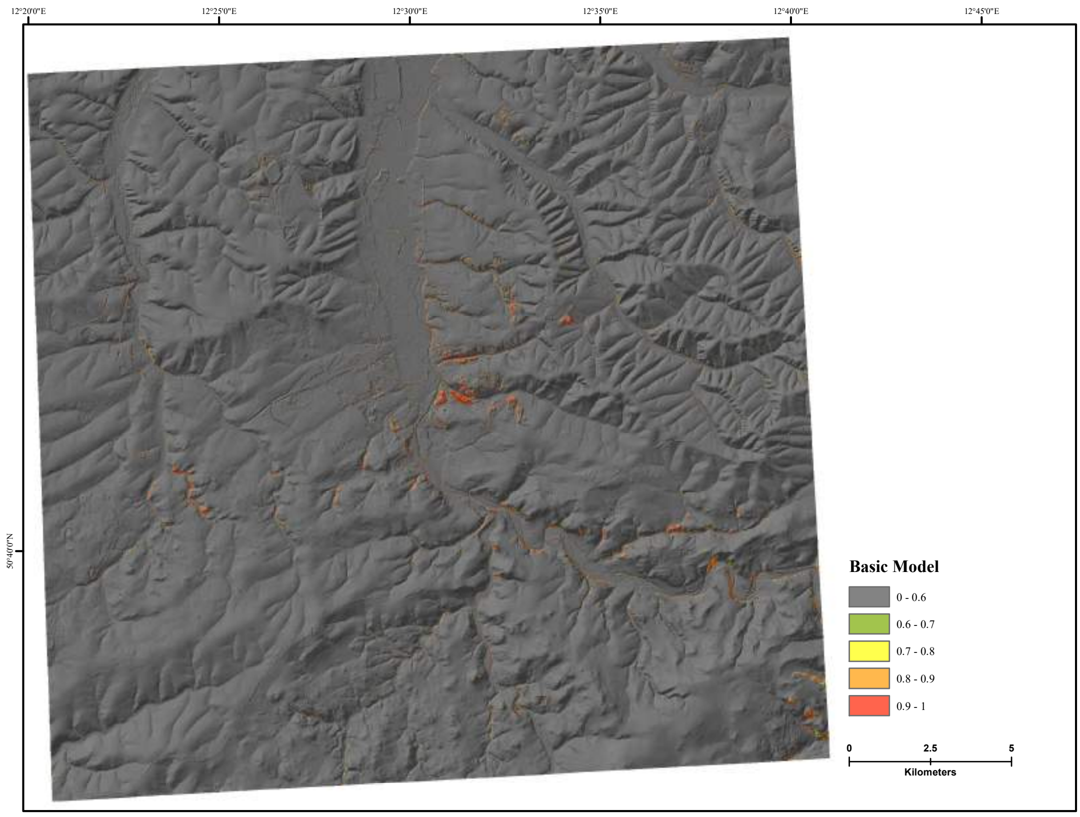




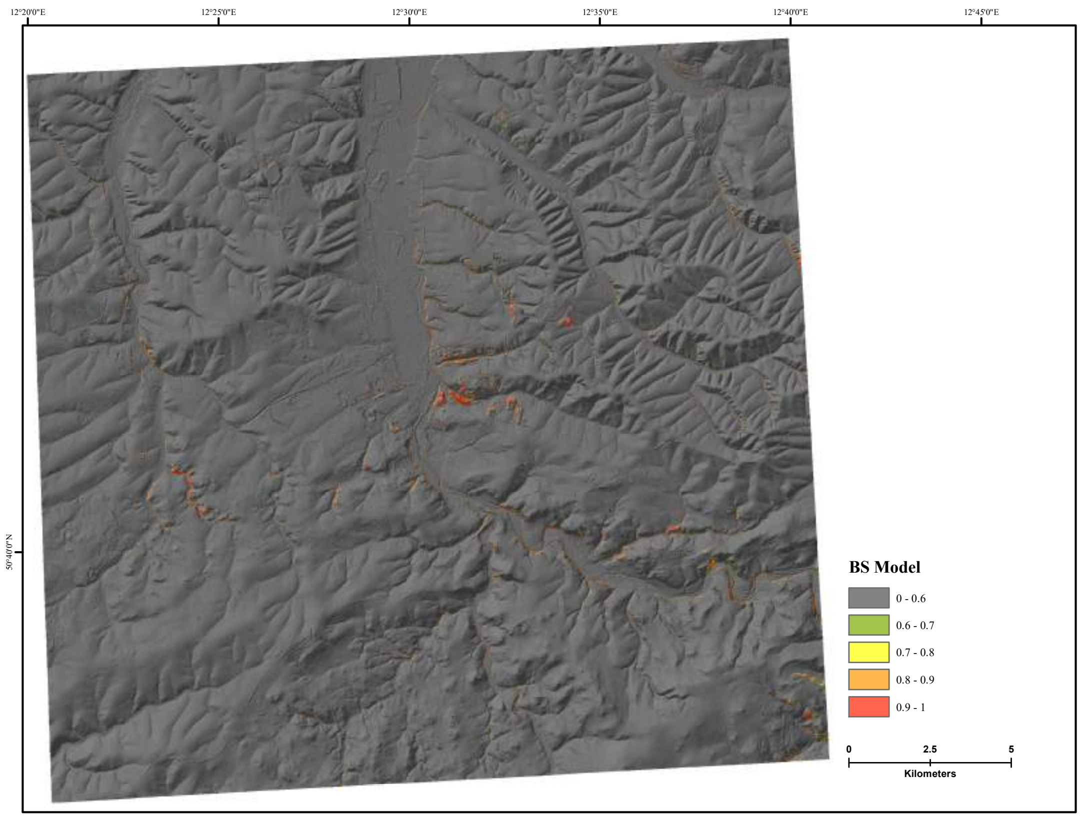
Appendix 5 Landslide occurrence plots per new parameters from inventory



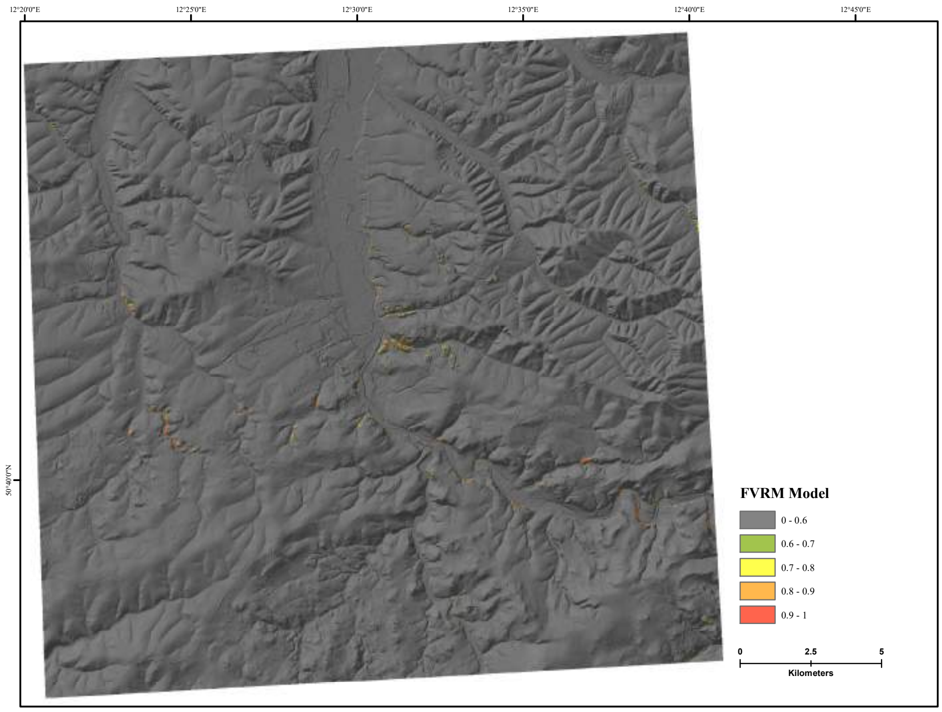
Appendix 6 Original Beak model in low resolution compared to basic higher res



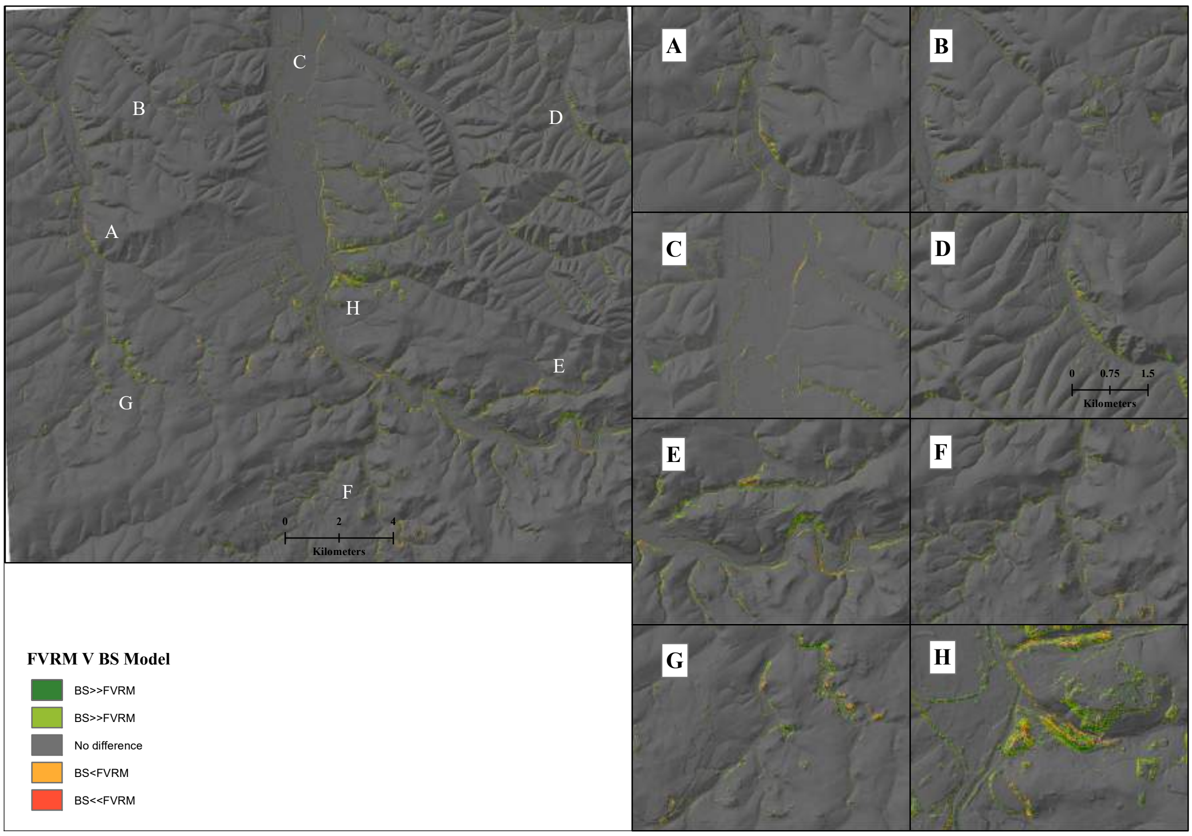
Appendix 7 Basic prediction model



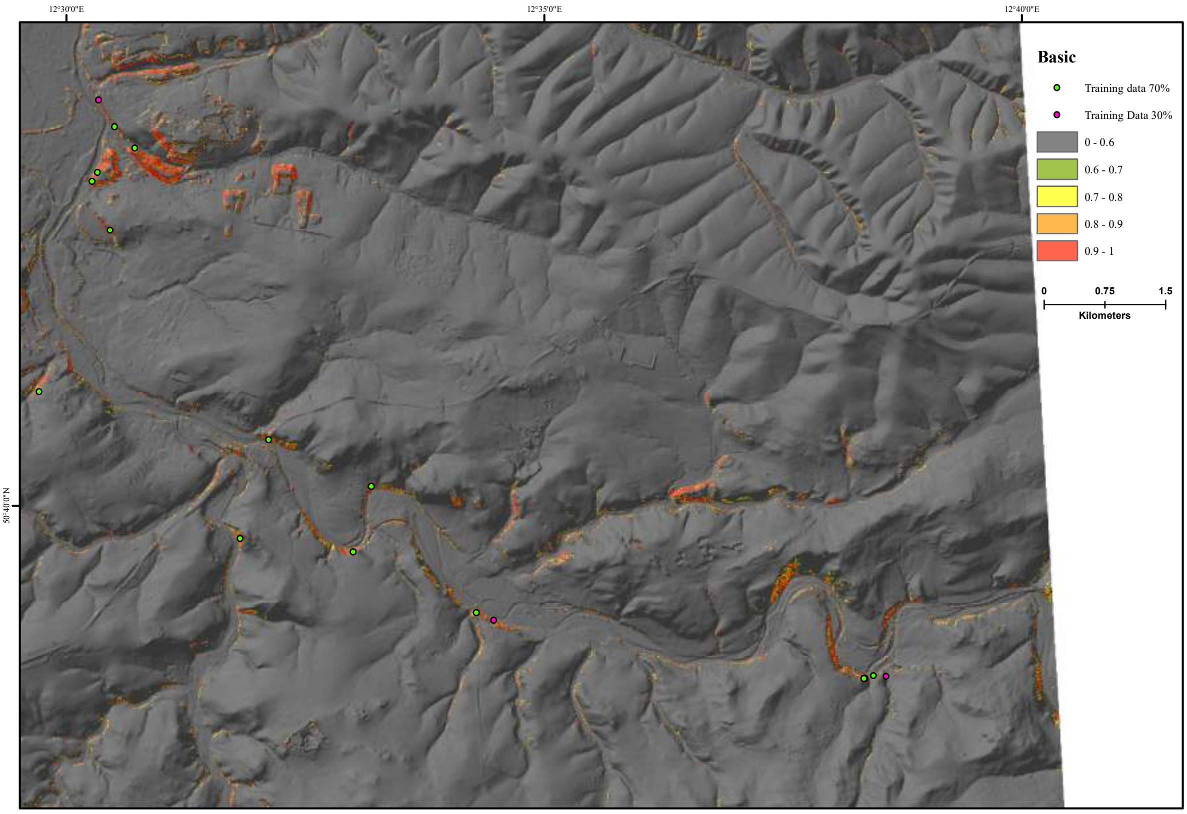
Appendix 8 BS prediction model

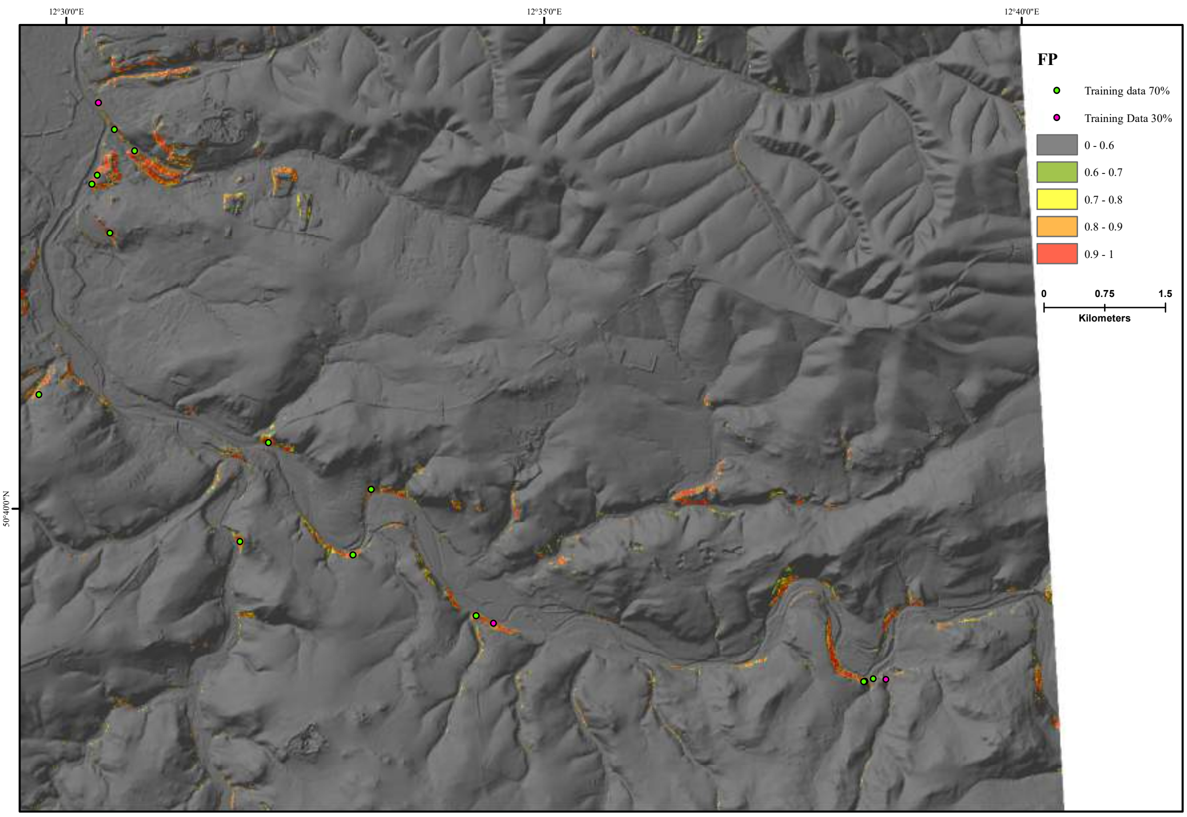


Appendix 9. FVRM prediction model

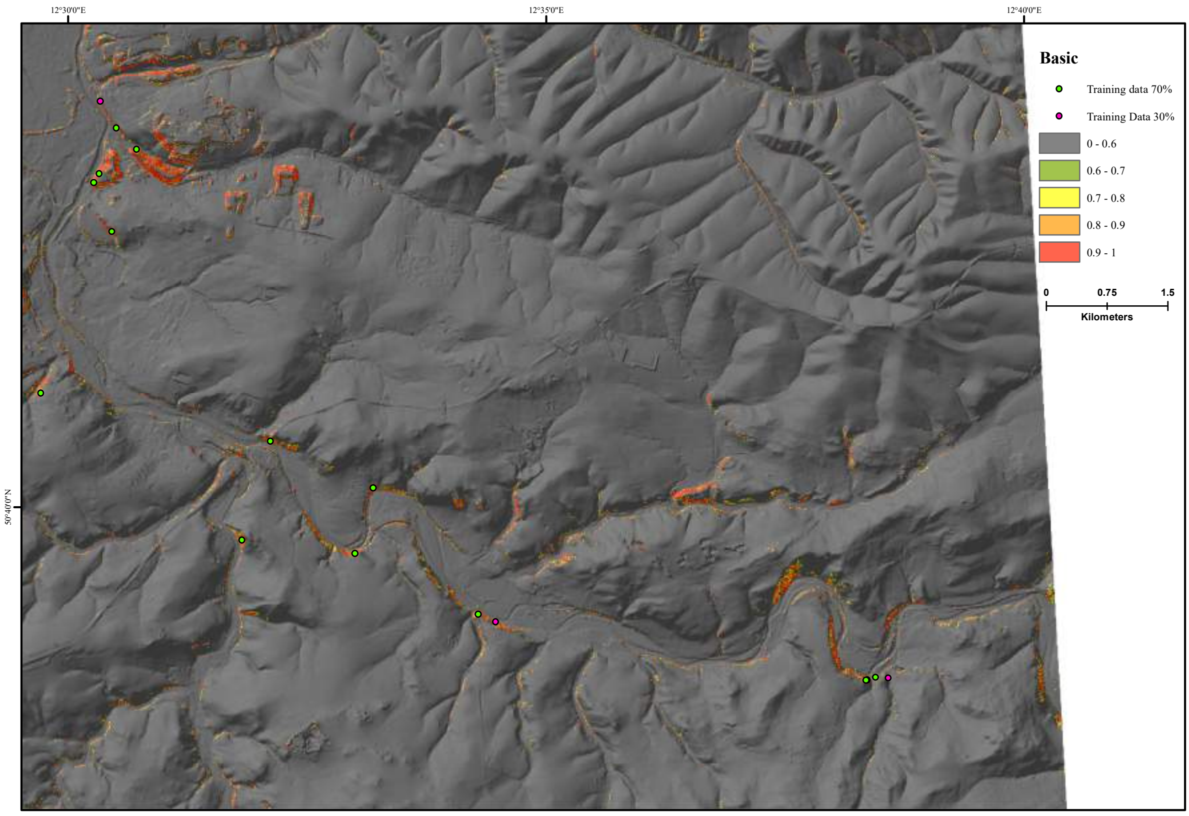


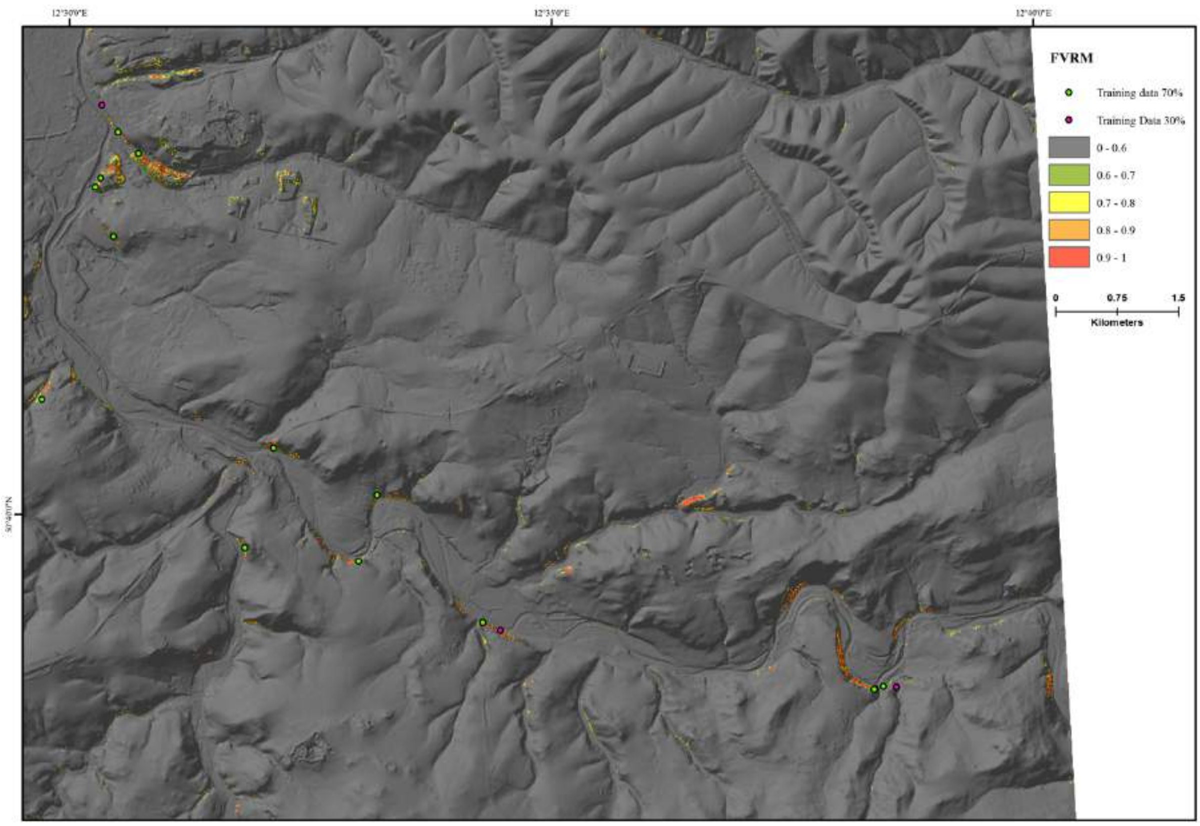
Appendix 10. FVRM - BS difference map





Appendix 11. Close up of E side of study area, basic & FP models.





Appendix 12. Basic, FVRM close up map