

## Thesis Landslide Forecasting Hollin Hill

Name student: Eduard Schoor

Student number: 4002040

Study: Water Science and Management (Msc)

Supervisors: drs. Jelle Buma (UU / Deltares), dr. Rens van Beek (UU)



(Image taken from BGS, 2014)

## Abstract

This master thesis main purpose was to assess regional forecasting of landslides around a known landslide location: Hollin Hill in the Howardian Hills, Yorkshire, United Kingdom. The idea was that since this is a known landslide location, the location could serve as a validation tool for the stability model used (PROBSTAB, which is coupled to the hydrological model STARWARS). Since the catchment of Hollin Hill was rather large and no river flow data was available for simulation, a nearby catchment (river Foss at Huntington) was used for validation of input parameters and Hollin Hill added as a sub catchment to this neighboring catchment. The model was first run on 100 m cell size for the period of 1988-2012 with historic rainfall and temperature data from MIDAS, to establish the best parameter combination. The model was then re-run on 25 m cell size with the optimal parameter combination. The modeled daily river flow was compared to the observed from the Centre of Ecology and Hydrology and its performance tested with the Nash-Sutcliffe Efficiency (NSE) score. The optimal score on 25 m was rather low: 0.21. This can be attributed to many factors: inaccurate rainfall patterns (all but one of the 8 rainfall stations used were outside of the catchment), recession constant and/or hydraulic conductivity of the bedrock too small which caused a lack of recession in the modeled river flow were likely the main factors however. The model was re-run for forecast purposes with a 8-day forecast of the European Centre for Medium Weather Forecasts (ECMWF). It was found that after two days the forecasts became rather unreliable in terms of safety factor, however these forecasts were uncalibrated. Possible calibration of forecasts is by means of calibrating them against accurate historic rainfall data from gauges, or against radar data. Another option is to calibrate the forecasts against the 0-day forecast.

It was hard to tell how well the various runs performed in terms of landslide forecasting, since most unstable area was found to be in forests. However, the forecast runs did predict instability around a known landslide location, which the historic could only partly predict or completely miss. Hollin Hill was missed completely as an unstable area in the regional model. This is mainly attributed to the fact that there was another erroneous model assumption: freely fluctuating water tables. An attempt was made to create a deep reservoir in the third layer of the model, simulating seepage into the upper layers or infiltration from the upper layers in to the third layer and therefore free fluctuation. This caused the transmissivity of the layer to be rather large, allowing for fast drainage of subsoil water and a fast drop in water level in case of no recharge. The water levels around Hollin Hill would therefore never reach critical heights to cause instability, since Hollin Hill is a rather moderate slope it needs a largely saturated soil to become unstable. This also shows the dominant role of slope angle (topography) in instability assessment.

Improvements of the study showed to be a better modeling of recession flow, soil depth estimation (the soil was too deep, allowing for fast drainage of water, causing more moderate slopes to not become unstable), classifications of instability according to safety factor and a more robust approach of parameterization if no information is available about spatial distribution in the form of combined soil and land-use maps, forming hydrological units.

## Acknowledgements

I would like to thank a couple of people who helped me with finishing my thesis. First I would like to thank Rens van Beek from Utrecht University. Rens took the time to help me out on countless occasions during my whole thesis, and helped me out a lot with Python. Even past regular office hours he always made sure that I was helped. He went beyond what you can expect from a supervisor in assisting me. I cannot thank him enough for this, because if it was not for this, I would not have been able to finish my thesis at all.

I would also like to thank Jelle Buma from Utrecht University / Deltares. As my second supervisor he was less directly involved than Rens van Beek, nonetheless the talks with him were always very refreshing and a great push in the right direction. Also his comments were always of great help to improving this thesis.

I would like to thank Tom Dijkstra and John Mackay of the British Geological Society. Tom Dijkstra especially, for the hospitality offered, the support at the office and the fieldwork, and the nice barbeque at his house. John Mackay I would like to thank for the company during the fieldwork and his support at the office.

I would also like to thank Albrecht Weerts from Deltares for hosting me at his office and helping me out with setting up FEWS, even though it did not work out in the end.

Finally, I would like to thank fellow student Hans van Woerden for some talks about this subject and accompanying me to England.

## Table of contents

List of figures.....	V
List of tables.....	VIII
1. Introduction.....	10
1.1. Background.....	10
1.2. Problem description.....	10
1.3. Main Objectives.....	2
1.4. Literature review.....	4
2. Methodology and description of site.....	6
2.1. Description of models.....	6
2.2. Area of study.....	11
2.2.1. Geological build-up.....	11
2.2.2. Soil types.....	14
3. Set up of a landslide early warning system.....	16
3.1. Design of an Early Warning System.....	16
3.2. Existing EWS.....	22
4. Data collection, processing and model setup.....	26
4.1. Geographic data.....	26
4.2. River flow data.....	27
4.3. Historical precipitation data.....	27
4.4. Historical temperature data.....	29
4.5. Forecast data.....	30
4.6. Land use data.....	31
4.7. Soil type data.....	32
5. Results.....	37
5.1. River flow results historical run on 100m.....	37
5.2. River flow results historical run on 25m.....	49
5.3. River flow results forecast run on 25m.....	53
5.4. Stability results.....	61
5.5. Recommendations.....	82
Appendix I.....	85
Appendix II.....	87
Appendix III.....	89
Appendix IV.....	90

References ..... 91

## List of figures

Figure 1-2: The location of Hollin Hill. From BGS (2014b).....	11
Figure 1-1: Rainfall and landslide statistics for England. From BGS (2014a).....	11
Figure 1-3: Position of the catchment and Howardian Hills. After CEH (sine anno).....	12
Figure 1-4: Flowchart of the forecasting chain. From Mercogliano <i>et al.</i> (2013).....	14
Figure 1-5: Different rainfall patterns used in the study. From Tsai & Wang (2011).....	15
Figure 1-6: Treshhold for landslide initiation with different rainfall patterns. From Tsai & Wang (2011).....	15
Figure 2-1: Description of STARWARS model. From Krzeminska <i>et al.</i> (2012).....	16
Figure 2-2: Forces acting on a slope according to the infinite slope model. From Wienhöfer <i>et al.</i> (2011).....	17
Figure 2-3: Schematization of hydrological model calibration. From Yu (2002).....	17
Figure 2-4: Schematization of the forecast model.....	19
Figure 2-5: Running a forecast with different initial states. From Met Office (2014).....	20
Figure 2-6: The geographical extent of the Lias Group. From Hobbs <i>et al.</i> (2012).....	21
Figure 2-7: Continental rifting in three steps. From Geoscience (sine anno).....	22
Figure 2-8: Horst and graben. From USGS (2012).....	23
Figure 2-9: The different formations at Hollin Hill as seen from above. From BGS (2014b).....	24
Figure 2-10: Cross-section of the landslide at Hollin Hill. From BGS (2014b).....	24
Figure 2-11: Soil types in the Foss catchment and Hollin Hill. After Cranfield University (sine anno).....	25
Figure 3-1: A near real-time system design. From Reid <i>et al.</i> (2008).....	26
Figure 3-2: Generic design of an EWS. From Intrieri <i>et al.</i> (2013).....	27
Figure 3-3: Framework for decision making. From Dai <i>et al.</i> (2002).....	27
Figure 3-4: The ALARP principle. From Wong <i>et al.</i> (1997).....	28
Figure 3-5: Flowchart for slope scale monitoring. From Intrieri <i>et al.</i> (2013).....	31
Figure 3-6: Organizational flow chart. From Intrieri <i>et al.</i> (2013).....	32
Figure 3-7: Flow chart showing the different actors involved. From Bazin (2012).....	33
Figure 3-8: Physical aspects of an EWS. From Bazin (2012).....	34
Figure 3-9: Number of criteria used. From Bazin (2012).....	34
Figure 3-10: Types of instruments used. From Bazin (2012).....	35
Figure 3-11: Means of alert. From Bazin (2012).....	35
Figure 3-12: Practical challenges. From Bazin (2012).....	36
Figure 4-1: Flow directions of material in a pcraster cell. From Utrecht University (2014).....	37
Figure 4-2: Catchment of Foss and Hollin Hill.....	38
Figure 4-3: Rainstations around Hollin Hill and the catchment.....	39
Figure 4-4: Spatially distributed rainfall pattern.....	39
Figure 4-5: Temperature stations used.....	41
Figure 4-6: Landuse map.....	45
Figure 4-7: Accuflux map.....	46
Figure 4-8: Ksat for each layer.....	46
Figure 4-9: Porosity for each layer.....	47
Figure 5-1: Modeled daily flow values with initial parameters from 01-01-09 until 31-12-2012 versus observed daily flow.....	48
Figure 5-2: Bedrock hydraulic conductivity values. From Ridell (2007).....	51
Figure 5-3: Values for ksat * 25 for layer 1 and 2.....	52
Figure 5-4: Porosity for layer 1 (left) and 2 (right).....	52
Figure 5-5: Modeled values on Ksat * 25 and porosity * 1.4 versus observed daily flow values from 01-01-09 until 31-12-2012.....	53

Figure 5-6: Annual precipitation in mm for 2009 (left) and 2010 (right).....	53
Figure 5-7: Annual precipitation in mm for 2011 (left) and 2012 (right).....	54
Figure 5-8: DEM of the catchment.....	54
Figure 5-9: Relative annual frequency of types of precipitation. From Cedar Lake Ventures Inc. (2015).....	56
Figure 5-10: Annual PET in mm for 2009 (left) and 2010 (right).....	56
Figure 5 10: Annual PET in mm for 2011 (left) and 2012 (right).....	53
Figure 5-12: Average PET in mm yr-1 for the Entire UK between 1961-1990. After NERC (2004). ....	57
Figure 5-13: Random field map showing parameter distribution.....	58
Figure 5-14: Slope angles for 10m, 25m, and 100m cell size .....	59
Figure 5-15: Results of the run on 25m for the Ksat *25 and porosity * 1.4 run for the years 2009-2012 .....	60
Figure 5-16: Results of the run on 25m for the initial run for the years 2009-2012 .....	62
Figure 5-17: Daily forecasted flow versus daily observed flow [m <sup>3</sup> s <sup>-1</sup> ] for the year 2012.....	65
Figure 5-5-18: NSE of day 1 versus day 0 .....	66
Figure 5-19: NSE of day 2 versus day 0 .....	66
Figure 5-20: Annual precipitation in mm for the historic run, forecast run day 0 and day 1 .....	67
Figure 5-22: PET totals [mm] for the historic run, forecast day 0, and day 1 from 10-01-2012 until 31-12-2012 .....	69
Figure 5-23: PET totals [mm] for the forecast days 5 and 8.....	69
Figure 5-24: Phase errors with increasing lead time. From ECMWF (2014). ....	70
Figure 5-25: The unstable cell at Hollin Hill.....	73
Figure 5-26: Potential area of instability.....	74
Figure 5-27: Unstable area around Hollin Hill for the historic run on initial values (left) and Ksat * 25 (right) for 1 day or more (upper) and 180 days or more (lower).....	75
Figure 5-28: Unstable area around Hollin Hill for the for the forecast run day 0 (left) and day 1 right for 1 day or more (upper) and 180 days or more (lower).....	76
Figure 5-29: Unstable area around Hollin Hill for the for the forecast run day 2 (left) and day 3 right for 1 day or more .....	77
Figure 5-30: Unstable area around Hollin Hill for the for the forecast run day 4 (left) and day 5 right for 1 day or more .....	77
Figure 5-31: Unstable area around Hollin Hill for the for the forecast run day 6 (left) and day 7 right for 1 day or more .....	78
Figure 5-32: Unstable area around Hollin Hill for the for the forecast run day 8.....	78
Figure 5-33: Instability in the upper part of the catchment for 1 day or more (upper) and 180 days or more (lower) for the initial run (left), the Ksat * 25 run (middle) and the forecast runs (right) .....	79
Figure 5-34: Possible landslide area .....	79
Figure 5-35: Landslide at Whinny Oaks Covert.....	80
Figure 5-36: Safety factor [-] and water level [m] for the historic run on initial values .....	80
Figure 5-37: Safety factor [-] and water level [m] for the historic run on Ksat * 25 and porosity * 1.4.....	81
Figure 5-38: Safety factor [-] and water level [m] for the forecast day 0 .....	81
Figure 5-39: Safety factor [-] and water level [m] for the forecast day 1 .....	82
Figure 5-40: Safety factor [-] and water level [m] for the forecast day 2 .....	82
Figure 5-41: Safety factor [-] and water level [m] for the forecast day 3 .....	83
Figure 5-42: Safety factor [-] and water level [m] for the forecast day 4 .....	83
Figure 5-43: Safety factor [-] and water level [m] for the forecast day 5 .....	84
Figure 5-44: Safety factor [-] and water level [m] for the forecast day 6 .....	84
Figure 5-45: Safety factor [-] and water level [m] for the forecast day 7 .....	85
Figure 5-46: Safety factor [-] and water level [m] for the forecast day 8.....	85
Figure 5-47: Phase shifts .....	86

Figure 5-48: Safety factor [-] and water level [m] for the cell around Hollin Hill with the historic run on initial values	86
Figure 5-49: Safety factor [-] and water level [m] for the cell around Hollin Hill with the historic run on $K_{sat} * 25$ and porosity * 1.4 .....	87
Figure 5-50: Safety factor [-] and water level [m] for the cell around Hollin Hill for the forecast day 0.....	87
Figure 5-51: Safety factor [-] and water level [m] for the cell around Hollin Hill for the forecast day 1.....	88
Figure 5-52: Safety factor [-] and water level [m] for the cell around Hollin Hill for the forecast day 2.....	88
Figure 5-53: Safety factor [-] and water level [m] for the cell around Hollin Hill for the forecast day 3.....	89
Figure 5-54: Six classes of instability. From Pack <i>et al.</i> (1998).....	92
Figure 5-55: Annual P [mm] for the day 2 forecast.....	92



## List of tables

Table 4-1: Latitude and Longitude of forecast data .....	42
Table 4-2: Corine landcover classes. From USGS (2008) and Hagemann <i>et al.</i> (1999).....	44
Table 4-3: Vegetation heights for different landcover classes. From Hagemann <i>et al.</i> (1999).....	44
Table 5-1: Maximum root depth for each land use type. From USGS (2008). .....	50
Table 5-2: NSE for different Ksat and porosity values .....	51
Table 5-3: number of rain and non rain days for the years 2009, 2010, 2011, and 2012 .....	55
Table 5-4: Results of runs with different hydraulic conductivities for NSE and total volume for 2009-2012 .....	61
Table 5-5: Different components of the water balance for the year 2012 for the driest and wettest run.....	61
Table 5-6: Daily and observed and precipitation flow for the different runs.....	62
Table 5-7: Daily and observed flow and precipitation for the different runs.....	63
Table 5-8: Daily and observed flow and precipitation for the different runs.....	63
Table 5-9: Daily and observed flow and precipitation for the different runs.....	64
Table 5-10: NSE scores for different forecast days and total volume of the year 2012 .....	65
Table 5-11: Results for day 0 .....	67
Table 5-12: Number of rain and non rain days for different forecast days .....	70
Table 5-13: Number of cells being unstable for various runs.....	72
Table 5-14: Influences of terrain properties on different cells.....	74
Table 5-15: Water levels and safety factors for day 0, 1, and 2 .....	82
Table 5-16: errors in the forecast.....	89

## 1. Introduction

In this chapter, an overview about the background and the problem description of this research will be presented to the reader. Additionally, a (relevant) literature review will be given, reviewing and assessing earlier research conducted in this field. Finally, the main and sub research questions, along with the objectives of this research are conveyed in this chapter.

### 1.1. Background

A landslide is a form of a mass movement. Varnes (1978) defined a landslide as “a downward and outward movement of slope forming material under the influence of gravity”. The resulting deformation of the slope is subsequently defined as slope instability and the onset of mass movement as failure (van Beek *et al.*, 2008). Next to gravity, water is the most important parameter regarding slope instability since slope stresses and strength of materials are governed by the interplay between rainfall and infiltration (Sirangelo & Braca, 2004). Causative factors of a landslide include among others toe cutting which is the removal of soil at the foot of a hill due to for instance construction activities and expanding (rail)road networks (Ramesh, 2014; van Beek *et al.*, 2008), soil destabilization and erosion due to vegetation removal and with more uncertainty climate change, for instance larger rainfall intensity (Sidle & Dhakal, 2002), steep slope angles and saturated soils (Ramesh, 2014). Triggers are rainfall and earthquakes (Ramesh, 2014). Landslides affect the following aspects of the natural environments: (1) topography and morphology of sub aerial and submarine surfaces, (2) rivers, grasslands and forests, (3) habitats of aquatic and terrestrial flora and fauna (Schuster & Highland, 2001).

Crozier (2010) argues that landslides be recognized as major formative processes within geomorphology and that ‘landslide geomorphology’ should be given a legitimate place within the wider geomorphology discipline. He defined a set of three criteria to characterize a *landslide geomorphology system*, i.e. a system where there are recurring relationships between dominating landslide activity, related geomorphic processes and land form, both spatial and temporal. The three criteria were extent of the landslide signature in the terrain, temporal persistence of landslide signatures and degree and effectiveness of coupling to a fluvial system. Using New-Zealand as his area of study, he identified six of those systems in New-Zealand.

In the United States of America it is estimated landslides cause 25 to 50 deaths and a 3 billion US dollar damage to the economy annually. In Japan, annual economic loss is up to 4 billion US dollar, in India and Italy between 1 and 2 billion dollar. Direct costs include repair, replacement or maintenance due to destruction or damage to properties. Indirect costs in affected areas include reduced real estate values, reduced agricultural, forest, industrial and tourist revenues and finally mitigation and prevention costs (Schuster & Highland, 2001). In the UK, landslides cost at least several millions of pounds on an annual basis due to (rail)road repairs (The Geological Society of London, 2001). In addition, the costs due to coastal erosion, a large part of which is caused by landslide activity, are 126 million pounds annually. Contrary to the lay mans popular belief, landslides have been the cause of fatalities in Britain. In most of these cases albeit not all, landslides have resulted in fatal railroad or traffic accidents (Pritchard *et al.*, 2013).

### 1.2. Problem description

Before describing the site of interest, an overview encompassing landslides in England will be given. The year 2012 has been an extremely wet year for England, second wettest after the year 2000 since rainfall recording began in 1910. Analyses done by the Meteorological Office revealed that England is facing more extreme rainfall events, possibly linked to climate change. This increased rainfall led to a significant increase in frequency of landslides, slope failures (on man-made slopes such as railroad embankments), especially during July and December 2012. Figure 1.1 (BGS, 2014a) shows rainfall and landslide statistics for England, in which the red-dashed lines represent the average number of landslides between 2006-2013 on a given month, the red solid line the number of landslides on a given month in 2012, the blue dashed lines the average rainfall on a given month between 2006-2013 and the blue bars the actual rainfall on a given month in 2012:

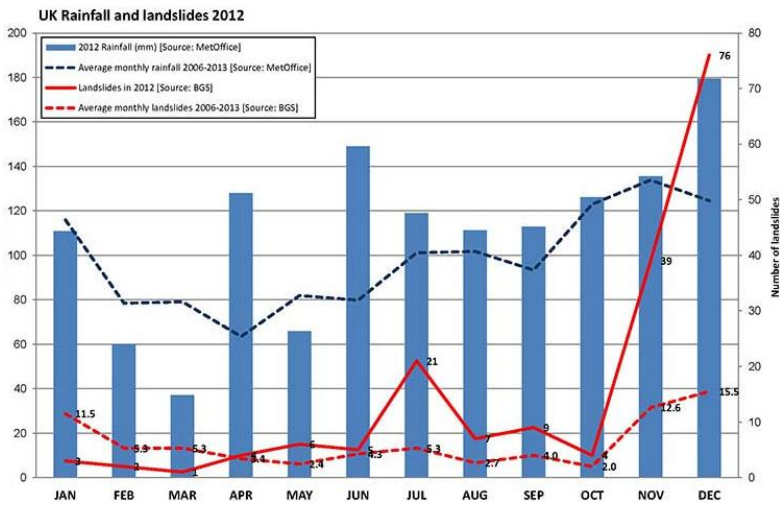


Figure 1-1: Rainfall and landslide statistics for England. From BGS (2014a).

Initial figures suggested a four-to five-fold increase in number of landslides for July and December compared to the average over the years 2006-2013. The British Geological Survey (BGS) started recording landslides in their national database in 2006, hence the years 2006-2013 for the average (BGS, 2014a). Because of this high number of landslides, the BGS desires more adequate regional landslide prediction and warning using physically based slope stability models (Tom Dijkstra, personal communication). Deltares, another party involved in this study, wants to see how these physically based slope stability models can be linked to their platform Flood Early Warning System (FEWS). FEWS is an open data handling platform, its initial design being tailored for hydrological forecasting and as an early warning system. FEWS can link various models and data to each other, as long as they are converted to the XML format. Because of this high flexibility, the platform is used in various fields such as water quality forecasting, operational sewer management optimization, peat fire prediction, and reservoir management (Deltares, sine anno). To this means, the Hollin Hill test site and surrounding catchment (Yorkshire, North-East England) could be used. Figure 1.2 (BGS, 2014b) shows where Hollin Hill is situated:



Figure 1-2: The location of Hollin Hill. From BGS (2014b).

In 2004, the British Geological Survey was informed about the presence of an active landslide at Hollin Hill, by the landowner. As of 2008, the site is constantly monitored and therefore deformation data about this site is available, making it an appreciable site for a study to test regional slope instability forecasting although the data was not yet available when writing this report (see section: description of field sites).

### 1.3. Main Objectives

In order to get a general appreciation of regional landslide forecasting, some light will be shed on how to implement a model into an overarching Early Warning System (EWS). EWS are defined as “monitoring devices designed to avoid, or at least to minimize, the impact imposed by a threat on humans, damage to property, the environment, or/and to

more basic elements like livelihoods” (Medina-Cetina & Nadim, 2008). The advantage of an EWS is that they generally speaking have less economical and environmental impact than structural interventions to prevent landslides, through risk mitigation and effective response, if properly implemented (Intrieri *et al.*, 2013).

The main aim of this research is to assess to which degree slope instability and/or failure can be forecasted at the regional scale around Hollin Hill, starting from sparse and incomplete data on a temporal scale of 1987-2012, being the time of recorded river discharge for the river Foss. Because Hollin Hill is situated in a catchment without river flow measurements (CEH, sine anno), calibrated catchment analysis of Hollin Hill proves to be impossible. Alternatively, a nearby catchment sharing a similar geology to Hollin Hill has been used. Discharge data closest to Hollin Hill is from the catchment of the river Foss at Huntington, a catchment partly situated in the Howardian Hills, like Hollin Hill, with a catchment area of 118 km<sup>2</sup> (CEH, sine anno). Hollin Hill has subsequently been added as a subcatchment to the existing catchment. This solution has a twofold advantage:

- Deformation data about Hollin Hill is available in the future, as well as some information about soil profiles of the landslide (Tom Dijkstra, personal communication). Modeled instability can then be compared to known deformation events;
- Regional parameters such as hydraulic conductivity et cetera can be calibrated against the recorded daily river flow of the gauging station from the Foss river, because the hydrological model used in this study also reproduces river flow. Assuming that these calibrated parameters are similar to a certain degree to the parameters of the Hollin Hill subcatchment, the subcatchment can be included in the model. In addition, it is known that the slope stability model should classify the Hollin Hill site as an unstable area (in the section methodology it is explained how an unstable area is classified), which is a spatial validation albeit not yet temporal, since deformation data will be available in the future.

No active landslide was found in the landslide database of the BGS within the Foss catchment. That does not exclude that landslides are present in the area, since this catchment has a similar geology and only recorded landslides are stored in the database. Modeled unstable areas will be crosschecked against google earth images to assess if there are active landslides present at the predicted site of instability. The catchment delineation and the position of Hollin Hill as a red dot can be seen on the left side figure 1.3, on the right side the position of the Howardian Hills with an unknown scale (CEH, sine anno):

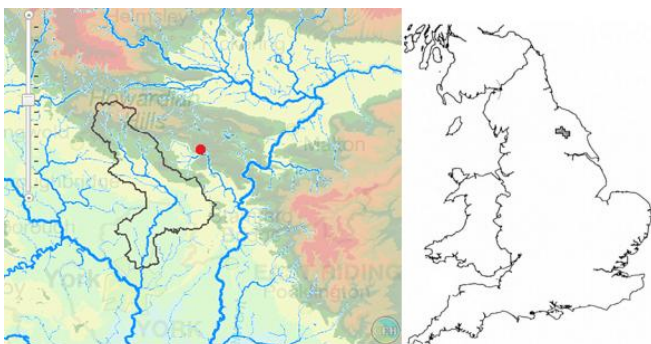


Figure 1-3: Position of the catchment and Howardian Hills. After CEH (sine anno).

Based on the calibrated and spatially validated model, the feasibility of implementing a regional operational landslide forecasting system will be assessed and key variables to improve performance of this forecasting system identified, for future application of the model in FEWS. Initially, the study was even aimed at setting up a forecasting environment in FEWS, but due to technical and time constraints this was not deemed possible within the timeframe of this research.

The model will first be run with historic rainfall and temperature data from 1988-2012 from the MIDAS database

from the Natural Environmental Research Council (Nerc, 2014) (see section data collection and processing and results) to calibrate and validate the model by using simulated daily river flow against recorded daily flow and find the best parameterization. After this, the forecasting capabilities of the model will be tested by using an eight day weather forecast from the European Centre for Medium-Range Weather Forecasts (ECMWF, 2014), see section data collection and processing and results. Using recorded weather data and subsequently forecasted weather data with the same parameterization allows for comparison between reproduction of historic and future events and therefore forecasting capabilities (or lack of).

#### 1.4.Literature review

A short literature review regarding some relevant studies will be dealt with in this section. Ramesh (2014) describes the set-up of an early warning system in Idukki, Kerala State (India), an area highly prone to rainfall-induced landslides. A network of 50 geological and 20 wireless sensors was installed, gathering data such as precipitation levels, soil moisture, pore pressures and movements, in conjunction with other soil properties and hydrological and geological data. Landslide probability is determined through modeling and simulation based on the field data at the Data Management Center 300km away. The modeling software compares the raw field data to data from soil test and lab setups, for a more accurate assessment of the FS. For the technical lay-out of the network, the reader is advised to read the study. A three system warning level was developed (Early, Intermediate and Imminent). Early depends on soil moisture, Intermediate on pore pressure, and Imminent on movement. Warning levels are correlated to different rainfall intensities and temporal distribution. The soil moisture sensor is the first to saturate during heavy rainfall, sending out an Early warning. Further in time, with equal or increased rainfall intensity, changes in pore pressure values with respect to infiltration rate can be detected. When this value also saturates, the Intermediate warning is sent. Local community and government are warned at this point, and citizens advised to evacuate if rainfall persists from this point. If the movement sensor detects movement along with the high pore pressure values, the Imminent warning is sent. Warning issues are removed if pore pressure values decrease dramatically, for instance due to a stop of rainfall. An Intermediate warning was issued when on July 18<sup>th</sup> 2009 200mm of rainfall was measured, with 150mm on July 19<sup>th</sup> and 20<sup>th</sup>. Local governments advised to evacuate people. Finally, rainfall levels reduced and a subsequent decrease in pore pressure was seen from the real-time streaming software, allowing the researchers to validate the system with an actual event. As a result of this, the Indian government is now considering to deploy this system across every landslide prone region in India.

Mercogliano *et al.* (2013) developed a forecasting chain for landslide initiation for a catchment of 3,103 km<sup>2</sup> in Tuscany, Italy, situated the northern Apennines. Global data were used as initial and boundary conditions for their limited area weather forecasting models, Cosmo LM. Short forecasts up to 48 hours were used for accuracy purposes. The spatial resolution for the Cosmo LM was 7km (at this resolution an adequate numerical forecast of near-surface weather conditions is feasible), but when the area of interest showed a complex orography, a resolution of 2.8km was used. The HIRESSS (High Resolution Slope Stability Simulator) model was used for landslide prediction, consisting of a hydrological module based on an analytical solution of an approximated form of Richards equation and a geotechnical module based on an infinite slope model that considers increases in strength and cohesion in unsaturated soils due to matric suction. The flow chart is shown in figure 1.4:

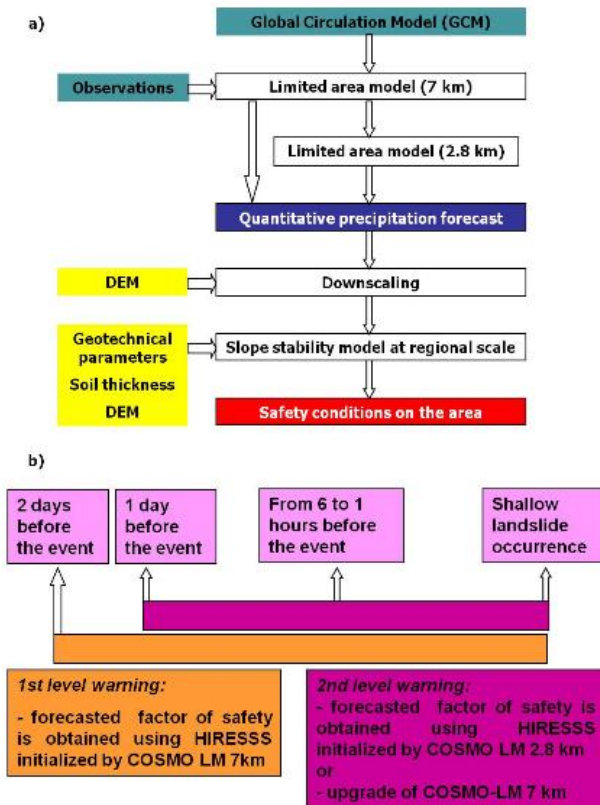


Figure 1-4: Flowchart of the forecasting chain. From Mercogliano *et al.* (2013).

The output was a series of instability maps from 30 October 2003 to 1 November 2003 in 58 hourly timesteps. The model successfully predicted the most vulnerable parts of the region (first western, followed by the northern part) in accordance with landslide occurrences, stability on October 30, initiation of slope instability on the beginning of October 31, increasing until they reach a peak at the first hours of November 1, after which conditions stabilized again. These results were consistent with reports of landslides triggered on the night of October 31 to November 1. The eastern part appeared to show overestimation of slope stability because no landslides were reported there, but the authors pointed out that the eastern part of the region is a mountainous area with low population density and that actual landslides may not have been reported. Their main recommendations for improvement of forecasting were: simulation of documented case-studies to calibrate for interpretation of probabilistic outputs (define probability values in correspondence with alarm thresholds) and a more deepened understanding of model parameters such as rainfall, soil thickness et cetera.

In a study by Kuriakose (2010) the hydrological model STARWARS and the slope stability model PROBSTAB will be used. Kuriakose chose these models over the models SINMAP, SHALSTAB and TRIGRS for his study since they were the best performing for three criteria namely capability for spatial parameterization, prediction accuracy of known landslides ( $FS < 1$ ) and total area predicted as unstable ( $FS < 1$ ) compared to the actual known landslide area. His study area was the Western Ghats mountain range in the state of Kerala, SW India, a state where impact of landslides is large since it is the state with the third highest population density (819 people per  $km^2$ ). The aim of the study was to map the susceptibility of the region to landslides, to determine the relationship between rainfall intervals and landslide occurrence, and to map long term and short term land cover changes and their effects on hydrology and slope stability. A key finding from this study was that land cover change from 1913 (pre-plantation) to 2008, in this area notably due to terracing, (resulting in changes in soil depth and root cohesion), resulted in an overall increase in potential area of failure by 43%, and temporally and spatially distributed probability of failure by 49%. In addition, an idealized land use scenario for 2058 would not bring the slope stability up to 1913 levels, due to lower root reinforcement levels. Resulting from this, the author concluded that transition probabilities of land cover change even outweighed rainfall quantities in establishing spatio-temporal probabilities of shallow landslide initiation for this area. This was contrary to common beliefs that only return probabilities of landslide triggering rainfall events are needed to assess spatio-temporal probabilities of landslide initiation. He therefore recommended transition probabilities of land cover change to be included in landslide hazard assessments on a regional scale. A

recommendation, relevant to this proposed study, is that STARWARS and PROBSTAB can be useful to forecast landslide initiation probabilities when hourly rainfall predictions are used as input. Finally, Tsai & Wang (2011) performed a study on the influence of rainfall patterns on shallow landslide initiation. Site of study was a landslide in Northern Taiwan. They used Richards equation in combination with the Mohr-Coulomb failure criterion to assess slope stability for different rainfall patterns (Advanced, Intermediated, Delayed, and Uniform). The rainfall patterns are shown in figure 1.5 (Tsai & Wang, 2011):

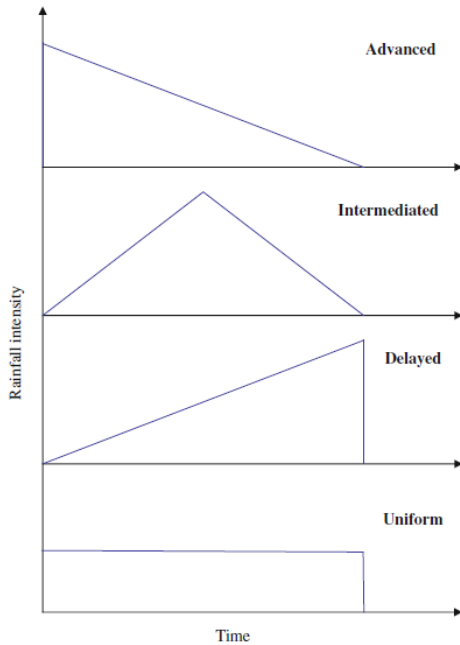


Figure 1-5: Different rainfall patterns used in the study. From Tsai & Wang (2011).

Resulting thresholds are shown in figure 1.6 (Tsai & Wang, 2011), with rainfall duration (hr) on the y-axis and rainfall amount (mm) on the x-axis, for the different rainfall patterns:

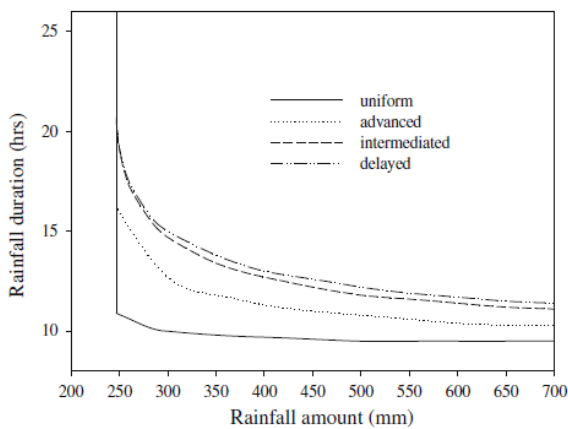


Figure 1-6: Treshold for landslide initiation with different rainfall patterns. From Tsai & Wang (2011).

The main conclusion was that shallow rainfall-induced landslides are dependent on rainfall patterns, because different patterns show a different decrease in matric suction in the unsaturated zone (when matric suction in the vadose zone dissipates due to infiltrating rain, the shear strength of the soil decreases). The delayed rainfall has the largest threshold, uniform the least. Landslide occurrence could be wrongly evaluated if not taking rainfall patterns into account. The researchers want to integrate their results with FEWS, to be able to make accurate landslide predictions (Chang, n.d.).

A final note on prediction on landslides is that it can be a complicated procedure, especially in remote areas because landslides that have been predicted and not have been reported could still have taken place in reality but simply not

have been registered (Mercogliano *et al.*, 2013 ; Jakob *et al.*, 2011). Jakob *et al.* (2011) state that a landslide prediction system behaves as desired when there are more warning levels than actual landslide events (in their case debris flows) whilst keeping the total number of warnings to a reasonable level.

## 2. Methodology and description of site

Use of appropriate models, input data and a proper methodology are paramount to conducting proper quantitative research. In this section it is therefore described (in a concise manner) which models will be used, data that will be collected and a lay-out for the methodology will be presented.

In order to forecast future landslide events, hydrological processes starting from rainfall such as elevated pore water pressure due to rising soil moisture content, resulting in changes in hill-slope pressures (Sirangelo & Braca, 2004), subsequently triggering landslide movement have been assessed using the physically-distributed hydrological model STARWARS (Storage And Redistribution of Water on Agricultural and Revegetated Slope) and the hill slope stability model PROBSTAB (PROBability of STABility), which (partly) uses the hydrological output of STARWARS (van Beek, 2002). The model is integrated in one script in the pcraster environmental modelling language, developed by the Utrecht University. The model will be run on a daily basis, since the data used are on a daily basis (see section: data collection and processing). Therefore all the output (river flow, slope stability and so forth) will also be on a daily basis.

### 2.1. Description of models

As mentioned, STARWARS will be used for the hydrological modeling. A conceptualization is given in figure 2.1 (Krzeminska *et al.*, 2012):

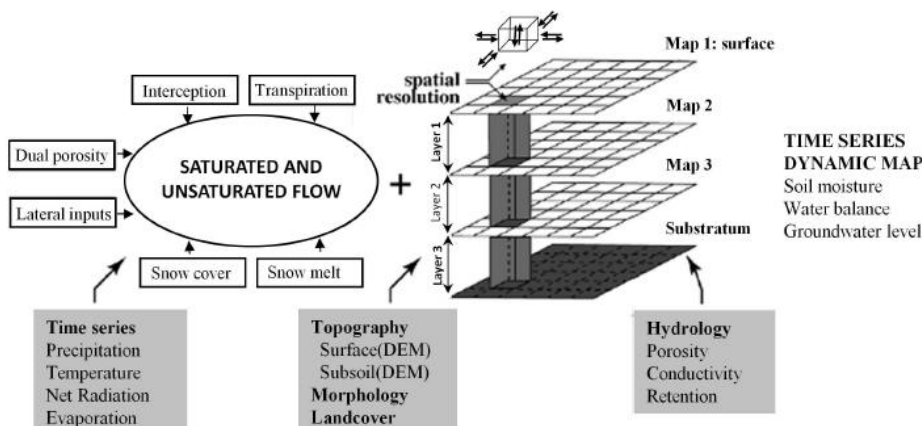


Figure 2-1: Description of STARWARS model. From Krzeminska *et al.* (2012).

STARWARS is a physically based, spatially distributed, and finite element model that runs in PCRaster. STARWARS models transient (un)saturated flow in vertical and lateral directions to simulate delay and loss of percolation in the vadose zone, which attenuates the response of groundwater levels to a major rainfall event. Transient flow is assumed to be only a function of elevation potential and not of matric potential as well, limiting percolation to vertical flow solely driven by gravity. The reason it is omitted is because matric potential in unsaturated flow is considered negligible when the soil is above field capacity (the soil moisture value at which the soil can hold water against gravity). The model assumes free draining and a direct response of the groundwater to net rainfall (precipitation minus interception and evaporation). This response is applied on a constant groundwater level or computed over a semi-impervious lower boundary, the latter preventing all soil moisture to drain into deeper strata (the fourth infinite layer in STARWARS). For the latter case the result is a perched groundwater level. As can be seen in figure 6, STARWARS is able to model using 3 different soil layers (van Beek, 2002).

PROBSTAB assesses the probability of failure using the Mohr-Coulomb failure criterion and the infinite slope model approach (van Beek, 2002). Slope stability can be assessed through the Mohr-Coulomb failure criterion, assuming the infinite slope shown in figure 2.2 (Wienhöfer *et al.*, 2011):



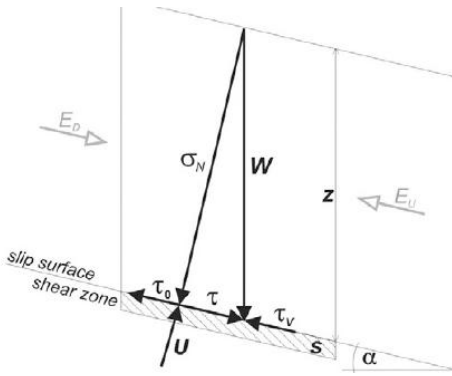


Figure 2-2: Forces acting on a slope according to the infinite slope model. From Wienhöfer *et al.* (2011).

$W$  is the weight per unit area of the element ( $M L^{-1} T^{-2}$ ),  $z$  denotes the depth of the slip surface ( $L$ ),  $\alpha$  is the slope angle of the slip surface,  $\tau$  and  $\tau_0$  are respectively the shear stress and strength acting on the slip surface ( $M L^{-1} T^{-2}$ ).  $\tau$  is calculated by  $W * \sin \alpha$  and  $\tau_0$  is given by  $c_r + (\sigma_n - u) * \tan \phi_r + F_{seep}$ ,

in which  $c_r$  is residual cohesion ( $M L^{-1} T^{-2}$ ),  $(\sigma_n - u)$  denotes the effective stress ( $M L^{-1} T^{-2}$ ) given by the total stress  $\sigma_n$  minus the pore water pressure  $u$  (the latter increasing with increasing soil saturation) and  $F_{seep}$  being a seepage term ( $M L^{-1} T^{-2}$ ). Using a Factor of Safety (FS), being dimensionless and defined as  $FS = \frac{\tau_0}{\tau}$  slope stability is calculated.

When the modelled  $FS = 1$ , i.e. when shear stress equals the shear strength (for instance due to increased pore water pressures) the probability of failure is 50% (because the real FS can be on either side of the error margin). When FS decreases to below 1, the probability of failure increases. Likewise, the probability of failure decreases when  $FS > 1$ . For more information on slope stability modelling one can read Wienhöfer *et al.* (2011).

A disadvantage of physically based models is that physical processes are often represented in a crude way, and that scales of measurements of many parameters are incompatible with model scales (Yu, 2002). Models therefore often need calibration. Calibration means that the model is able to reproduce data observed in the field such as soil moisture content and hydraulic heads to a satisfactory (defined by the user) extent. This requires adjustment of model parameters, usually occurring on a basis of trial and error by keeping all parameters constant while changing one. Constraints have to be placed on the range of parameter values by making use of in situ data such as hydraulic conductivity and stream flow. The process of model calibration is depicted graphically in figure 2.3 (Yu, 2002):

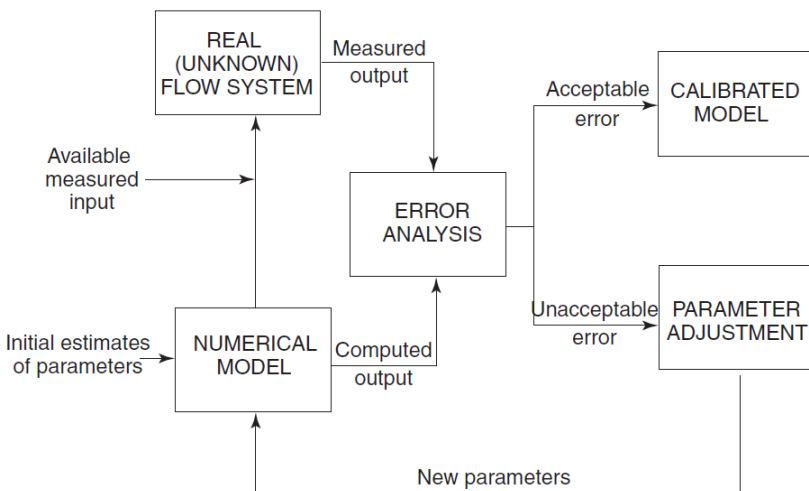


Figure 2-3: Schematization of hydrological model calibration. From Yu (2002).

As mentioned before, to improve model performance, the model will be calibrated using historic meteorological data against the stream flow data from the gauging station from the river Foss at Huntington. This is for the hydrological module of the model. For the slope stability module validation is twofold:

- The Hollin Hill site with known coordinates serves as a spatial validation. Since Hollin Hill is a recorded landslide, the slope stability model should predict slope instability around Hollin Hill;
- Areas declared unstable can be crosschecked with google maps. On google maps satellite images active landslides are visible. Vice versa, areas declared stable but with a steep slope (say, 15 degrees or steeper) can be crosschecked if they are indeed stable in reality. However, this is only a spatial validation method and not a temporal validation method, since google maps does not have temporal information about landslide movement. The aerial photo validation method was also employed by Guimarães *et al.* (2003).

The model was run with initial values with a cell size of 100 square meter for the duration of 1988-2012, the timeframe of the historic meteorological data. This resolution allows for a fast run (1.5 hours for the entire timeframe) and subsequent river flow validation. This resolution does not allow for reasonable slope stability analysis however, the reason being that the DEM is smoothed out to much on this resolution, significantly underestimating slope angles and therefore overestimating slope stability. Therefore the optimal parameter combination validated against the known river flow data will be used for the run on 25m, saving time. The years of 1988-2008 have been used as spin up of the model, since it takes time for the model to reach an equilibrium situation with regards to water level build up etcetera. The model seems to require a couple of years of spin up before it runs stable was discovered during the runs, therefore this timeframe is long enough, according to a study by Merz *et al.* (2009). They simulated the water balances of 269 catchments in Austria, with a size range of 10 to 130,000 km<sup>2</sup>, by means of a semi distributed model and using NSE. Their conclusion was that the change in calibration efficiencies decrease most after 5 years, and vary most between 1 and 5 years. They also discovered that the correlation between some parameters increases even beyond 5 years of calibration, therefore although 5 years already yields good results, some parameters are still associated with larger uncertainties than with a calibration period extending 5 years.

The years 2009-2012 are used for hydrological and parameter validation by using simulated daily discharge against recorded daily river flow data [m<sup>3</sup> s<sup>-1</sup>]. This timeframe is arbitrary, but since the years 2007 and 2008 did not have any recorded river flow data, it is a rather convenient one to use and using four years at least some annual variation with regards to wet and dry years can be expected. In order to assess the model accuracy, the Nash-Sutcliffe Efficiency (NSE) statistical test has been performed (Nash & Sutcliffe, 1970):

$$NSE = 1 - \frac{\sum(Q_{calc} - Q_{obs})^2}{\sum(\bar{Q}_{obs} - Q_{obs})^2}$$

In which:

Q <sub>calc</sub> = calculated discharge	[L <sup>3</sup> T <sup>-1</sup> ]
Q <sub>obs</sub> = observed discharge	[L <sup>3</sup> T <sup>-1</sup> ]
$\bar{Q}_{obs}$ = mean observed discharge	[L <sup>3</sup> T <sup>-1</sup> ]

Testing of catchment modeling performance by means of NSE is also done in a study by Merz *et al.* (2009). The lower limit of the NSE is  $-\infty$ , the upper limit is 1. If the value is below zero, it means that the mean observed discharge (i.e. summation of daily flow values divided by total number of days considered) is a better predictor during the timeframe considered than the modeled daily flow. If NSE equals zero, it means the model performs as good as the mean observed discharge in predicting observed flow. If the value is between zero and one it means that the model has better predicting capabilities than the mean, where a value closer to one denotes a more accurate prediction. If NSE equals one, it means that all calculated values for the timeframe considered are equal to the observed values.

One can intuitively understand that for daily catchment modeling over a period of years the latter will very likely not be the case. The reason the NSE is employed, is because this is a way to test the parameterization of the model. An assumption of the distribution and range of parameters in the model leads to a certain hydrological output such as simulated daily river flow, and by comparing this to recorded daily river flow one can get some understanding about the underlying assumptions.

After establishing the best fit parameters during the validation period using the historic dataset of MIDAS, the model will be tested for its forecasting capabilities for the final year of recorded river flow, 2012. This was a relatively wet year. The forecasting capabilities will be tested by means of a forecast dataset with temperature and precipitation forecasts from the ECMWF. The model will be run as following:

- Since the model starts at 02-01-2012 the endstates of 01-01-2012 of the historic run are used as initial conditions for 02-01-2012. From January 10 onwards, the model accumulates 9 forecasts per day. To illustrate this: on January 2<sup>nd</sup>, there is a forecast for January 2<sup>nd</sup>. On January 3<sup>rd</sup>, there is a forecast from January 2<sup>nd</sup> for January 3<sup>rd</sup> and a forecast from January 3<sup>rd</sup> for January 3<sup>rd</sup>. On January 4<sup>th</sup>, there is a forecast from January 2<sup>nd</sup> for January 4<sup>th</sup>, from January 3<sup>rd</sup> for January 4<sup>th</sup> and from January 4<sup>th</sup> for January 4<sup>th</sup> and so forth until January 10, from then on there are forecasts from day n for day n until day n-8 for day n. For each day the results will be stored in a folder containing the date and 9 subfolders, ranging from 0 (forecast from day n for n) to 8 (forecast from day n-8 for day n). This means that the model will do 9 runs for the year under consideration, using the following structure: on each day n it will use the end states of day n-1 as initial conditions and for x ranging between 0 and 8 it will use the forecasts of day n-x as meteorological input for day n. This process is graphically depicted in figure 2.4 using the forecast at day zero and day 2 as an example:

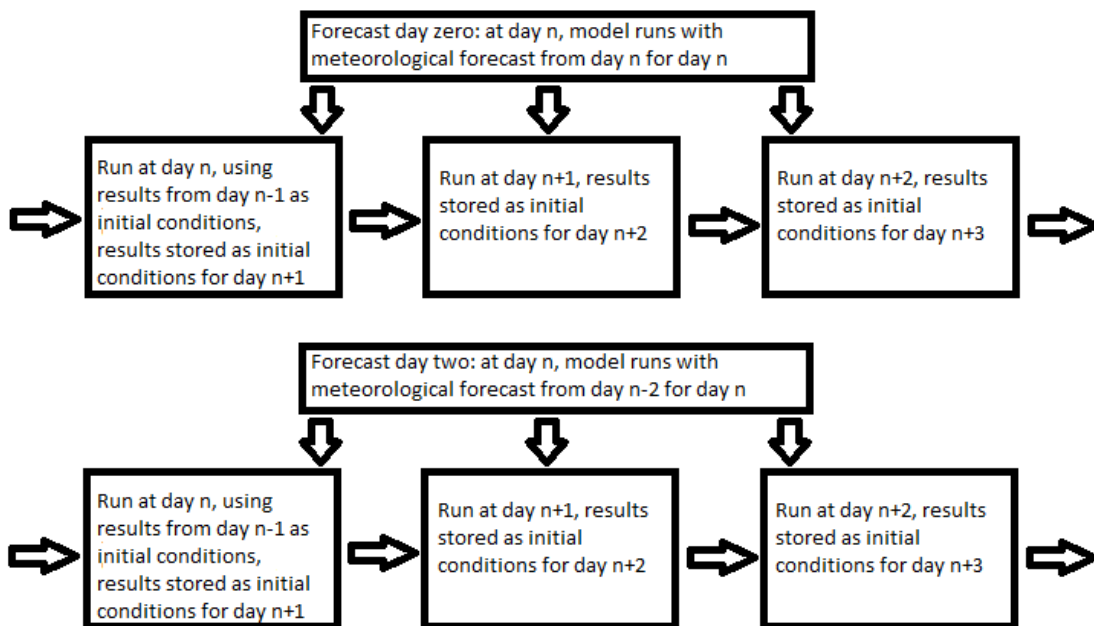


Figure 2-4: Schematization of the forecast model

This means that from January 10 to 31-12 2012 there will be 9 daily values, based on the forecast of day n until day n+8. The following method is employed:

- The forecast at day zero (forecast at day n for day n) will be used as reference forecast, i.e. the most accurate. Forecasts generally lose some prediction power and accuracy the further you look into the future because the inherent chaotic structure of Earth’s atmosphere can become increasingly apparent. For instance, small events over the Atlantic ocean can potentially have significant consequences for the weather outlook in the UK several days later. Therefore, even very small errors in the initial state can lead to large errors later in the forecast. A perfect forecast can only be produced when all initial conditions are known and the model is perfect. Since neither is the

case, above mentioned errors can occur. This is graphically depicted in figure 2.5 (Met Office, 2014):

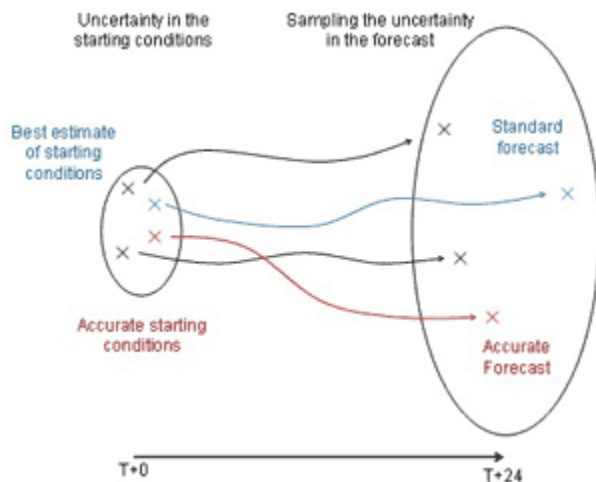


Figure 2-5: Running a forecast with different initial states. From Met Office (2014).

The red line shows the path starting from accurate initial states using a perfect model leading to an accurate forecast. The blue line shows the best guess of the initial states leading to a (sometimes) inaccurate forecast and the black lines show different ensembles (running the model with slightly different initial states, representing uncertainty in observed initial states). By running the model with ensembles an appreciation of the general uncertainty level of the forecast can be obtained. In addition, these ensembles help meteorologists assess how likely a certain weather event will be because they have different outcomes at hand. If there is a lot of variation in the forecasts the meteorologist knows that the prediction is highly uncertain. If the forecasts show a (high) degree of similarity the meteorologist can be fairly confident about a certain weather event occurring (Met Office, 2014)

Using the baseline forecast at day zero, a comparison can be made with forecasts further into the future to answer certain questions: Are rain storms missed the further you look into the future (i.e. on the 11<sup>th</sup> of October there is a forecast of 10mm d<sup>-1</sup> rainfall but on the 6<sup>th</sup> of October there was a prediction of 0 mm d<sup>-1</sup> rainfall for the 11<sup>th</sup> of October) or are they significantly under / overestimated and how does this affect the safety factor, probability of failure, groundwater levels and modeled discharge? Do the forecasts need a correction (i.e. a multiplication factor based on monthly totals of precipitation and subtraction / adding based on mean temperature anomalies).

- Therefore, all 9 scenario's (zero day forecast to forecast 8 days ahead): total (almost) unstable area, which is a summation of the total number of cells showing instability or near instability, i.e.  $SF < 1.1 [-]$ . Furthermore, the daily discharges [m<sup>3</sup> s<sup>-1</sup>] and its fluctuations will be compared. Finally, probability of failure PF [-] and groundwater levels [m] will be compared for one cell in three different unstable area's (so three cells in total), one of them being Hollin Hill. By using one cell in three different unstable areas, it can be assessed how the changes affect cells with different properties (i.e. slope angle, hydraulic conductivity, porosity et cetera). The number three is an arbitrary number, of course more cells or less cells could be compared, but by picking three different areas there is at least some variation in properties such as slope angle, hydraulic conductivity and so on. The time series will be compared (variation of SF, PF, water level and discharge throughout the whole year).

After discussing these results and taking uncertainties and findings into account, it will be discussed how to implement the forecasting model into an overarching EWS.

## 2.2. Area of study

Both Hollin Hill and the catchment of the river Foss at Huntington are part of a network of low-lying hills known as the Howardian Hills, which rise to a maximum of 170 meters above sealevel. The Howardian Hills form a bridge

between the higher grounds of the North York Moors northwards and the Yorkshire Wolds southwards. The western part is divided by many stream tributaries, flowing north or northeastward towards the river Rye, the Holbeck, the Wath Beck or south and southeastward towards Spital Beck, the river Derwent and the river Foss (Carter, 1995).

### 2.2.1. Geological build-up

The Howardian Hills are part of the Lias Group which denotes the first of three sub-divisions from the Jurassic Period: the Lias, Dogger and Malm (Hobbs *et al.*, 2012) and was formed during the Lower Jurassic some 200-170 Ma ago (Gunn *et al.*, 2013). The group is mainly composed of argillaceous sediments such as marine limestones, shales, marls, and clays. The Lias Group forms a continuous band extending from Dorset (South England) through Wales to the coast in North Yorkshire, NE England (Chambers *et al.*, 2011). The root of the word Lias can likely be traced back to the Old French word 'Liais' describing hard layered limestone. The extent of the Lias group is shown in figure 2.6 (Hobbs *et al.*, 2012):

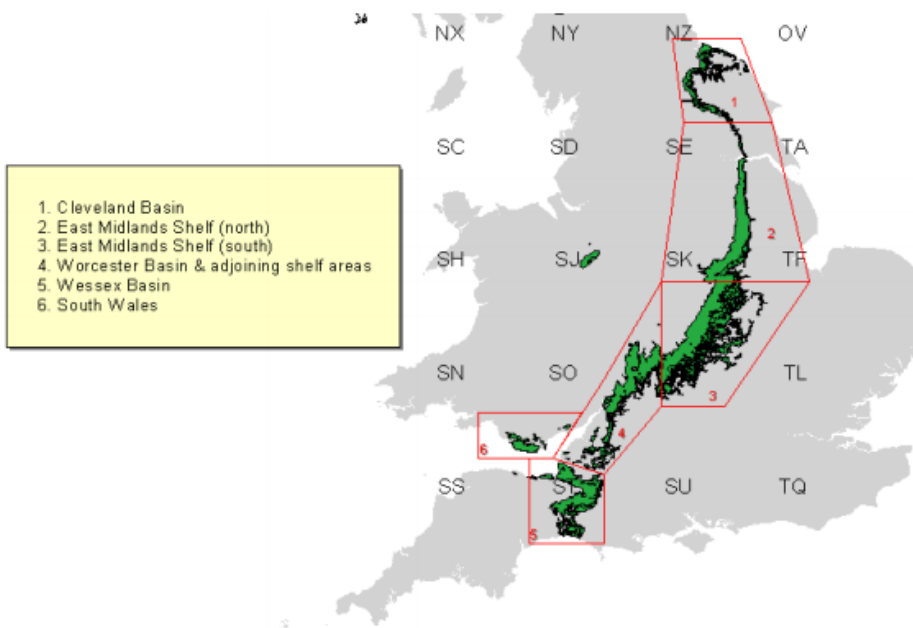


Figure 2-6: The geographical extent of the Lias Group. From Hobbs *et al.* (2012).

Both Hollin Hill and the Huntington catchment are situated in the so called 'Cleveland Basin'. A short geological introduction to the Cleveland Basin formation will be given. For an extended period of time during the last part of the Palaeozoic period (541 to 252.17 million years ago) mountain formation due to continuous tectonic plate movement uplifted the crust in contemporary North-West Europe to such an extent, that at the beginning of the Permean Period (approximately 290 million years ago) modern day Great-Britain became dry land. Subsequently, orogenic stresses tempered and tensional forces became the dominant driving force behind the formation of the North Atlantic Ocean Basin and continental rifting in western Great-Britain (Hobbs *et al.*, 2012). Rifting is the process of Lithosphere deformation (stretching), a continental rift is the zone where this process occurs. Eventually, rifting can lead to the formation of ocean basins. The process of rifting is graphically depicted in figure 2.7 (Geoscience, sine anno):

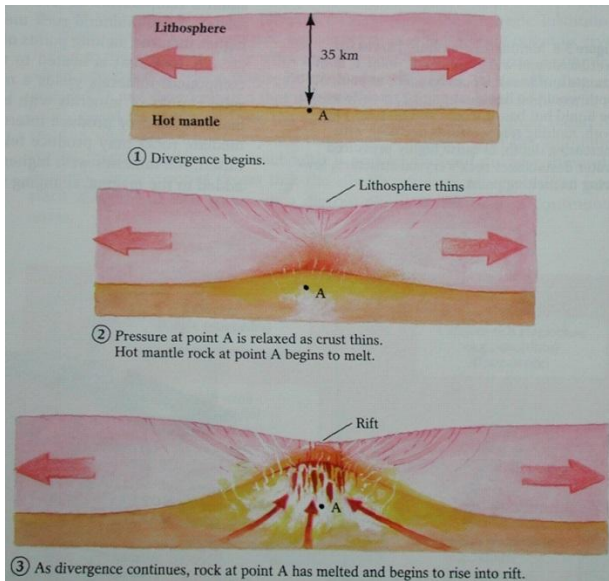


Figure 2-7: Continental rifting in three steps. From Geoscience (sine anno).

As one can see in figure 14, during phase 3 a depression is formed above point A, due to the forces stretching the lithosphere both left and right bound of point A. The combined processes of rifting and stress relaxation resulted in gradual regional subsidence, eventually multiple graben (fault-bounded basins) were formed where subsidence took place at a faster rate than the adjoining higher grounds, namely Wessex, Bristol Channel, Worcester and Cleveland Basin. A graben, according to the USGS, is a “down-dropped block of the earth's crust resulting from extension, or pulling, of the crust” and borders a horst which is formed by “the upthrown blocks that lie next to the graben” and is shown in figure 2.8 (USGS, 2012):

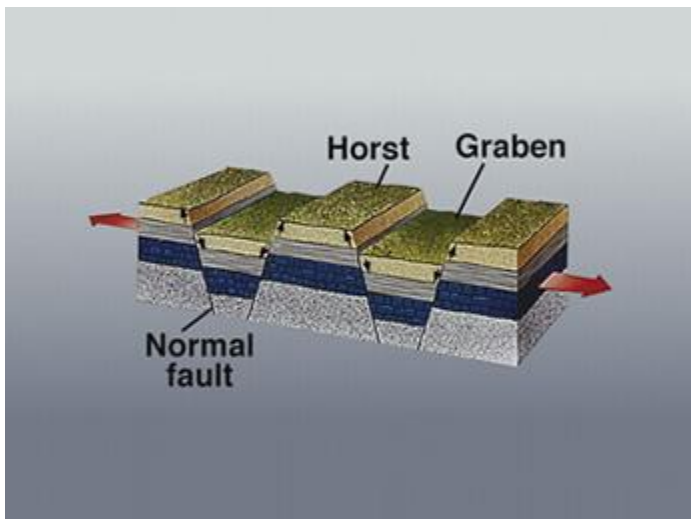


Figure 2-8: Horst and graben. From USGS (2012).

At the final stage of the Triassic period (210 million years ago) the sea began flooding into these basins, depositing marine mudstones and limestones referred to as the Penarth Group. As mentioned before, during the subsequent Jurassic between 200 and 170 million years ago, the sea continued to flood these basins and the Lias was deposited on top of the fairly eroded, bored surface of the Penarth. Subsidence continued to take place in these basins albeit at reduced rates and interruptions in between continuing the buildup of Lias sediments, the Cleveland Basin being 400 to 500m thick.

The Cleveland Basin Lias is comprised of five successions:

- Blea Wyke Sandstone Formation (BWSF)
- Redcar Mudstone Formation (RMF)
- Staithes Sandstone Formation (SSF)
- Whitby Mudstone Formation (WMF)
- Cleveland Ironstone Formation (CIF)

The CIF has a maximum depth of 2m in the area of interest and will therefore not be discussed (Powell, 2010). The BWSF is cut out throughout the whole basin except for coastal areas and will therefore not be discussed as well. First the regional buildup will be discussed, subsequently the Hollin Hill geology will be elaborated upon. The RMF represent the Lower Lias in the Cleveland Basin. The upper 150m are visible at Robin Hood's bay, a coastal place approximately 30 kilometers northeast of the Howardian Hills. The thickness is around 220 to 250m on the coast, and around 300 meters more inland, thinning out westwards to 194 meters at Thirsk, approximately 15 kilometers northwest from the Howardian Hills. Nearing the southward margin of the Cleveland basin, the RMF thins out to around 30 meters at Market Weighton, approximately 20 kilometers southeast of the Howardian Hills. The Redcar Mudstone is characterized by fossiliferous, fissile grey mudstones and siltstones with secondary thin layers of limestone at some levels, and occasional patches of argillaceous limestone.

The SSF forms the Middle Lias group and is around 30 meters thick inland and around 26 meters at the coast, where the formation is visible at Staithes Harbour (40km northeast of the Howardian Hills) and in an area 10 kilometers above Robin Hood's bay. Like the RMF, the formation thins out southwards. Throughout the basin, the RMF is composed of argillaceous silty sandstone and some units of non-argillaceous laminated sandstone. Texture is fine-grained, especially in middle and upper parts, and the RMF has various bedding structures illustrative to deposition during shallow sea conditions. The lower boundary is formed by the transition to the underlying RMF. The upper boundary is marked by a swift transition to the mudstones of the CIF.

The WMF, being the Upper Lias clays, is exposed and at its thickest at 115m at Whitby Harbour (around 50 kilometers northeast of the Howardian Hills), thinning out to around 40 meters at Thirsk. Southwards nearing Market Weighton the WMF is completely eroded. The WMF comprises fossiliferous medium to dark grey mudstone and siltstone, sometimes with thin silty mudstone or siltstone beds. At some horizons hard nodules of argillaceous limestone are also quite frequently present. This Upper Lias Group is very prone to landslides having one of the highest landslide densities in the UK with 42 landslides per 100 km<sup>2</sup> (Gunn *et al.*, 2013 ; Chambers *et al.*, 2011).

The Hollin Hill landslide itself is a complex landslide, extending around 200 meters down slope and a couple of hundred meters in the horizontal plane with a slope of angle of around 12 degrees (BGS, 2014b). At the base of the slope (Lower Lias) at Hollin Hill is the RMF, consisting of grey, silty, calcareous and sideritic mudstone and thin shelly limestones (BGS, 2014b ; Chambers *et al.*, 2011). The layer is up to 200m thick (BGS, 2014b) and is characterized by a poorly drained dayey soil (Chambers *et al.*, 2011).

At the boundary uphill of the RMF, there is a transition to the SSF which is at maximum 20 meters deep and consists of fossiliferous, micaceous, calcareous, and very fine- to fine-grained sandstone (BGS, 2014b). This Middle Lias formation is characterized by a well drained loam soil.

At the transition zone uphill is the WMF, consisting of (dark) grey mudstones and siltstone with calcareous and sideritic concretions, in scattered bands. The formation is about 25m thick and shows a high degree of weathering and poor consolidation, resulting in a poorly drained clay soil. The final formation at the very top of the slope is the Dogger Formation (DF) and is around 2.7m thick. It consists of massive sandstone, fine to medium grained with a

calcareous matrix and a basal part of sandy limestone (BGS, 2014b). The free-draining sandy soil is rich in rock fragments, varying in size from gravel to cobble. The layering as seen from above is shown in figure 2.9 (BGS, 2014b):

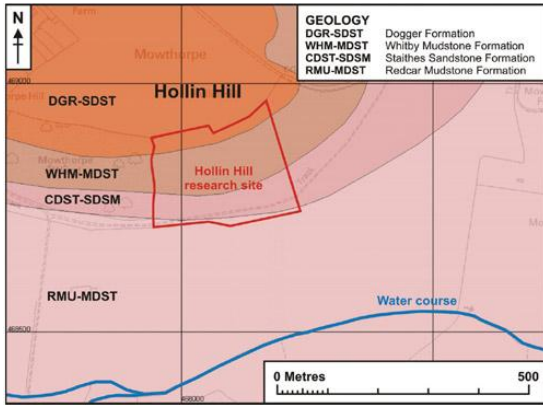


Figure 2-9: The different formations at Hollin Hill as seen from above. From BGS (2014b).

The landslide is primarily triggered by the movement of the WMF (BGS, 2014b). At the top of the landslide the WMF experiences rotational failures (mass moving along a concave plane), moving into translational failures (mass moving along two parallel “plates”) further downslope, dropping over the SSF, subsequently feeding four large slow moving (order of magnitude: tens of centimeters per year) lobes at the toe (BGS, 2014b). A cross-section from the landslide can be seen in figure 2.10 (BGS, 2014b):

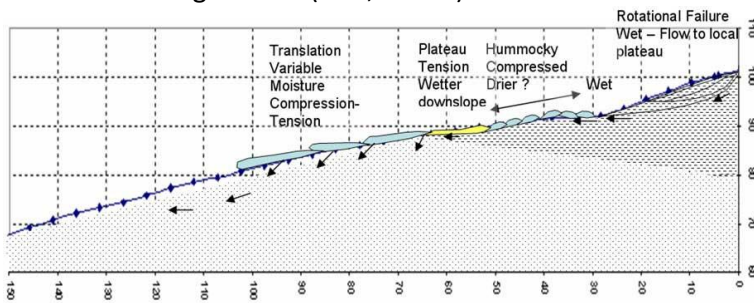


Figure 2-10: Cross-section of the landslide at Hollin Hill. From BGS (2014b).

### 2.2.2. Soil types

Within the catchment of the river Foss and the Hollin Hill subcatchment as can be seen on figure 2.11 (Cranfield University, sine anno), well-drained sandy and loamy soils are present in the catchment as well as slowly permeable clayey soils. A description of soil data can be found in the data chapter.



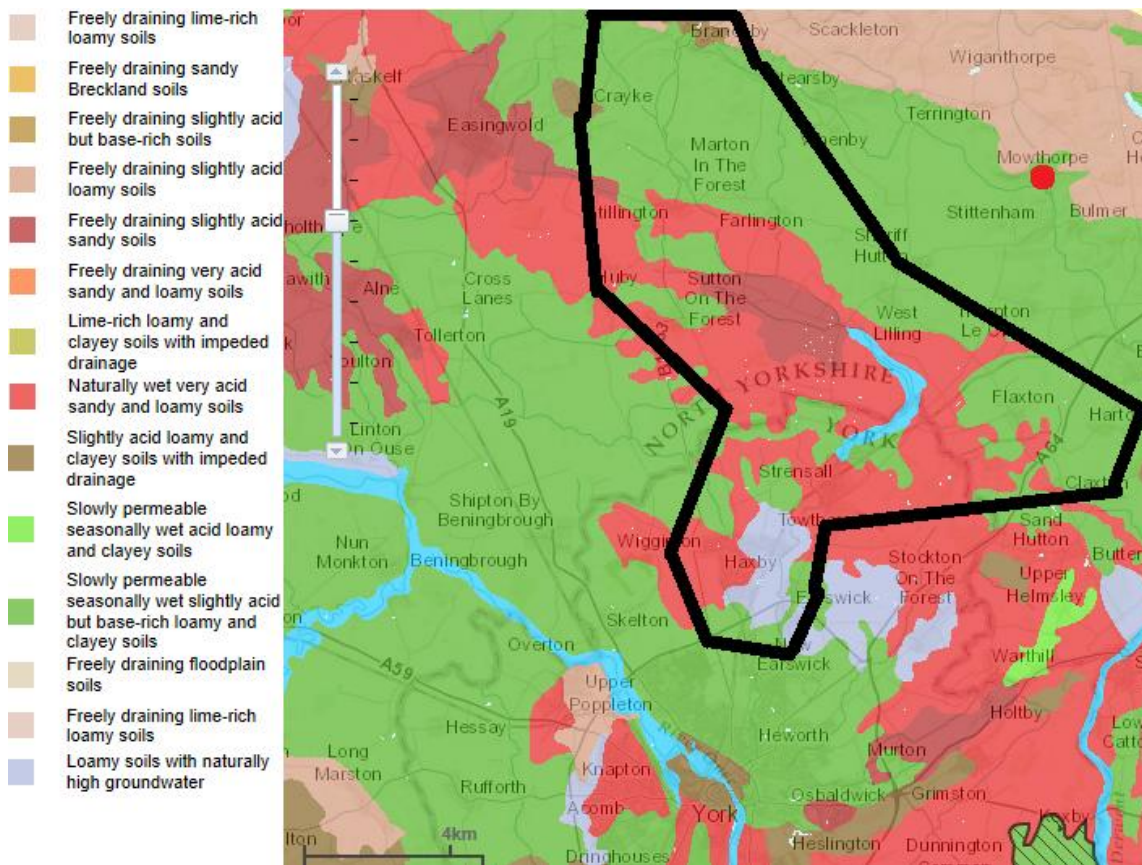


Figure 2-11: Soil types in the Foss catchment and Hollin Hill. After Cranfield University (sine anno).

### 3. Set up of a landslide early warning system

This chapter involves a literature research into how to set up an Early Warning System (EWS). The aim of this chapter is also to show the various aspects of an EWS, in order to get an appreciation of all the components of an EWS, outside of the modeling. First a generic design will be discussed, afterwards real live cases are presented.

#### 3.1. Design of an Early Warning System

When setting up an EWS, it is important to bear in mind that systems are always site-specific (due to different geological and topographical settings) and also purpose-dependent. The design of systems intended for scientific data collection and research (an experimental EWS) differ from an operational EWS intended for public safety, in particular with regard to the frequency of data collection and monitoring of the data itself (Reid *et al.*, 2008). Reid *et al.* show a typical scheme for near-real time landslide monitoring, shown in *figure 3.1*:

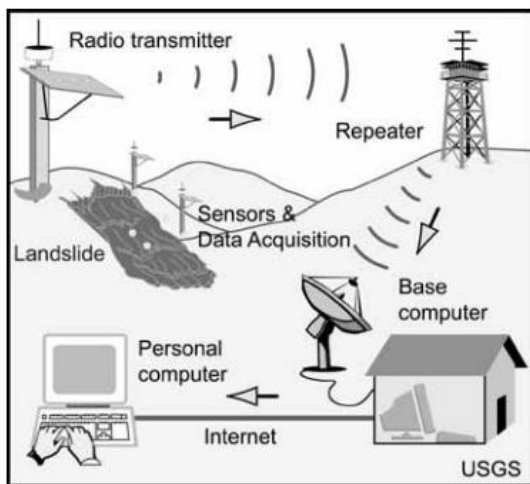


Figure 3-1: A near real-time system design. From Reid *et al.* (2008).

These systems are called near-real time, because they send data to a computer (for instance a groundwater level) that has already been measured, with a slight delay. These monitoring systems are of course a powerful tool in testing forecasting capabilities, especially if motion sensors detecting landslide movement are implemented, because it allows for matching model simulations against recorded events and models can be updated with the acquired data. In addition, they provide the researcher with knowledge about the (geo)hydrological setting of the area under consideration (Reid *et al.*, 2008).

According to the United Nations International Strategy For Disaster Reduction (UNISDR, 2006) and stated by Intrieri *et al.* (2013) the key tasks in the design phase of a landslide EWS are:

“(1) determining the needs and vulnerabilities of the population at risk, (2) identifying any impediments to the population taking action if a warning is issued, and (3) characterizing the geologic and meteorological setting and conditions that lead to landslide initiation. These conditions are referred to as the geo-indicators.”

The research focuses mainly on the third part, but the other two parts cannot be ruled out when setting up a landslide EWS. Therefore these parts will also be addressed in this chapter. EWS may be less effective when an imbalance between these three components is present, thus a lack of instruments, inadequate communication to the public or lack of knowledge about the occurrence and trigger mechanisms of landslides in a certain area (Intrieri *et al.*, 2013). To this end, Intrieri *et al.* developed a generic EWS design, which is shown in *figure 3.2*:

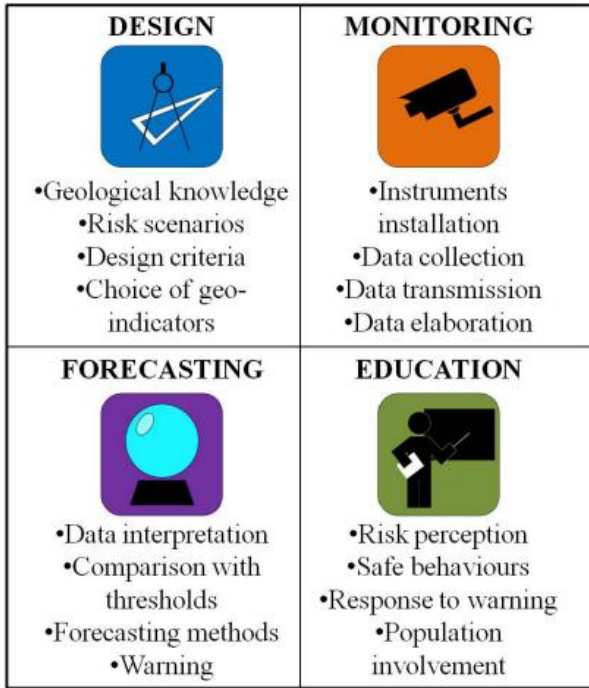


Figure 3-2: Generic design of an EWS. From Intriери *et al.* (2013).

The importance of the first phase and especially risk scenarios is underlined by the study of Opadeyi *et al.* (2003), studying landslide hazards in the Caribbean. One of their findings was that ineffectiveness of past hazard reduction programmes could at least partially be explained by implementing them before a technical appreciation of risks was achieved with all stakeholders involved. It is highly advised to read the findings on this specific point of a joint collection of scientists, professionals, research centers and institutes from the Safeland project, compiled by Bazin (2012).

Also, Dai *et al.* (2002) stated that landslide hazard reduction ideally be integrated in a risk management (based on decision-making) framework, in which levels of acceptable, tolerable and intolerable risk are defined. A general layout for such a framework is graphically depicted in figure 3.3 (Dai *et al.*, 2002):

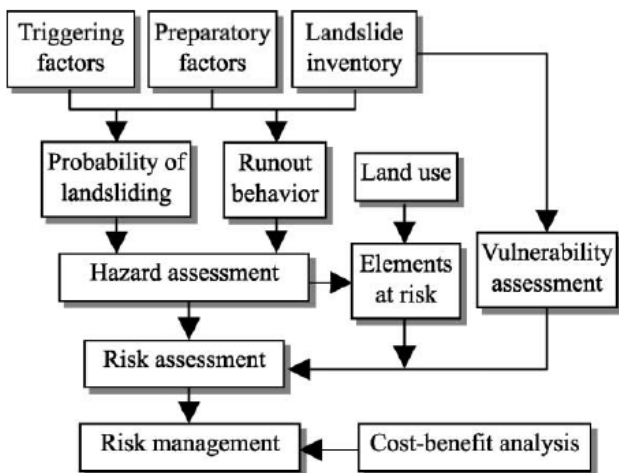


Figure 3-3: Framework for decision making. From Dai *et al.* (2002).

A tolerable risk could be a risk that society accepts in order to secure certain benefits given that it is properly controlled and kept under review (Anderson *et al.*, 2011), an acceptable risk is one that everyone usually accepts

because one can never completely eradicate all the risk (Bazin, 2012). The risk levels could be based on (statistical) estimations of hazards, decisions whether or not these are acceptable and designing controlling measures if not deemed acceptable. This could be based on the ALARP principal (As Low As Realistically Possible). One way of applying this principle is by developing an F-N (frequency versus number of victims, but number of victims can also be replaced with monetary numbers, which is probably more relevant to the UK) curve as shown in *figure 3.4* (Wong *et al.*, 1997):

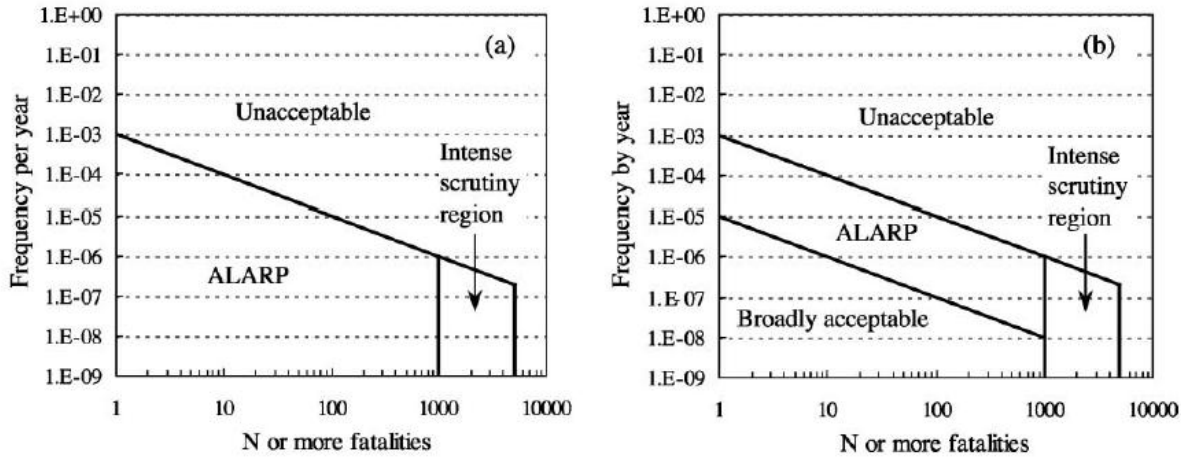


Figure 3-4: The ALARP principle. From Wong *et al.* (1997).

Generally speaking, the idea behind figure 5.4 is that the number of casualties (N) on the x-axis is allowed to increase when the annual frequency F decreases, representing a more serious event. This is based on the ALARP principal, meaning that everything at hand should be done to prevent more damage or fatalities, but would that happen nonetheless, it could realistically not have been prevented. An intense scrutiny region is defined as a region somewhere between the maximum allowed number of casualties or damage and zero. In reality, this could reflect the scrutiny of higher priority regions, which are responsible for the number of victims in the intense scrutiny region of the graph (right hand number on the x-axis minus the left hand number on the x-axis). The unacceptable part is a part where a certain amount of casualties or damage is somehow deemed realistically preventable, due to the higher frequency of the landslide, meaning it is a less devastating landslide. For instance, the graph shows that for an annual landslide, 1,000 victims are unacceptable. The right hand side of the figure shows an alternative scenario, in which a broadly acceptable part is added as an interim solution. It basically implies that fatalities / damages within this region are more or less accepted, because they correspond to more heavy landslides. The consequences of these landslides cannot be mitigated / foreseen entirely, resulting almost certainly in fatalities / damages. The ALARP part is a transition zone, in which it is said again that everything which can be done realistically speaking should be done (corresponding to more frequent landslides). The unacceptable part again corresponds to even more frequent landslides in which realistically speaking a large amount of fatalities / substantial financial damage can be prevented.

Parameters taken into consideration are: landslide probability, routes of debris flow, vulnerability and landslide risk of people and property to landslides, and management strategies and decision making (Dai *et al.*, 2002), Bazin (2012) mentions to also take the displacement rate of the landslide into account. Landslide risk of people for instance is calculated with the following equation (Dai *et al.*, 2002):

$$R(DI) = P(H) * P(S|H) * P(T|S) * V(L|T)$$

In which:

R(DI)= the risk, defined as the annual probability that an individual dies due to a landslide

P(H)= annual probability of a certain landslide event (H)

P(S|H)= probability of the spatial impact given H (S)

$P(T|S)$  = probability of the temporal impact given S (T)

$V(L|T)$  = vulnerability of an individual of an individual given T

For more details, the reader is advised to read Dai *et al.* (2002). As Anderson *et al.* (2011) stated that risk levels should be kept under review, figure 5.2 is best not seen as a static figure, but rather as an iterative figure. For instance, if the forecast phase would reveal a greater risk to society than previously thought, risk scenarios and design criteria could be revised (or less desirable, if a real time event provides new insights). It has to be noted that the methods of Wong *et al.* and Dai *et al.* reflect just a possible means of how to give shape to the first phase of figure 3.3. They could give rise to some ethical questions. One could argue for instance that when speaking about human lives, no fatalities are acceptable or that you cannot simply define an acceptable level because there are no criteria for that. As a counterargument, one could argue that it is impossible to prevent fatalities at all time.

The monitoring part of figure 3.2, including setting up of appropriate data collecting devices (such as pore pressure meters, piezometers, motion sensors, camera's, rain gauges etc.) should take place throughout the lifetime of the EWS. In the initial design phase monitoring should be used to study the landslide behavior, occurrence and (hydro)geological conditions of the area under study. Also, criteria for early warning should be developed (e.g. thresholds). Once the appropriate data has been gathered and analyzed, the monitoring activities become operational, for instance to feed a hydrological and slope stability model, in this case STARWARS and PROBSTAB.

This is the forecasting phase, which is the most difficult phase of the project. Not only model wise, but also socially and legally. A non-conservative threshold (e.g. a certain rainfall intensity forecast, or an alarm that is raised based on model results) could leave too little time for emergency plans, or even miss (the start of) the event at all. A conservative threshold could lead to false alarms, with the public losing trust in the forecasting capabilities. Reliability can also be improved by cross-checking a miss in the field. As Guimarães *et al.* (2003) state, their model seemingly predicted a landslide at an area which was not declared unstable based on their aerial scans. When they crosschecked the area in the field, they discovered that the slope was actually a steep bedrock outcrop. Since the soil depth was overestimated from the interpolated values, the model correctly predicted instability because in reality this bedrock was not able to hold soil as well.

In the forecasting phase, in order to deal with the aforementioned scale problems (mismatch between scale of triggering mechanisms and scale of mapping) one could chose to use coarse regional mapping and zoom in (increase model resolution) on areas of high(er) priority, by means of a priority ranking system. Criteria can be probability and consequence of failure, which if multiplied reflect the relative risk of a landslide (Dai *et al.*, 2002). For instance a high probability landslide with low consequences could receive a moderate relative risk. The areas with the highest relative risk could be given more priority, for instance with regards to increasing model resolution and / or extra monitoring with various instrumentation. This is also stated by Bazin (2012), who concludes that priority should be given to areas with higher risk with regards to monitoring and observing. In addition, this is in line with respects to the study from Anderson *et al.* (2011), mentioned in the discussion of the hydrological results. They argued that in order for a design and construction of landslide hazard reduction to be fully effective, spatial resolution of landslide hazards needs to match the scale of the instability processes. Intriati *et al.* (2013) developed a flow chart with regard to the instrumentation for various landslides and budgets (figure 3.5) and organization (figure 3.6), based on existing EWS around the world. Note that these were designed for slope scale EWS, but when doing regional slope stability assessment one can use this scheme for areas with high priority for validation of the coarser regional model and/or more detailed slope stability modeling. Intriati *et al.* themselves state that this flow-chart is highly flexible and modular, components can be added when moving from slope to regional scale or the flowchart can be implemented in an overarching framework of a regional approach. The attractiveness of such an approach lies in the fact that when zooming in on individual slopes it can be possible to use high resolution DEM, such as used in the earlier

mentioned slope stability study of Guimarães *et al.* (2003). In such a high resolution DEM, site specific slope stability aspects such as curvature of concave hill slopes are accurately reflected, meaning the mismatch between mapping and the scale of slope instability mechanisms as stated by Anderson *et al.*(2011) is removed. This can possibly improve the slope stability assessment. An advantage of zooming in on specific sites is that you do not add detail where it is not needed for slope stability assessment, such as in low-lying parts with no slopes. Also, if a more detailed level of slope stability modeling is required for a certain slope, the data collection could also be matched with the desired level of accuracy: samples providing information about soil depth, hydraulic conductivity, porosity and angle of internal friction can be taken around and on the slope, also serving as validation for some spatial interpolation into the regional model. Regional modeling can also reveal which areas should be paid more attention to (Intrieri *et al.*, 2013). But, as Bazin (2012) and Intrieri *et al.* (2013) state, regional slope stability modeling is a) the most cost-effective measure and b) also a more reliable measure (in case of shallow landslides and debris flows) because landslides are diffuse events, usually not recurrent to one given site and you cannot monitor each separate site.

### Instrumentation flow-chart

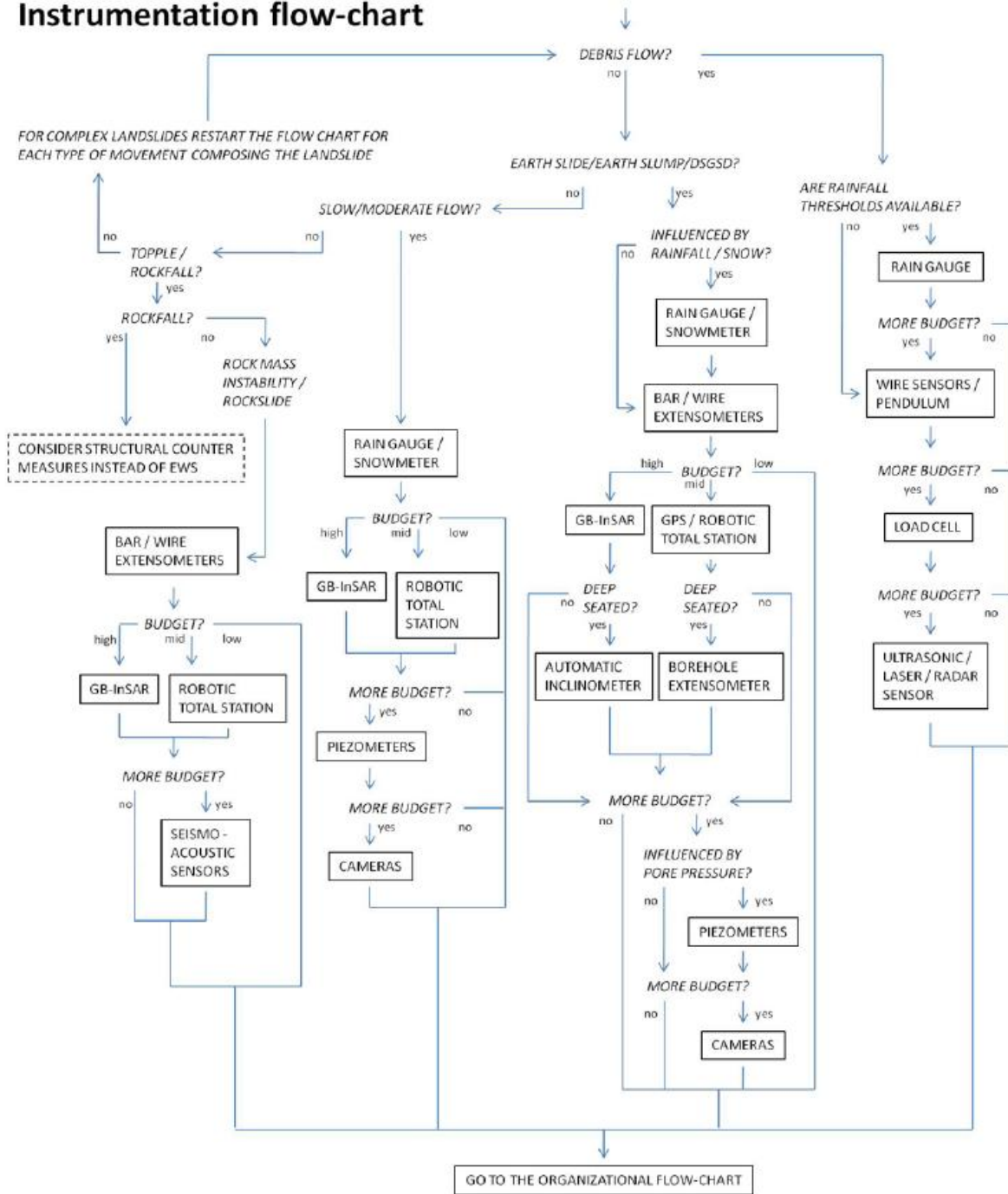


Figure 3-5: Flowchart for slope scale monitoring. From Intrieri *et al.* (2013).

# Organizational flow-chart

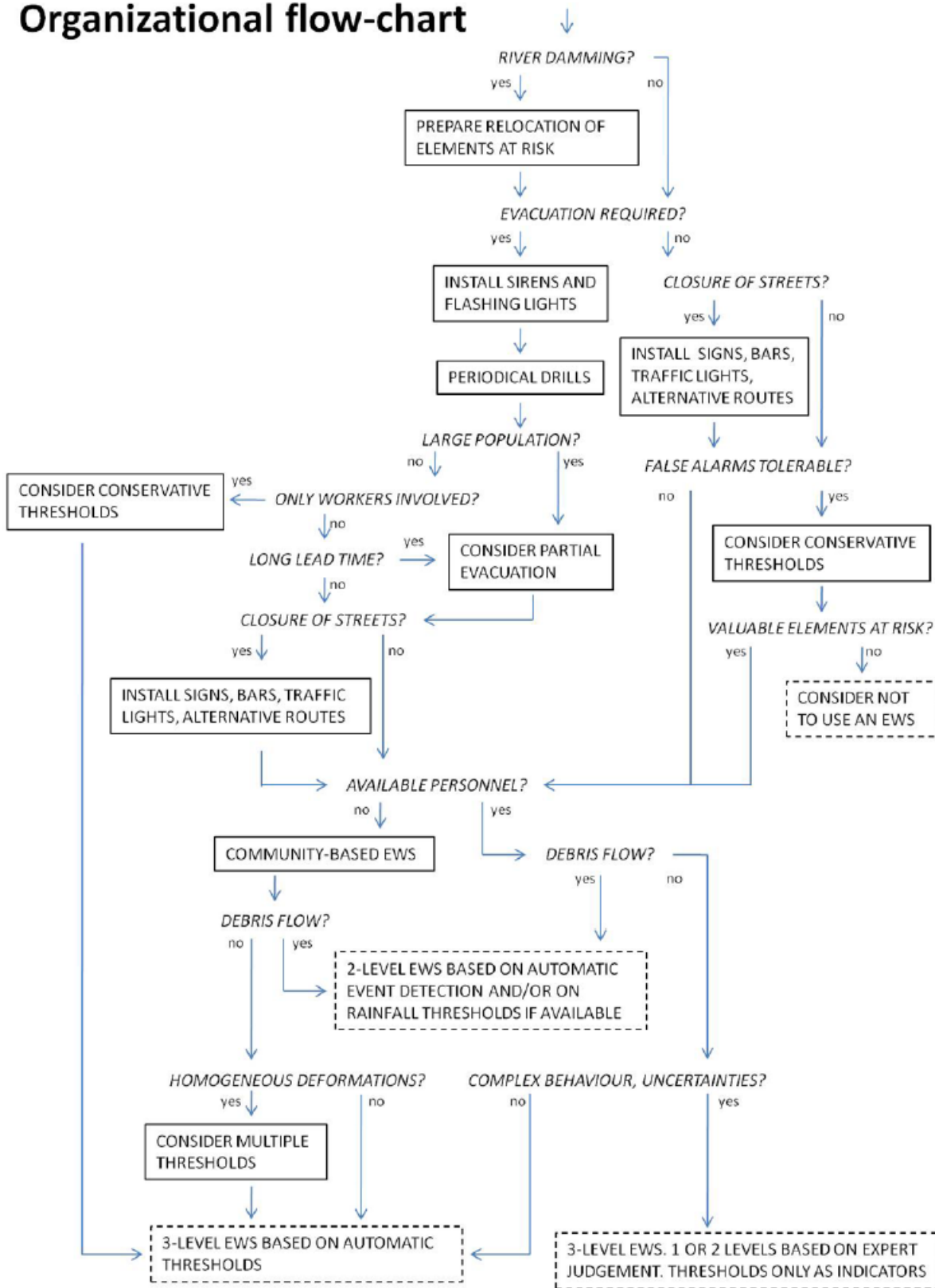


Figure 3-6: Organizational flow chart. From Intrieri et al. (2013).

For the education phase, scale can be an advantage. If regional landslide modeling is conducted, the authorities can also use regional means: SMS, radio / TV broadcasts, or social media such as Twitter. However, when electric supply is cut off these measures fail to produce the desired results or they are not always specified for the area which is under threat (Garcia & Fearnly, 2012). A flow chart involving the public is shown in figure (Bazin, 2012):



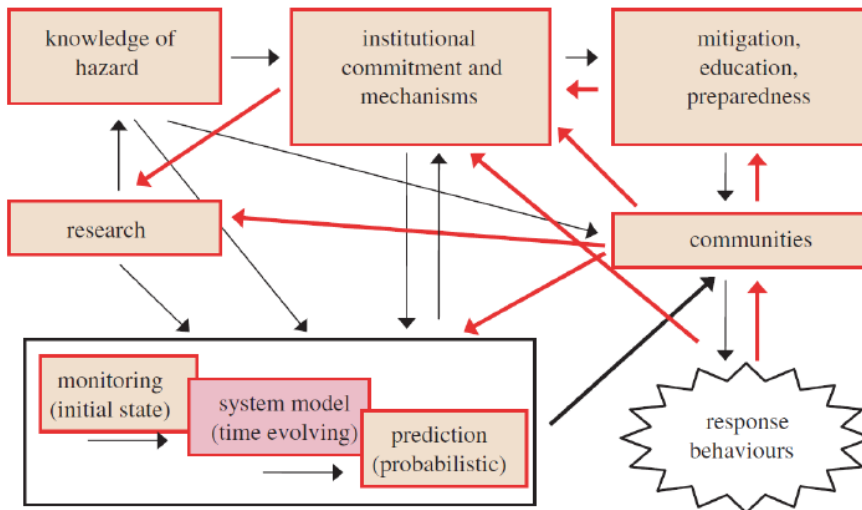


Figure 3-7: Flow chart showing the different actors involved. From Bazin (2012).

This figure shows the interrelatedness of the community in setting up an EWS. The most important aspects to consider:

- Researchers have in depth knowledge about hazards and are responsible for the functionality of an EWS. They can advise on design parameters, but also be integrated into the EWS by means of collecting and analyzing data.
- Government bodies like police, emergency services and (local) authorities, should be included in the system: by providing input to the system such as minimum response time, ways of providing information, and receiving information from the system on the other hand. They should therefore have knowledge about the hazards.
- Mitigation, education, and preparedness are some keywords for the EWS. Information ideally should trigger a response from the people, matching the seriousness of the early warning given. People therefore have to know how to respond by means of education and instruction, and this skill has to be regularly revised. One way of doing this is by conducting full-scale tests, to learn about shortcomings of the system and gain practical knowledge.
- The structure of the community has to be taken into account when designing an EWS but also when educating the people: What are the geographic and (technical) infrastructure? Is the community remote or densely populated? Also, local knowledge and experience should be involved into the EWS. The added bonus of this is that local and anecdotic knowledge about historic events, landslide susceptible areas, geological setting, weather events and such, is often unaware to professionals or researchers “out of town”. Also, locals can come with refreshing ideas and involving local people can also help delineate responsibilities and risk mitigation better (who bears which risk and who is responsible for what?). Also, data should be made publically available and stored in databases.

### 3.2. Existing EWS

Bazin (2012), compiled statistics about the physical part of the EWS and it was concluded that in 82% of the 23 surveyed cases, geo(morpho)logy was a criterion for design of an EWS, followed by displacement data in 63.6 % of the cases, hydrogeology and modeling both in 54.5%, as shown in figure 3.8:

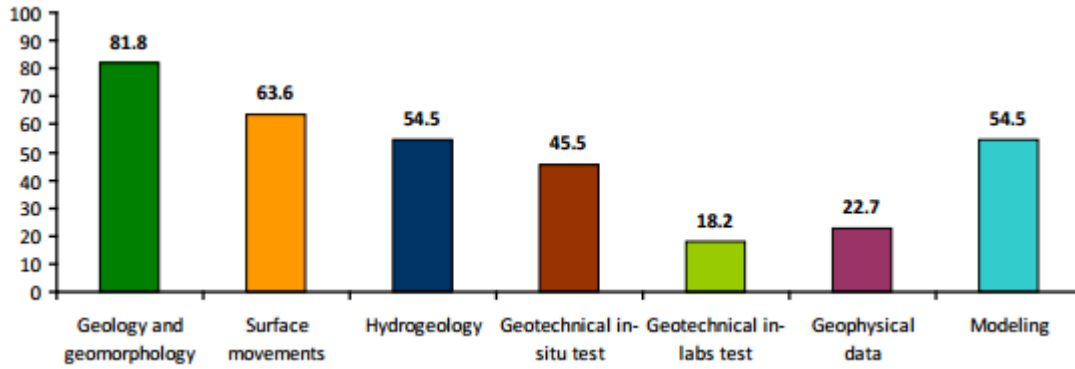


Figure 3-8: Physical aspects of an EWS. From Bazin (2012).

Statistics about the use of the above mentioned criteria were also compiled (note that if for instance 4 criteria are used, it can be a combination of any four out of the seven above mentioned criteria) and shown in figure 3.9:

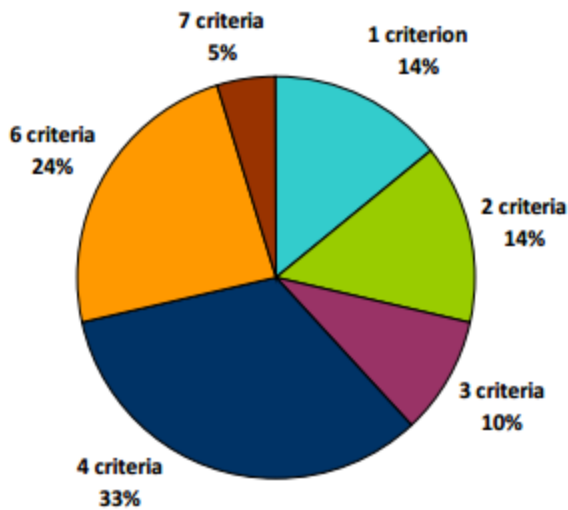


Figure 3-9: Number of criteria used. From Bazin (2012).

The statistics about which devices and their purpose are used are shown in figure 3.10:

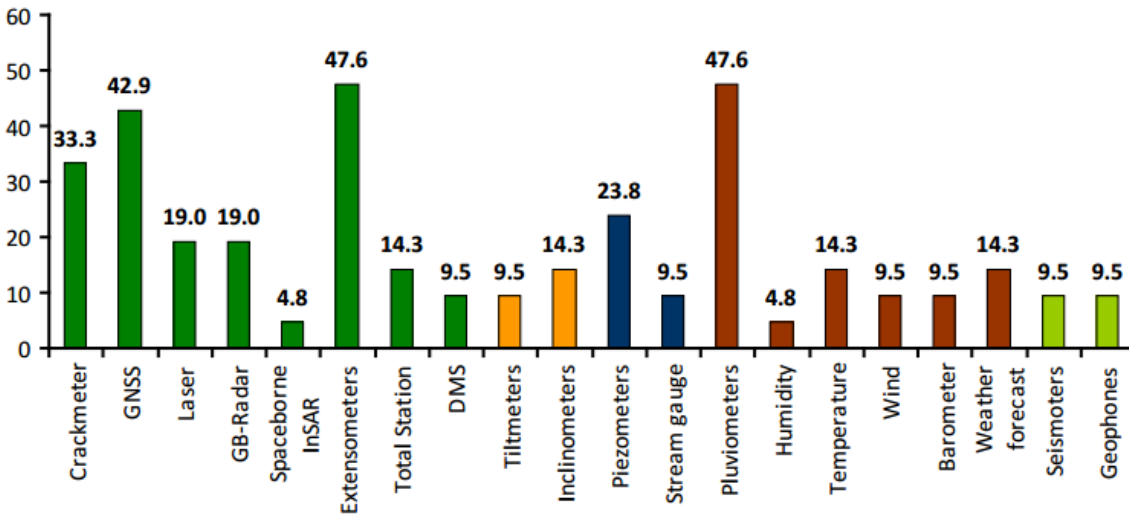


Figure 20: Percentage of the 21 landslides which use those instruments in order to monitor displacement (in dark green), angle changes (in orange), water level conditions (in blue), weather (in brown), and seismic activity (in light green). The total percentage is over 100, because a landslide can be monitored using several instrument types.

Figure 3-10: Types of instruments used. From Bazin (2012).

Criteria for means of alert are shown in figure 3.5:

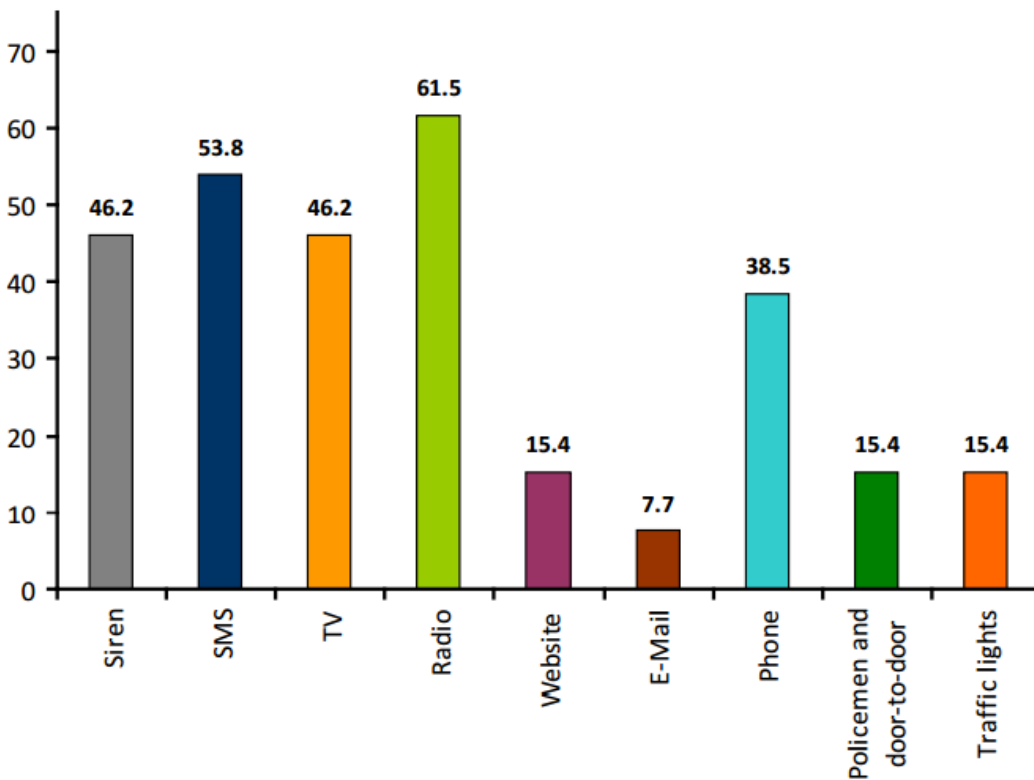


Figure 3-11: Means of alert. From Bazin (2012).

Curiously, even though several of the EWS use multiple criteria to alert people, only in a minority of the cases a real associated scenario to a certain threshold level was present, in line with the earlier mentioned findings of Opadeyi *et al.*

Finally, and very important, the practical challenges with regards to the EWS are shown in figure 3.12:

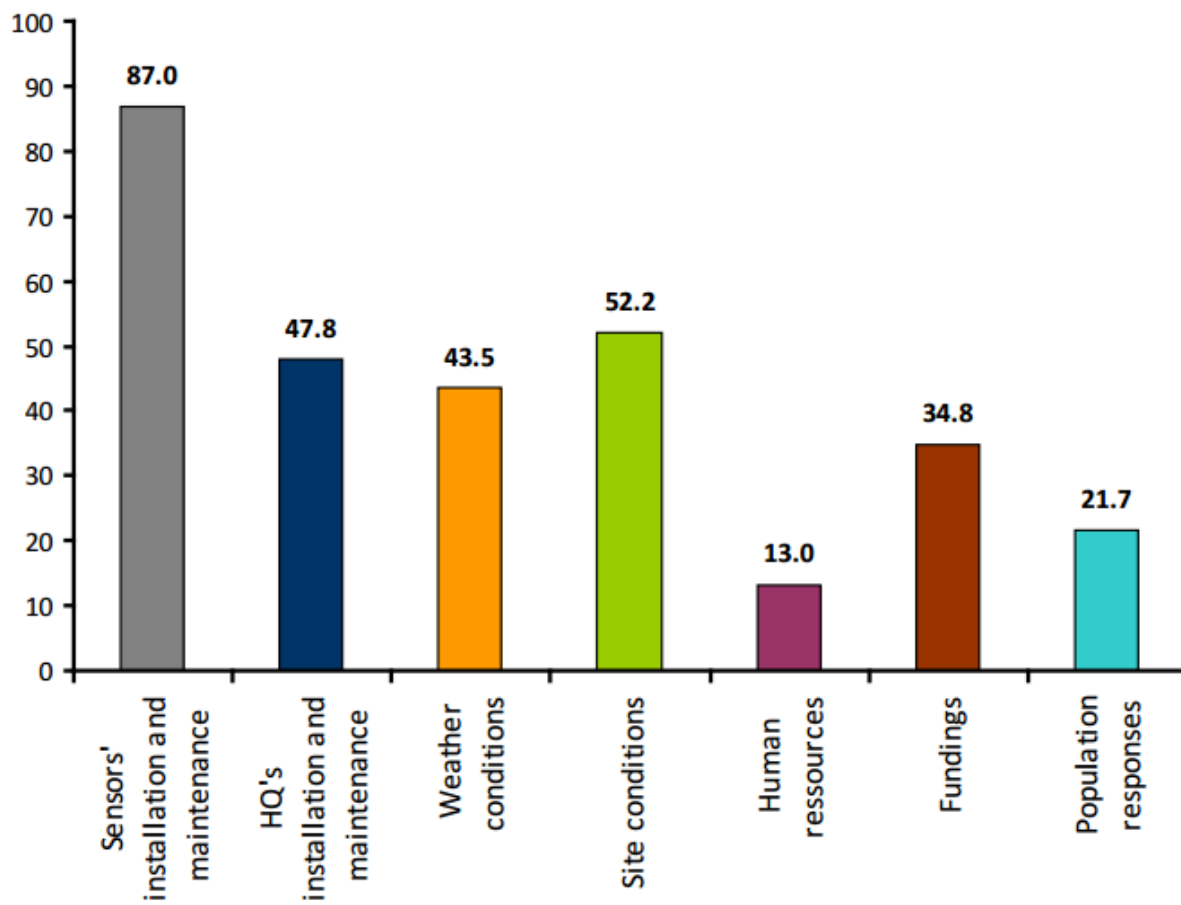


Figure 3-12: Practical challenges. From Bazin (2012).

## 4. Data collection, processing and model setup

Data serve as input for models and to gain insight in the characteristics of the area of study, such as soil properties and land use. New data can also be derived from existing data, for instance using pedotransfer functions one can estimate soil hydrological properties such as hydraulic conductivity based on textural information such as silt, sand and clay fractions. Data can also be used to validate a model by means of comparison of observed data and simulated (model) data. This chapter deals with the various type of data that were collected for this study, why they were collected and how they were processed.

### 4.1. Geographic data

Geographical data serve as input for the STARWARS model. The following geographical data were used: a 10m resolution Digital Elevation Model (DEM) of a larger part of the Yorkshire region and a 10m resolution river network provided by the BGS. Using the DEM and the river network, the Local Drainage Direction (LDD) can be created using PC Raster. The LDD calculates for every cell to which neighboring steepest downslope cell a certain material (in this case water) will flow, hence creating the flow direction over the entire area of interest and shown in figure 4.1 (Utrecht University, 2014):

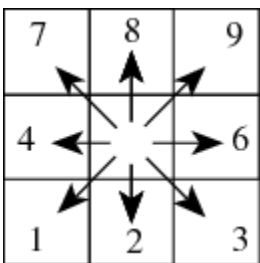


Figure 4-1: Flow directions of material in a raster cell. From Utrecht University (2014).

The numerical keypad numbers represent the flow direction, the number 5 represent a pit (a cell with no flow direction because all its neighboring cells have an higher elevation). The middle cell does not represent a pit, this cell can have any value, indicating that water can flow to any direction where elevation is lowest. So if the middle cell has value 2 the water will first flow downwards and if the cell right below the middle cell has the lowest elevation, water will continue to flow in the direction of 3 and so forth, until it meets a pit. Usually these pits are caused by grid mismatches, data (processing) errors or a poor resolution. Commonly, one wants to remove a pit since in most cases these are not physical but digital entities. This can be done when creating the LDD, the command for creating the LDD in PC Raster can be found in appendix 1.1.

The LDD does not always follow the real river pattern, since it is created based on a DEM which may possess above mentioned intrinsic flaws. To ensure the LDD follows the river pattern, the “burning” method has been employed (Wesseling, van Deursen, & de Wit, 1997). First, the River grid has been converted from an ArcGIS file format to a Boolean (True or False) PCRaster .map file using GDAL where a river is true and no river is false. Subsequently, a relative DEM .map file with values ranging between 0 and 1 is created. The procedure for this is elaborated on in appendix 1.2.

With the proper LDD created, the total catchment delineation (as good as possible) can be reproduced in PC Raster, since this cannot be converted directly to a PC Raster .map file. This is done by using the *catchment* statement in PCRaster (Appendix 1.3). In addition, the subcatchment of Hollin Hill was added and both the catchments were merged (appendix 1.4). This resulted in the nominal map as seen in figure 4.2 in which 1 is the subcatchment of Hollin Hill, 2 is the catchment of Foss at Huntington, 3 is Hollin Hill (the green dot with the upper black circle around it), and 4 is the outer extent, i.e. the area outside of the catchment. The black circles in the lower corner and the

right corner represent where the outflow points are situated. The outflow point in the lower corner will be used for the validation of recorded discharge and modeled discharge, since for this outflow point there is measured river flow data available. The outflow point of the Hollin Hill subcatchment will not be used for validation, since this is an arbitrary outflow point to create a subcatchment that includes Hollin Hill, hence no flow data exists for this point. The idea behind adding the Hollin Hill subcatchment to the Foss catchment is that one can assume some degree of similarity between soil and elevation characteristics for the unstable parts of both the Foss catchment and the Hollin Hill subcatchment, since they are both part of the Howardian Hills and close to each other. The river flow validated parameters from the Foss catchment can then assumed to be valid for the Hollin Hill subcatchment, with a degree of uncertainty due to spatial variability. In addition, extra uncertainty can be expected due to the fact that recorded river flow can be influenced by human interferences on the river system such as abstractions, sewer discharges and so forth which are not included in the model, hence the best fit parameters may mistakenly be validated against such processes which add to river flow fluctuations but are independent of any parameterization.

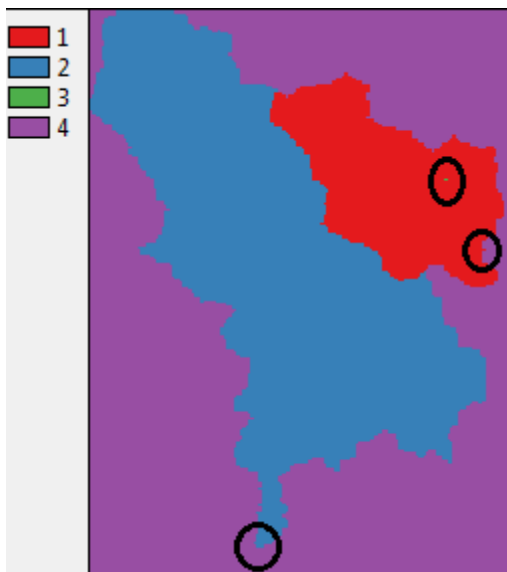


Figure 4-2: Catchment of Foss and Hollin Hill

#### 4.2. River flow data

Daily river flow data from 15-04-1987 until 31-12-2012 from the Centre of Ecology and Hydrology (CEH) has been used from the outflow point of the catchment (the Foss at Huntington). Daily river flow data are a means to validate and calibrate the STARWARS model. STARWARS calculates outflow from the outflow point in [ $\text{m}^3 \text{s}^{-1}$ ] after a model run, the same unit as the measured daily discharge. If the modeled outflow captures the pattern of recorded river flow fluctuations, or even better, nearly matches the recorded discharges, the model performs (reasonably) well.

#### 4.3. Historical precipitation data

Daily precipitation data were used as historical meteorological input for the STARWARS model for the historic run. Precipitation data was taken from the MIDAS database from NERC (NERC, 2014), from the same timeframe as river flow records. NERC data is labeled with a quality check, in which 0 stands for unchecked data (unreliable) and 1 for confirmed and checked data (reliable). Only precipitation values with value 1 have been used, to ensure accurate meteorological input for the model. Data from all eight stations from 01-01-1988 until 31-12-2012 were used. The stations were selected because of their complete records (some stations nearby did not have complete records) and because they had quality checked data. In addition, they are fairly close to the catchment, which is important because uncertainty increases with distance as shown in figure 4.3. Note that this is a GIS picture without the Hollin Hill subcatchment, because in PCRaster point locations in a nominal map are quite small and on a larger picture like this they would simply not be visible.

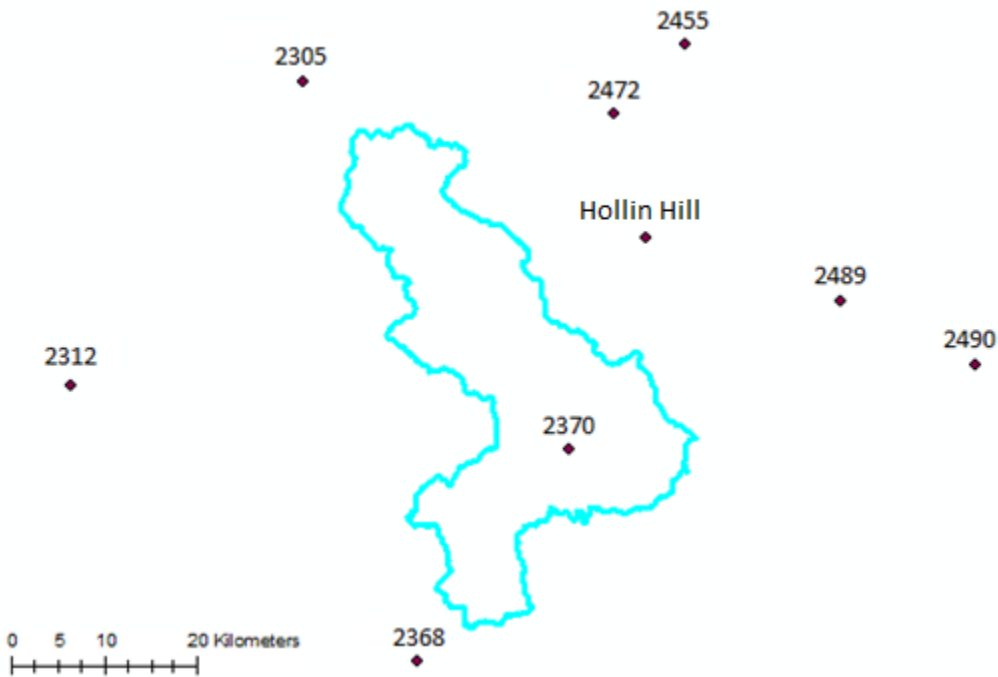


Figure 4-3: Rainstations around Hollin Hill and the catchment

The point data were spatially distributed by means of an Inverse Distance Weighing (IDW) procedure described in appendix 2.1. The spatially distributed rainfall values were resampled to the extent of the clone map, where the catchment is true and the area's that are not part of the catchment are false. This is because the geographical extent of the rain station's exceeded the catchment geographical extent (most stations were actually outside the catchment) and the area of interest for daily precipitation values is only the catchment itself. In figure 4.4 a spatially distributed rainfall pattern is shown for some visual clarification, from 01-01-1988 in  $\text{mm d}^{-1}$  (in the model this will be converted to  $\text{m d}^{-1}$ ):

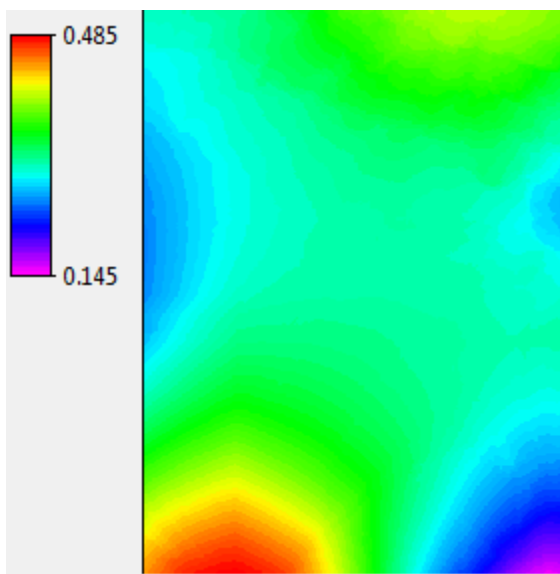


Figure 4-4: Spatially distributed rainfall pattern

This picture has the same spatial extent (number of columns and rows and coordinates) as the clone map and only the area within the clone map will be used as input for the model. The spatially distributed rainfall pattern is the

result of the input of all eight rainstations, each having a different weight value in a certain cell depending on their distance to that cell (inverse correlation, weight decreases with increasing distance).

#### 4.4. Historical temperature data

Recorded temperature data were used to calculate seasonal (and land use) dependent parameters such as as Leaf Area Index (LAI) [-], being the ratio of the one-sided leaf tissue area over the total soil surface area. In addition, precipitation interception/storage  $S$  [L] being a fixed maximum amount of precipitation that can be stored by the canopy was calculated. Vegetation fraction cover  $C$ , [-] being the fraction of surface area that is covered by a certain vegetation type and precipitation through fall ratio FTR [-] being the fraction of rainfall that reaches the surface ( $1-C$ ) were determined as well. Finally, the crop factor  $k_c$  [-] being a crop dependent correction factor used in the following equation (Allen *et al.*, 1998):  $ET_c = k_c \cdot ET_0$  in which  $ET_c$  is the crop corrected evapotranspiration [ $L T^{-1}$ ] and  $ET_0$  is the reference potential evatranspiration [ $L T^{-1}$ ], based on a standardized grass crop.

$ET_0$  was calculated using the Hargreaves equation (Shahidian, Serralheiro, Serrano, Teixeira, & Haie, 2012 ; Samani, sine anno):

$$ET_0 = 0.0135 * K_T * (T_{avg} + 17.8) * (d_t)^{\frac{1}{2}} * R_a$$

In which:

$R_a$  = Total receiving extraterrestrial sun radiation, using same units as evaporation [ $L T^{-1}$ ]

$K_T$  = An empirical coefficient, 0.162 for land mass regions and 0.190 for coastal regions[-]

$d_t$  = Mean maximum temperature – mean minimum temperature per timestep [ $^{\circ}C$ ]

$T_{avg}$  = Mean temperature per timestep [ $^{\circ}C$ ]

How  $R_a$  is shown in appendix 2.2. The same procedure was followed as for the precipitation input (filter to use only label 1 values, inverse distance weighing). In addition, missing values for a given day  $x$  were estimated from an interpolation of known values between an interval of  $x-2$  and  $x+2$ . If those values were also missing values, the long term daily average for that given day (expected value) was used to replace the missing value. So for instance, if a particular station had a missing value on 03-04-1978, the mean known value from every value for the date of 03-04 for any year were taken in order to replace the missing value for the expected value of that day (i.e. the long term mean). If an entire month or year was missing, it was left out because using long term mean to replace missing values would become a fairly inaccurate procedure due to temperature anomalies in the long term. Temperature stations which have been used are shown on figure 4.5:



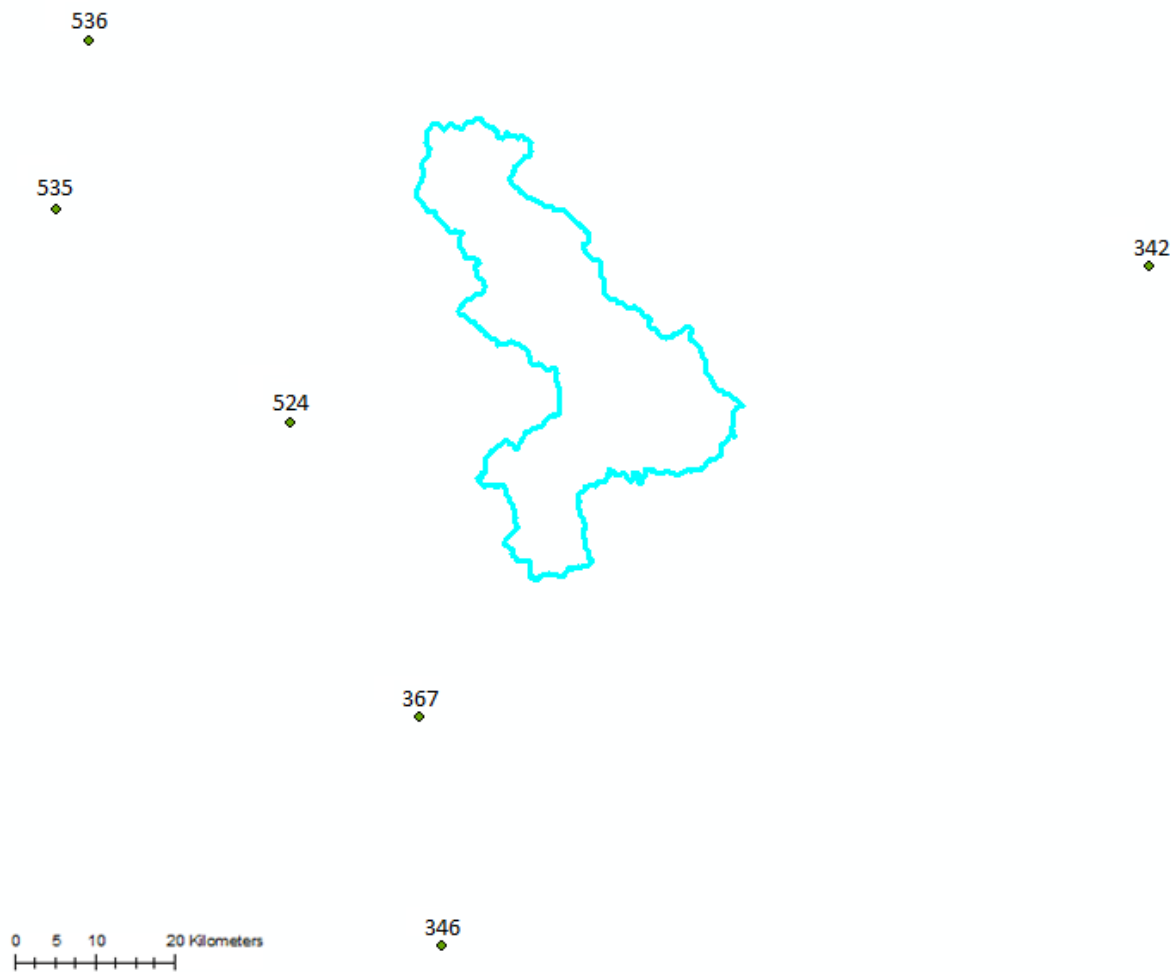


Figure 4-5: Temperature stations used

Also for the spatially distributed temperature values the resample procedure was applied, since only the temperature distribution within the catchment is of interest.

#### 4.5. Forecast data

For the eight day forecast run forecast data were used from the EMCWF. The forecast is a medium-range (3-10 days) 9 day forecast, starting from 01-01-2012 at 12 noon. From 01-01-2012 12:00 until 02-01-2012 0:00 there will be a 12 hour increment for precipitation, followed by 36, 60, 84 and so forth. For the temperature, on the 1<sup>st</sup> of January there is a forecast for the 1<sup>st</sup> of January consisting of two six hour increments, and one for the 2<sup>nd</sup> of January with four six hour increments, and for the 3<sup>rd</sup>, 4<sup>th</sup>, 5<sup>th</sup>, 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup>. The forecast for the 1<sup>st</sup> of January will not be processed, since this is not a complete forecast (from noon to midnight instead of midnight to midnight). With respect to precipitation, the values are summations, meaning that on day n the value of day n+1 has to be subtracted. With respect to temperature, the maximum and minimum temperature are required for the Hargreaves equation to estimate PET. The maximum temperature is simply the maximum value of one of the four blocks and is expected during the day. The minimum temperature is the minimum value of the four blocks and is likewise expected during the night.

The forecasts have a resolution of  $0.125^\circ$  or  $13.5 \text{ km}^2$  and are derived from a set of model equations, of which some are diagnostic, describing static relationships between input such as pressure, density, temperature and height. Other equations are prognostic and describe temporal changes on the level of an air cell for input such as horizontal wind, temperature, surface pressure and water vapor content. An additional set of equations describe changes in hydrometeors (snow, rain, cloud ice content, liquid water). Processes occurring on a smaller scale than the model

resolution, such as radiation, vertical turbulence, gravity wave drag, surface interaction and cloud formation are included statistically as parameterization processes, hence their effect is included. For example, dose surface air exchanges heat with the earth surface through a process of turbulent diffusion or convection, driving air towards a state of neutral stability. For an elaborate description of the physical and numerical modeling used the reader is advised to read the ECMWF user guide (ECMWF, 2013). A total number of 9 of these 13.5 km<sup>2</sup> grid cells were used, some within the catchment and some (barely) outside of the catchment, 2 in the lower part, 4 in the middle part and 3 in the upper part (this has to do with the shape of the catchment, being very small in the lower part and wide in the middle part). There is no GIS map, but the coordinates are shown in table 3.1 in Latitude / Longitude (for comparison, the coordinates of Hollin Hill are 54.112 -0.959):

Table 4-1: Latitude and Longitude of forecast data

	Lat	Lon	Lat	Lon	Lat	Lon	Lat	Lon
<b>Lower</b>	54.000	-1.125	54.000	-1.000	54.000	-0.875		
<b>Middle</b>	54.125	-1.250	54.125	-1.125	54.125	-1.000	54.125	-0.875
<b>Upper</b>	54.250	-1.250	54.250	-1.125				

#### 4.6. Land use data

Land use information about the catchment is needed as an input for STARWARS to obtain land use/vegetation and time/temperature dependent characteristics such as Leaf Area Index [-] (LAI), which is a ratio of the total area of leaves [L<sup>2</sup>] above a certain ground area [L<sup>2</sup>], rainfall interception by vegetation S [m], vegetation cover C [-], rainfall throughfall ratio FTR [-] (1-C), and crop factors  $k_c$ .

To obtain land use or cover the Corine Landcover 2006 dataset was accessed through the server of the Utrecht University. For the LAI, S, C, FTR and  $k_c$  for each land use type the mean temperature of one station has been taken as input for simplification reasons, being Linton on Ouse (524). This station was closest to the catchment (and there was very little temperature variation anyway). LAI is calculated by the following equation (Hagemann *et al.*, 1999):

$$LAI = LAI_d + LAI_{rel} * (LAI_g - LAI_d)$$

In which:

$LAI_d$  = LAI during dormancy [-]

$LAI_{rel}$  = relative LAI throughout the year as a function of temperature between 0 and 1 [-]

$LAI_g$  = LAI during growing season [-]

And (Dickinson, Henderson-Sellers, & Kennedy, 1993):

$$LAI_{rel} = 1 - \left( \frac{T_{max} - T}{T_{max} - T_{min}} \right)^2$$

In which:

$T_{max}$  = Maximum monthly temperature [°C] under condition  $T_{max} \geq 5$  °C

T = Monthly temperature [°C]

$T_{min}$  = Minimum monthly temperature [°C] = 5 °C

The threshold of 5 assumes that growth is not occurring whenever  $T < 5$  °C for extended periods. Henceforth,  $LAI_{rel}$  equals 0 when  $T \leq T_{min}$  and 1 when  $T = T_{max}$ . Subsequently  $S$  [m] can be obtained by multiplying LAI by 0.0002. The procedure for parameters  $C$  and  $k_c$  is detailed in appendix 2.3.

In order to obtain the correct values for  $LAI_d$ ,  $LAI_g$  and  $h$ , serving as input for the calculations, land use classifications of Olsson (USGS, 2008) with matching aforementioned values per land use as given by Hagemann *et al.* (1999) have been used. Land use classifications have been matched with the land use classifications obtained from the Corine dataset. The original ten classes of the Corine landcover dataset were replaced by eight new land use classes, since the two merged classes of the Corine dataset were fairly similar in description to the classes of Olsson. This resulted in the following classes (with the original Corine landcover classes, and the new classes:

Table 4-2: Corine landcover classes. From USGS (2008) and Hagemann *et al.* (1999).

Corine Landcover Class	Class	Subclass	Olsson land use description	New Landcover Number
1	Artificial surfaces	Continuous urban fabric	Urban	1
2	Artificial surfaces	Discontinuous urban fabric	Crops and town	2
10	Artificial surfaces	Green urban areas	Crops and town	2
11	Artificial surfaces	Sport and leisure facilities	Urban	1
12	Agricultural areas	Non-irrigated arable land	Corn and Beans Cropland	12
18	Agricultural areas	Pastures	Grass Crops	18
23	Forest and semi natural areas	Broad-leaved forest	Deciduous Broadleaf Forest	23
24	Forest and semi natural areas	Coniferous forest	Coniferous Forest	24
25	Forest and semi natural areas	Mixed Forest	Mixed Forest	25
27	Forest and semi natural areas	Moors and heathland	Marsh Wetland	27

These new soil classes had the following  $LAI_g$ ,  $LAI_d$ , and vegetation height values to be used as input for the calculation of  $LAI$ ,  $S$ ,  $C$ ,  $FTR$ ,  $K_{c\ full}$ , and  $k_c$  for the different land cover classes cover:

Table 4-3: Vegetation heights for different landcover classes. From Hagemann *et al.* (1999).

Landcover Class [-]	$LAI_g$ [-]	$LAI_d$ [-]	Vegetation Height [m]
1 (urban)	0	0	10
2 (crops and town)	4.5	1.1	1
12 (corn and beans)	4.5	1.1	1
18 (grass)	0	0	10
23 (deciduous broadleaf forest)	2.5	0.4	1
24 (coniferous forest)	2	0	0.7
25 (mixed forest)	5.2	0.1	10
27 (moors and heathland)	9.2	9	10

And the land use map:

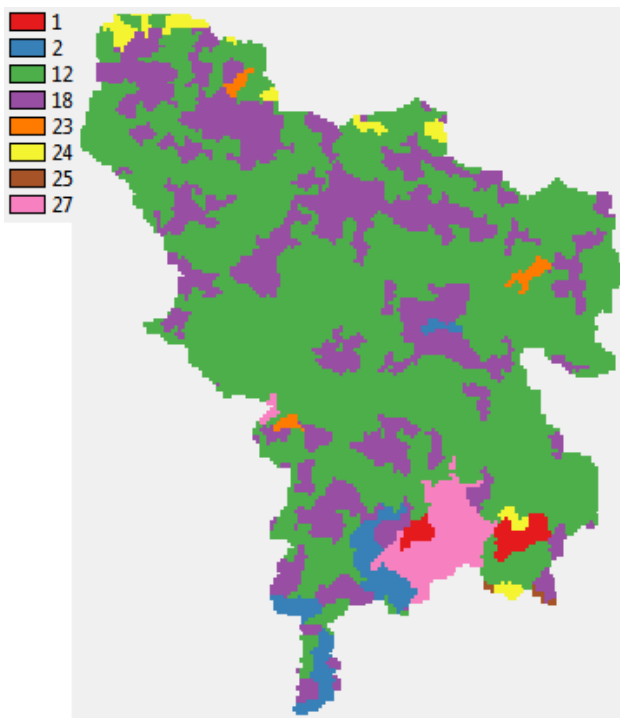


Figure 4-6: Landuse map

#### 4.7. Soil type data

In order to determine the soil types present in the area, the SGDBE Europe soil map on the database of the Utrecht University has been accessed via ArcGIS. The resolution of this soil map is 1 kilometer. The soil map has an attribute table in which every soil type (represented by pixels with the same color) in which attributes can be looked up using the number for the given soil type from the legend of the map. Through the accompanying manual on the database and using the obtained number textural information such as sand and clay fractions have been found for each soil class, as well as changes in these fractions throughout depth. The fractions are usually given in ranges (e.g. clay fraction 0.18-0.35) and therefore you have a minimum and a maximum fraction. The changes are important because STARWARS uses three layers, therefore it can occur that layers possess different soil properties. By making use of a pedo transfer script provided by van Beek (2014), soil parameters such as the slope of the soil water retention curve (moisture plotted against suction), soil bulk density, cohesion, saturated hydraulic conductivity, angle of internal friction (critical angle at which a given soil will start to slide), the residual moisture content and the saturated moisture content have been determined, using sand, clay, and silt fractions as input.

It was assumed that sand and clay fractions were uniformly distributed along the catchment area. Using a script called `gstat` provided by van Beek (2014) a `randomfield.map` was generated with values ranging from 0 to 1, where 1 denotes the highest fraction and 0 the lowest fraction. It was assumed that the radius in which there is spatial correlation between soil properties such as hydraulic conductivity is equal to the average accumulated amount of material flowing downstream [ $m^2$ ] towards the two outflow points of the catchment. In addition, a river mask with the river network was applied applying missing values to the river network, assuming that channels / rivers account for different depositions so areas in between rivers are segmented. The `accuflux` function is the following:

$$pcrcalc\ accuflux.map = accuflux(ldd, cellarea())$$

This creates a map `accuflux.map` which counts the number of upstream cells flowing towards the two outflow points

and multiplies it times the cell area (which is  $625 \text{ m}^2$  on a 25 m cell size). The accuflux map is shown in figure 4.7:

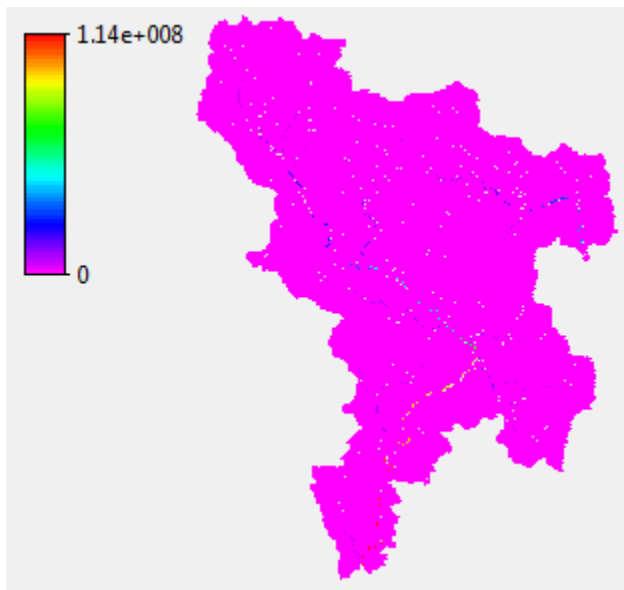


Figure 4-7: Accuflux map

The result is divided by the cell width of 25 m, to get the slope length in meters (376), implying the maximum length to which spatial correlation is assumed between cells. This is used as input for the script which makes use of a Gaussian Simulation algorithm described in Gomez-Hernández and Journel (1993). How the spatially distributed minimum and maximum soil fraction maps were created is described in appendix 3. The most important soil parameters derived from the pedo transfer script, hydraulic conductivity and porosity are shown for each layer (layer 1 most left, layer 3 most right):

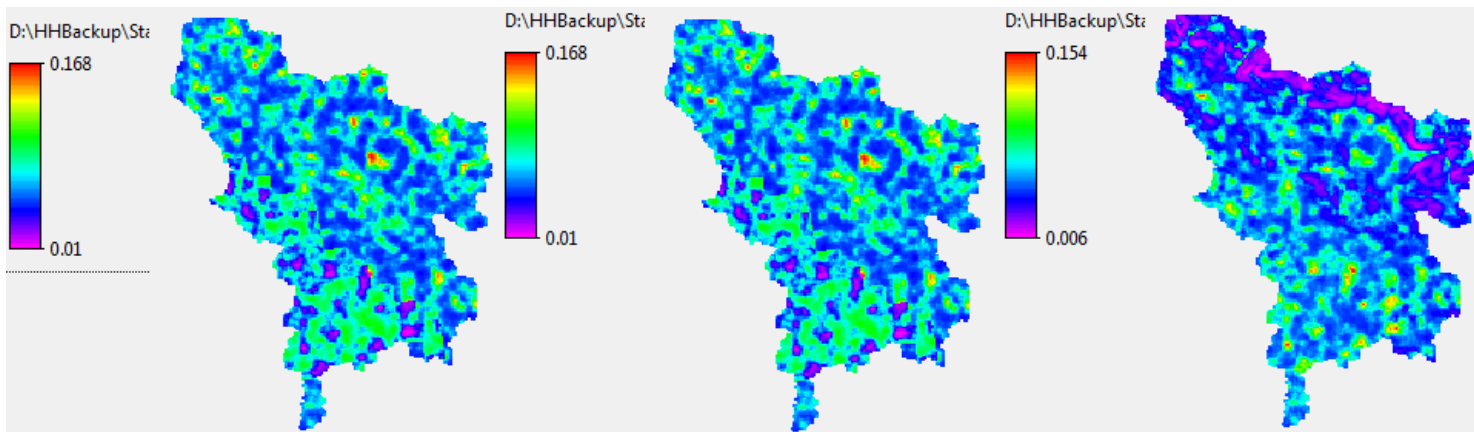


Figure 4-8: Ksat for each layer

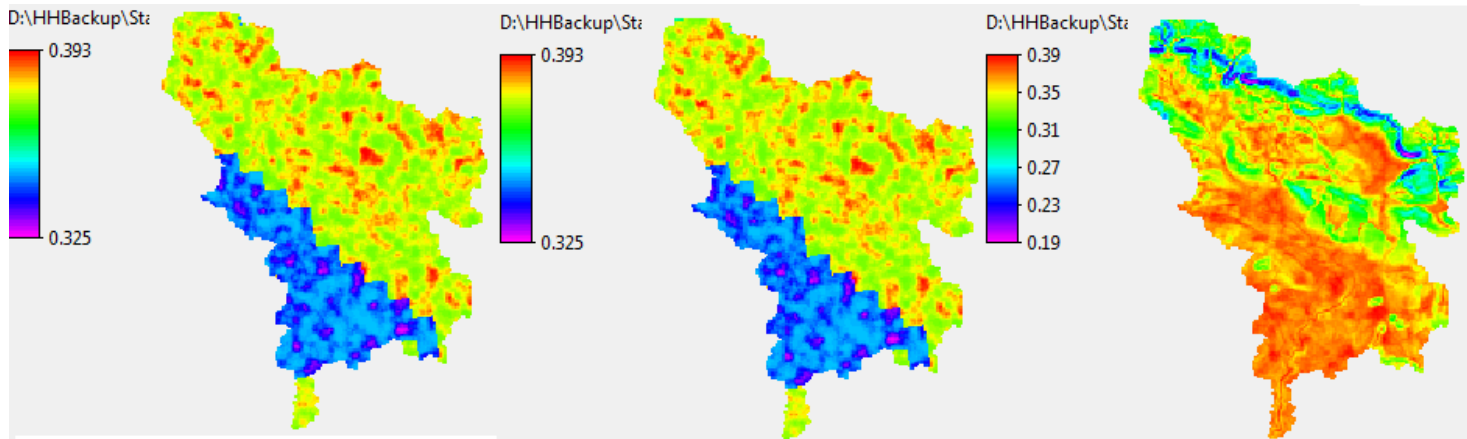


Figure 4-9: Porosity for each layer

Since Starwars uses three soil layers, the following setup has been used:

- It was assumed that the first layer has a depth of 0.25m based on the data from the SGDBE soil map;
- It was assumed that the second layer has a depth of 0.95m and therefore extends to -1.2m underground, based on the same soil map;
- In order to simulate the role of seepage from deeper groundwater and fluctuations in these levels, the depth of the third layer was assumed to be the height of the DEM minus layer 1 and 2. Hydraulic conductivity and porosity for this third layer were assumed to decrease with depth (as the soil is becoming more “stony” with depth and finally bedrock can be expected), according to the following equation (van Beek, personal communication 2014):

$$K(z) = K_{\text{sat}} * \exp(-\beta * z)$$

In which:

$K(z)$  = hydraulic conductivity as a function of depth [ $\text{m d}^{-1}$ ]

$K_{\text{sat}}$  = saturated hydraulic conductivity of layer 1 [ $\text{m d}^{-1}$ ]

$\beta = 0.01$  [ $\text{m}^{-1}$ ]

$z$  = depth of layer 3 [m]

The result is a giant water storage in which water levels can fluctuate due to seepage upwards and recharge to deeper groundwater tables downwards, as could happen in reality.

The depth of the shear plane was known to be around 1.5m at Hollin Hill (Dixon *et al.*, 2012). This depth was multiplied using the randomfield map to get an estimate of the depth of the shear plane around the whole catchment. Since there was almost no spatial variability with regards to the angle of internal friction and cohesion from the pedotransfer script, literature values from the Whitby Mudstone were taken as input with cohesion at 1.0 kPa and angle of internal friction at 23 degrees (Foster *et al.*, 2007).

## 5. Results

This chapter describes the outcomes of the various model runs. First the hydrological results (simulated versus observed daily discharge) of the model run with a cell size of 100 m will be discussed. Second, the hydrological results of the run with a cell size of 25 m will be discussed. The model is first run on 100 m because this allows for a fast run, and the best fit parameter combination of the 100 m run will be used to re-run the model on 25 m. The aforementioned runs are all done with historical (measured) meteorological data. The run with the zero to eight days ahead forecast on 25 m will be discussed after the runs with the historical data. Finally, a slope stability analysis comparing the results of the 25 m historic run and the 25 m forecast run will be made.

### 5.1. River flow results historical run on 100m

The results of the validation years 2009-2012 with regards to simulated daily discharge with initial parameters are shown in figure 5.1 where the daily observed discharge is also shown:

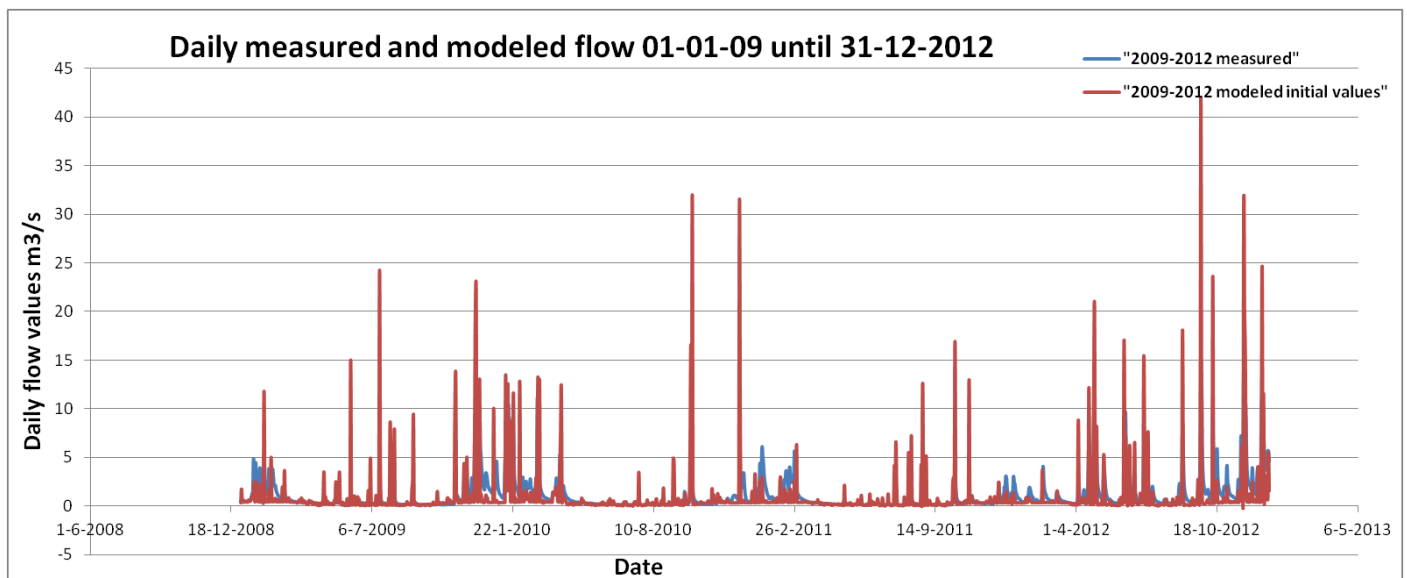


Figure 5-1: Modeled daily flow values with initial parameters from 01-01-09 until 31-12-2012 versus observed daily flow

The results of the run with the initial parameters show that further calibration of hydraulic conductivity and porosity is required: the NSE is -1.20, meaning the mean observed flow is a better predictor for daily flow values than the model. When studying this graph it becomes apparent that peak flows are severely overestimated using initial parameterization, suggesting that hydraulic conductivity and porosity values are too low. This also explains the poor NSE, since the equation assumes a quadratic relationship a mismatch between modeled and observed peak flow significantly outweighs a mismatch between modeled and observed base flow (order of magnitude in differences:  $2 \text{ m}^3 \text{ s}^{-1}$  and  $0.1 \text{ m}^3 \text{ s}^{-1}$ ).

An explanation for the mismatch is that when hydraulic conductivity values are low (as is the case with the initial run), this results in limited infiltration capacity of water into the soil during a rainstorm, in addition the soil drains at a slower rate after a rainstorm. This has two outcomes: soils show higher degrees of saturation between rainstorms and combined with less drainage capacity during rainstorms it takes less time for them to become fully saturated, hence a larger proportion of rainwater during a storm becomes quick flow. The latter process subsequently results in the high modeled peak discharges as seen in the graph. When porosity values are low, it means there is less storage capacity in the soil compared to higher porosity values. This means that soils become saturated quicker, resulting in more quick flow during storm events. There is also an interplay between the two parameters: when porosity values are high and hydraulic conductivity is low, it means that hydraulic conductivity is the more dominant parameter for quick flow creation, because during a high intensity rain storm the rainfall rate would exceed the



infiltration capacity of the soil even before the soil is saturated. Vice versa, with high a high hydraulic conductivity and low porosity, soil saturation is reached in less time even when rainfall rate does not exceed infiltration capacity of the soil, simply because the amount of storage in the soil is limiting. However, with a high hydraulic conductivity the soil drains faster, so initial degree of saturation before a rainstorm event may be lower in the case of a higher hydraulic conductivity.

Krzeminska (2012) performed a sensitivity analysis on STARWARS and her conclusions were the following, in line with van Beek (2002): *“The porosity (nmat, nfis) and saturated hydraulic conductivity (ksat,mat, ksat,fis) are the parameters with the largest influence on the hydrological model (modelled storage). This is not surprising, since those two parameters control the soil moisture percolation with depth, groundwater recharge and saturated lateral flow. Changing these two parameters by adding or subtracting 25% and 50% of absolute values for porosity and 50% and 100% of the absolute value for saturated hydraulic connectivity (for both matrix and fissures fractions at one time) results in maximal 10% (for n) or 12% (for ksat) variation in 110 Chapter 5 - modelled storage”*

Due to the relative insensitivity of the model to changes in ksat and porosity a multiplication factor of 2 will probably not suffice, because the initial run showed a significant lack of storage and therefore high peak flows. Calibration to obtain a reasonable fit took place with the following multiplication factors:

Ksat \* 10: Resulting in a Ksat between 0.005 (for urban area, standard value was 0) and 1.36 m d<sup>-1</sup> for layer 1 and 0.005 and 0.63 m d<sup>-1</sup> for layer 2

Ksat \* 15: Resulting in a Ksat between 0.005 and 2.04 m d<sup>-1</sup> for layer 1 and 0.005 and 0.95 m d<sup>-1</sup> for layer 2

Ksat \* 25: Resulting in a Ksat between 0.005 and 3.40 m d<sup>-1</sup> for layer 1 and 0.005 and 1.58 m d<sup>-1</sup> for layer 2

And porosity multiplied by 1.2 and 1.4. These factors are random, but by using these factors once can quickly assess whether the model performs better with a much higher hydraulic conductivity (Ksat \* 25) or whether a lower multiplication factor suffices (lower than Ksat \* 10). It should be noted that there were no soil samples available to determine a possible range of maximum and minimum values of said parameters, however the multiplication factors may be on the high end. The multiplication occurred for layer 1 and 2, since these extend to 1.2m depth and the hydraulic conductivity and porosity of layer 3 were calculated using an assumption of decrease with depth (see data collection). In order to obtain a land use dependent multiplication factor (vegetation has different root depth depending on which type of vegetation) the hydraulic conductivity was also multiplied by the root fraction of layer 1 and 2. The root fraction is the fraction of the total amount of roots one can expect to find in a given layer, ranging between 0-1 and summing up to 1 for layer 1, 2 and 3. This root fraction is land use dependent and is a function of maximum root depth. The following root depth per land use class was found in the Olsson table (USGS, 2008):

Table 5-1: Maximum root depth for each land use type. From USGS (2008).

Landcover Class [-]	Depth [m]
1 (urban)	0
2 (crops and town)	3.707
12 (corn and beans)	2.058
18 (grass)	1.104
23 (deciduous broadleaf forest)	2.147
24 (coniferous forest)	0.988
25 (mixed forest)	2.544
27 (moors and heathland)	7.152

In order to obtain the rootfraction, a rootbeta.map had to be created. Rootbeta is the parameter which the determines the minimum fraction of roots under the maximum root depth. This means that you assume that under the maximum root depth there is still a very small fraction of roots (van Beek, personal communication), for instance because the maximum root depth covers 99.99 % of all roots . This parameter is calculated by means of the following equation (van Beek, personal communication 2014):

$$pcrcalc\ rootbeta.map = \frac{\ln(1E - 4)}{z}$$

In which:

z = maximum root depth [m], if maximum root depth < soil depth, then use soil depth instead

The root fractions are subsequently calculated by using the following equation:

$$pcrcalc\ rootfrac1.map = 1 - \exp(\text{rootbeta.map} * \text{layer1.map})$$

$$pcrcalc\ rootfrac2.map = 1 - \exp(\text{rootbeta.map} * \text{layer2.map})$$

$$pcrcalc\ rootfrac3.map = 1 - \exp(\text{rootbeta.map} * \text{layer3.map})$$

$$pcrcalc\ rootfrac3.map = \text{rootfrac3.map} - \text{rootfrac2.map}$$

$$pcrcalc\ rootfrac2.map = \text{rootfrac2.map} - \text{rootfrac1.map}$$

Logically, rootfrac3.map and rootfrac2.map have to be calculated twice, because the root fraction of layer 3 is dependent on the root fraction in the overlying layer 2, which is dependent on the thickness on layer 2, and likewise for layer 2 with layer 1. This is not yet included in the initial calculation of rootfrac2 and rootfrac3.map.

In addition, the model was run with an initial value for bedrock hydraulic conductivity of  $0.75 * 10^{-3} \text{ m d}^{-1}$ . This is based on the following assumption: the value of bedrock hydraulic conductivities for sandstone and limestone present in the area (see section: area of study) show hydraulic conductivity values between  $0.86 * 10^{-5} (10^{-10} \text{ cm s}^{-1})$  and  $0.86 * 10^{-1} \text{ m d}^{-1} (10^{-6} \text{ cm s}^{-1})$ , although they are based on a Canadian study as shown in figure 5.2 on the next page (Riddell, 2007), no values for UK were found. The value of  $0.75 * 10^{-3}$  is somewhere in the middle. This value can also be calibrated later on. It does have some importance to the value of the modeled discharges, since the base flow (slow recession) is dependent on this conductivity value. But for comparison, all the runs with adjusted hydraulic conductivities and porosities were run with this value.

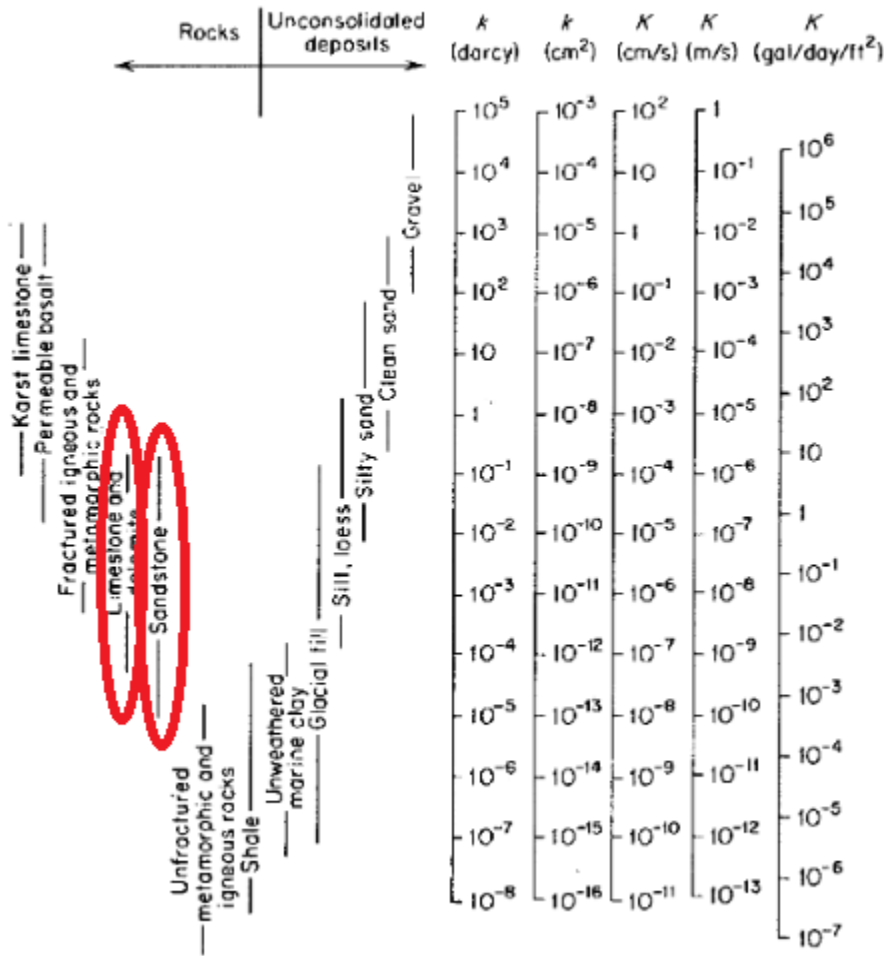


Figure 5-2: Bedrock hydraulic conductivity values. From Ridell (2007).

The results of the re-run are shown in table 5.2 where the NSE is shown for the different combination of porosity and hydraulic conductivity multiplication, using the afore mentioned bedrock hydraulic conductivity of  $0.75 \cdot 10^{-3} \text{ m d}^{-1}$ :

Table 5-2: NSE for different Ksat and porosity values

Porosity multiplication	Hydraulic conductivity multiplication		
	NSE 10	NSE 15	NSE 25
NSE 1.2	0.21	0.27	0.31
NSE 1.4	0.34	0.35	0.40

These results show that at a low porosity multiplication factor (20% increase in porosity) the model performance is quite sensitive to an increase in hydraulic conductivity. The higher hydraulic conductivity results in a soil that drains at a faster rate, resulting in decreasing peak flows. It is interesting to see that these values converge more or less when porosity is increased with 40%. The hydraulic conductivity multiplication factor of 25 still gives the best NSE, however values are closer to each other for the different multiplication factors of hydraulic conductivity. This can probably be explained by the fact that with increased soil storage, hydraulic conductivity is of less influence on the time it takes to reach full saturation. One can imagine that with increased soil storage, low intensity rainstorms can be stored independent of hydraulic conductivity values. In addition, one can imagine that during high intensity rainstorms rainfall rate exceeds infiltration capacity anyway and as a result soils become saturated during these events. A higher hydraulic conductivity only has a delaying factor in this process. The values for hydraulic conductivity for multiplication factor 25 are shown in figure 5.3 for layer 1 on the left and layer 2 on the right:

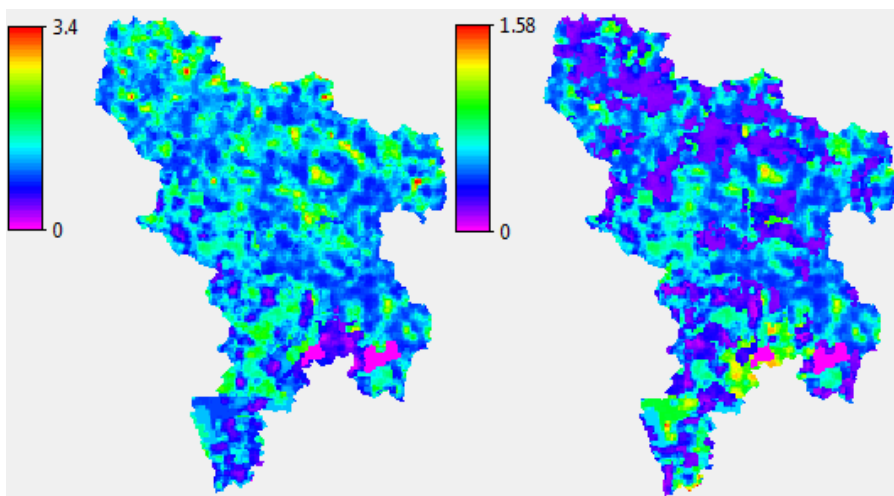


Figure 5-3: Values for  $ksat * 25$  for layer 1 and 2

And porosity in figure 5.4, again for layer 1 on the left side and layer 2 on the right side:

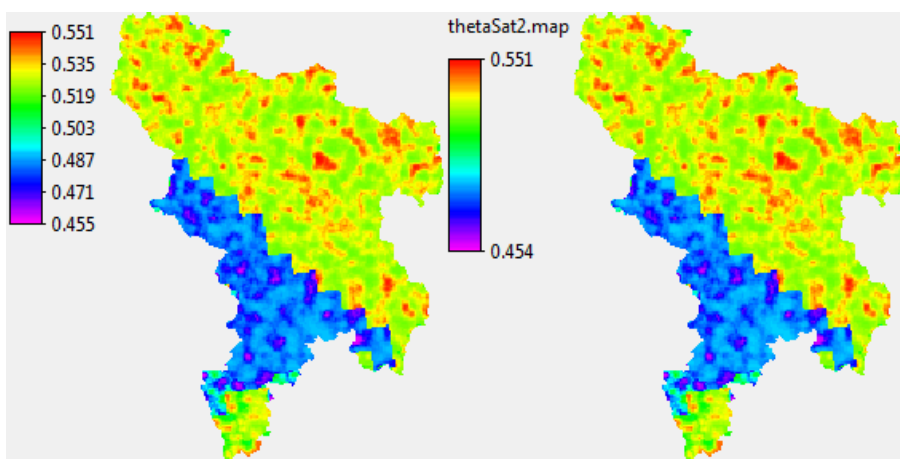


Figure 5-4: Porosity for layer 1 (left) and 2 (right)

The results of the run with an increase in porosity by 40% and the hydraulic conductivity times 25 are shown in figure 5.5:

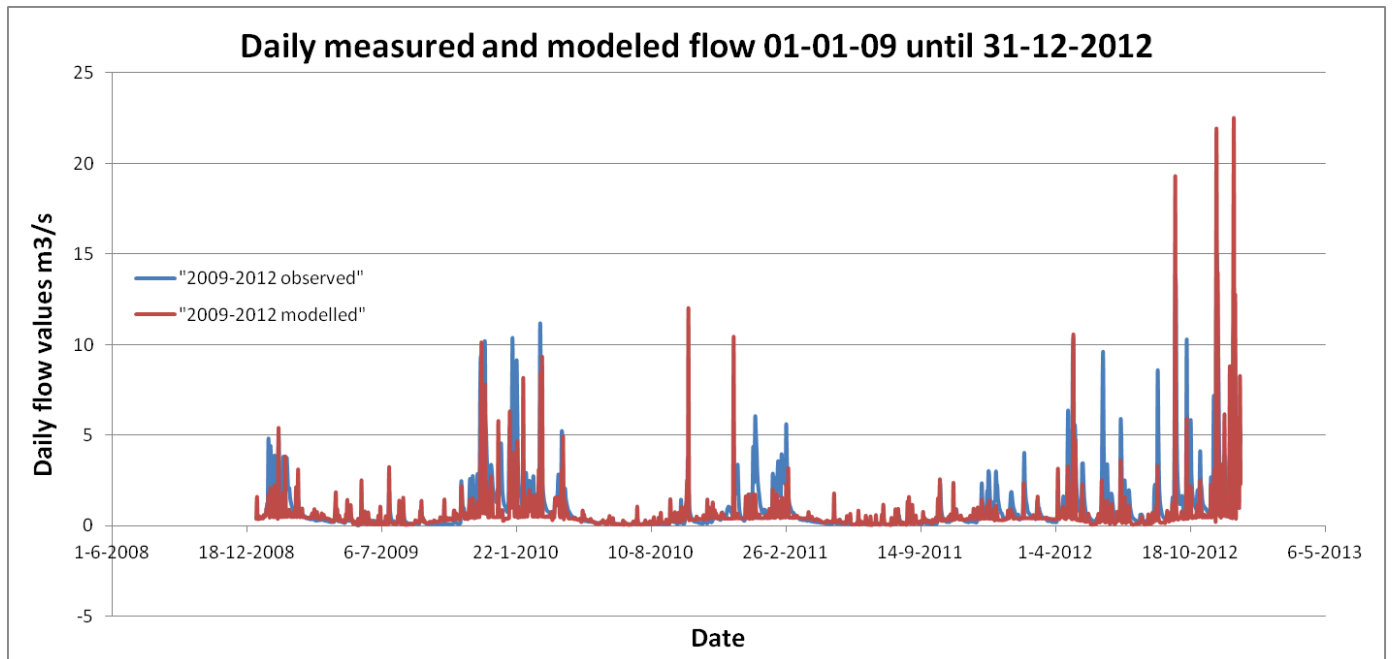


Figure 5-5: Modeled values on  $K_{sat} * 25$  and porosity  $* 1.4$  versus observed daily flow values from 01-01-09 until 31-12-2012

Most peak flows that occur in the initial run are damped, and the maximum peak flow value is halved (from around  $45 \text{ m}^3 \text{ s}^{-1}$  around 18-10-2012 to  $23 \text{ m}^3 \text{ s}^{-1}$ ). In addition, the timing is reasonably accurate. The validity of this parameter combination can be questioned however, the  $K_{sat}$  seems to have been overestimated for a clay soil type. An important remark about the graph is that although the timing is reasonable (usually the model predicts peak flows when peak flows occur in reality, albeit overestimated) around 6-7-2009, 10-8-2010 and after 18-10-2012 peaks occur in the modeled flow which are not observed in reality, also peaks are both over and underestimated. This may have to do with the quality of the forcing of the model, e.g. the meteorological input. When examining the rainfall data from the different stations, a high variability in rainfall values was found. This is shown in figure 5.6, showing the annual rainfall in mm for 2009-2010 and figure 5.7 for 2011-2012:

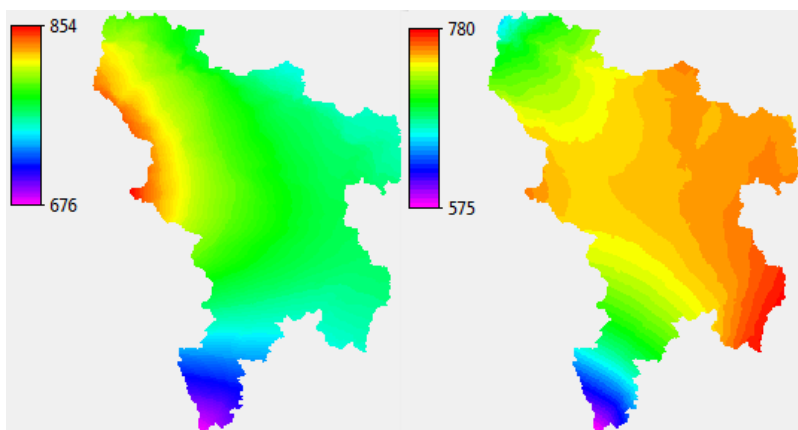


Figure 5-6: Annual precipitation in mm for 2009 (left) and 2010 (right)

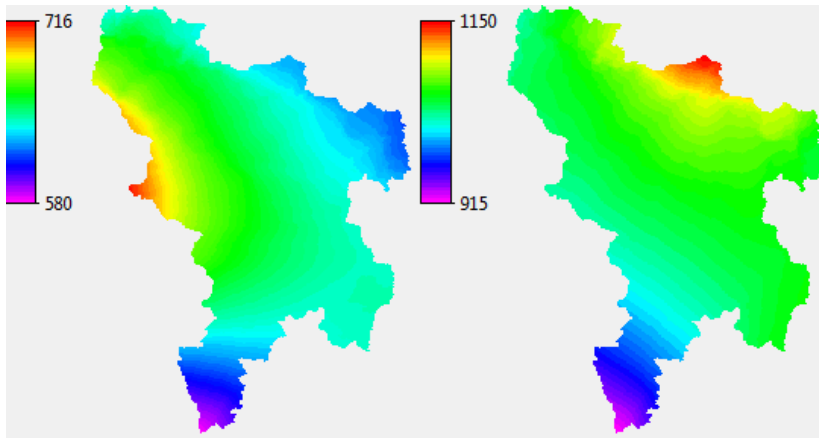


Figure 5-7: Annual precipitation in mm for 2011 (left) and 2012 (right)

It is interesting to see that there is a) a large difference in annual rainfall, as can be expected between drier and wetter years and b) that places of high total rainfall differentiate throughout the years. Each year the place with highest rainfall differs as can be seen by the distribution of yellow and red colors in figures 4.6 and 4.7. In general, rainfall patterns more or less coincide with the DEM, because air forced to ascend by the presence of hills may cool down to dew point and subsequently produce clouds and rainfall (Met Office, 2015) because cooler air can hold less vapor (Eden, 2015). This effect is called rain shadow, and it causes the windward side of the hill (where the air containing moisture is blown towards), to show higher precipitation levels than the leeward side (Eden, 2015). Figure 5.8 shows the DEM:

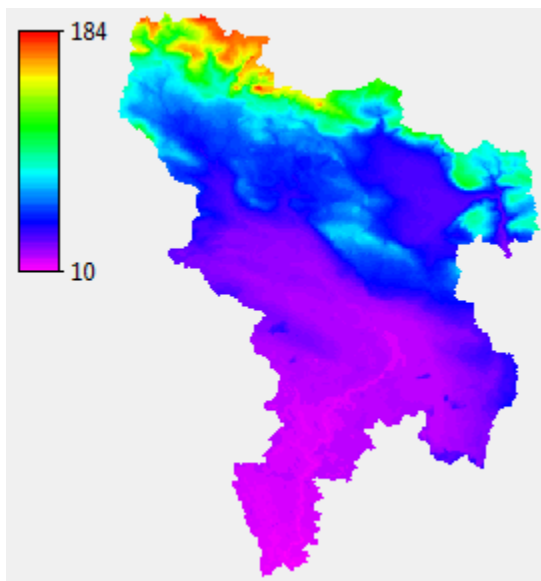


Figure 5-8: DEM of the catchment

There appears to be no direct correlation with height and maximum rainfall patterns but this has to be appreciated in the sense that a) the height differences are not dramatic and b) regionally speaking the Pennine mountain range with elevations up to 893m causes a more pronounced rain shadow effect, with an average annual precipitation of 1,500 mm, this number decreasing eastwards towards the coast having less than 600 mm annual rainfall on average and increasing to 1000mm around the higher North York Moors north of the catchment (Met Office, 2015). The Yorkshire Vale is also a dry area (Met Office, 2015) which is reflected in the low annual rainfall in the southern part of the catchment, where the Yorkshire Vale is situated. Over the entire UK, there is an even higher variability in rainfall patterns, with some parts of the Scottish Highlands showing annual precipitation totals of 4,000 mm and some parts of Eastern England only 500 mm (Met Office, 2015).

To illustrate how variable rainfall in the vicinity of the catchment can be the statistics of two stations are used: Cawood which is 10 km south of the southern tip of the catchment and at 8 m above mean sea level (amsl) and Bradford which is 20 km southeast of the southern tip of the catchment and 134 m amsl. Cawood has an average annual rainfall of 601 mm and Bradford 878 mm (Met Office, 2015). Since the rain stations used are at the same distance or even slightly further away from the catchment, this high variability in rainfall (as shown from the two stations and the regional statistics from the Pennines, North York Moors and the coast) could have led to an overestimation or underestimation of rainfall up to a certain degree as input to the model. Another problem with using interpolated data is that if one station records a rain shower and the others do not, there will always be some rain over the entire catchment because each station contributes a certain weight. In reality, this would mean merely a locally restricted rain shower. To illustrate this problem, the statistics with regards to rain and non rain days for the years 2009, 2010, 2011, and 2012 are shown in Table 5.3:

**Table 5-3: number of rain and non rain days for the years 2009, 2010, 2011, and 2012**

Year	#of rain days	#of non rain days
2009	269	96
2010	282	83
2011	251	114
2012	278	88

According to the statistics, the average number of rain days per year in Yorkshire amounts to 108 and the average amount of non rain days to 257 (Meteorologisk Institutt, 2014), quite the opposite of the statistics from the years 2009, 2010, 2011, and 2012, illustrating the problem about the interpolated values.

Unfortunately, these types of model mismatches cannot be removed in the calibration process because they are independent of parameterization. The mismatches may account for the NSE of 0.4, which is still not close to a preferable level of close to 1. The total number of rain days could have been decreased by applying a threshold precipitation value for the interpolation, effectively cancelling out one local rain shower. However, when reviewing the data it was also found that sometimes large local rain shower occurred at one station. A possible solution to that problem would have been to alter the interpolation in such a way that if only one station records some precipitation, the interpolated values are only valid within a given radius of the station and outside (i.e. the larger part of the catchment) they are put to 0. This would not actually decrease the number of rain days *somewhere* in the catchment, but would simulate a local rain shower (which of course gives a different hydrological effect to the model than a rain shower covering the entire catchment).

Another factor of uncertainty is the actual duration of the rainfall. The model assumes that rainfall is distributed over a 24 hour block with an equal intensity, so if the total recorded rainfall amounts to  $1 \text{ mm d}^{-1}$ , this means that the hourly intensity will amount to  $1/24 \text{ mm h}^{-1}$ . Since no information is known about the actual intensity of the rainfall because the rain stations only show total daily rainfall, intensities could be underestimated (they would always be underestimated if wrong, since the model already assumes the lowest intensity). However, for the region under study this underestimation may be less dramatic than when studying a tropical region, with very high intensity, short duration rainfalls. Figure 5.9 (Cedar Lake Ventures Inc., 2015) shows the annual type of rain for the station Linton On Ouse, situated in the catchment and some 15 kilometers north, slightly westwards of York:

### Types of Precipitation Throughout the Year

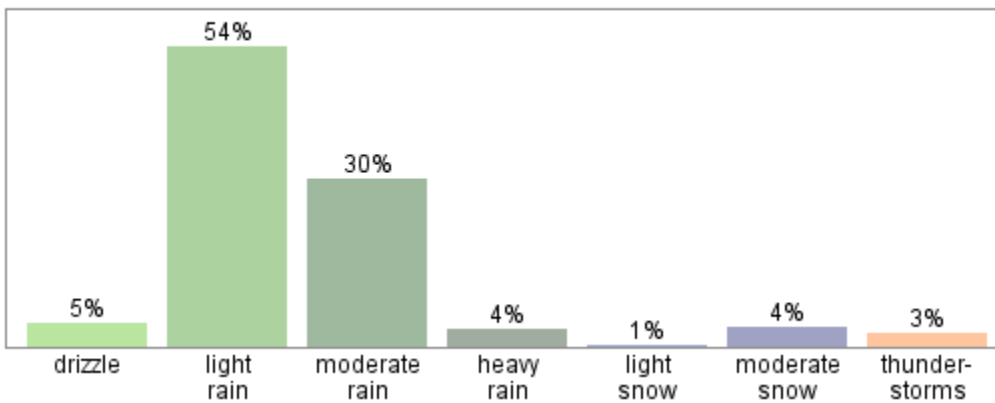


Figure 5-9: Relative annual frequency of types of precipitation. From Cedar Lake Ventures Inc. (2015).

From figure 5.9 it can be seen that the most prevailing types of rain are drizzles (5%), light rain (54%), and moderate rain (30%). These types of rainfall have a far lower intensity than heavy rain (4%) and thunderstorms (3%). Would the latter be the dominating types of rainfall, river flows may have been underestimated, because more quick flow would occur.

The annual PET in mm for the years 2009-2010 is shown in figure 5.10 and for 2011-2012 in figure 5.11:

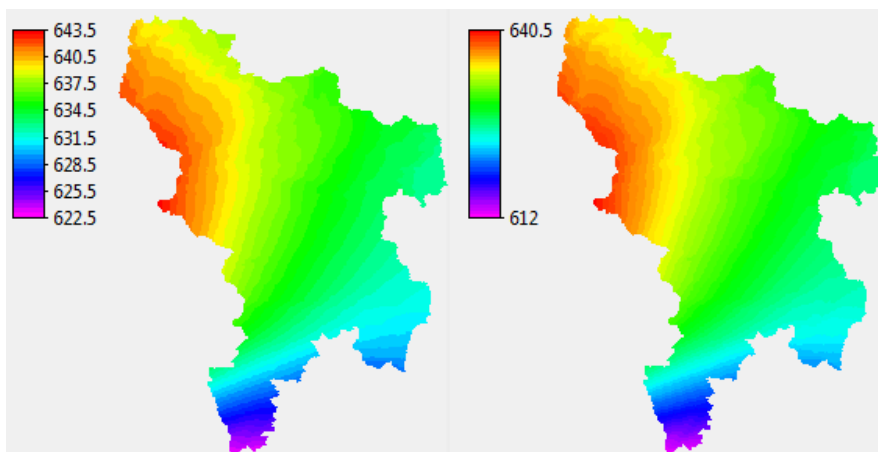


Figure 5-10: Annual PET in mm for 2009 (left) and 2010 (right)

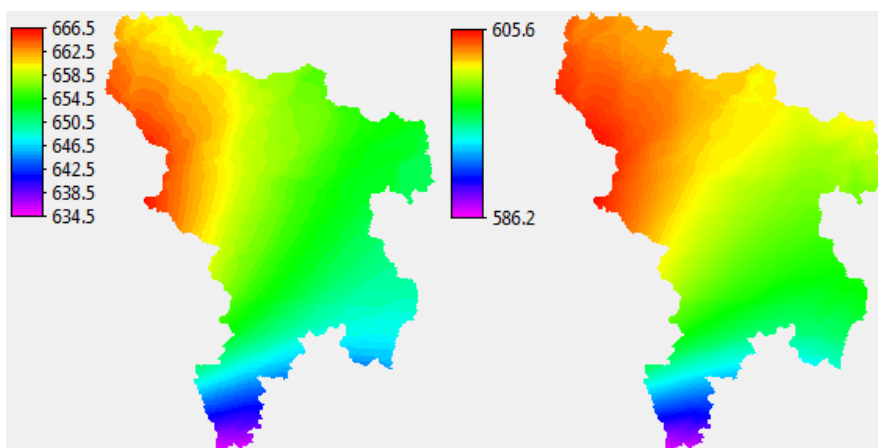


Figure 5-11: Annual PET in mm for 2011 (left) and 2012 (right)



These figures appear to be slightly higher (10%) around 10% than the annual average of 1961-1990 shown in figure 5.12 in which the catchment is encircled in black (Environment Agency, 2001) but appear to be more in line with stats of the NERC showing that in a more recent years (1995, 2003 and 2004), PET was considerably higher, up to 15% in some areas of the UK and that PET has risen in general (NERC, 2004):

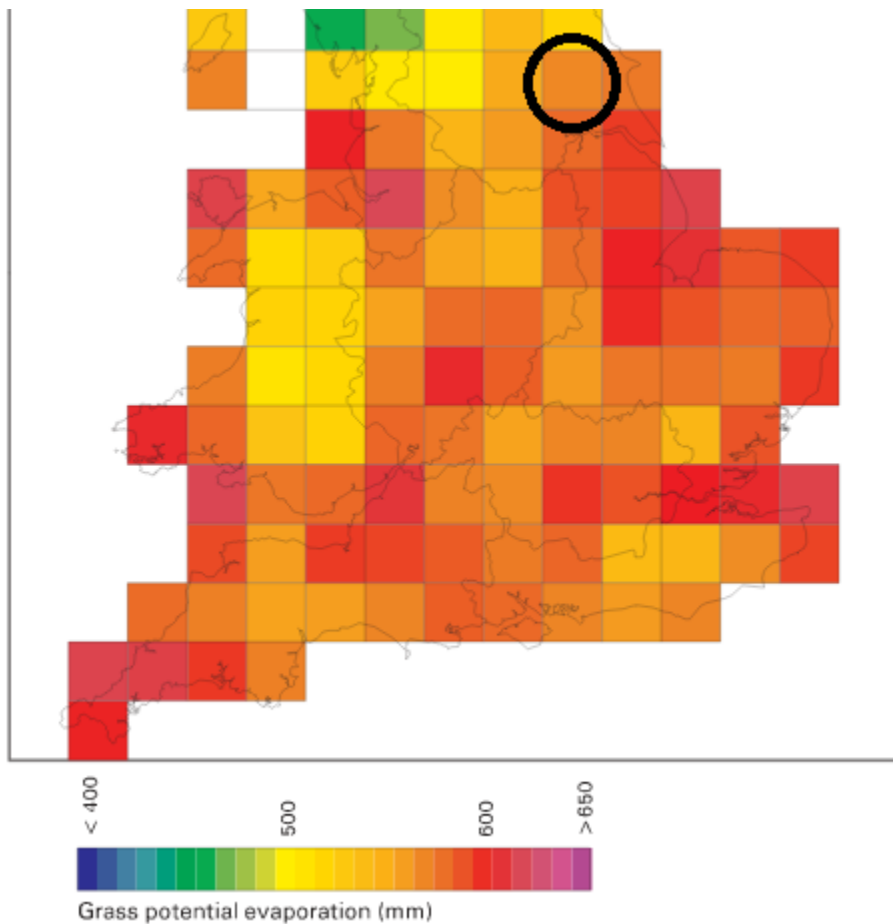


Figure 5-12: Average PET in mm yr-1 for the Entire UK between 1961-1990. After NERC (2004).

While PET seems reasonably correct, the crop factors may hold uncertainties as well. PET as simulated and shown in figures is based on an open grass evapotranspiration (Environment Agency, 1991) and crop factors correct for land use other than open grass, such as forest, agricultural, urban et cetera. The Corine land cover map used for this study has a cell size of 250 \* 250 m and may therefore not be accurate enough for this study, especially since crop factors as calculated for the different land uses significantly deviate (they can be as low as 0.6 and as high as 1.2). Also the transition of land classes from the Corine dataset to the Olsson classes as described in the section land use data of the chapter data collection and processing was based on a best match between the description of the two different classes merging was based on a judgment as to which classes matched the best, but holds some uncertainty because there was significant difference in LAI between classes, altering crop factors. The transition from Corine classes to Olsson classes was needed because for Olsson classes there was data on LAI and such, whereas for Corine land cover classes data was not available. But being that the Corine land cover map was the only map available for this study, it had to be used to determine regional land cover. All in all, the effects of spatially improperly distributed crop factors or under / overestimation of crop factors over the entire area can lead to an under/ overestimation of evapotranspiration at confined places or over the entire catchment, altering simulated river flow.

Another explanation of the hydrological mismatch can be that the way in which the parameters are spatially distributed in the model does not accurately capture the real soil and hydrological conditions (which is determined

by the random field map shown in figure 5.13 which determines whether parameters get the minimum or maximum value or a value in between, see section data collection and processing). Parameterization of large catchments can prove to be difficult (Guimarães *et al.*, 2003) because there is no extended knowledge on how parameters are distributed throughout the catchment and the variability between the parameters.

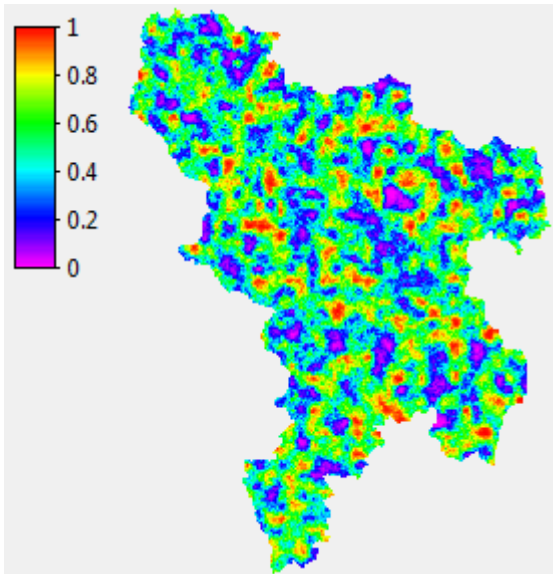


Figure 5-13: Random field map showing parameter distribution

The random field map determines whether parameters like Ksat get a maximum or a minimum value. This is determined by taking the minimum assumed value of Ksat and add the difference between the maximum and minimum Ksat to this and multiply this difference with the randomfield value:

$$\text{Ksat (per cell)} = \text{Ksat min} + (\text{Ksat max} - \text{Ksat min}) * \text{randomfield}$$

This equation means that where randomfield is 0 (meaning no spatial correlation to any neighboring cells), Ksat gets the minimum value and where randomfield is 1 (maximum spatial correlation) Ksat gets the maximum value. Any value randomfield value between 0 and 1 means Ksat will be between the minimum and maximum assumed value. It has to be noted that by increasing or decreasing hydraulic conductivity, one can expect a more or less linear increase or decrease of peak flows because the underlying assumption of spatial distribution of minimum and maximum values as shown in figure 7 does not change.

Soil depth is another factor of uncertainty in the model. Soil depth was based on the information from the SGDBE database. Soil depth directly influences the available subsurface storage, affecting residence times and influencing the water balance (Nicótina *et al.*, 2011). An example of a different approach is an approach used by Penížek and Boruvka (2006). They used three primary terrain attributes – elevation, slope and aspect- to predict soil depth. They employed ordinary kriging, co-kriging, regression-kriging and linear regression and compared the results with validation data obtained from the field. They obtained the best results with co-kriging with respect to the mean depth, variance and range. Interestingly, use of auxiliary terrain data did improve the model but only very little because of the low correlation to the primary terrain data. Another approach is one used by Lanni *et al.* (2012). They measured soil depth at a total of 410 points in their multiple study catchments in Italy to come up with an empirical relation of  $L(x, y) = 1.01 - 0.85 \tan \beta(x, y)$  in which L is the soil depth [m] and  $\beta$  the slope angle [°], valid when  $\beta < 45$  degrees and elevation  $< 2000$  m and  $\beta < 40$  degrees when elevation  $> 2000$  m. They found no significant relationship between soil depth and variables such as curvature and specific catchment area. When testing their findings with the NSE, they obtained a score of 0.4 with a mean error of 7%. Such an approach is of course highly localized, however it

is an attainable approach since it only requires a DEM and soil depth measurements as input. In addition, Tesfa *et al.* (2009) stated that other statistical or physically based approaches only achieve partial success at best.

Finally, the quality of the DEM is also of importance. The original resolution of the DEM was  $100 \text{ m}^2$  ( $10 * 10 \text{ m}$ ). This was interpolated to  $100 * 100 \text{ m}$  and  $25 * 25 \text{ m}$ . This interpolation of course has an effect on the steepness of the slopes, which decreases with increasing resolution. This effect translates to the distribution of soil moisture, the pattern of the LDD and slope stability calculation. This is shown in figure 5.14 where the slope angles on 10m, 25m and 100m cell size are shown (note that these need to be reversed with  $\tan^{-1}$ , to get the slope angle in degrees and that the catchment differs slightly for the 100 m cell size, due to scale of cells):

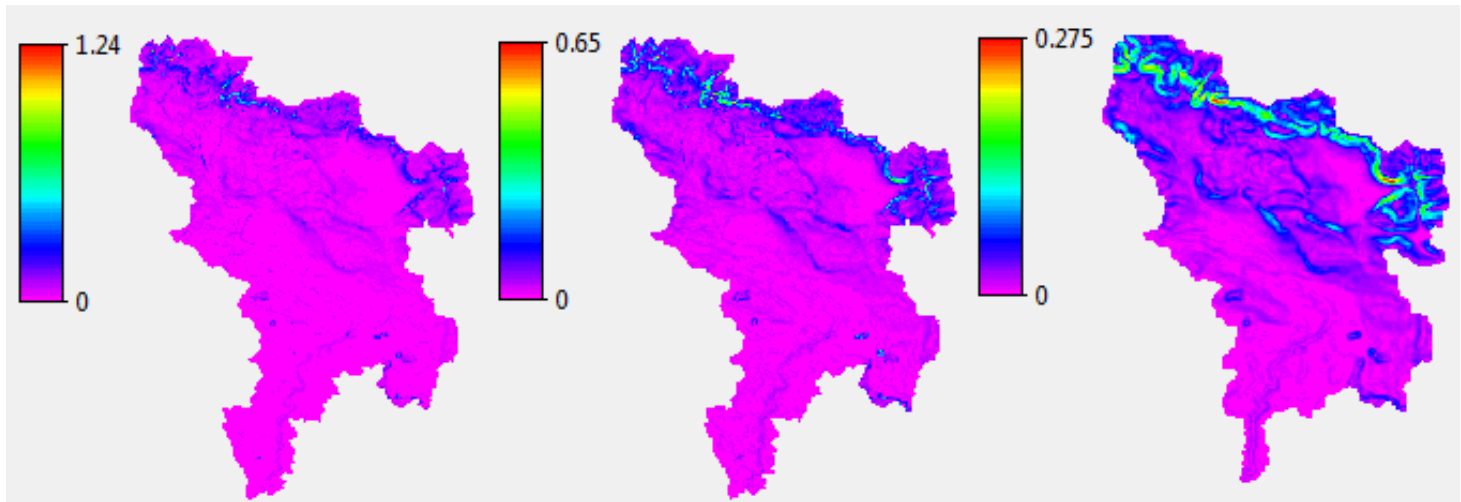


Figure 5-14: Slope angles for 10m, 25m, and 100m cell size

Guimarães *et al.* (2003) address this as well. They state that their best-fit model did not correspond to typical soil property values. They suspected that internal friction angles may be higher in low-confining stress environments related to superficial soils, and cohesion on the field-scale is low due to translating the properties obtained from small samples to an entire landslide (the scale problem which was mentioned before). They concluded that using a high resolution DEM (they used a 2 by 2 m DEM), and therefore accurate reflection of steep converging terrain and other high resolution topographic data is a) more achievable than using accurate soil parameters, both in value and spatial distribution and b) yields better results, if the landslides are topographically driven to a certain extent. The obvious practical problem arising with this conclusion is the following: whereas they simulated a small catchment of  $5.4 \text{ km}^2$ , this catchment is roughly 25 times that size, meaning a lot larger maps and more computation time. Using maps on a  $10 * 10 \text{ m}$  cell size already resulted in practical problems such as storage and computation time. A  $100 * 100 \text{ m}$  run for the entire calibration period lasted around 1.5 hours. If it would be run on  $10 * 10 \text{ m}$  resolution it would last  $10 * 10 * 1.5 = 150$  hours or  $\sim 6$  days per run. If the entire model would be run on  $2 * 2 \text{ m}$  it would cost  $150 * 5 * 5 = 3750$  or  $\sim 156$  days. It does need to be taken into account that these figures are based on the run time of a 4 year old laptop. Especially in the future, with more expensive office based powerful computers, these computational times could be significantly reduced. Other authors also mentioned the scale problem. In their paper about landslide reduction in the Caribbean, Anderson *et al.* (2011) stated that the scale of mapping of landslides is not on par with the scale of the human and physical triggering mechanisms of said landslides. They further state that this point is widely ignored because mapping hazards based on widely and readily available GIS-data is easier, however GIS-data does not usually match the scale of the mechanics of slope failure. These mechanics involve very site-specific surfaces and subsoil convergence of water due to concave curvature of hill slopes, which can be dominant landslide triggers but are found on finer resolutions than one GIS-based map cell. They conclude that this causes the low effectiveness of hazard maps. This (apparently still relevant) problem was already stated by Opadeyi

*et al.* (2003) in their assessment about hazard in the Carribean. Obviously, these problems are not constricted to the Carribean. It does need to be taken into account that one usually lacks information about spatial distribution of parameters (Pechlivanidis *et al.*, 2011 ; Guimarães *et al.*, 2003). Because of this, one can say that the increase in resolution is not backed by proper knowledge about parameterization. Mulungu & Munishi (2007) concluded that correct forcing is of primary importance compared to parameterization, which has a second-order modifying effect. This seems to make intuitive sense because if rainfall or PET is not accurately distributed over the catchment the pattern of river flow with regards to peak and low flows cannot be correctly simulated which can also be seen in figure 5.5, with simulated peak flows that do not occur in reality. However, changing the parameters does have a profound effect on the values of peak flows as can be seen when comparing figure 5.1 to 5.5.

The order of importance with regards to hydrology is hard to quantify exactly but as demonstrated forcing has the most profound effect. The DEM seems to be second in order: derivative products from the DEM are used in the model such as the LDD, the catchment delineation (which is based on the accumulated material upstream to each cell determined by the LDD). The original catchment delineation on 100 m compared to the one on 10 and 25 m can also be seen in figure 5.14, which is a worse representation than the catchment delineation from the CEH seen in figure 4.3. In addition, the DEM can also be used to simulate soil depth as done by Penížek and Boruvka (2006) based on the primary terrain attributes elevation, slope, and aspect derived from the DEM. Since both soil depth and parameter combination chosen have an effect on the amount of subsoil storage and water levels, they could both be placed on the lowest end of importance. Finally, the DEM is an important (if not the main) product for slope instability assessment since its derivative is the slope angle, a parameter for the safety factor.

Taking these uncertainties into account, the best combination being Ksat times 25 times the root fractions and the porosity times 1.4 was used for the run on 25 m. The run of 25 m was subsequently used for slope stability analysis (the model was run with a higher multiplication factor above 25 to be sure this was the optimal multiplication factor and the NSE decreased with higher multiplication factors). In addition, the value of  $0.75 \cdot 10^{-3} \text{ m d}^{-1}$  turned out to be the most optimal value for bedrock hydraulic conductivity, lowering or raising this value did not lead to improvements in the NSE score.

## 5.2. River flow results historical run on 25m

Using the afore mentioned “optimal” parameter combination of Ksat \* 25 and porosity \* 1.4 the model was re-run on 25m. This lead to a significantly decreased NSE of 0.21. The results are shown in figure 5.15:

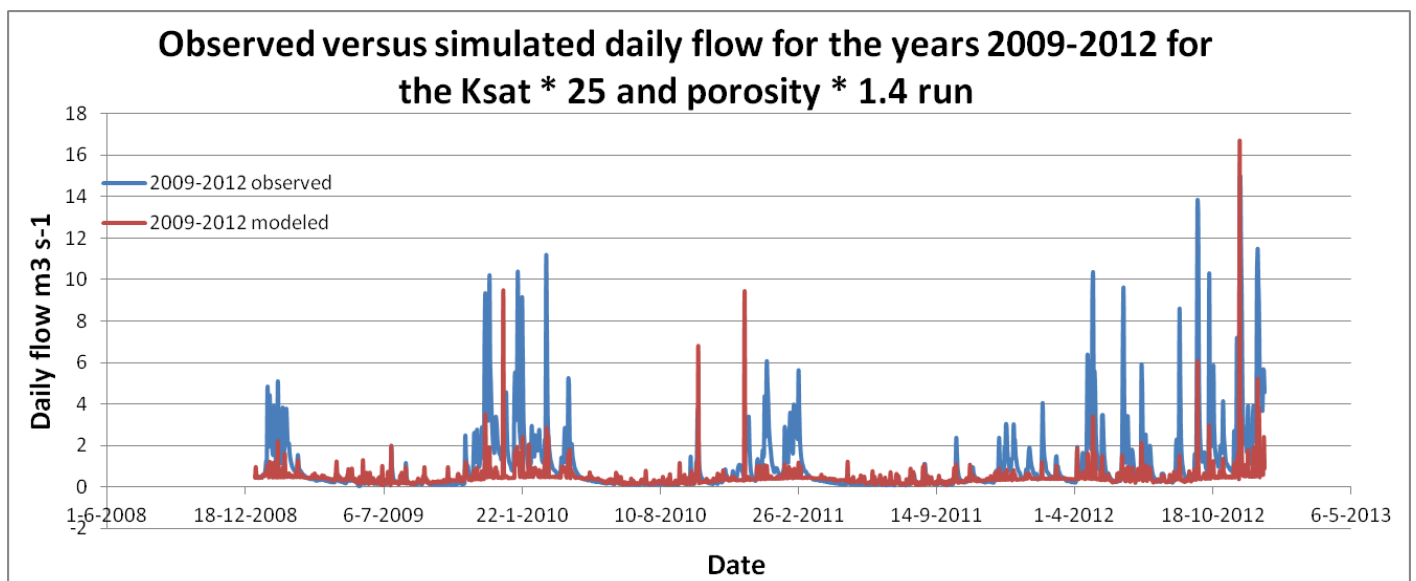


Figure 5-15: Results of the run on 25m for the Ksat \*25 and porosity \* 1.4 run for the years 2009-2012

The peaks are far more damped than on 100m and there is also a water budget shortage. The total observed daily flow over the modeling period (summation of daily flow) amounts to  $1524 \text{ m}^3 \text{ s}^{-1}$ . The simulated volume of water amounts to  $754 \text{ m}^3 \text{ s}^{-1}$ . The pattern of the background flow coincides quite well. Peaks are quite underestimated though, they are either completely missed (around 26-2-2011) or much underestimated (start of 2009, around 22-1-2010, 1-4-2012 and 18-10-2012). Therefore, the model was re-run with different hydraulic conductivity multiplications (Initial, Porosity \* 1.4, Ksat \* 5 and porosity \* 1.4, Ksat \* 10 and porosity \* 1.4, Ksat \* 15 and porosity \* 1.4, Ksat \* 20 and porosity \* 1.4, and Ksat \* 25 and porosity \* 1.4). The results of these run with respect to NSE score and water budget (summated daily flow in  $\text{m}^3 \text{ s}^{-1}$  from 2009-2012) are shown in Table 5.4:

Table 5-4: Results of runs with different hydraulic conductivities for NSE and total volume for 2009-2012

Run	NSE [-]	Flow [ $\text{m}^3 \text{ s}^{-1}$ ]
Initial	-1.13	1342
Por * 1.4	-0.38	1071
Ksat*5 Por * 1.4	0.05	920
Ksat*10 Por * 1.4	0.14	872
Ksat*15 Por * 1.4	0.21	803
Ksat*20 Por * 1.4	0.15	787
Ksat* 25 Por * 1.4	0.21	754

In order to look at the differences between the driest run (Ksat \* 25) and the wettest run, the initial run, the following water balance can be applied:

$$\Delta S = \text{Precipitation}(P) - (\text{Evapotranspiration}(ET) + \text{River flow}(Q))$$

In which all are taken in mm. This means that the incoming precipitation flux is divided in two components: ET and Q, Q being divided into background flow and quick flow. If both the outgoing fluxes exceed or are lower than the incoming flux P, the difference must be explained by the change in storage delta S (water released from storage or stored). Table 5.5 show the different components for the driest and the wettest run for the year 2012:

Table 5-5: Different components of the water balance for the year 2012 for the driest and wettest run

Run	P [mm]	ET [mm]	V[mm]	$\Delta S$ [mm]	Difference [mm]
Initial	1036	442	369	130	95
Ksat*25	1036	451	179	119	287

Oddly enough, the dry run shows and even higher evaporation total (451 mm). Furthermore, it is notable that the different components do not add up to a closing water balance at all. This may have to do with the fact that when delta storage is reported to a map pc raster, rounding errors start to occur. A lot of different fluxes are associated to the calculation of the change in storage, and rounding errors can accumulate (van Beek, personal communication 2015). The closing difference is also reported to a map stack after all the components have been added up and then the closing error is only 0.7% for the Initial run, and 1.19% for the Ksat \* 25 run. These errors may have to do with how the model loses its water to the 4<sup>th</sup> layer along the path of the drainage network, as only at these points water leaves the system.

Although the initial parameterization is closest to the observed flow in a quantitative manner, its NSE score is lowest. Furthermore it can be seen that both increasing the porosity (as is already demonstrated when increasing porosity by 1.4, that the summated daily flow decreases with almost  $271 \text{ m}^3 \text{ s}^{-1}$  to 1071 from 1342) and increasing hydraulic conductivity has a damping effect on the flow. This makes sense, since subsoil storage is increased when increasing the porosity and infiltration capacity is also increased when increasing hydraulic conductivity so flows are damped more. The graph from the initial run is shown in figure 5.16:

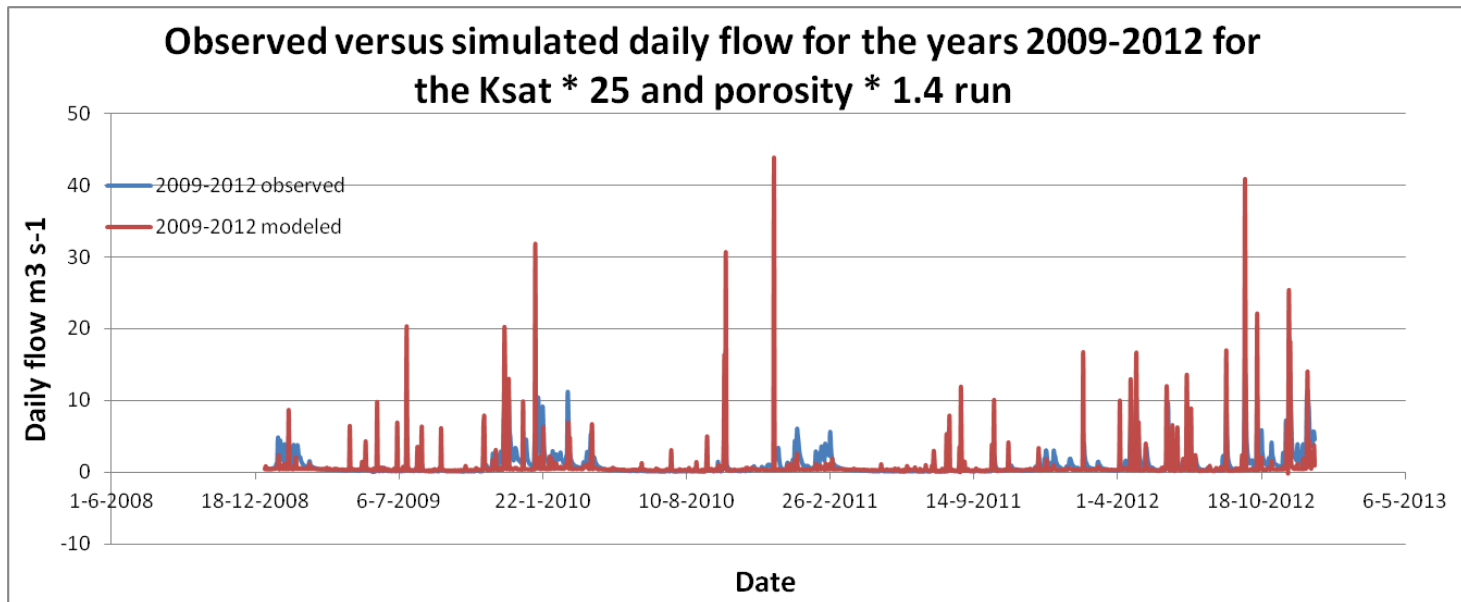


Figure 5-16: Results of the run on 25m for the initial run for the years 2009-2012

In this run the peak flows are overestimated largely, and other peak flows are underestimated. The NSE scores can give a somewhat distorted picture, because they reflect the score which is best based on the used meteorological forcing (which may not be that accurate) and perhaps more important: they do not implicitly take into account deviations from the observed annual flow or a correct pattern of river flow. For instance, a correct pattern with regards to peaks and low flows but over / underestimated values may get a lower score than an incorrect pattern which simply deviates less from the observed flow on average. In order to investigate the point made, it is interesting to look at some of the misses in the model (over or underestimation of observed flow compared to observed flow) and see how they affect the score. Therefore it is worthwhile to examine some meteorological patterns on days of misses (over / underestimation of flow). A great series of misses with regards to peak flow are the months of January and February 2009. Although there are no major peak flows, daily flows are reasonably high all the time with a mean value of around  $2 \text{ m}^3 \text{ s}^{-1}$ , compared to a typical background flow value of around  $0.4 \text{ m}^3 \text{ s}^{-1}$ . The results are shown in Table 5.6, showing the observed flow, all the simulated flows [ $\text{m}^3 \text{ s}^{-1}$ ] and the catchment accumulated average rainfall at the out flow point [ $\text{mm d}^{-1}$ ]:

Table 5-6: Daily and observed and precipitation flow for the different runs

Date	Obs [ $\text{m}^3 \text{ s}^{-1}$ ]	Ini [ $\text{m}^3 \text{ s}^{-1}$ ]	Por *1.4 [ $\text{m}^3 \text{ s}^{-1}$ ]	Ksat*5 [ $\text{m}^3 \text{ s}^{-1}$ ]	Ksat*10 [ $\text{m}^3 \text{ s}^{-1}$ ]	Ksat*15 [ $\text{m}^3 \text{ s}^{-1}$ ]	Ksat*20 [ $\text{m}^3 \text{ s}^{-1}$ ]	Ksat*25 [ $\text{m}^3 \text{ s}^{-1}$ ]	P [ $\text{mm d}^{-1}$ ]
12-2-2009	1.70	0.49	0.63	0.50	0.49	0.49	0.49	0.48	0.26
<b>13-2-2009</b>	<b>2.02</b>	<b>2.86</b>	<b>1.54</b>	<b>1.77</b>	<b>1.80</b>	<b>1.71</b>	<b>1.69</b>	<b>1.63</b>	<b>10.95</b>
14-2-2009	2.13	0.58	0.52	0.54	0.52	0.53	0.51	0.52	0.24
15-2-2009	3.74	0.46	0.50	0.52	0.51	0.52	0.50	0.51	0.34
16-2-2009	3.19	0.48	0.49	0.49	0.50	0.50	0.49	0.50	0.00

2009									
------	--	--	--	--	--	--	--	--	--

Table 5.6 proves the point made about the NSE score reflecting the best result based on a sometimes inaccurate meteorological forcing. In reality, observed flows were much higher than the simulated flows, except on 13-2-2009, which shows some heavy rainfall (10.95 mm). It is likely that on the other days there was more rainfall than simulated, which is missed due to only one rain gauge being present in the catchment, in the lower southern part. This is in line with the earlier point made about the large annual rainfall variability in Yorkshire, which extremes are as high as 1500 mm yr<sup>-1</sup> around the Pennines and as low as 600 mm yr<sup>-1</sup> at the coast. How about a miss with an overestimation of flow? The results of this situation are shown in Table 5.7, showing the observed flow, all the simulated flows [m<sup>3</sup> s<sup>-1</sup>] and the catchment accumulated average rainfall at the out flow point [mm d<sup>-1</sup>]:

Table 5-7: Daily and observed flow and precipitation for the different runs

Date	Obs [m <sup>3</sup> s <sup>-1</sup> ]	Ini [m <sup>3</sup> s <sup>-1</sup> ]	Por *1.4 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*5 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*10 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*15 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*20 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*25 [m <sup>3</sup> s <sup>-1</sup> ]	P [mm d <sup>-1</sup> ]
2-7-2009	0.12	0.28	0.20	0.24	0.25	0.19	0.21	0.14	0.00
3-7-2009	0.46	0.26	0.27	0.26	0.16	0.13	0.18	0.18	0.00
<b>4-7-2009</b>	<b>0.29</b>	<b>4.06</b>	<b>0.98</b>	<b>1.04</b>	<b>1.04</b>	<b>0.99</b>	<b>0.94</b>	<b>1.03</b>	<b>15.95</b>
5-7-2009	0.25	0.28	0.27	0.17	0.28	0.24	0.23	0.12	0.05
6-7-2009	0.18	0.29	0.29	0.28	0.20	0.22	0.16	0.24	0.94

The same pattern as in table 5.6 is observed: a substantial amount of precipitation on the day of peak flow (which is not observed in reality). The tables also reveal a problem related to the discharge. If a large precipitation event causes a peak flow, it is discharged on the same day and the next day there is no recession. Normally, some recession before background flow is reached should be expected, with daily flow values gradually and not instantly decreasing, which can be seen for instance from 26-2-2010 until 5-3-2010 in tables 5.8 and 5.9. This could have some possible causes: the hydraulic conductivity of the 4<sup>th</sup> layer is too low, not releasing enough water from storage. The standard recession period of a month may also have been too long. The recession period determines how fast the aquifer releases water from storage. It could also be that on consecutive days of observed peak flow, there are more rainfall days or at least the effect is "felt" longer, due to more water being released from storage. Unfortunately, due to time constraints it is not possible to run the model again with different parameterizations for the hydraulic conductivity of the fourth layer. This probably does explain why the simulated river flow is too dry compared to the observed, because unlike in reality, peak flows are instantly discharged and river flows drop back to normal levels the day after. In addition, when less water is released from storage, it also means there is less subsoil storage because of wetter soil conditions leading to overestimation of peak flows, as seen in figure 5.5.

Based on older results, which were inaccurate (due to a resampling error when resampling the layer 3 map from 100 m to 25 m, layer 1, 2 and 3 combined actually exceeded the DEM at some points, causing water levels to be above ground), Ksat \* 25 and porosity \* 1.4 was chosen for the forecast run. Due to time constraints, it was not possible to re-run the forecast with another parameter combination. Possibly, parameters could have been estimated much lower when taking recession into account. Chapter 5.4 will reveal how these hydrological problems affect the stability on the slope scale, because the scales of summated river flow and slope hydrology are incomparable. To this end, the results of the wettest run (initial) and the driest run (Ksat \* 25 and porosity \* 1.4) will be compared to the forecasted slope stability results.

Table 5-8: Daily and observed flow and precipitation for the different runs

Date	Obs [m <sup>3</sup> s <sup>-1</sup> ]	Ini [m <sup>3</sup> s <sup>-1</sup> ]	Por *1.4 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*5 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*10 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*15 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*20 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*25 [m <sup>3</sup> s <sup>-1</sup> ]	P [mm d <sup>-1</sup> ]
25-2-2010	2.16	0.50	0.53	0.51	0.53	0.53	0.53	0.50	0.25
<b>26-2-2010</b>	<b>11.00</b>	<b>3.43</b>	<b>2.23</b>	<b>2.51</b>	<b>2.24</b>	<b>2.18</b>	<b>2.12</b>	<b>2.04</b>	<b>10.42</b>
<b>27-2-2010</b>	<b>8.06</b>	<b>9.36</b>	<b>3.25</b>	<b>2.73</b>	<b>3.49</b>	<b>2.66</b>	<b>2.77</b>	<b>2.82</b>	<b>10.9</b>
28-2-2010	3.85	0.51	0.52	0.51	0.53	0.51	0.52	0.50	0.03
<b>1-3-2010</b>	<b>2.57</b>	<b>7.34</b>	<b>4.33</b>	<b>2.89</b>	<b>2.55</b>	<b>2.28</b>	<b>2.61</b>	<b>2.47</b>	<b>12.46</b>
2-3-2010	1.97	0.46	0.70	0.53	0.51	0.51	0.52	0.51	0.11
3-3-2010	1.65	0.50	0.45	0.52	0.51	0.51	0.49	0.49	0.11
4-3-2010	1.36	0.46	0.51	0.51	0.51	0.51	0.51	0.50	0.00
5-3-2010	1.20	0.55	0.49	0.51	0.50	0.52	0.50	0.50	0.01

Table 5-9: Daily and observed flow and precipitation for the different runs

Date	Obs [m <sup>3</sup> s <sup>-1</sup> ]	Ini [m <sup>3</sup> s <sup>-1</sup> ]	Por *1.4 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*5 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*10 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*15 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*20 [m <sup>3</sup> s <sup>-1</sup> ]	Ksat*25 [m <sup>3</sup> s <sup>-1</sup> ]	P [mm d <sup>-1</sup> ]
1-10-2010	1.17	4.61	1.30	0.53	0.67	0.55	0.62	0.58	8.00
<b>2-10-2010</b>	<b>1.39</b>	<b>15.89</b>	<b>7.17</b>	<b>6.46</b>	<b>3.25</b>	<b>2.20</b>	<b>1.92</b>	<b>1.15</b>	<b>18.71</b>
3-10-2010	3.69	0.77	0.88	0.57	0.60	0.53	0.46	0.56	5.56
<b>4-10-2010</b>	<b>2.89</b>	<b>31.68</b>	<b>27.30</b>	<b>13.41</b>	<b>14.14</b>	<b>7.22</b>	<b>9.41</b>	<b>6.81</b>	<b>30.49</b>
5-10-2010	1.23	0.24	0.31	0.28	0.23	0.26	0.27	0.21	0.02

All in all, it can be stated that a better representation of the correct meteorological conditions is needed. Firstly, the amount of rain days is over estimated (which can partly be solved by applying a threshold value if only 1 station records rainfall or by setting a maximum radius in which 1 station can influence the surroundings when only 1 station records rainfall in order to simulate local rainfall, to prevent 1 station from simulating rainfall over the entire catchment). Secondly, the pattern with regards to peak flows is not correctly simulated as well. The river flow is also too dry, with peak flows being instantly discharged. This is caused by a constraining 4<sup>th</sup> layer and a too long recession period, allowing too little water to be released from storage to simulate gradual recession of peak flows towards a background flow level. This also causes a likely overestimation of hydraulic conductivity. If more water is released from storage, hydraulic conductivity can be estimated lower because there is more subsoil storage.

### 5.3. River flow results forecast run on 25m

The results of the eight day forecast with respect to NSE scores [-] and total annual daily flow Q [m<sup>3</sup> s<sup>-1</sup>] are shown in Table 5.10 (for comparison, the recorded daily flow amounted to 659 m<sup>3</sup> s<sup>-1</sup>). Note that for forecast day 0 the values are starting at 02-01-2012 and end at 31-12-2012, the day 1 values range from 03-01-2012 until 31-12-2012, day 2



from 04-01-2012 until 31-12-2012 and so forth. The NSE scores relative to the day 0 forecast (which reflects how well the various forecast days perform compared to the day 0 forecast) are also shown.

Table 5-10: NSE scores for different forecast days and total volume of the year 2012

Forecast day [d]	NSE [-]	NSE rel. to day 0	Q [m <sup>3</sup> s <sup>-1</sup> ]
0	0.02		414
1	-0.01	0.29	354
2	-0.03	0.45	359
3	-0.10	-0.05	354
4	-0.18	-0.23	359
5	-0.15	-0.06	347
6	-0.07	-0.24	355
7	-0.18	-0.67	377
8	-0.15	-0.24	352

The results of the day 0 forecast are shown in figure 5.17:

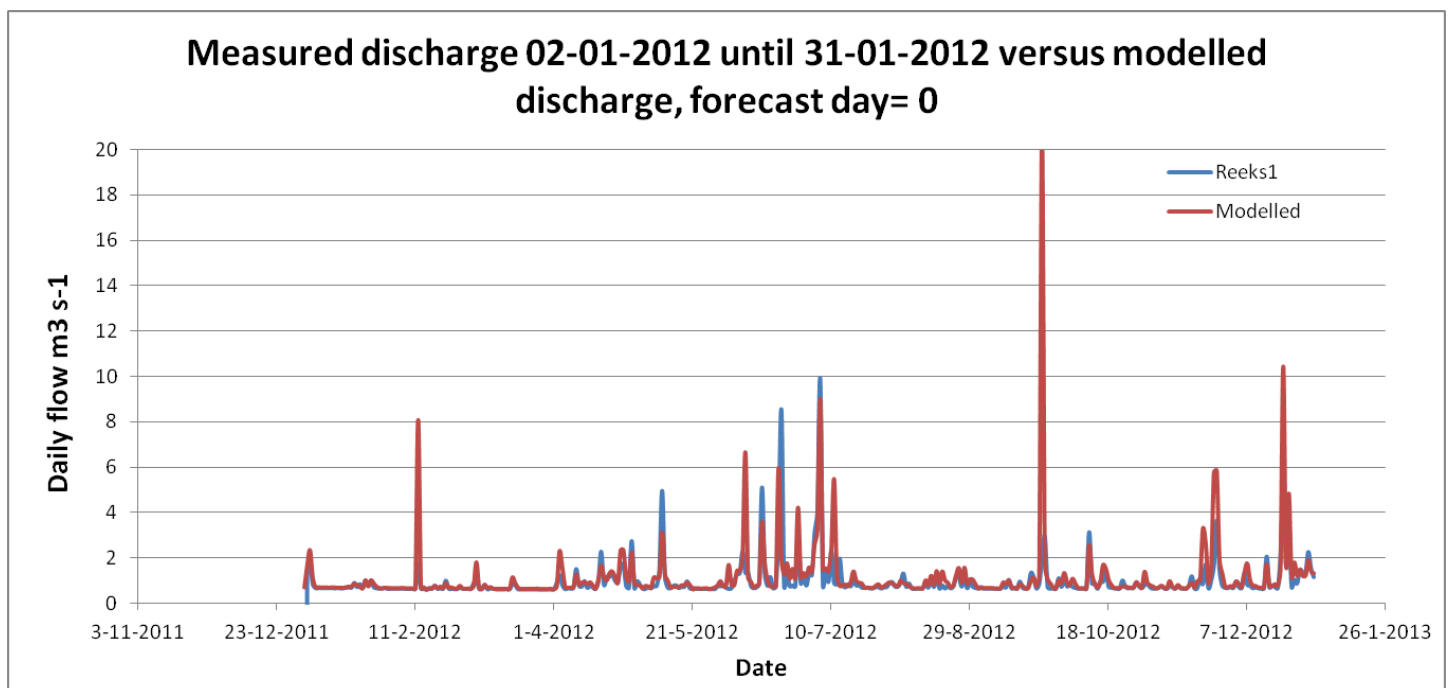


Figure 5-17: Daily forecasted flow versus daily observed flow [m<sup>3</sup> s<sup>-1</sup>] for the year 2012

From the NSE it appears that the model has no forecasting capabilities with respect to daily river flow forecasting. The graph contradicts this to a certain extent. The forecasted pattern is quite accurate, albeit with some misses that clearly influence the score. The first miss is around 11-2-2012, the second (smaller) miss around 10-7-2012 (overestimation of low flow), the third miss is between 29-8-2012 and 18-10-2012 and the last miss is at the right side of the graph, slightly overestimating the low flow. The overall patterning is much better than with the historic run. This may lead to the conclusion, which was already mentioned, that the historic run precipitation patterns may have been not so accurate because of the stations used for interpolation being outside of the catchment. Although the NSE scores were better, they produced a pattern which was much worse. Also, the historic run with the best score ( $K_{sat} * 25 \text{ porosity} * 1.4$ ) was much too dry for the year 2012, showing a summated daily flow of  $252 \text{ m}^3 \text{ s}^{-1}$  versus an observed of  $658 \text{ m}^3 \text{ s}^{-1}$ . The forecast run is better with regards to that, showing  $414 \text{ m}^3 \text{ s}^{-1}$  for the 0 day forecast. The reason that the day 2 forecast has a better correlation towards the day 0 forecast than the day 1 forecast, is mainly because of some specific peak flow values, as can be seen in figures 5.18 and 5.19:

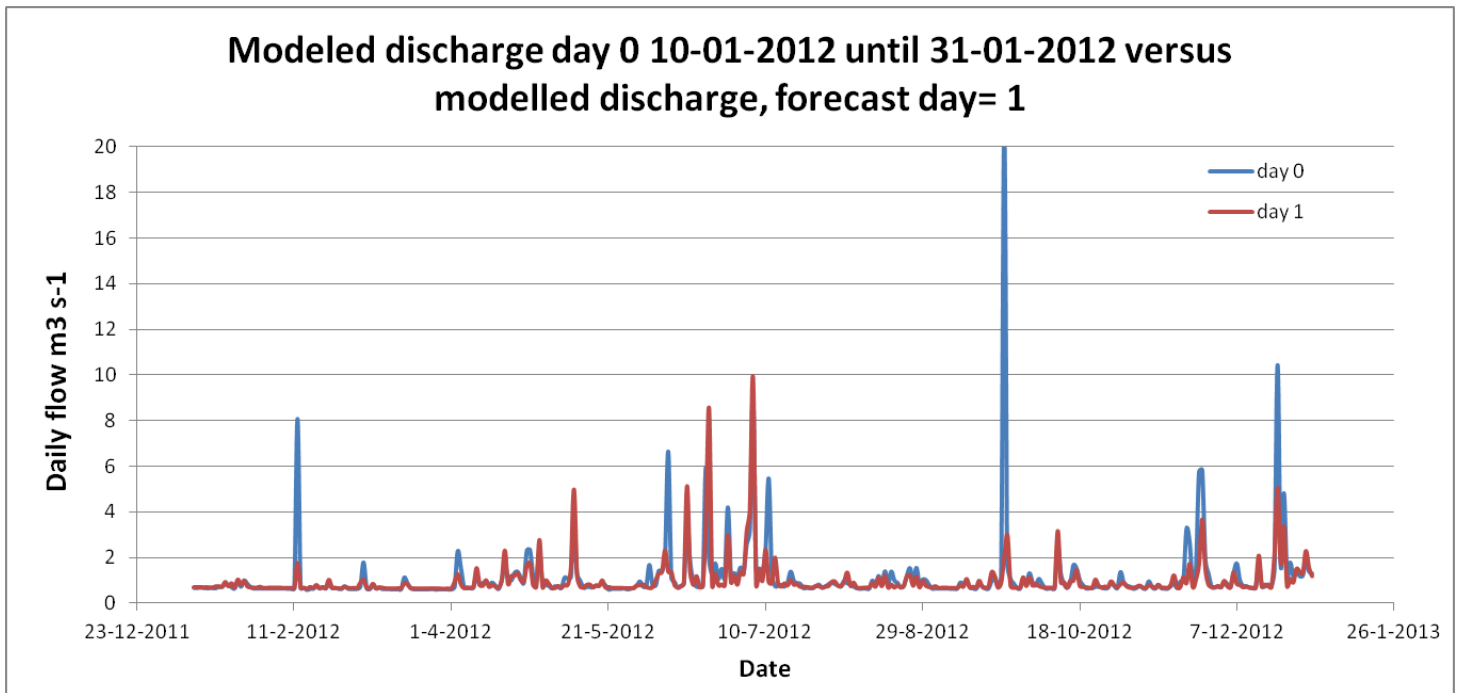


Figure 5-5-18: NSE of day 1 versus day 0

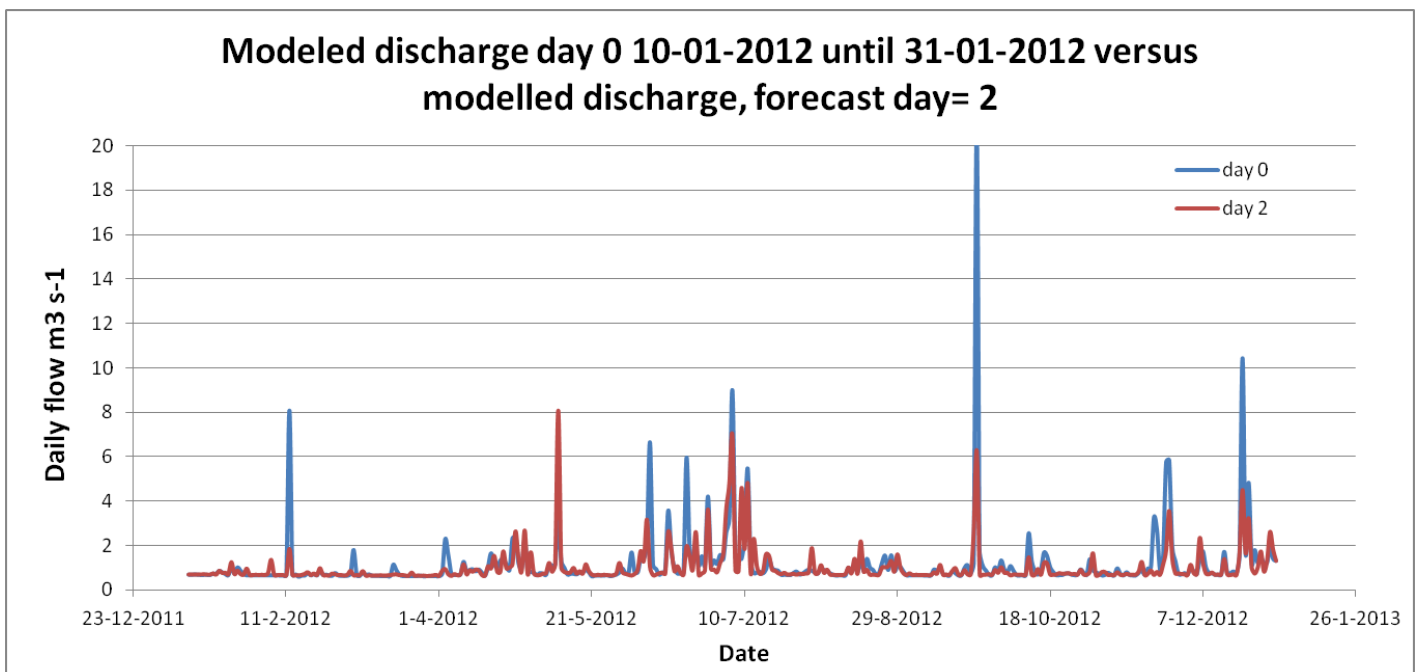


Figure 5-19: NSE of day 2 versus day 0

The misses that are occurred are shown in Table 5.11, showing the observed flow on that day, the forecasted flow, and the forecasted precipitation:

Table 5-11: Results for day 0

Date	Q Obs [m <sup>3</sup> s <sup>-1</sup> ]	Q Day 0 [m <sup>3</sup> s <sup>-1</sup> ]	P Day 0 [mm]
12-2-2012	0.52	8.07	0.00
9-6-2012	6.98	6.65	21.00
10-6-2012	9.57	1.12	3.49
11-6-2012	4.37	0.94	0.80
12-6-2012	2.3	0.69	0.00
13-6-2012	1.46	0.77	1.99
24-9-2012	5.85	20.55	34.01
25-9-2012	13.8	1.70	7.85
26-9-2012	13.2	1.06	2.70
27-9-2012	5.73	0.87	1.00

The first miss on 12-2-2012 shows no rainfall. This is caused by snowmelt however, as 11-2-2012 showed a mean temperature of -3.46 and 10-2-2012 a mean temperature of -4.51, and 12-2-2012 a mean temperature of 1.51. This was likely a miss in the model, even though the observed temperatures show a similar pattern. The other problem which the model is showing is that there is not enough background flow, also shown in figure 5.16. The model responds to precipitation (for instance on 24-9-2012) with increased river flow, but it drops significantly the day after instead of gradually. As said in chapter 5.2, this is likely caused by a too low recession coefficient [d<sup>-1</sup>] (which is set at one month or 0.033 as a standard value) and / or too low Ksat of the 4<sup>th</sup> layer. If the Ksat of the 4<sup>th</sup> layer is increased, the model can release more water from storage. This has to be supplied however, therefore the recession coefficient could be increased as well, allowing the model to release more water from storage. When the model was re-run with the recession period halved to 15 days and the hydraulic conductivity of the 4<sup>th</sup> layer doubled to 0.0015 m d<sup>-1</sup> on the historic run with Ksat \* 25 and porosity \* 1.4 for the years 2009-2012, the NSE immediately improved to 0.33 and the summated daily flow amounted to 1720 m<sup>3</sup> s<sup>-1</sup> with most notably an overestimation of background flow, compared to an earlier NSE of 0.21 and 754 m<sup>3</sup> s<sup>-1</sup>, and an observed total daily flow of 1524 m<sup>3</sup> s<sup>-1</sup>. This shows that calibrating the 4<sup>th</sup> layer and the recession period (which could be fine tuned even more of course) partially solves the “dry” problem and the related low NSE score. The lack of a really good score can still likely be attributed to poor rainfall patterning.

If the total volumes are observed, it becomes apparent that the model has an extra dry bias after 24 hours (day 1 and after). To explain this, the annual precipitation totals [mm] of the historic run, forecast day 0 and day 1 are shown in figure 5.20 and of day 5 and day 8 in figure 5.21 (these values are shown because they are start, mid level and end level values):

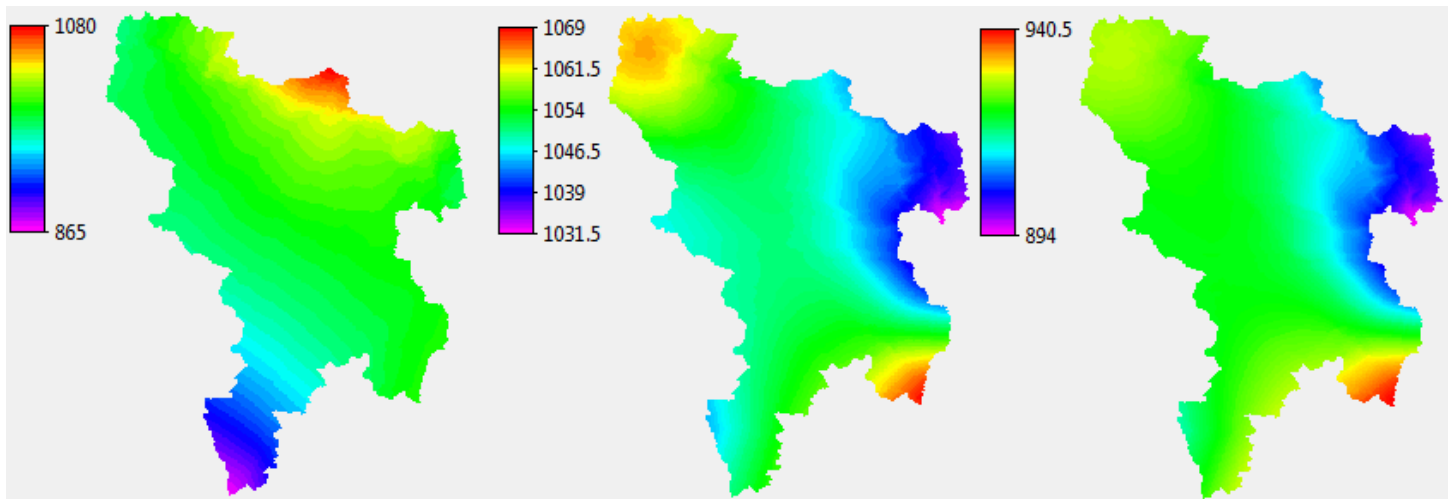


Figure 5-20: Annual precipitation in mm for the historic run, forecast run day 0 and day 1

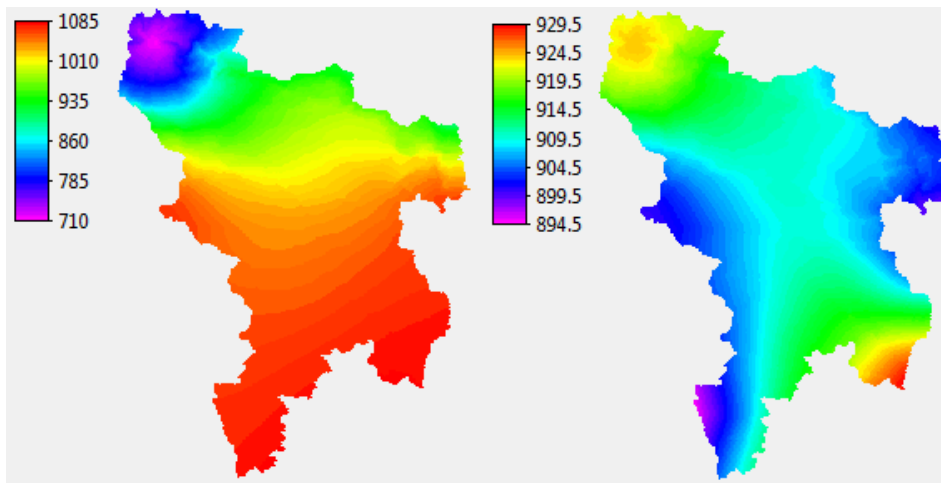


Figure 5-21: Annual precipitation in mm for forecast run day 5 and 8

For fair comparison, the precipitation totals from all the figures are the totals from 10-01-2012 until 31-12-2012 (the start date of the day 8 forecast). When looking at these figures it can be concluded that the lower volumes of day 1, day 5 (although day 5 has a higher maximum precipitation, it also has a far lower minimum precipitation, cancelling each other out), and day 8 can be explained by lower precipitation totals, in comparison to day 0 and the historic run. The variability for all forecast days seems to be quite low (range between minimum and maximum rainfall value), except for forecast day 5. However, the variability from day 5 is a total mismatch to the variability from the historic run. The extremes seem to be reversed (in the historic run, highest rainfall values are in the top part of the catchment and lowest in the low part, for forecast day 5 this is the complete opposite). In addition, while the maximum rainfall is very comparable (1080 mm for the historic run versus 1085 mm for the day 5 run), the minimum rainfall is not comparable (865 mm for the historic run versus 710 mm for the day 5 run, a difference of 155 mm). Also, for the day 5 forecast there is a much larger area having near maximum rainfall values than for the historic run.

The total PET [mm] for the historic run, forecast day 0, and day 1 is shown in figure 5.22 and for day 5 and 8 in figure 5.23, starting on 10-01-2012 and ending on 31-12-2012:

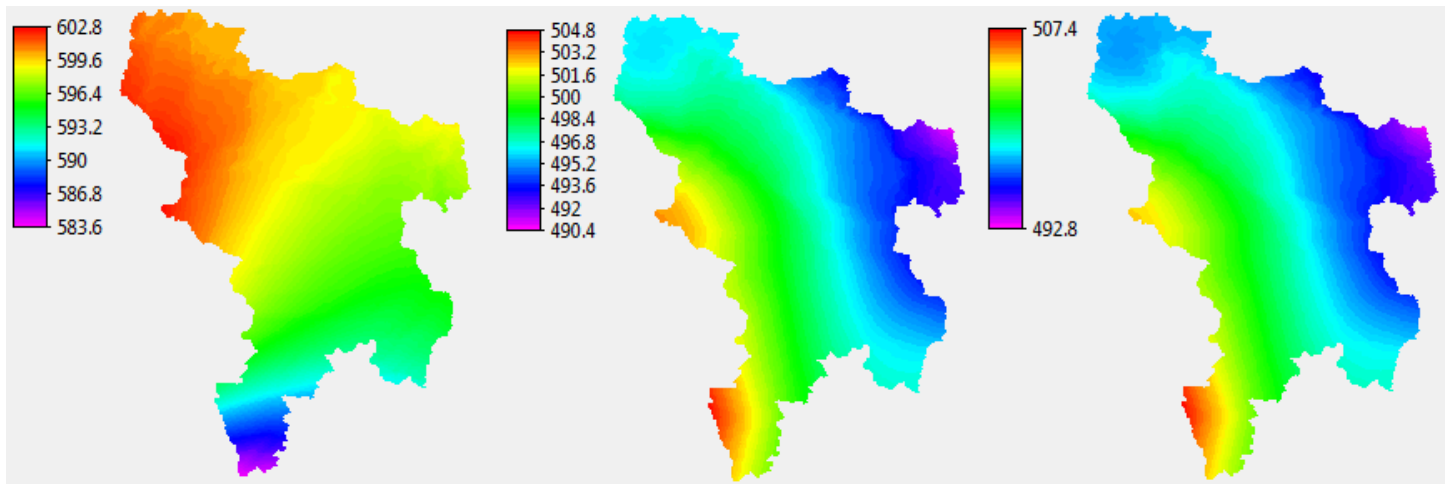


Figure 5-22: PET totals [mm] for the historic run, forecast day 0, and day 1 from 10-01-2012 until 31-12-2012

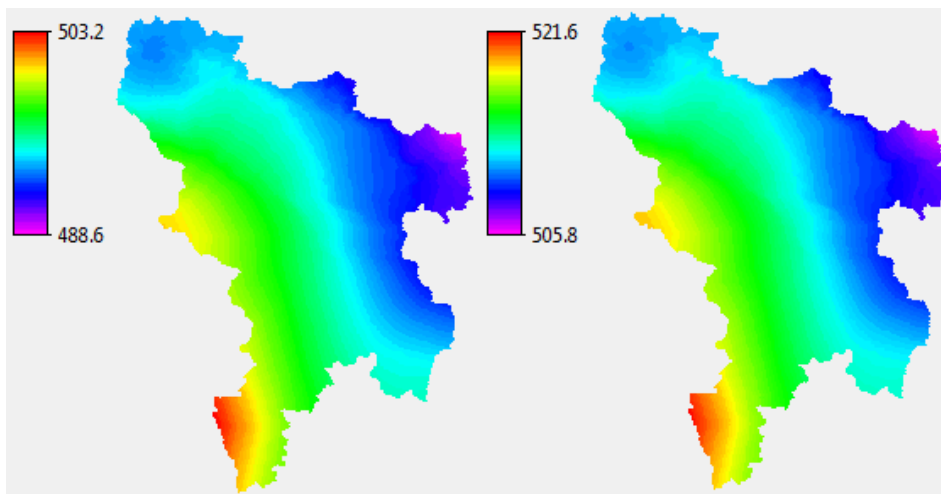


Figure 5-23: PET totals [mm] for the forecast days 5 and 8

A few interesting observations can be made when studying both figures 4.16-4.20:

- The range of PET values is more or less comparable for the historic run and the various forecast runs, ~15-20 mm;
- The historic run assumes a significantly higher PET of around 100 mm higher, which is quite significant (around 20%). Crop factors can make this total absolute difference even higher. Assume that an average crop factor throughout the year is 1.2. The result would be a local PET of  $600 * 1.2 = 720$  mm for a cell with total PET value 600 mm and a total of  $500 * 1.2 = 600$  mm for a cell with total PET value 500 mm for the forecast run, meaning the absolute difference became 120 mm, i.e. 20% of the previous absolute difference. Higher PET values means more water (if available) can evaporate, altering hydrological results and explaining the drier model with the chosen parameter combination for the historic run. Also, the mentioned lack of recession explains this. The annual PET values of the forecast run are more or less the same, meaning the average yearly temperature is the same for the forecast runs. This does not imply that the monthly average temperatures match as well, because with a similar yearly average temperature there can be alternating monthly temperature fluctuations of course.
- The figures explain why each forecast day run has a quite different NSE score. Precipitation values differ significantly in their spatial distribution and in some cases (day 0 and day 5 compared to day 1 and day 8)

also quantitatively, whereas PET values are very similar, both quantitatively and spatially. Likewise, the model gives a different hydrological response to these variations in forcing.

Another explanation for the differences between the different forecast days can be found in the total number of rain days. The total number of rain and non rain days from 10-01-2012 until 31-12-2012 are shown in Table 5.12:

Table 5-12: Number of rain and non rain days for different forecast days

Forecast day [d]	# of rain days	# of non rain days
0	221	136
1	230	127
2	238	119
3	235	122
4	244	113
5	239	118
6	245	112
7	228	129
8	234	123

Table 5.12 shows that the number of rain days is different for each forecast day, effectively altering the pattern and intensity of daily precipitation. Finally, phase errors can cause different hydrological outcomes. Phase errors denote a situation where anomalies are correct in frequency, but not in time, usually occurring with increasing lead time. This is situated in figure 5.24, which is a general example of such a situation (ECMWF, 2014):

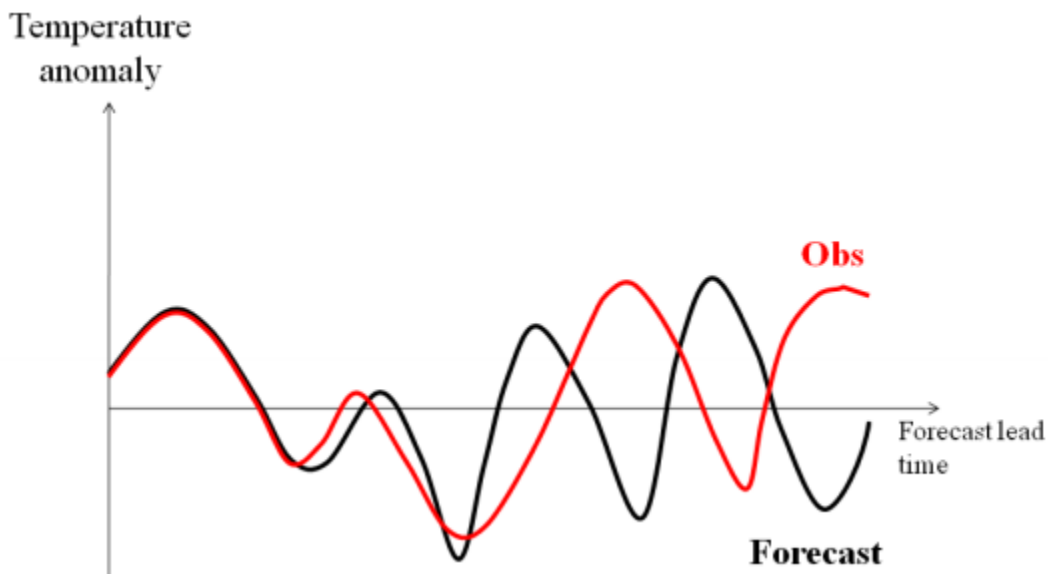


Figure 5-24: Phase errors with increasing lead time. From ECMWF (2014).

Figure 5.24 clearly shows that when lead time increases towards the right on the x-axis the differences between observed and forecast temperature on the y-axis increase drastically, even though the anomalies and timing between the anomalies are similar. This is caused by the phase error, causing the black graph to be similar in shape to the red graph, but erroneously shifted towards the left. This can lead to a situation shown in figure 5.20 where a peak in the observed temperature (red) coincides with a minimum of the forecasted temperature (black), enhancing the difference. In this situation it is assumed that anomalies and timing between them are comparable, but there is just a phase error, however, anomalies and timing in between anomalies can also differ and coexist with phase

errors (ECMWF, 2014).

In the user guide of the ECMWF forecast it is also specifically mentioned that the convective model has extra uncertainty regarding advecting winter rain showers over the inner part of Scotland and Northern England (where Hollin Hill is situated) originating from North Sea, being relatively warm. The diagnostic convection scheme works per model column and therefore cannot simulate large amounts of precipitation falling over the (relatively) dry and cold winter inland areas of Northern England and Scotland. In reality, large amounts of precipitation can occur through upper-level warm anomalies induced by convection leading to air lifting and increased air saturation (ECMWF, 2013). Advection can cause the total amount of rainfall to be higher due to continuous transport towards a certain location (moisture convergence), which can occur along cold fronts (Haby, sine anno). As mentioned in the methodology, the chaotic structure of earth atmosphere is a difficult aspect with forecasting, because small deviations in initial states (meaning measurement errors or simply missing data) can lead to large(r) anomalies later in the forecasting period (Met Office, 2014). Other reasons can be that meso- or microscale circulation systems are not detected entirely, resulting in different weather patterns and therefore a different forecast. Finally, numerical models used in forecasting can use imprecise equations (AMS Weather Studies, 2014). For instance, as mentioned in the methodology, the ECMWF includes processes such as cloud formation which are too small to cover one cell of  $13.5 \text{ km}^2$ , as statistical parameterization processes (ECMWF, 2011). Likewise, a clear physically based coupling is absent in such an approach, but such concessions need to be made since physically based models demand recorded input which is not available on such small scales and computation wise this would also demand much more processing power and time.

## 5.4. Stability results

The total number of cells being unstable are shown in Table 5.13, showing this instability ( $SF < 1$ ) for an interval of 1 day and more, 30 days and more, 90 days and more, 180 days, 270 days and more and 357 days (unconditionally unstable), starting from 10-01-2012 and ending at 31-12-2012, for all the 9 forecast runs, the historic run on the  $K_{sat} * 25$  and porosity \* 1.4 parameterization (driest run) and the historic run with the initial parameterization (wettest run). The chosen intervals are random, but their purpose is to show to what extent (if at all) unstable areas grow or decline throughout the year.

Table 5-13: Number of cells being unstable for various runs. Total number of cells ~200,000.

Run	#Cells SF < 1 1-357 days	#Cells SF < 1 30-357 days	#Cells SF < 1 90-357 days	#Cells SF < 1 180-357 days	#Cells SF < 1 270-357 days	#Cells SF < 1 357 days
Historic initial	227	160	128	116	101	40
Historic Ksat25	495	272	208	186	160	56
Day 0	616	372	265	206	150	56
Day 1	642	370	266	206	149	55
Day 2	639	366	267	203	150	52
Day 3	633	362	269	205	150	53
Day 4	648	367	269	207	150	50
Day 5	634	368	270	209	150	51
Day 6	626	367	270	208	151	50
Day 7	737	367	270	208	150	48
Day 8	732	362	269	209	149	49

This table clearly shows some differences between the various runs. However, there are also similarities. All runs show a decrease from cells being unstable one day or more as the number of unstable days increases. This is what one would expect, given that cells can be seasonally or occasionally unstable. Only a smaller portion of the cells is *unconditionally unstable* (unstable throughout the year). This is what you could expect in areas with steep slopes and low soil depths, as they saturate easily. The number of unconditionally unstable or near unconditionally unstable (270 days or more) is also more or less comparable, except for the initial run. There are two explanations for this: the initial run has a low hydraulic conductivity and a low porosity. The former causes more overland flow and less infiltration, the latter decreases the weight of the wet soil. Also, the historic runs showed an higher PET. This means more water can evaporate, likely leading to (slightly) drier conditions, leading to less unstable cells. This could explain why initially the forecast runs have much more unstable cells, but as the number of unstable days increases this more or less converges to the same number of cells as the historic runs. The cells that are unstable throughout the year or the major part of the year, are less dependent on hydrologic conditions. The cells that are occasionally unstable are, hence the major differences.

The last two days of the forecast run show even more unstable cells. This can be attributed to the different spatial distribution of rainfall patterns as shown in figures 5.20 and 5.21. Perhaps more rain is predicted to fall on the cells



prone to instability during the last days of the forecast, causing more instability. Figure 5.20 and 5.21 clearly show that the distribution of rainfall for the last day of the forecast is completely different to that of day 0.

This could imply that hydrological characteristics on the catchment scale (when looking at daily river flow) do not have a direct relationship to much smaller scale hydrology (looking at unstable cells). These scale problems are well known for (partly) physically based models and are addressed in various studies (Pechlivanidis *et al.*, 2011 ; Merz *et al.*, 2009 ; Blöschl and Sivapalan, 1995). Problems can be that the physics behind these models are often based on small scale laboratory or in-situ experiments and translated to catchment scale, often with the assumption that these physics are independent of scale (Beven, 2004). Catchments also typically show high levels of spatial heterogeneity which is almost impossible to observe and/or represent entirely in the model. In subsurface modeling this becomes even more apparent, because of the enhanced difficulty to measure / detect soil/aquifer heterogeneity. Usually, the scale of spatial heterogeneity even exceeds the scale of the cell size of the model itself (Pechlivanidis *et al.*, 2011), in this case for example, one would expect spatial variation of parameters in the cell of 625 m<sup>2</sup> itself. The parameters can be estimated through calibration or from experiments and databases, but the uncertainties are large enough to an extent that they are able to encompass a wide range of hydrological responses (Stephenson and Freeze, 1974), as can be seen in this study as well. Finally, to lessen the burden with respect to computational time and/or data requirements, simplified physics are used sometimes, such as the simplified St. Venant equations or Green-Amp equation, raising questions about the applicability (Pechlivanidis *et al.*, 2011).

The stability results with respect to cells being unstable one day or more and 180 days or more are displayed in figures 4.27-4.32. Figure 5.27 represents the initial run being unstable 1 day or more (upper left) and Ksat25 run (upper right), and being unstable 180 days or more (lower left and right). Figure 5.28 represents the same for forecast runs day 0 and day 1. The unstable area of 180 days and more did not change after forecast day 1, therefore only the results of day 1 and more are shown for forecast pairs 2 and 3, 4 and 5, 6 and 7, and 8 alone. Where the two black parts intersect is the lower end of Hollin Hill, which is shown in the red circle in figure 5.25:



Figure 5-25: The unstable cell at Hollin Hill

All runs show the same agglomerations of unstable areas, with a notable exception of the area in the right upper corner of figure .. (small black circle). Unfortunately, when it was crosschecked on google maps these cells turned out to be right in the middle of Mowthorpe Dale Wood (a local forest), so visual verification was not possible. The other patches of instability are mainly situated in a forests as well, Stittenham Wood (lower left circle) and Ox Pasture Wood (middle circle). However, in between the two forests (upper circle) and below Ox pasture Wood (right circles), there seems to be a potential area for instability:



Figure 5-26: Potential area of instability

Hollin Hill is not correctly classified in all runs, with the area towards the left of Hollin Hill (the forest) being correctly interpreted as unstable at some time intervals. When walking through this forest during field work it was established that many landslide scars were present there. The lower left cell of Hollin Hill is the most right part of the unstable patches. That means that the other landslides at Hollin Hill are not correctly predicted by the model. This is a clear miss. When investigating these cells it was revealed that water levels were too low below the surface (around 1.5 meters below surface at most) combined with a less steep slope angle of around  $12^\circ$ . This is close enough to the real slope angle, however due to no or little water being present in the shear zone with this slope angle the cell cannot reach instability. The question arises how accurate these results are then.

All runs show similar unstable areas, although slightly varying in size and decreasing when the number of unstable days increases. This means that the model is sensitive to topographical features such as slope angles and shear planes, alongside hydrological conditions. Hydrological conditions are also implicitly seen in the results, because of the declining number of unstable cells when the number of unstable days increases. An example of both the influence of hydrology and terrain properties is shown in is shown in Table 5.14 for two neighboring cells. Note that the principle demonstrated in this example is valid in general, because it gives insight into the calculation of the safety factor based on water being present in the shear zone.

Table 5-14: Influences of terrain properties on different cells

Cell#1: Lat:54.162 Lon: -1.127	Water level below surface [m]	Slope angle [°]	Depth shear plane below surface [m]	Surface height rel. to 0 [m]	SF [-]
<b>Timestep 1</b>	0	25.62	-1.35	111.37	0.62
<b>Timestep 180</b>	-1.15	25.62	-1.35	111.37	0.96
<b>Timestep 357</b>	-1.14	25.62	-1.35	111.37	0.96
Cell#2: Lat:54.162 Lon: -1.128					
<b>Timestep 1</b>	-1.58	27.98	-0.89	115.47	0.89
<b>Timestep 180</b>	-1.23	27.98	-0.89	115.47	0.93
<b>Timestep 366</b>	-1.20	27.98	-0.89	115.47	0.94

Table 5.14 gives some insight into the factors affecting instability. For cell #1, the perched water table is constantly in the shear zone, meaning as soon as water levels go up, safety factor decreases. This is seen in time step 1, when the soil is fully saturated, with a safety factor of 0.62 \*(failure). As soon as the water levels drop, the safety factor increases drastically to 0.96 (albeit still unstable). This can be explained by the high slope angle and the fact that there is still some water in the shear zone, increasing the weight of the column. Both factors combined are enough to cause this cell to be unstable. Cell #2 shows that the water table is constantly below the shear zone, yet is unstable. This can be explained by the steep slope. The only matter, besides slope angles, affecting the safety factor then is the degree of saturation in the unsaturated zone (inverse relationship to safety factor due to increased weight of the soil column when soil wettens). It has to be noted that cohesion for both cells is constant at 2 kPa, as well as the angle of internal friction, which is 23.6°. Cell #1 also became stable during some time steps (slightly above 1.00).

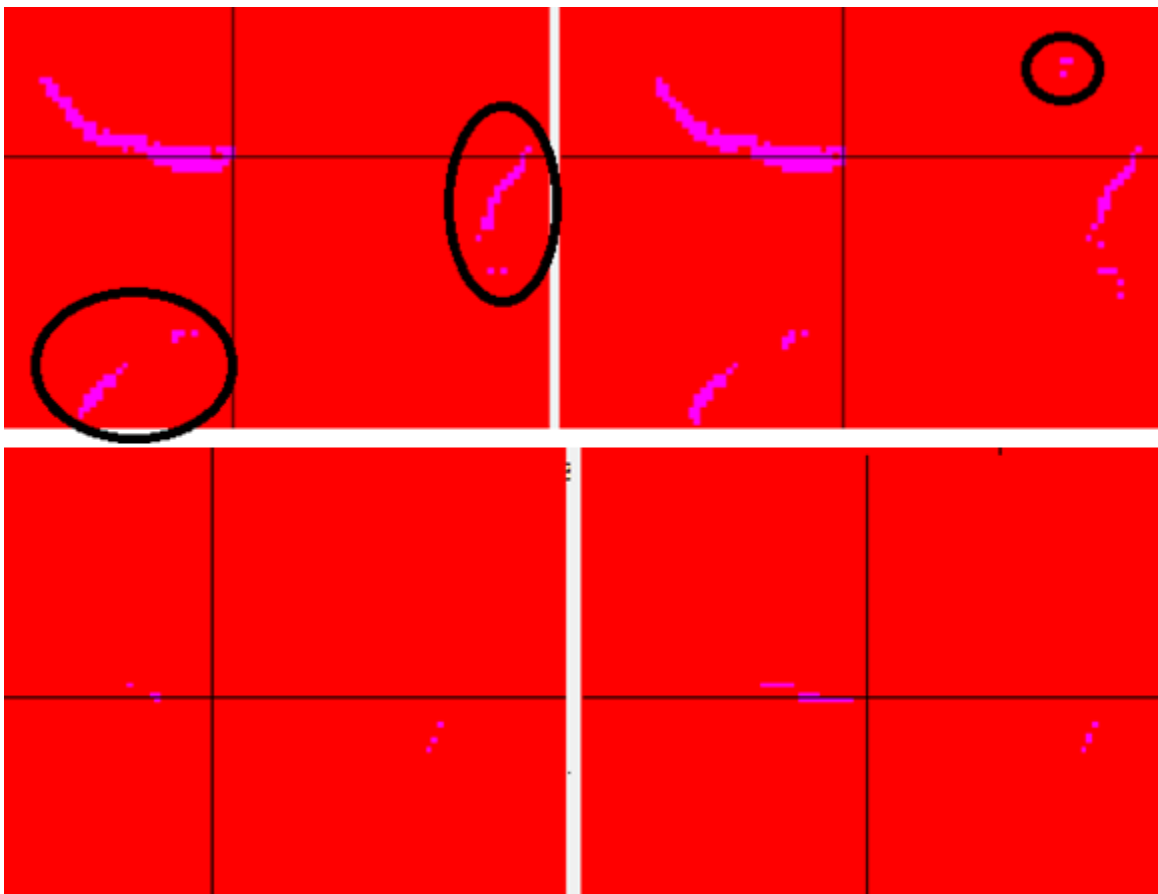


Figure 5-27: Unstable area around Hollin Hill for the historic run on initial values (left) and Ksat \* 25 (right) for 1 day or more (upper) and 180 days or more (lower)

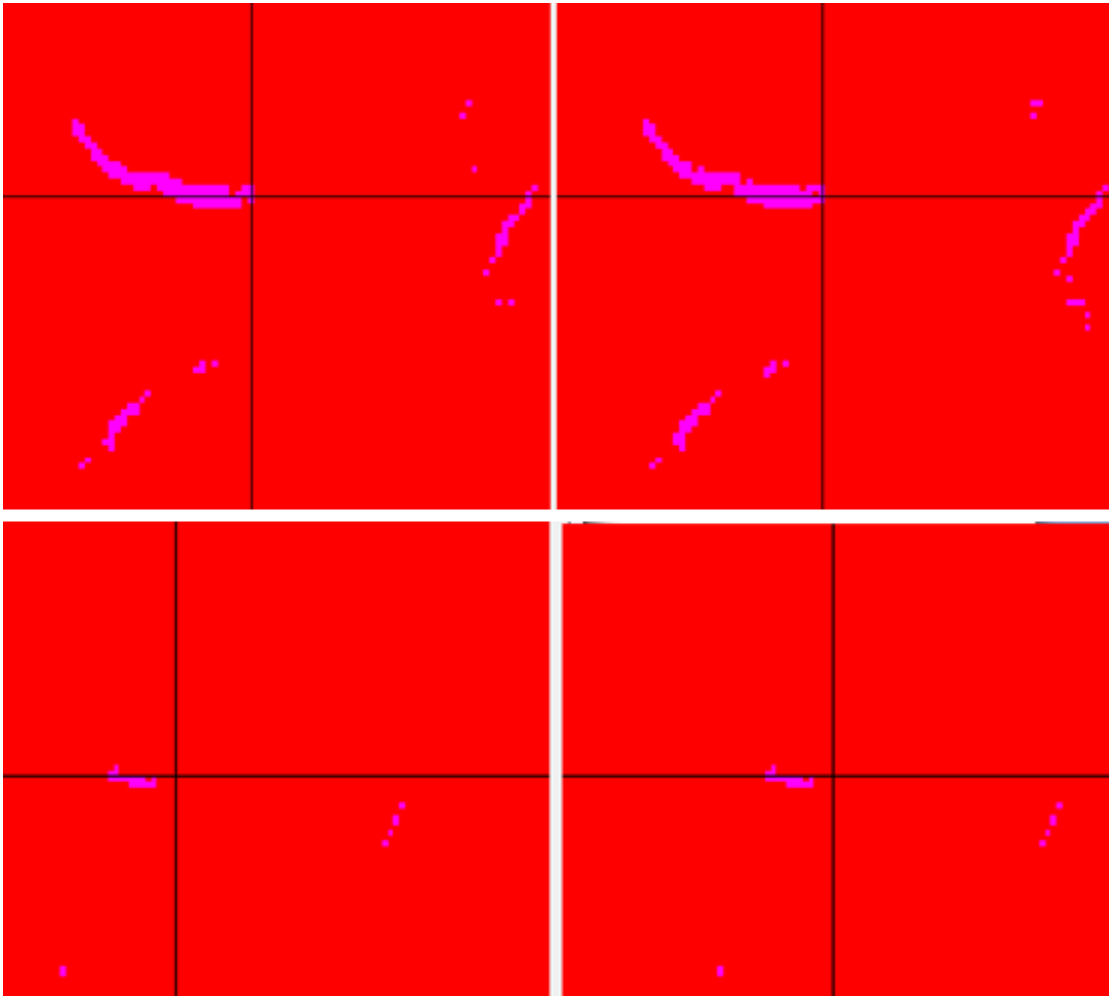


Figure 5-28: Unstable area around Hollin Hill for the for the forecast run day 0 (left) and day 1 right for 1 day or more (upper) and 180 days or more (lower)

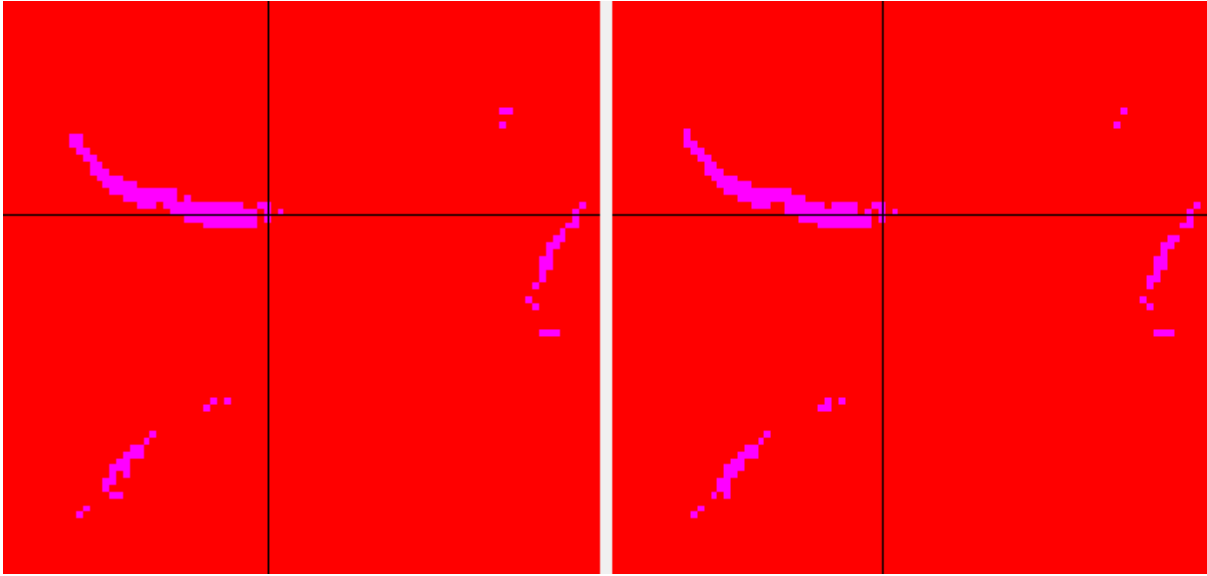


Figure 5-29: Unstable area around Hollin Hill for the for the forecast run day 2 (left) and day 3 right for 1 day or more

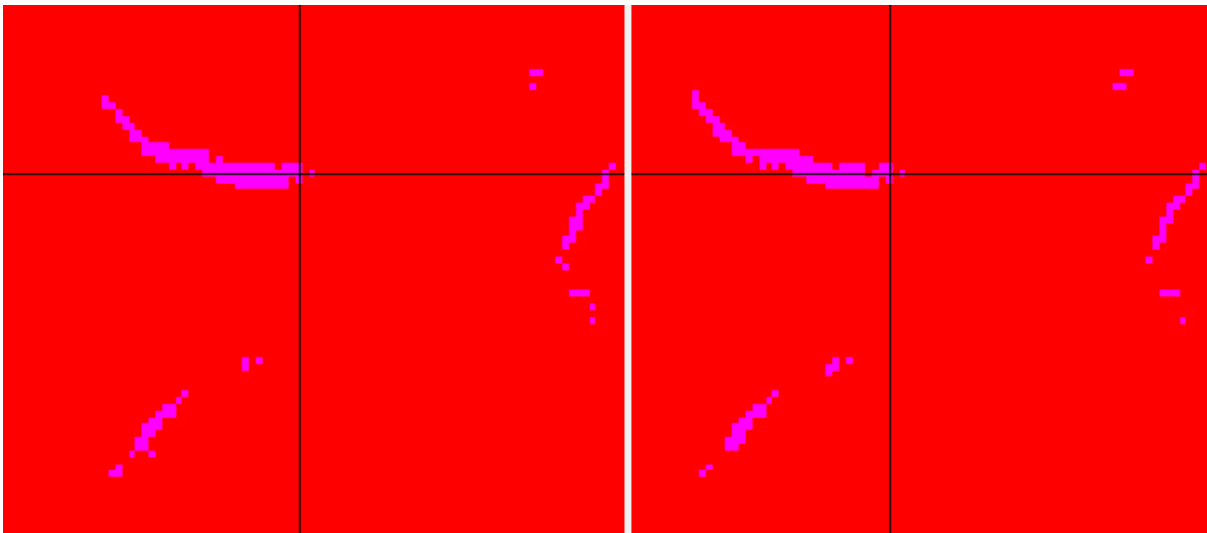


Figure 5-30: Unstable area around Hollin Hill for the for the forecast run day 4 (left) and day 5 right for 1 day or more

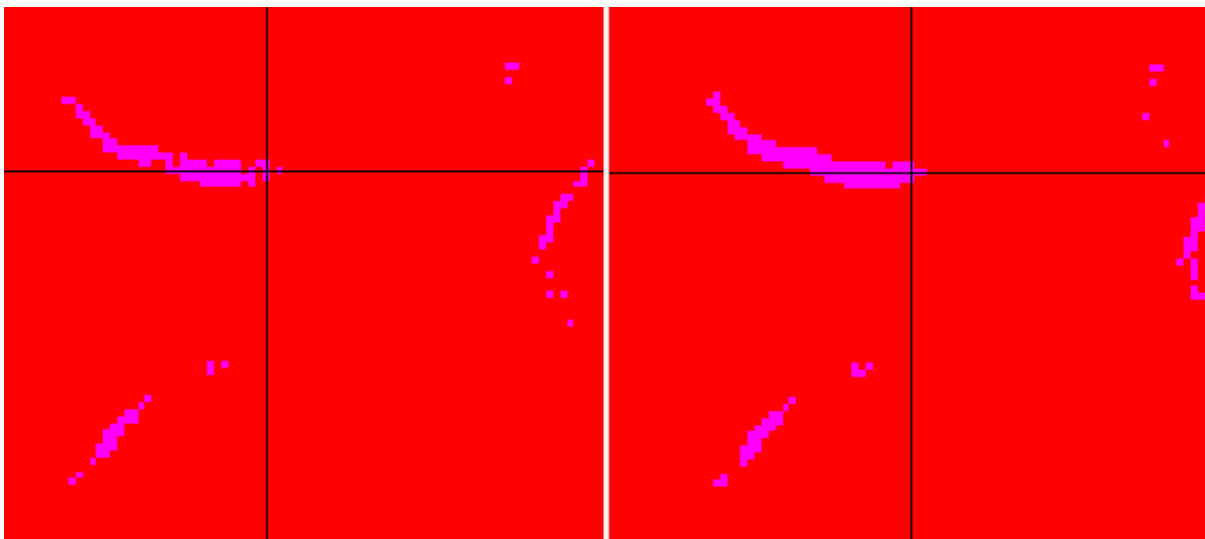


Figure 5-31: Unstable area around Hollin Hill for the for the forecast run day 6 (left) and day 7 right for 1 day or more

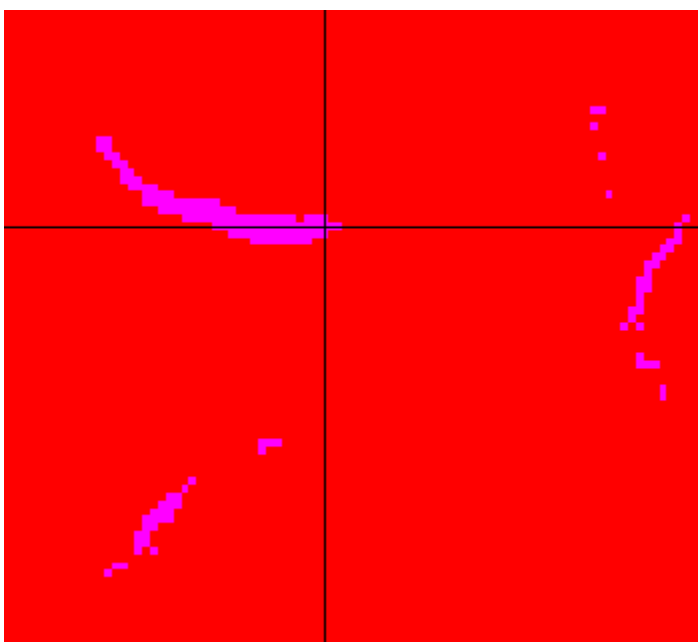


Figure 5-32: Unstable area around Hollin Hill for the for the forecast run day 8

Another patch of instability was found in the upper corner of the catchment. The result is shown in figure 5.33 for the initial run (left), the historic run with  $K_{sat} * 25$  (middle) and the forecast runs (which are almost completely similar and therefore merged into one picture on the right for more than one day of instability and for more than 180 days):

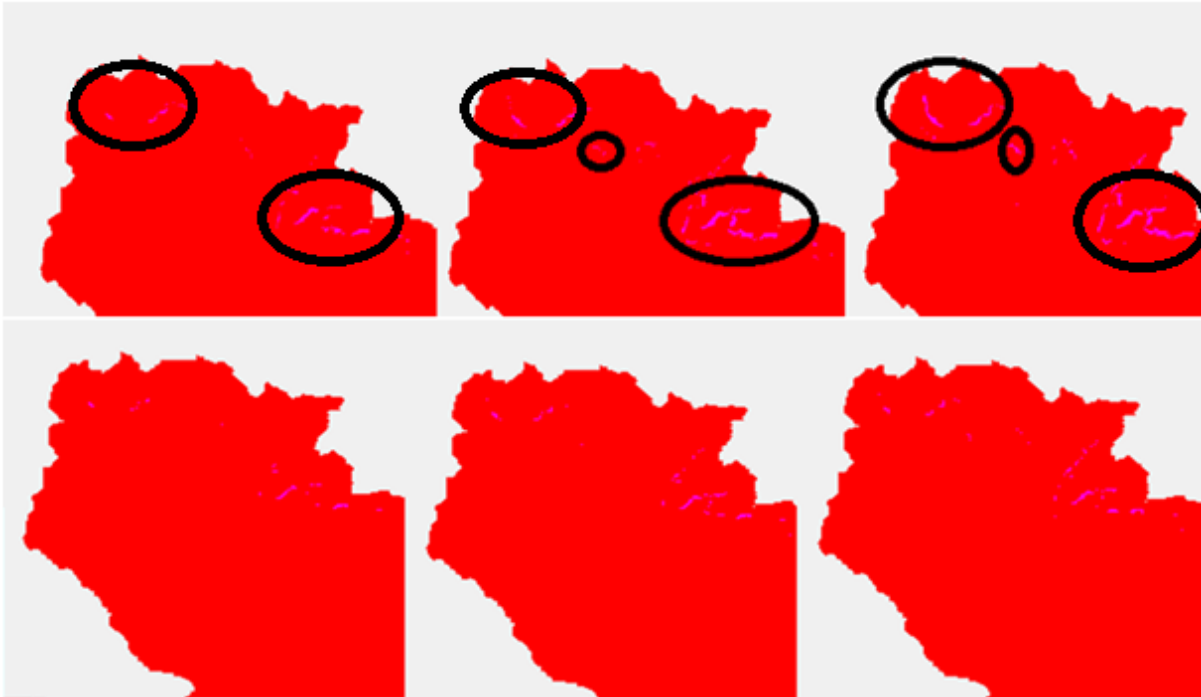


Figure 5-33: Instability in the upper part of the catchment for 1 day or more (upper) and 180 days or more (lower) for the initial run (left), the Ksat \* 25 run (middle) and the forecast runs (right)

Most areas within the big black circles turn out to be forest areas when crosschecked. The terrain does show some sign of possible (old) landslides around the forests and within the reach of the cells. This is shown in figure 4.34:

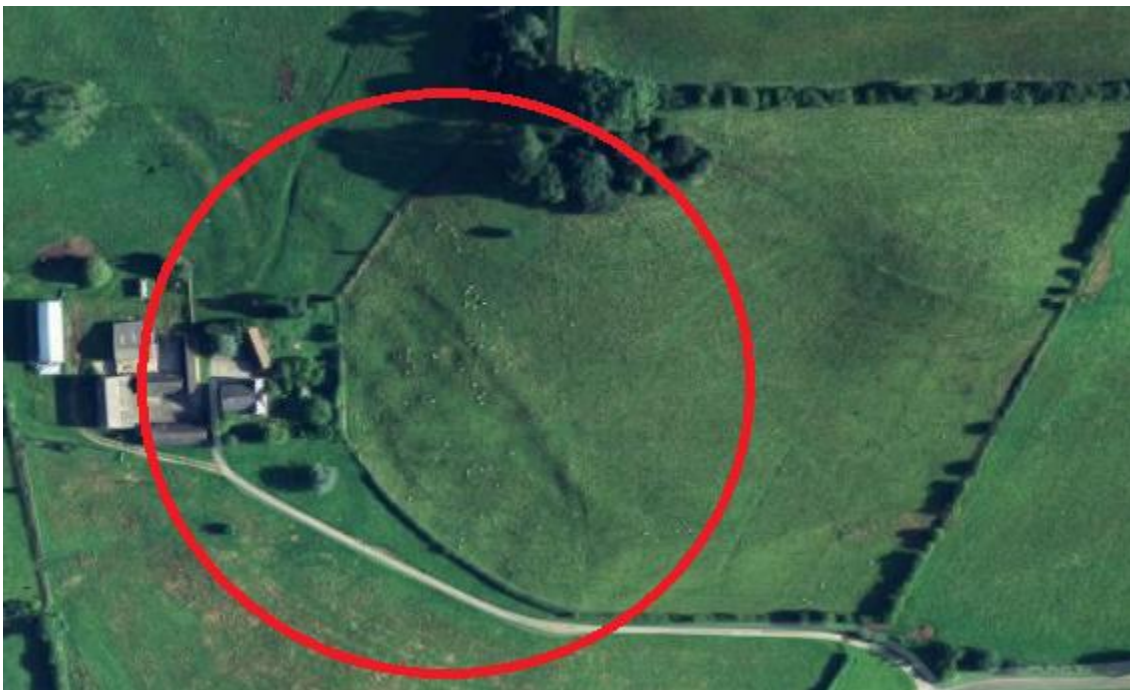


Figure 5-34: Possible landslide area

However, there is a clear landslide for all runs except for the initial run, which is shown only as a couple of cells and therefore not visible, but within the small black circles. The left small circle is near Whinny Oaks Covert. When checking google maps the following figure is revealed on the next page. The areas in red clearly show landslides:



Figure 5-35: Landslide at Whinny Oaks Covert

This landslide is completely missed when running the model on initial values. Therefore it is worthwhile to investigate this cell more thoroughly. The cell with coordinates 54.15625 (latitude) and -1.12090 (Longitude) is taken for this. This cell is exactly in the upper circle of figure 5.33.

The cell is 140.61 m above MSL, has a slope angle of 17.11° and a shear zone depth of 1.83m. The results for each run, starting with the historic run on initial values, historic run on Ksat \* 25, and all the forecast runs are shown in the figures 4.36 - 4.45, showing the safety factor [-] on the left hand side, and the water levels [m] on the right hand side of the figures (time period being 10-01-2012 until 31-12-2012).

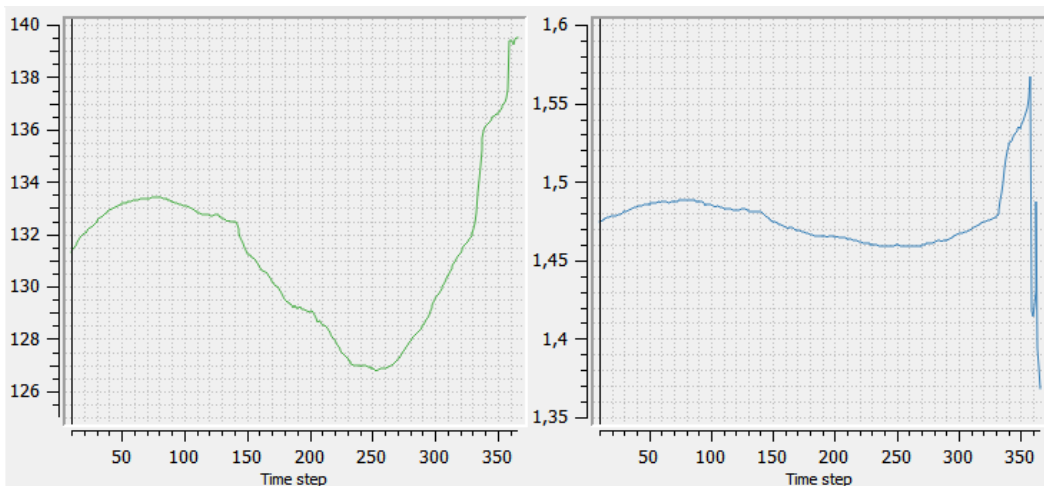


Figure 5-36: Safety factor [-] and water level [m] for the historic run on initial values

In the last 7 time steps, the water level is finally in the shear zone (only time step), being at maximum 1.09 m below surface. The shear zone is 1.82 m below surface, meaning the shear zone is ~40% fully saturated and ~60% unsaturated. The safety factor, although dropping as can be seen in figure 5.34 is still above 1 with the lowest value being 1.368 in the last time step. For this cell to become unstable the percentage of fully saturated shear zone needs to be increased, because the slope angle is not steep enough to cause instability at lower saturation levels.



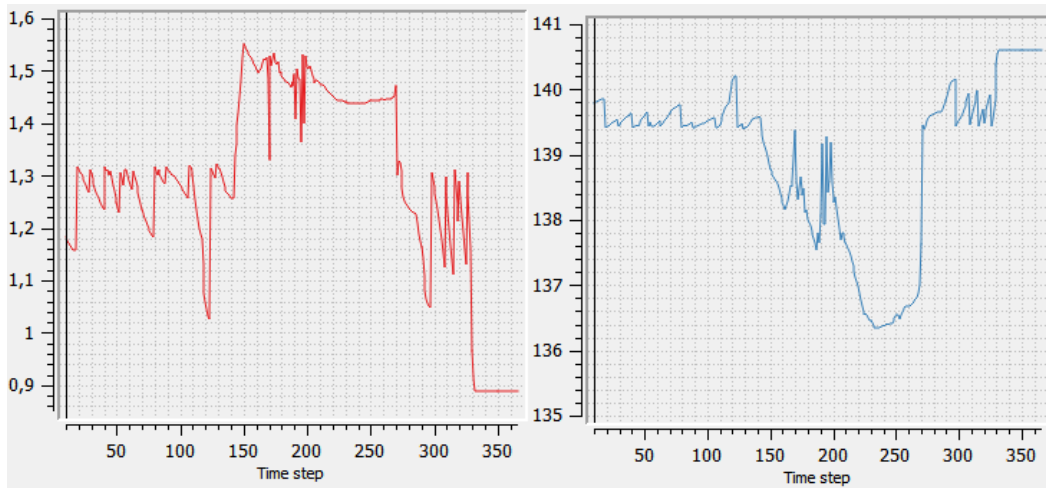


Figure 5-37: Safety factor [-] and water level [m] for the historic run on  $K_{sat} * 25$  and porosity  $* 1.4$

In the  $K_{sat} * 25$  run, the landslide becomes unstable in 36 time steps (26-11-2012 until 31-12-2012). On 25-11-2012, the water level is 76 cm below surface or 1.07 m above the shear zone. This means the shear zone is 58 % fully saturated, with a safety factor of 1.159, still being stable. On 27-11-2012 the water level is 0.22 m below surface, meaning the shear zone is 88% fully saturated. The cell becomes unstable now, with a safety factor of 0.965. After 27-11-2012, the water level reaches the surface (full saturation) and the safety factor drops to a minimum of 0.889.

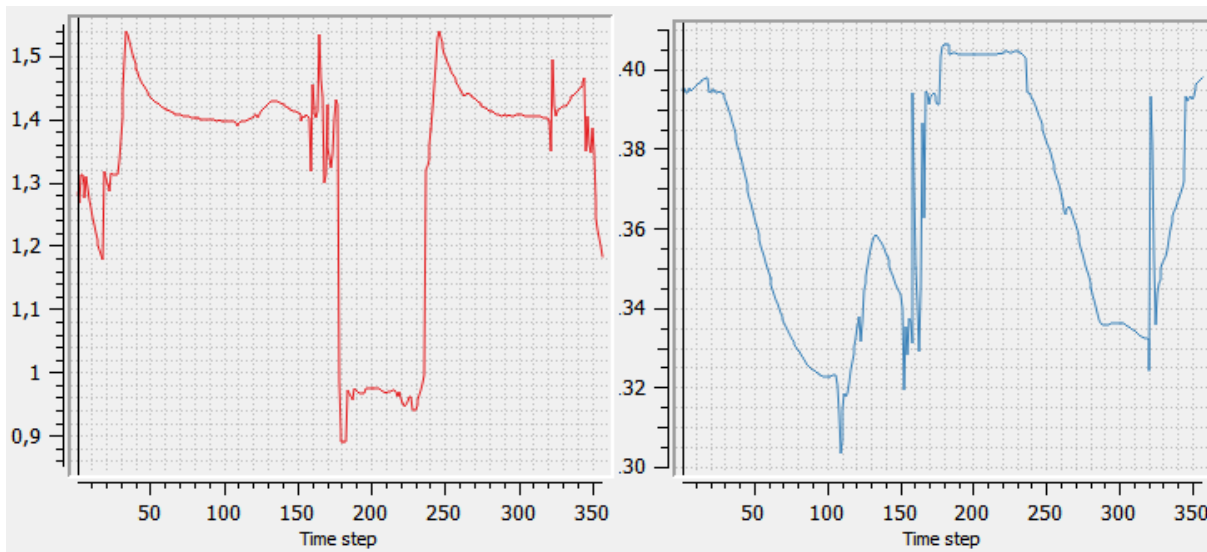


Figure 5-38: Safety factor [-] and water level [m] for the forecast day 0

The forecast run of day 0 shows a completely different picture, which could be expected given the different spatial distribution of rain and evaporation patterns. The cell becomes unstable between 5-7-2012 and 1-9-2012, but never reaches the lowest safety factor due to never being fully saturated.

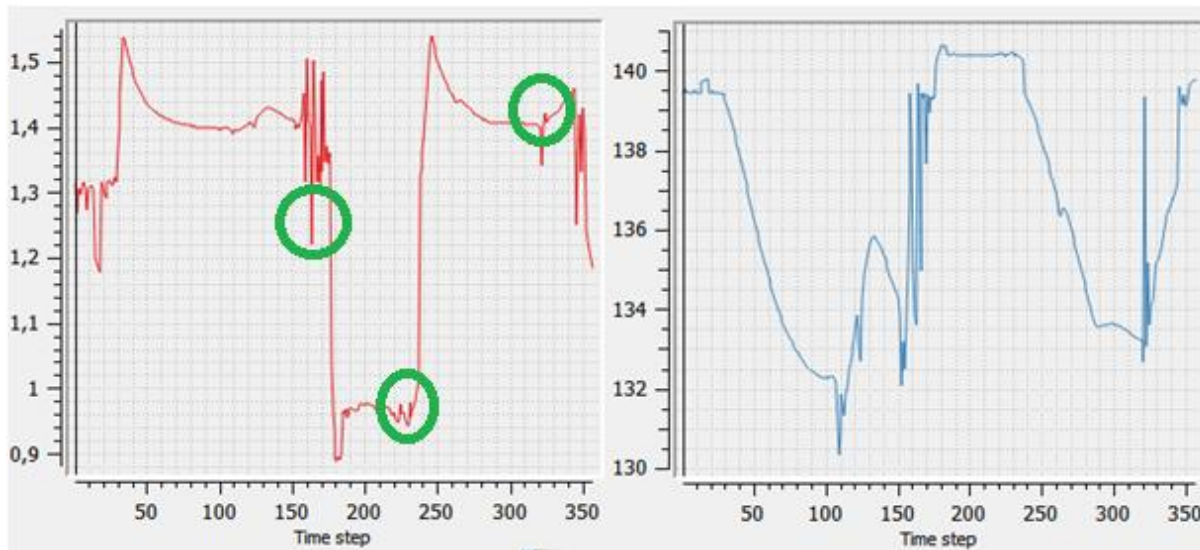


Figure 5-39: Safety factor [-] and water level [m] for the forecast day 1

The forecast run of day 1 interestingly shows a very similar picture to the day 0 forecast. The cell also becomes continuously unstable between 5-7-2012 and 1-9-2012, with minor differences in values. The large increase in safety factor around time step 300 seems to have been missed. There are some minor differences in values around the lowest circle, and there is a drop in safety factor which was not predicted in the day 0 forecast (left circle). All in all, the pattern is similar.

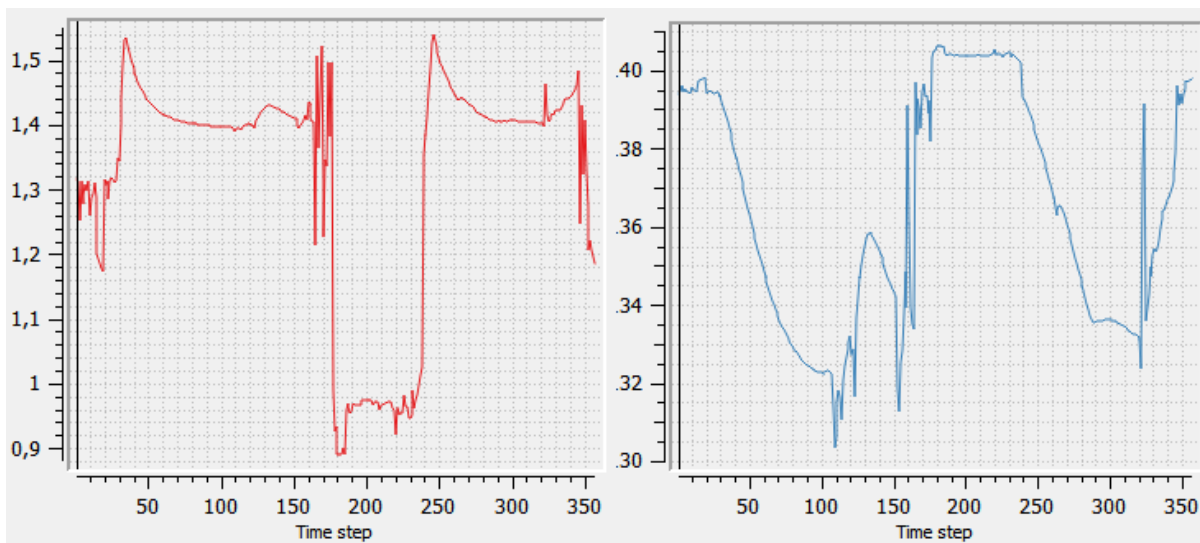


Figure 5-40: Safety factor [-] and water level [m] for the forecast day 2

The forecast day 2 run still shows quite a similar pattern, although slight deviations start to occur. First of all, the period of instability has slightly increased with one day, the cell now being unstable from 4-7-2012 until 1-9-2012. Also, safety factors and water levels start to differ slightly more. For instance, on 16-8-2012 (around 225 in the graph) and 28-8-2012 (236 in the graph) the following safety factors and water levels were reported:

Table 5-15: Water levels and safety factors for day 0, 1, and 2

	Water level [m]	SF [-]	Water level [m]	SF [-]
<b>Day 0</b>	140.416	0.957	140.416	0.957
<b>Day 1</b>	140.398	0.963	140.418	0.956
<b>Day 2</b>	140.521	0.920	140.322	0.990

On other days the levels are closer to each other, but still slightly deviating.

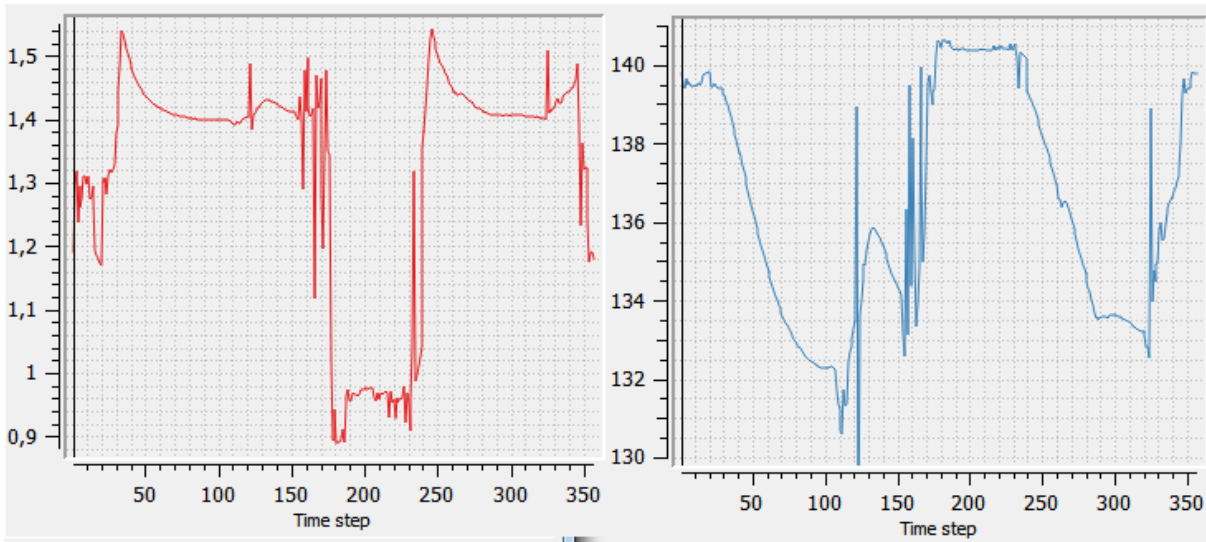


Figure 5-41: Safety factor [-] and water level [m] for the forecast day 3

At the day 3 forecast run, the deviation becomes more apparent. The unstable period is also in between 5-7-2012 until 1-9-2012, but not continuous anymore. On the 29<sup>th</sup> and 30<sup>th</sup> of august, (time steps 233 and 234), the slope becomes stable again, showing respective safety factors of 1.038 and 1.319. This seems to coincide with the patterns of day 0, day 1 and day 2. The overall patterning is still reasonably close to day 0, although it can be seen that more deviations start to occur. There start to be more sudden jumps in safety factors (less smooth transitions)

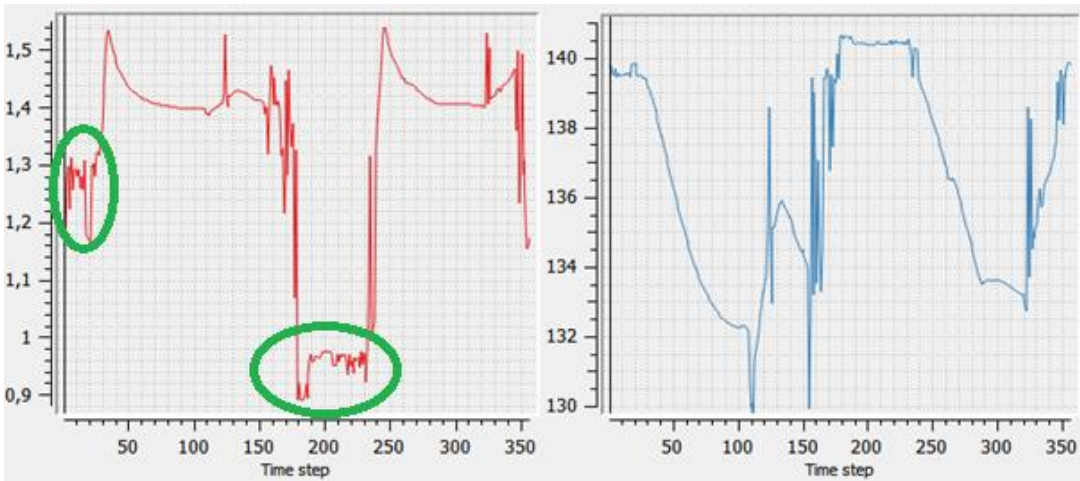


Figure 5-42: Safety factor [-] and water level [m] for the forecast day 4

The deviations start to become larger again for the day 4 run. First of all, the unstable period has decreased by four days compared to day 0, starting at 6-7-2012 and ending at 28-8-2012. The sudden increase in safety factor on the 28<sup>th</sup> and 29<sup>th</sup> of August can be seen as well. Also on other days sharper differences can be seen: on time steps 184-188 (11-7-2012 until 15-7-2012) the differences in safety factor are quite substantial: 0.971 (day 0) versus 0.891 (day 4), 0.966 versus 0.898, 0.961 versus 0.919, 0.957 versus 0.893, and 0.956 versus 0.934. After this period, levels become quite similar again. The pattern of less smooth transitions is also visible here. Also, the pattern of instability starts to deviate quite clear from the day 0 pattern, as seen in both the circles.

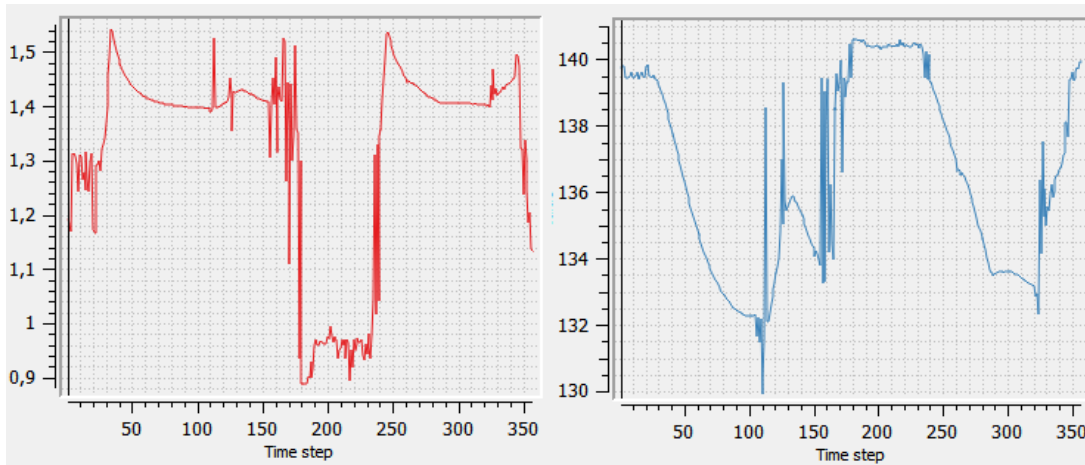


Figure 5-43: Safety factor [-] and water level [m] for the forecast day 5

For the day 5 forecast run there are more pronounced differences again. The timeframe of instability is between 5-7-2012 until 30-8-2012. However, on 6-7-2012, the safety factor suddenly increases to 1.300, meaning the cell has become stable again, which did not occur previously. Also, the jump in safety factor towards stability occurs with a slight delay. On time steps 235 (31-8-2012) and 236, the safety factor becomes stable showing 1.044 and 1.311 (0.978 on time step 234). Although in general patterning there is still resemblance to the day 0 forecast, on a more detailed level (forecasting of precise instability) it does not match day 0 at all anymore. For instance, on time step 186 (13-7-2012), the safety factor is 0.899, compared to 0.961 (day 0 and 1), 0.960 (day 2), 0.889 (day3), 0.919 (day 4), 0.901(day 6 and 7), and 0.914 (day 8).

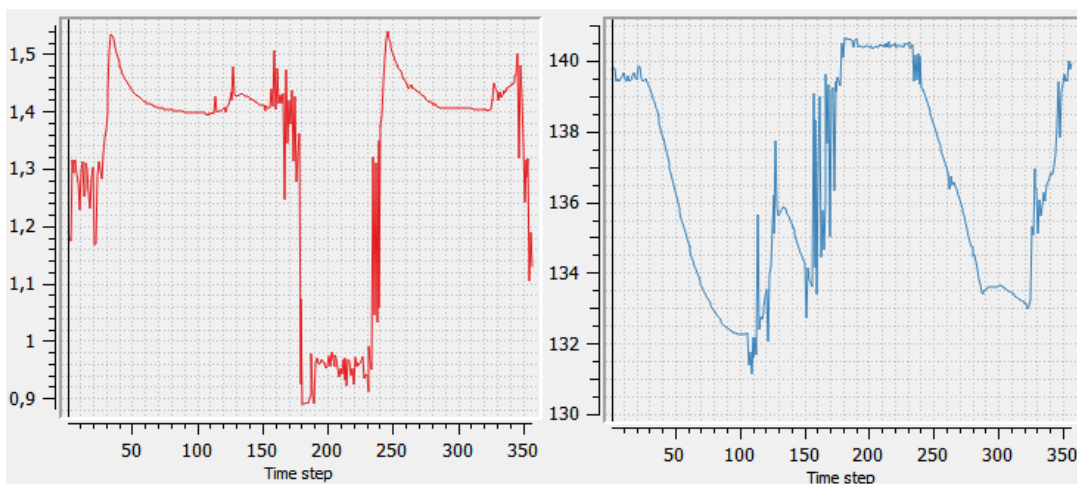


Figure 5-44: Safety factor [-] and water level [m] for the forecast day 6

The unstable period for the forecast day 6 run now starts on 6-7-2012 and ends on 31-8-2012. On 7-7-2012, the cell is stable with a safety factor of 1.073. This implies that the pattern seen from forecast day 5 has exactly shifted 1 day ahead. This will be elaborated later on.

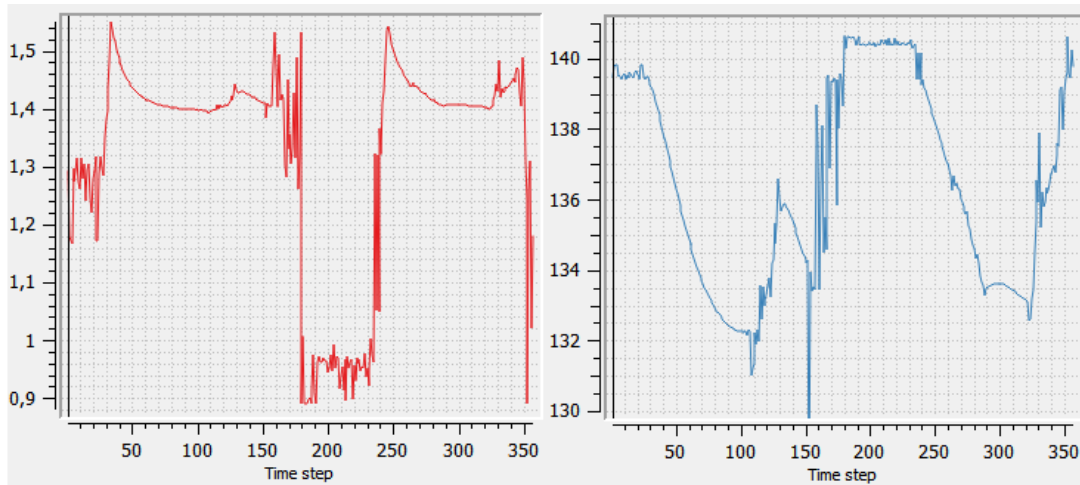


Figure 5-45: Safety factor [-] and water level [m] for the forecast day 7

The unstable period starts at 6-7-2012 and until 31-8-2012. On 8-7-2012, the cell becomes stable again with a safety factor of 1.006. On 29-8-2012 (233), the cell becomes stable with a safety factor of 1.003, unstable the two days after (234 + 235), and becomes stable again at 1-9-2012, with a safety factor of 1.323.

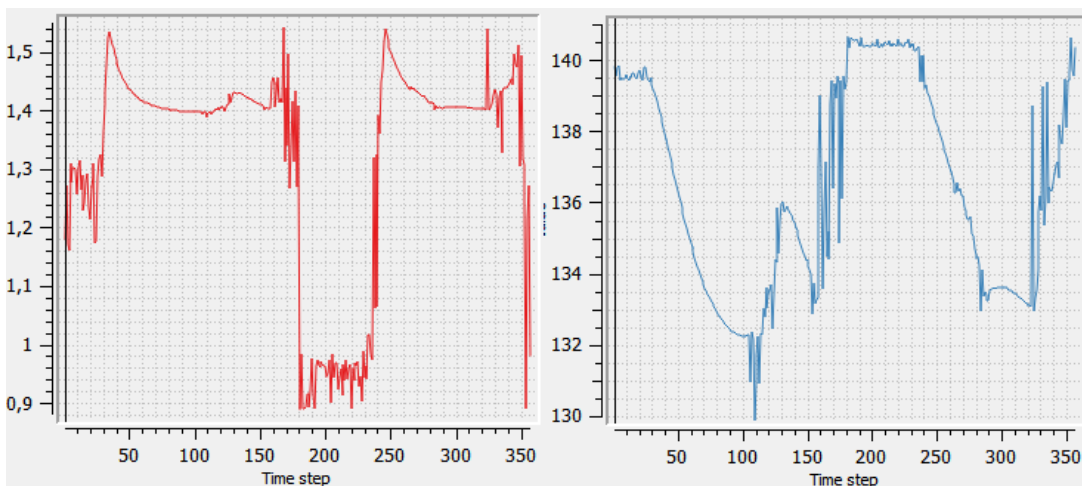


Figure 5-46: Safety factor [-] and water level [m] for the forecast day 8

The unstable period starts at 8-7-2012 and ends at 2-9-2012. It is continuous until 29-8-2012 and 30-8-2012 (233 + 234), where the safety factor on both dates becomes 1.016. On 31-8-2012 and 1-9-2012 the cells become unstable again, after which the safety factor increases to 1.321 on 2-9-2012 (237). There has been a shift again with regards to the end and the beginning of the unstable period. However, the cell becoming stable on 6-7-2012 (day 5), 7-7-2012 (day 6) and 8-7-2012 (day 7) did not occur on 9-7-2012, but this can be explained because the increase in safety factor decreased gradually: 1.300 (day 5), 1.073 (day 6), 1.003 (day 7), 0.985 (day 8).

In general, some phase shift seems to occur. This is demonstrated in figure 5.45 showing the forecast days (1=0, 2=1 etc.) from time steps 236-241 (1 until 6-9 2012):

1	0.996242	1.32134	1.32486	1.33565	1.35615	1.38246
2	0.996654	1.01098	1.32473	1.33509	1.35984	1.38088
3	0.996289	1.01126	1.02646	1.33331	1.3523	1.38548
4	0.997968	1.01099	1.02616	1.04286	1.35314	1.38152
5	1.00081	1.01369	1.02597	1.31904	1.32551	1.37803
6	1.3115	1.0155	1.33064	1.04177	1.34422	1.35618
7	1.04447	1.30929	1.03165	1.34789	1.05801	1.37535
8	1.32337	1.05059	1.319	1.04762	1.36593	1.32357
9	0.972809	1.32091	1.06139	1.32517	1.06378	1.39007

Figure 5-47: Phase shifts

The green circles show a clear pattern shifting occurring and moving one day further in time. This is the so called phase error described in chapter 4.3. It is likely that these are phase errors in the forecast, since forecast day 5 to 9 showed some similar phase error at the start of the unstable period. The red circles show a possible example of this too, however these could also be coincidental because a) it stops after time step 239 and b) all other values of forecast day 1, day 2, day 3 are quite close to each other. The distribution of rainfall is different for all days, as shown in chapter 4.3, which can lead to day to day differences. All in all, the correct timeframe of instability can be predicted 3 days ahead. More factual forecasts towards day zero are seen 2 days ahead, and certainly 1 day ahead. This could lead to the conclusion that the 3 day forecast (72-96 hours) can be used for pattern studying, and the 48-72, 24-48, and 0-24 hours forecasts for more detailed instability forecasting. However, this is based on studying one cell. Therefore another cell will be studied: The cell of Hollin Hill which proved to be unstable for the forecast run, with coordinates 54.1103 -0.9616, already shown in figure 5.23. This cell has a height of 70.34 m, a slope angle of 18.23°, and a shear plane depth of 1.63 m. Figures 4.46-4.51 show the historic initial run, the historic run on Ksat \* 25, and the forecast runs.

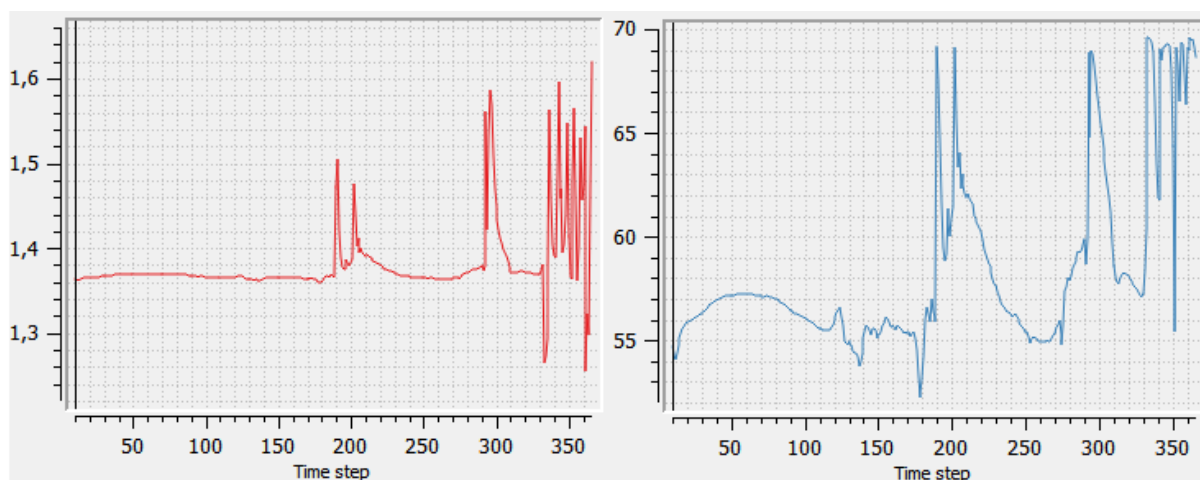


Figure 5-48: Safety factor [-] and water level [m] for the cell around Hollin Hill with the historic run on initial values

As can be seen, the safety factor does not really correspond to the water level, because the water level does not reach the shear zone in most cases. The safety factor is then influenced by the degree of saturation of the soil. On

time step 353 (27-12-2012) the water level does reach the shear zone, the water level is then 0.70 m below the surface or 0.93 m above the shear zone, meaning the shear zone has a saturated zone of 57%. The safety factor subsequently drops from 1.544 in time step 351 to 1.254 in time step 35. This does not prove to be enough to cause instability however.

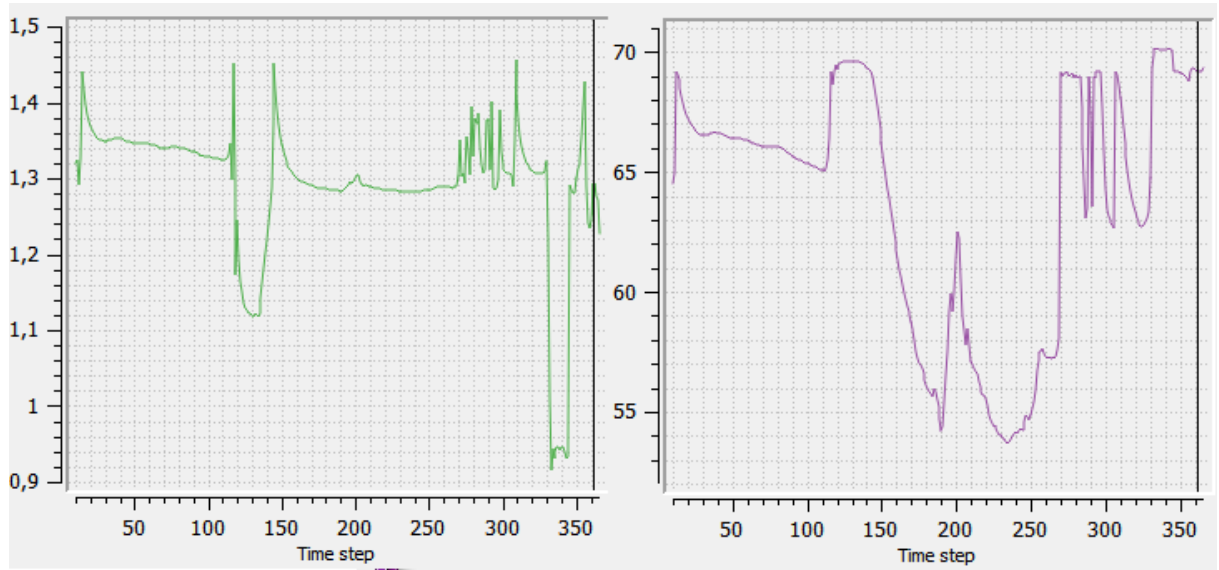


Figure 5-49: Safety factor [-] and water level [m] for the cell around Hollin Hill with the historic run on Ksat \* 25 and porosity \* 1.4

In the Ksat \* 25 run, the slope does reach instability. This happens at the time steps 324-336 (28-11-2012 until 10-12-2012). For instance, on time step 324 the water level is 0.16 m below surface, or 1.47 m above the shear zone. This means that the shear zone is 90% saturated, causing the safety factor to drop to 0.91. It is also interesting to see that the model is in some regard sensitive to the parameterization, because in the historic run meteorological forcing is the same, as well as most other parameters. The only parameters that differ are the hydraulic conductivity and the porosity, giving rise to different patterns. This could mean that calibration of parameters can also take place based on instability patterns, as done by Guimarães *et al.* (2003). Porosity of course does have a small influence on slope stability, since the weight of the soil can increase slightly with higher porosity and therefore the safety factor can drop slightly. For instance, if porosity is increased from 0.4 to 0.6, there is  $2 \text{ kg m}^{-3}$  of extra density for a saturated soil (around 10% of the total density).

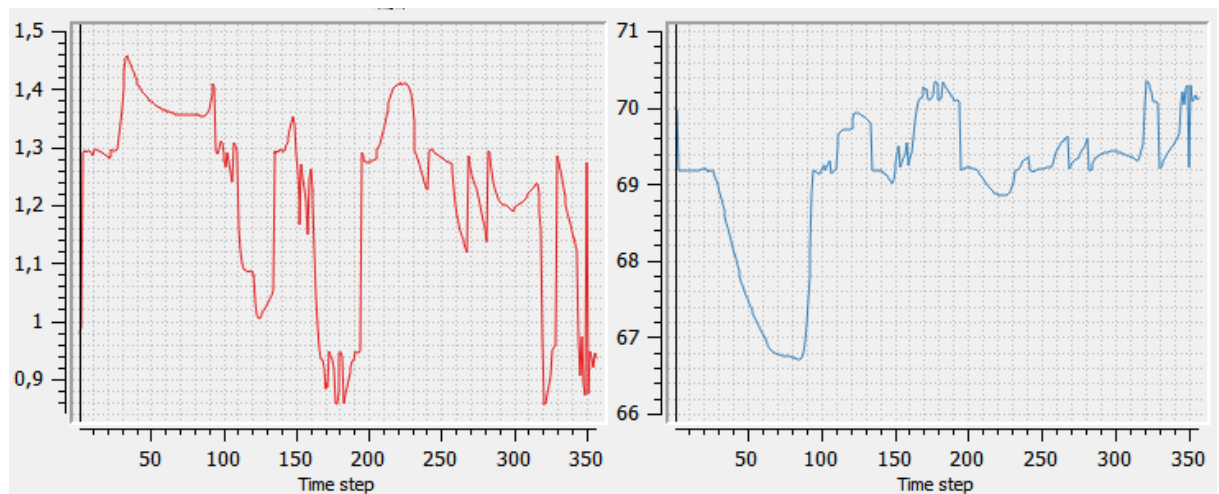


Figure 5-50: Safety factor [-] and water level [m] for the cell around Hollin Hill for the forecast day 0

The day 0 forecast shows the following pattern: instability at time steps 1 and 2 (10 and 11-1-2012), non-stop instability at 165-194 (22-6 until 21-7-2012), at 320-329 (24-11 until 3-12-2012), and 345-357 (18-12 until 31-12-2012), being interrupted at 350 (24-11-2012).

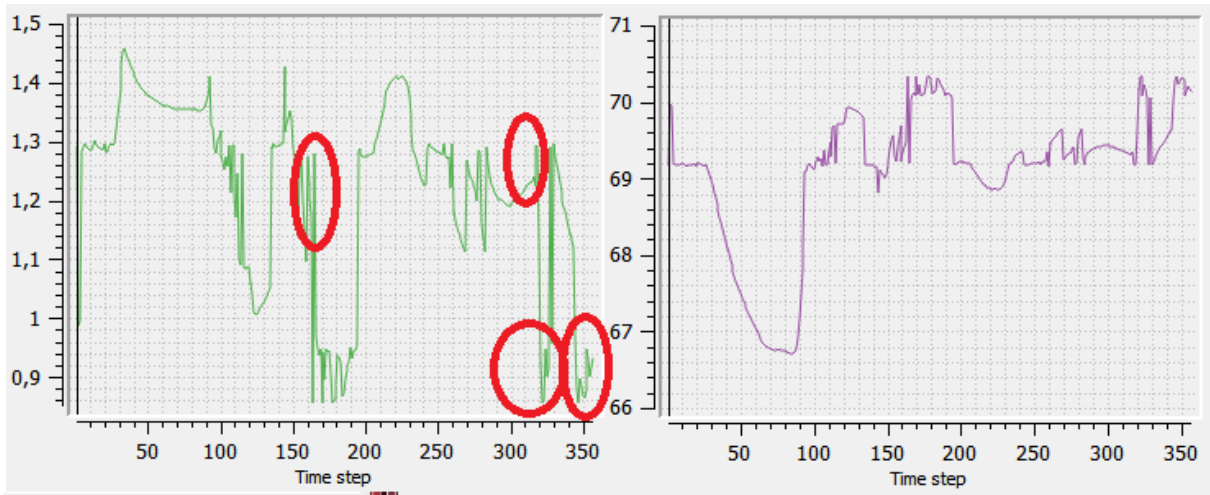


Figure 5-51: Safety factor [-] and water level [m] for the cell around Hollin Hill for the forecast day 1

The day 1 forecast is more or less comparable to day 0, but there are some differences. Instead of time step 1 and 2, time step 2 and 3 are unstable, although in all cases instability is barely reached with safety factors  $\sim 0.99$ . The model is quite sensitive to fluctuations in water levels, so this is not a major miss. The other instability period is between time steps 164-194 (21-6 until 21-7-2012), being interrupted at 165 (22-6). The safety factor rises to 1.280 suddenly, compared to 0.994 for the day 0 run (left circle). When checking the catchment averaged precipitation, it shows that for the day 0 forecast the forecasted rainfall amounted to 8.37 mm, compared to 0.48 mm for day 1. The next unstable period is between time steps 321-329 (25-11 until 3-12-2012), with an interruption at time step 328. The safety factor suddenly rises to 1.291 then (upper right circle). The lower left circle shows a small deviation from day 0 as well. The last unstable period is also 345-357, but unlike on day 0, uninterrupted (lower right circle).

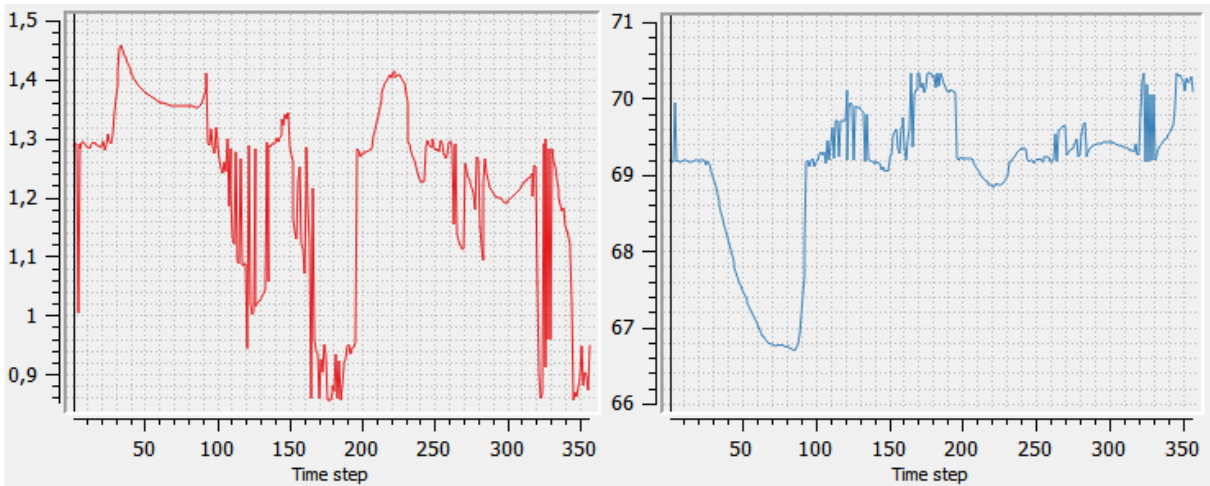


Figure 5-52: Safety factor [-] and water level [m] for the cell around Hollin Hill for the forecast day 2



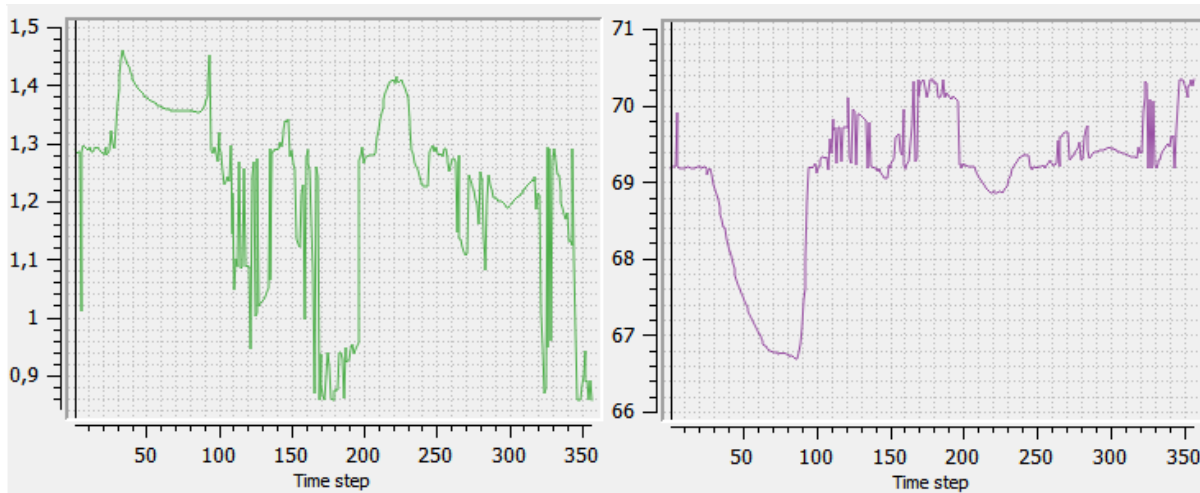


Figure 5-53: Safety factor [-] and water level [m] for the cell around Hollin Hill for the forecast day 3

Both forecast day 2 and day 3 runs miss the instability at the initial time steps. At time step 121 (9-5-2012), the day 2 run shows an instability with a safety factor of 0.944, compared to 1.078 (day 0), 1.058 (day 1), and 1.086 (day 3). On time step 122, day 3 shows an instability with a safety factor of 0.945, compared to 1.033 (day 0), 1.039 (day 1), and 1.287 (day 2). Also on other days, values are more deviant. There seems to be a phase shift occurring again, if Table 5.15 is checked:

Table 5-16: errors in the forecast

	Time step	Day 0	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8
SF [-]	119	1.087	1.087	1.087	1.089	1.085	1.089	1.087	1.081	1.172
SF [-]	120	1.087	1.086	1.087	1.089	1.085	1.089	1.216	1.085	1.079
SF [-]	121	1.078	1.058	0.944	1.086	1.086	1.287	1.063	1.200	1.080
SF [-]	122	1.033	1.039	1.287	0.945	1.081	1.082	1.257	1.291	1.181
SF [-]	123	1.001	1.001	1.284	1.007	0.945	1.070	1.076	1.220	1.121
SF [-]	124	1.004	1.005	1.000	1.266	1.285	0.946	1.060	1.070	1.269
SF [-]	125	1.005	1.006	1.005	1.001	1.241	1.248	0.946	1.049	1.060
SF [-]	126	1.001	1.001	1.284	1.007	1.005	1.220	1.277	0.938	1.035
SF [-]	127	1.013	1.014	1.014	1.271	1.009	1.009	1.203	1.233	0.932
SF [-]	128	1.018	1.018	1.019	1.017	1.261	1.011	1.014	1.190	1.265
SF [-]	129	1.022	1.022	1.022	1.023	1.020	1.251	1.013	1.020	1.181
SF [-]	130	1.027	1.026	1.026	1.028	1.034	1.028	1.033	1.190	1.265

It is clearly demonstrated that at day 2 the forecasts become significantly different from the day 0 and 1 forecasts, propagating further into the future. This is caused by deviations in the precipitation. For instance, on time step 121 (9-5-2015) the day 0 forecast amounted to 4.03 mm, the day 1 forecast to 4.08 mm and the day 2 forecast to 6.32 mm. This leads to the conclusion that the forecasts after day 1 are becoming increasingly reliable for slope stability modeling.

There is room for improvement however. First of all, these results are based on non-corrected forecasts. One way to improve forecasts is by correcting monthly mean and extremes. Usually this is done based on correcting observed values to forecasted values, however the historic run did not produce accurate rainfall patterns in the catchment, since the rain gauges were mainly outside of the catchment. Alternatively, the forecasts can be corrected to the day 0 forecast. Kumik *et al.* (2012) describe a method of correcting mean and extremes in forecasted precipitation,

based on the assumption that (forecasted) precipitation follows a gamma distribution. Only in 1.5% of the grid cells their correction method failed for all 8 forecast models they used, and in 4.5% for six models. Radar data can also be used to correct the forecast. In the UK, there is a 1 km rain-rate observation system called NIMROD, which period of record is 2002-present. It is a short term weather forecast system which also observes rainfall rate every 5 minutes. The observed precipitation rate is based on processed radar (15 radars in the UK) and satellite data, surface reports, and Numerical Weather Prediction (NWP) fields (Centre for Environmental Data Archival, 2015). These 5 minute rainfall-rates can be converted to daily rainfall rates of course. If proper rainfall data from rain gauges is present, the radar data itself can also be corrected. As Edwards (2006) states, the strength of rain gauges is determined by its ability to measure rainfall at a point and the weakness to measure the rainfall in between those points. The strength of a radar is that it provides information about spatial distribution of rainfall, the weakness is that it cannot give exact rainfall at a given point. In his study, he corrected radar data in two ways: correcting for mean difference between recorded gauge values and radar values and correcting for mean difference between IDW interpolated gauge values and radar values. He found the former method to have smaller differences and attributed this to an error of a mathematical nature embedded in the IDW method, because of the inherent difficulties to predict rainfall in areas without rainfall observation.

The slope stability results clearly show another aspect of the modeling assumption: the assumption that the water levels can fluctuate freely in the thick third layer proved somewhat erroneous. The water levels can indeed fluctuate freely, but the model is very quick to respond to changes in input. As soon as the precipitation levels drop, there can be a very large increase in safety factor compared to the previous day. This can also be seen in the water levels. For instance, in figure 5.47 the water level drops from 70 to 55 m in ~50 time steps. This is likely caused by the large transmissivity [ $L^2 T^{-1}$ ], being an indicator for the horizontal permeability, and the product of  $K_{sat}$  [ $L T^{-1}$ ] times the thickness [ $L$ ]. In reality, either separate confined water tables could be expected, and / or lower  $K_{sat}$  values for the bedrock. The quick transport of water is an explanation for why Hollin Hill does not reach instability in the model. Other studies done with STARWARS, for instance Kuriakose (2010) and van Beek (2002) only used the soil depth based on interpolated field measures as input for the layers (order of magnitude: 0-2m), meaning the 4<sup>th</sup> infinite layer is bedrock. Also the way the shear plane was distributed showed some room for discussion. The values differed quite significantly on a cell to cell basis and were based on one value of Hollin Hill, and then interpolated. To improve this, more soil data could be gathered and soil depth could be estimated based on the method described by Penižek & Boruvka (2006) or Lanni *et al.* (2012), see chapter 5.1. The shear plane can then be estimated as being the soil depth. The influence of parameters like hydraulic conductivity and porosity on the slope stability therefore remains somewhat hard to tell, because it is possible that with a more correct assumption on shear plane depth and soil depth their effects would be less profound. Also, the assumption about the recession coefficient and hydraulic conductivity of the 4<sup>th</sup> layer needed improvement, as is shown in chapters 5.2 and 5.3.

With another regard to the parameterization, Hwang *et al.* (2015) state that compared to the DEM, the spatial precision and knowledge of distribution about parameters such as (root) cohesion, friction angle, soil depth, hydraulic conductivity and porosity is low. As a result, landslide hazard assessment often becomes dependent on topographic features such as slope angle, and derivatives such as drainage area, combined with precipitation patterns. This is in line with the findings of Ulmer *et al.* (2009), who state that the increased availability of high resolution DEM calls for generalization of parameter results. However, they do state that high resolution DEM such as a 2m<sup>2</sup> DEM which they used could violate the main assumption behind many slope stability models (also PROBSTAB), the assumption of the infinite slope (slope width  $\gg$  soil depth). Also, Pack *et al.* (2005), the authors of the slope stability model SINMAP (an infinite slope model calculating stability based on pore pressures obtained from a steady-state DEM-input based hydrological model), state that their models purpose is to map more general stability classes from which scrutiny regions can be identified. This is in line with the study of Intrieri *et al.* (2013) described in chapter 3.1, stating that regional modeling serves the purpose of identifying regions where more

detailed assessment is needed. Scrutiny can be based on the ALARP principle of Wong *et al.* (1997), also described in chapter 3.1. Another approach with regard to regional parameterization is that of the Hydrologic Response Unit (HRU) from the commonly used Soil & Water Assessment Tool (SWAT) model, where the modeler can define HRU in a catchment for parameterization based on soil and land-use properties (Texas A&M University, sine anno). This approach can be simulated for further calibration of the parameters in PCRaster by making a manual HRU, combining a land-use map with the soil type map. Each field can then get an own adjustment for parameters, based on land-use type. This is similar to another model used in slope instability, SINMAP (Pack *et al.*, 2005). This model allows for calibration regions in which parameters can be uniformly calibrated within pre-defined upper and lower limits based on soil and land use type. Another well known slope stability model, SHALSTAB, is effectively parameter free based on the assumption that the ratio of effective precipitation over soil transmissivity ( $q/T$ ) can be derived from the drainage area per width of subsurface flow and the local slope, both of which are DEM derivatives. It is therefore relatively easy to use for validation, because the model can be rejected when the predicted instability patterns do not match with the observed ones (Dietrich *et al.*, 1998). Rosso *et al.* (2006) state this model has been successful in various instances, because of the dominant role of topographic control in shallow landsliding. However, the model does not allow transient movement of water because of the steady-state assumption and only considers soil density, angle of internal friction, and slope angle to calculate slope instability. In addition, it neglects soil moisture content in the unsaturated zone. Therefore, Rosso *et al.* came up with improvements in their approach. The approach is based on altering the equation for the factor of safety and algebraic manipulation to derive the equation for calculating the critical relative saturated soil depth  $\omega_r$  [-]. This procedure is shown in appendix 4, the equation for calculating  $\omega_r$  is shown underneath:

$$\omega_r = \frac{(G_s + e * Sr) * \left(1 - \frac{\tan \phi'}{\tan \theta}\right)}{1 + e - e * (1 - Sr) * \left(1 - \frac{\tan \phi'}{\tan \theta}\right)}$$

In which:

$\omega_r$  = Critical relative soil saturated depth [-]

$G_s$  = specific gravity of solids [-], being defined by the ratio of the saturated soil density over the density of water  $\gamma_{sat} / \gamma_w$

$Sr$  = The average degree of saturation  $V_w / V_v$  [-] (In which  $V_w$  is the volume of water and  $V_v$  the volume of voids)

$e$  = The average void ratio above the groundwater table  $V_v / V_s$  [-] (In which  $V_s$  is the volume of solids)

$\phi'$  = angle of internal friction [°]

$\theta$  = slope angle [°]

It has to be noted that this approach is based on cohesionless soils. This approach can also be used to make instability classes, which was not done in this study. Such an approach can provide an insight into which slopes are more susceptible to fail. Real landslides commonly occur when a threshold is exceeded for an amount of time (Le *et al.*, 2014). Le *et al.* (2014) in their study about a catchment in Vietnam close to the borders of both Laos and Cambodia, therefore used critical relative soil saturation values and classified the landslides in three classes (which they did not further specify), but an example of such classes is conditionally stable, conditionally unstable and partly (in)stable. Landslide potential was only defined for stable / unstable cells where critical relative soil saturated depth ( $\omega_r$ ) ranged between 0 and 1. Those (clusters of) cells can be the intense scrutiny regions from Wong *et al.* Pack *et al.* (1998) used a specified six classes for instability:

Table 1. Stability Class Definitions

Condition	Class	Predicted State	Parameter Range	Possible Influence of Factors Not Modeled
$SI > 1.5$	1	Stable slope zone	Range cannot model instability	Significant destabilizing factors are required for instability
$1.5 > SI > 1.25$	2	Moderately stable zone	Range cannot model instability	Moderate destabilizing factors are required for instability
$1.25 > SI > 1.0$	3	Quasi-stable slope zone	Range cannot model instability	Minor destabilizing factors could lead to instability
$1.0 > SI > 0.5$	4	Lower threshold slope zone	Pessimistic half of range required for instability	Destabilizing factors are not required for instability
$0.5 > SI > 0.0$	5	Upper threshold slope zone	Optimistic half of range required for stability	Stabilizing factors may be responsible for stability
$0.0 > SI$	6	Defended slope zone	Range cannot model stability	Stabilizing factors are required for stability

Figure 5-54: Six classes of instability. From Pack *et al.* (1998).

The link between river hydrology and the slope stability remains somewhat unclear. Although the forecast runs showed a poorer NSE score, the pattern of instability was slightly better. This can be explained by better matching precipitation patterns for the first days of the forecast run, because the forecasted precipitation was actually in the catchment itself, whereas the historic precipitation was interpolated from outside the catchment. This possibly resulted in better slope stability forecasting, however this cannot be said conclusively because most unstable area was found in small patches of forest. This may not be coincidental, since all the stable area around it was used for agricultural purposes and / or living. There was more unstable area with the forecast runs around the landslide shown in figure 5.33, which was completely missed using initial values and only partially covered with the historic Ksat \* 25 run. Interestingly, to mention the aforementioned forecasts: the link between hydrology and forecast accuracy is not true for the day 2 forecast. This forecast had a higher NSE to the day 0 forecast than the day 1 forecast (this had mainly to do with the peak flow values), but showed poorer slope stability results. This can be explained by the following principle: the spatial distribution of the day 0 and day 1 precipitation pattern was very similar (although day 0 was slightly wetter), leading to very similar slope stability results. The spatial distribution of the day 2 forecast precipitation was different to the day 0 forecast, which seemed to have less impact on the catchment scale, but more impact on the slope scale. From day 2 on onwards errors started to propagate. This can be seen when showing the annual P in mm for the day 2 forecast and comparing them to figures 5.20 and 5.21:

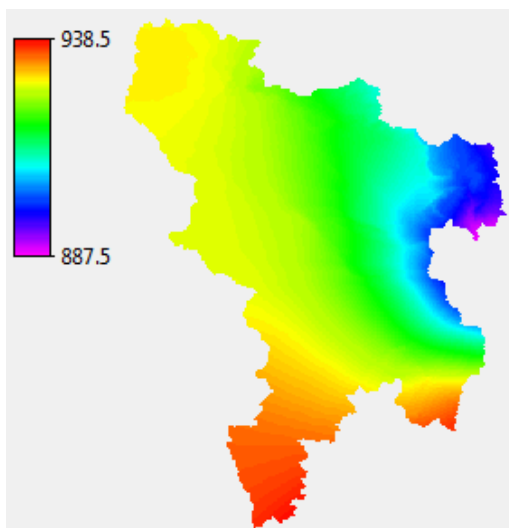


Figure 5-55: Annual P [mm] for the day 2 forecast

With regards to Hollin Hill, the model was lacking. Both the historic Ksat \* 25 run and the forecast runs only predicted instability at one cell of the entire Hollin Hill landslide, comprising of 16 cells. There is no data to check if the timing of instability is correct (the forecast runs predict more instability). The study does show that the slope stability results are a) dependent on meteorological forcing, as mentioned before but also on b) parameterization. The differences in slope stability results between the historic runs (with the same forcing, but different Ksat and porosity) in slope stability results was quite drastically, quantitative but also in pattern. The difference between the historic runs and the forecast run with regards to slope stability was also significant, with the forecast runs showing a different forcing than the historic Ksat \* 25 and porosity \* 1.4 run, and different forcing and parameterization compared to the historic initial run. However, with other factors also being highly uncertain and erroneous, such as the depth of the third layer, the recession period and hydraulic conductivity of the 4<sup>th</sup> layer, and the shear plane depth, it is uncertain to what extent the parameterization in terms of hydraulic conductivity and porosity do / did influence the model. However, in various studies (Ulmer *et al.*, 2009 ; Mulungu & Munishi, 2007 ; Rosso *et al.*, 2006 ; Dietrich *et al.*, 1998) state that either topography or meteorology is dominating over parameterization.

## 5.5. Recommendations

Based on this study and literature, the following recommendations can be given with regards to the modeling part, how to set up a model for a new area and how to set-up a forecast system for an environment like FEWS.

Mulungu & Munishi (2007), in their study about the Simiyu River Catchment, discovered that the use of higher quality data did not improve their results in terms of river flow simulation. Their findings were similar with the results from this study: inaccurate rainfall distribution led to measured peak flows not being represented or being over / under represented. It is even stated in some studies that rainfall distribution has severe impacts on the spatial distribution of soil moisture and production of runoff (Kustas & Goodrich, 1994 ; Schmutge *et al.*, 1994 ; Michaud & Sorooshian, 1994). Spatial variability in soils and vegetation are said to possess a second-order modifying effect on the spatial variability from the assumed rainfall distribution. Therefore Mulungu & Munishi (2007), conclude that more attention should be paid to the spatial distribution of rainfall data than spatial land data. Obtaining a high resolution DEM and placing rainfall meters in areas of interest for regional landslide modeling is certainly more attainable than obtaining high resolution soil data. The study revealed that precipitation levels in Yorkshire can differ up to 900 mm on an annual basis, showing the importance of proper rain measurements. These historic rain measurements can also serve to correct and improve the forecasts, as done by Kurnik *et al.* (2012). It is therefore recommended that, when setting up a regional forecasting environment for a new region, for instance in FEWS the following procedure can be followed:

- Assess whether historical rainfall records are available for calibration of the forecast data (removing bias with a correction factor), and how many measurements there are within the catchment itself;
- If deemed insufficient, simple tipping bucket rain gauges can be placed on various points in the catchment, preferably the higher places susceptible to slope failure. These are buckets that show response to rain, tipping every time a 1 mm increment is reached and sending an electronic signal to the control station. Usually, results are quite accurate. However, wind may deflect precipitation away from the bucket under consideration. Wind speeds of 63 km h<sup>-1</sup> or larger have been known to cause a loss of ~50% of precipitation towards the rain gauge. Plants, insects and animals are also known to cause troubles, next to electronic malfunctions. Finally, during intense rainstorms, buckets have difficulties keeping up, causing precipitation to be lost during the 0.5 second it tips (it is emptied during the tip and then fills up again). An error of ~20% should be expected for convective rainfall. However, in general they provide very reasonable exposure (Edwards, 2006).
- Next to rain gauges, the 1 km radar data from NIMROD can be used in order to calibrate the forecast data, for instance if historical data from rain gauges is not available yet. Finally, and this is especially handy for FEWS, on days of intense scrutiny (for instance when a large amount of precipitation is expected) the model can be run with a 3-hour forecast. The Met Office has 3-hour site specific forecast data in XML format (the input format for FEWS) for over 5.000 sites in the UK, with regards to both rainfall and temperature (Met Office, 2014). These 3-hour forecasts, which you can expect to be more reliable, can also be summated of course and the difference between the 3-hour and daily forecast can be removed by removing the bias with a correction factor.

Data acquisition as is the next area of improvement. With regards to the slope stability and hydrological assessment, the DEM is very important as show in figure 5.14 and mentioned in multiple studies (Ulmer *et al.*, 2009 ; Mulungu & Munishi, 2007 ; Rosso *et al.*, 2006 ; Dietrich *et al.*, 1998). Therefore it is recommended to run a model with a high resolution DEM. In this case a 25 \* 25 m resolution was used, but a 10 \* 10 m (original resolution) or even higher resolution would have been more preferable for more accurate slope representation, mentioned by Anderson *et al*

(2011) and Guimarães *et al.* (2003). However, this is only feasible with high-end computers and as Ulmer *et al.* (2009) state, it could violate the assumption of the infinite slope. It also needs to be taken into consideration that the quality of many other important data, such as soil data, land use data etc. is not on par with this resolution. Land use data can be improved by making use of LandSat TM 30 \* 30 m land use data, which is done in a study by Mulungu & Munishi (2007).

For the parameterization, it can be considered to first run the model with less assumptions about the spatial distribution such as the random field assumption in this study (as adding details can decrease or increase uncertainty), especially if nothing is known about this distribution such as in this study. For instance, it can be considered to use LandIS soil map instead of the SGDBE. In this study the soil parameters were derived from a pedotransfer script, based on Organic Carbon content of the soil polygon on a 1 km<sup>2</sup> resolution. The LandIS soil map uses this same resolution, but also has an Organic Carbon content correction for different land uses derived from the Corine Landcover map (Cranfield University, 2015a). This is comparable to the HRU approach from SWAT or the approach of calibration regions in SINMAP. Another option is to make use of an already parameterized polygon map from LandIS on a 1 km<sup>2</sup> resolution (Cranfield University, 2015b). Calibration for the combined land and soil regions can then take place, based on (if available) river records and resulting NSE and mapped landslide scars. The latter is a method used in many studies (Gorsevski *et al.*, 2015 ; Le *et al.*, 2014 ; Rosso *et al.*, 2006 ; Guimarães *et al.*, 2003 ; Pack *et al.*, 1998) and is actually more preferable: not all catchments have daily river flow modeling records, and it allows for greater freedom if one is only interested in parts of a catchment for regional assessment. Since PROBSTAB simulated many unstable areas to be present in forested areas, it is recommended to use light detection and ranging (LiDaR)-derived DEM for landslide mapping up front. According to a study by van den Eckhout *et al.* (2007) in the Flemish Ardennes LiDaR allowed for detection of old landslides under forested areas. In addition, LiDaR helped them detect 10 new landslides, correct boundaries of 11 out of the 77 total mapped landslides, and they were able to detect 70% of the mapped landslides from the field survey. Schulz (2007) discovered that four times more landslides were discovered using LiDaR than with Aerial Photo Investigation (API) or field mapping. The other advantage of LiDaR is that it allows for soil and land-use classification (Anderson & Croft, 2009 ; Liu, 2008). This can be an advantage if the modeler wants a more detailed assessment of a certain area and wants to increase the resolution of the model.

Calibration of recession period and the hydraulic conductivity of the 4<sup>th</sup> layer can also be recommended. As a quick test showed, by halving the recession period from a month to 15 days and by doubling the hydraulic conductivity of the 4<sup>th</sup> layer from 0.00075 m d<sup>-1</sup> to 0.0015 m d<sup>-1</sup>, the NSE from the historic run with Ksat \* 25 and porosity \* 1.4 improved from 0.21 to 0.33 and the problem of the model being too dry was also solved). The still relatively poor NSE score was attributed to inaccurate meteorological forcing. The effects of this improvement of NSE score on slope stability could unfortunately not be assessed, due to a lack of time.

Other improvements for the study are: improving recession of the model, soil depth and shear plane assumption. The assumption of freely fluctuating water levels proved to cause the water levels to drop to such levels that instability could never be reached for some cells. Also, shear plane varied too much on a cell to cell basis. Interpolating soil depths based on field samples, such as done by Penížek & Boruvka (2006), could help solve this problem. This approach is also followed by Kuriakose (2010) and van Beek (2002). Alternatively, if such samples are not available (yet), one can start with estimating soil depth from literature values.

Another recommendation is to make use of the classification per time step as described by Pack *et al.* (1998) in figure 5.54. This can easily be done in a set-up like FEWS, as FEWS allows for classification of threshold levels in between the minima and maxima, unlike PCRaster but also like ArcGIS. This allows for better appreciation of which sites are conditionally unstable and which sites can become unstable due to certain rainfall events. Classification approach was also used by Gorsevski *et al.* (2015), who used five susceptibility classes. The critical relative soil depth

approach from Le *et al.* (2014) and Rosso *et al.* (2006) also allows for better appreciation of which sites are more susceptible to fail.

For the region, Hollin Hill proves to be a good learning site. The site is actually monitored and therefore served as a validation for this study. However, in order to assess the model performance more accurately, data about displacements times, temperature and rainfall on site should be used. A meteorological station is present on site. These can all be used when zooming in as mentioned before, as well as soil samples collected from the site. In general, when zooming in on a certain area for more detailed assessment, one can also think of using piezometers for groundwater level calibration, soil moisture meters, and pore pressure meters (Dios *et al.*, 2010). Brocca *et al.* (2012) argue that antecedent soil moisture conditions defined by the Soil Water Index (SWI), which is the moisture condition of the soil at various depth and obtained from remote sensing and some equations based on a parameter called characteristic time length  $T$ , and Antecedent Precipitation Index (API) improved their landslide prediction rate from a rock landslide in central Italy. For instance, a landslide is not likely to occur with a significant rainstorm after a dry period, whereas a moderate rainstorm during a wet period can cause a landslide. Their regression coefficient  $r$  from the linear regression model improved from 0.4 when only assessing rainfall and to 0.68 when using API and to 0.85 when using SWI, when modeling the re-opening of cracks in this rock-slide, proving that in this case antecedent moisture conditions play a larger role than antecedent precipitation patterns. Even though this is an empirical method, the method improved their prediction rate.



## Appendix I

1:

$$pcrcalc\ x.map = lddcreate(elevation, outflowdepth, corevolume, corearea, catchmentprecipitation)$$

In which x can be replaced for the preferred name of the LDD PC Raster map, elevation is a .map file containing the DEM and the four variables after elevation are set to 1 \* 1031. This ensures that the pit is being filled up and no longer exists. The DEM.map file was created using GDAL, a standalone software package that can handle raster data, to convert it from an ArcGIS file format to a PC Raster .map file.

2:

$$pcrcalc\ relativedem.map = \frac{dem.map - mapminimum(dem.map)}{mapmaximum(dem.map) - mapminimum(dem.map)}$$

This command subtracts for any cell on the DEM the lowest height from the DEM and divides it by the maximum height minus the minimum height and writes the result to a new .map file called relativedem.map. From this calculation it can readily be seen that on the cell(s) with the maximum height, relative height becomes 1 and on the cell(s) with the lowest height relative height becomes 0. For cells between the minimum and the maximum height the relative height will range in between 0 and 1. Subsequently, from any cell where a river flows through relative height has been subtracted with 1. This has been done using the following PC Raster command:

$$pcrcalc\ burneddem.map = relativedem.map - if(river.map(boolean(1)),1,0)$$

This creates a new .map file called burneddem.map based on the following statement: if the river map equals the Boolean value 1 (meaning true so a river is flowing there) then 1 is subtracted from the relative DEM and if the river map equals Boolean 0 (meaning false so no river is flowing there) 0 is subtracted from the relative DEM. The general code for an if statement in PC Raster is if(expression, then, else). Finally, the new LDD can be created using the aforementioned command for creating a LDD map, this time with the burned DEM as input.

3:

First, the spatial coordinates of the outflow point were estimated in ArcGIS using the catchment delineation from the catchment map from CEH and the river grid of the BGS, the outflow point is where the lowest river point meets the catchment border. A point map was created in PC Raster using the coordinates of this outflow point. One way of doing this is creating a .txt file with the name of the outflow point (e.g. 1) in column 0, the x coordinate in column 1, and the y coordinate in column 2. Finally, using the following command:

$$pcrcalc\ outflowpoint.map = col2map\ outflowcoordinates.txt\ dem.map$$

This creates a point map with the spatial coordinates of the outflow point with the same spatial extent (cell size, rows, columns) as the DEM map.

4:

The coordinates of Hollin Hill were added in the same way as the technique described in 3, and using the path command the drainage direction of Hollin Hill was determined, this function calculates based on the ldd map where water would drain using the Hollin Hill coordinates as a starting point:

$$pcrcalc\ path.map = path(ldd.map, hhcoordinates.map)$$

Since Hollin Hill is the main site of interest for the slope stability analysis and not the entire catchment around Hollin Hill, a logical point along the drainage path was

chosen as the arbitrary outflow point. The spatial coordinates of this point were added to the outflowcoordinates text file and the outflowpoints map was created. Now both catchments could be calculated using the catchment function:

```
pcrcalc catchment.map = catchment(ldd.map, outflowpoint.map)
```

This creates a map with the catchment delineation based on the position of the outflow points and the LDD map.

## Appendix II

### 1:

In case of missing precipitation values for a given station due to no record being present or a 0 label, this particular station was not used for a period as long as the time frame of missing values. Missing values were given the value -999.9. A rain station map was created by creating a text file `rainstationcoordinates.txt` and by creating an empty map `empty.map` with the upper left maximum XY coordinates and the lower right minimum XY coordinates. Then using the command

```
col2map - - clone empty.map rainstationcoordinates.txt rainstations.map
```

The nominal `rainstations.map` was created with the same spatial extent as the `empty.map` and the location of all the eight rain stations on it.

Input from this stations was spatially distributed over the catchment (i.e. point precipitation values became catchment covering precipitation values) using the following method (van Beek, 2014):

First the weighed distance for any given cell has been determined using the Python equation

$$distanceWeight = pcr.cos(pcr.atan(pcr.slope(DEM))) ** -1$$

This means that the weighed distance for any given cell is equal to the arc tangent of the slope (height difference per cell) of the Digital Elevation Model to the power of minus 1, and then the cosine of that result.

Using the `spread` command and the location of the rain stations as input, the friction distance (meaning that cells at a shorter distance to the interpolant cell of interest have more importance) was calculated:

$$spread = pcr.spread(stationLocations == sCnt, 0, distanceWeight)$$

In which `sCnt` equals the number of stations and `stationLocations` are the locations of the stations, 0 is the initial friction and `distanceWeight` is the weighed distance as explained above.

Finally, the weight of each station is determined using the following command:

$$weight = (0.5 * pcr.celllength / (spread + 0.5 * pcr.celllength())) ** 2$$

Meaning that the weight of a given station is equal to  $0.5 * \frac{\text{cell length}}{\text{spread} + 0.5 * \text{cell length}}$  and that result to the power of 2.

To account for missing values, the following code was employed:

```
values= stationData[dayCnt,:].ravel()
```

```
totalWeight= 0
```

```
totalValue= 0
```

```
for key, weight in weights.iteritems():
```

```
    value= values[key-1]
```

```
    if value <> MV:
```

```
        totalWeight+= weight
```

```
        totalValue+= weight*value
```

```
totalValue= pcr.ifthen(totalWeight > 0, totalValue/totalWeight)
```

This means that first the values were initialized and taken from the stationData (a data array which was defined before and contains all the daily precipitation data for the eight stations). The totalWeight is equal to 0 initially but takes the weight value as calculated in the weight = command above during the iteration, and the totalValue equals the precipitation values times the weights. This statement is only valid when a given value is not equal to a missing value (MV) or -999.9. In that case the rain station is left out of the equation and will not be used as input for the calculations of the weighted precipitation distribution. Finally, the (estimated) precipitation value for a given cell is obtained by updating the totalValue by using the if then statement where the totalWeight has to be larger than zero (meaning there have to be stations as input) and if that condition is met then the totalValue is equal to the totalValue (sum of all precipitation values times the sum of all weights) over the total Weight.

**2:**

$R_a$  can be calculated using the following equation:

$$R_a = \frac{24(60)}{\pi} * G_{sc} * d_r * [\omega_s * \sin(\phi) * \sin(\delta) + \cos(\phi) * \cos(\delta) * \sin(\omega_s)]$$

In which:

$R_a$  = daily extraterrestrial radiation [ $\text{MJ m}^{-2} \text{d}^{-1}$ ], can be converted to [ $\text{m}^{-2} \text{d}^{-1}$ ] by dividing by 0.408

$G_{sc}$  = solar constant:  $0.082 [\text{MJ m}^{-2} \text{d}^{-1}]$

$d_r$  = relative distance of the earth from the sun [-]

$\omega_s$  = sunset hour angle [°]

$\phi$  = latitude of site in radians

$\delta$  = declination in radians

And:

$d_r = 1 + 0.033 * \cos\left(\frac{2*\pi}{365} * J\right)$  where J represents the day of the year with January 1<sup>st</sup>=1

$\delta = 0.4093 * \sin\left(2 * \pi * \left(\frac{284 + J}{365}\right)\right)$

$\omega_s = \arccos[-\tan(\phi) * \tan(\delta)]$

**3:**

C is calculated by the following equation:

$$C = 1 - \exp(-k * LAI)$$

In which:

k = empirical coefficient, -0.4 for crops and - 0.6 for natural vegetation

FTR is then being calculated by 1-C. Finally,  $k_c$  is obtained by using the following equation (Allen, Pereira, Raes, & Smith, 1998) :

$$k_c = k_{c \min} + (k_{c \text{ full}} - k_{c \min}) * (1 - \exp(-0.7 * LAI))$$

In which:

$k_{c \min}$  = minimum crop factor for bare soil [-], ranging between 0.15-0.20, mean of 0.175 was used

$k_{c \text{ full}}$  = estimated crop factor under full conditions

And:

$$k_{cfull} = 1.0 + 0.1h$$

In which:

$h$  = vegetation height [m],  $k_{cfull}$  is limited to 1.2m if  $h > 2m$ .

## Appendix III

The spatially distributed minimum sand fraction map for the first layer was created using the soil map and a text file for the minimum fractions, where each soil id from the soil map is in column 0, and the minimum sand fraction in column 1. This can be done by using the following command:

```
pcrcalc sandminfraction1.map = lookupscalar(sandminfraction1.txt, soilclasses.map)
```

In a similar fashion, the sandmaxfraction1.map was created using a text file with the maximum fraction of sand for the first layer per soil class, as well as the clayminfraction1.map and the claymaxfraction1.map. Now the spatially distributed sand map for the first layer can be created using the following command:

```
pcrcalc sand1.map
      = sandminfraction1.map
      + (randomfield.map * (sandmaxfraction1.map – sandminfraction1.map))
```

The outcome is that the minimum sand fraction is assigned to the sand1.map where the randomfield.map has a value of 0, and the maximum fraction is assigned when the randomfield.map has a value of 1, and for values in between 0 and 1 proportional fractions are assigned. The same procedure was applied to obtain the clay1.map and now the silt1.map could be produced using:

```
pcrcalc silt1.map = 1 – clay1.map – sand1.map
```

Since only clay and sand fractions were given and these do not amount to 1 for each soil type, the remainder has to be the silt fraction hence the 1 minus the clay and the sand fractions. The same procedure was applied for layer 2 and layer 3, so finally spatially distributed maps for three layers with sand, clay and silt fractions were used as input for the pedo transfer script to obtain all the above mentioned soil parameters.

## Appendix IV

$$FS = \frac{c' + [(1 - \omega) * \gamma + \omega * \gamma'] * z * \cos^2 \theta * \tan \phi'}{[(1 - \omega) * \gamma + \omega * \gamma_{sat}] * z * \sin \theta * \cos \theta}$$

In which:

FS = Factor of Safety [-]

$c'$  = Soil cohesion [kPa]

$\gamma$  = Average bulk unit weight of soil above groundwater level [ $M L^{-1} T^{-2}$ ]

$\omega$  = Groundwater level index [-]

$\gamma' = \gamma_{sat} - \gamma_w$ , better known as the submerged unit weight of soil [ $M L^{-1} T^{-2}$ ]

$z$  = depth of shear plane [m]

$\theta$  = slope angle [°]

$\phi'$  = angle of internal friction [°]

$\gamma_{sat}$  = saturated unit weight of soil [ $M L^{-1} T^{-2}$ ]

The following components can be redefined:

$$\frac{\gamma}{\gamma_w} = \frac{G_s + e * S_r}{1 + e}$$

$$\frac{\gamma_{sat}}{\gamma_w} = \frac{G_s + e}{1 + e}$$

$$\frac{\gamma'}{\gamma_w} = \frac{G_s - 1}{1 + e}$$

In which:

$G_s$  = specific gravity of solids [-],  $\gamma_{sat} / \gamma_w$

$S_r$  = The average degree of saturation  $V_w / V_v$  [-] (In which  $V_w$  is the volume of water and  $V_v$  the volume of voids)

$e$  = The average void ratio above the groundwater table  $V_v / V_s$  [-] (In which  $V_s$  is the volume of solids)

If rewritten and  $c'$  is neglected (but can of course also be included) to get a FS independent of  $z$  FS now becomes:

$$FS = \frac{[G_s + e * S_r - \omega * (1 + e * S_r)] * \tan \phi'}{[G_s + e * S_r + \omega * e * (1 - S_r)] * \tan \theta}$$

If  $\omega = 0$  i.e. the groundwater lies at the slip surface  $FS = \frac{\tan \phi'}{\tan \theta}$  and if  $\omega = 1$  (i.e. fully saturated soil)  $FS =$

$$\frac{(G_s - 1) * \tan \phi'}{(G_s + e) * \tan \theta} = \frac{\gamma' * \tan \phi'}{\gamma_{sat} * \tan \theta}$$

The critical depth  $\omega_r$  can now be calculated:

$$\omega_r = \frac{(G_s + e * S_r) * \left(1 - \frac{\tan \phi'}{\tan \theta}\right)}{1 + e - e * (1 - S_r) * \left(1 - \frac{\tan \phi'}{\tan \theta}\right)}$$

## References

- Allen, R., Pereira, L., Raes, D., & Smith, M. (1998). *Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56*. Rome: FAO.
- AMS Weather Studies (2014). *Introduction to Atmospheric Science, 5<sup>th</sup> edition*. Retrieved June 8, 2015, from <http://webcache.googleusercontent.com/search?q=cache:a33g900-efgl:www.geo.mtu.edu/~scarn/teaching/GE2640/Ch13Pres.ppt+&cd=1&hl=nl&ct=clnk&gl=nl>
- Anderson, K., Croft, H. (2009). Remote sensing of soil surface properties. *Progress in Physical Geography*(33), 457–473.
- Anderson, A.G., Holcombe, E., Blake, J.R., Ghesquire, F., Holm-Hielsen, H., & Fisseha, T. (2011). Reducing landslide risk in communities: Evidence from the Eastern Carribean. *Applied Geography*(31), 590-599.
- Beek, L.P.H. van. (2002). *Assessment of the influence of changes in land use and climate on landslide activity in a Mediterranean environment*. Utrecht, Netherlands: Labor Grafimedia b.v. .
- Beek, L.P.H. van., Cammeraat, E., Andreu, V., Mickovski, S., & Dorren, L. (2008). HILLSLOPE PROCESSES: MASS WASTING, SLOPE STABILITY AND EROSION. In J. Norris, A. Stokes, S. Mickovski, E. Cammeraat, L. van Beek, B. Nicoll, & A. Achim, *Slope Stability and Erosion Control: Ecotechnological Solutions* (pp. 17-64). Springer-Verlag New York Inc.
- Beek, L.P.H. van. (2014, various dates). Meetings about model approaches. (E. Schoor, Interviewer)
- Beven, K. (2004). Robert E Horton's perceptual model of infiltration processes. *Hydrological processes*(18), 3447-3460.
- BGS. (2014). *Increased incidence of landslides in 2012*. Retrieved March 3, 2014, from <http://www.bgs.ac.uk/landslides/November2012.html>
- BGS. (2014). *The Hollin Hill Landslide Observatory, Yorkshire*. Retrieved March 1, 2014, from <http://www.bgs.ac.uk/science/landUseAndDevelopment/landslides/hollinHill.html>
- Blöschl, G., & Sivapalan, M. (1995). Scale Issues in Hydrological Modelling: A Review. *Hydrological processes*(9), 251-290.
- Bouttier, F., & Courtier, P. (1999). *Data assimilation concepts and methods*. Retrieved March 10, 2014, from [http://msi.ttu.ee/~elken/Assim\\_concepts.pdf](http://msi.ttu.ee/~elken/Assim_concepts.pdf)
- Brocca, L., Ponziani, F., Moramarco, T., Melone, F., Bemi, N., & Wagner, W. (2012). Improving Landslide Forecasting Using ASCAT-Derived Soil Moisture Data: A Case Study of the Torgiovanetto Landslide in Central Italy. *Remote Sensing*(4), 1232-1244.
- Caine, N. (1980). The rainfall intensity-duration control of shallow landslides and debris flows. *Geografiska Annaler*(62), 23-27.
- Carter, A. (1995). *HOWARDIAN HILLS MAPPING PROJECT: A REPORT FOR THE NATIONAL MAPPING PROGRAMME*. Accessed December 4, 2014, on [https://www.english-heritage.org.uk/publications/howardian-hills-nmp/HOWARDIAN\\_HILLS\\_NMP\\_web.pdf](https://www.english-heritage.org.uk/publications/howardian-hills-nmp/HOWARDIAN_HILLS_NMP_web.pdf)



Cedar Lake Ventures Inc. (2015). *Types of Precipitation Throughout the Year*. Retrieved June 8, 2015, from <https://weatherspark.com/averages/28791/Linton-on-Ouse-England-United-Kingdom>

Centre for Environmental and Data Archival (2015). *1 km Resolution UK Composite Rainfall Data from the Met Office Nimrod System*. Accessed July 20, 2015, on <http://catalogue.ceda.ac.uk/uuid/27dd6ffba67f667a18c62de5c3456350>

Chambers, J., Wilkinson, P., Kuras, O., Ford, J., Gunn, D., Meldrum, P., . . . Ogilvy, R. (2011). Three-dimensional geophysical anatomy of an active landslide in Lias Group mudrocks, Cleveland Basin, UK. *Geomorphology*(125), 472-484.

Chang, L.-Y. (sine anno). *Integration of Rainfall Triggered Shallow Landslide Model in FEWS\_Taiwan*. Retrieved March 3, 2014, from [http://oss.deltares.nl/c/document\\_library/get\\_file?uuid=6e2e4b72-b557-4147-9048-7864e5c43769&groupId=145641](http://oss.deltares.nl/c/document_library/get_file?uuid=6e2e4b72-b557-4147-9048-7864e5c43769&groupId=145641)

Cranfield University. (sine anno) *Soilscapes map*. Retrieved November 28, 2014, from <http://www.landis.org.uk/soilscapes/>

Crozier, M. (2010). Landslide geomorphology: An argument for recognition, with examples from New-Zealand. *Geomorphology*, 3-15.

Dai, F.C., Lee, C.F., & Ngai, Y.Y. (2002). Landslide risk assessment and management: an overview. *Engineering Geology*(64), 65-87.

Dickinson, R., Henderson-Sellers, A., & Kennedy, P. (1993). *Biosphere-atmosphere Transfer Scheme (BATS) Version 1e as Coupled to the NCAR Community Climate Model*. Boulder, Colorado: National Center for Atmospheric Research . Accessed on July 23, 2014, on <http://nldr.library.ucar.edu/repository/assets/technotes/TECH-NOTE-000-000-000-198.pdf>

Dietrich, W.E., Asua, R.R. de, Orr, J.C.B., & Trso, M. (1998). *A validation study of the shallow slope stability model, SHALSTAB, in forested lands of Northern California*. Accessed on July 20, 2015, on [http://www.krisweb.com/biblio/gen\\_ucb\\_dietrichetal\\_1998\\_shalstab.pdf](http://www.krisweb.com/biblio/gen_ucb_dietrichetal_1998_shalstab.pdf)

Dios, R.J.C. de., Enriquez, J., Victorino, F.G., Mendoza, E.A., Talampas, M.C., Marciano Jr., & J.J.S. Tilt, Soil Moisture, and Pore Water Pressure Sensor System for Slope Monitoring Applications. *Science Diliman*(21), 15-27.

Dixon, N., Spriggs, M.P., Meldrum, P., & Haslam, E. (2012) *Field trial of an acoustic emission early warning system for slope instability*. Accessed January 20, 2015, on <https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/10239/8/ISL-Dixon%20et%20al%20Hollin%20Hill%20paper%20Final%2017%201%2012.pdf>

Dijkstra, T. (2014, January 24). First meeting discussing research details. (E. Schoor, Interviewer)

Eden, P. (2015). *The rain-shadow effect*. Accessed July 8, 2015, on <http://www.weatheronline.co.uk/reports/philip-eden/Sheltered-by-mountains.htm>

Edwards, B. (2006). *GIS-based Radar Rainfall Verification*. Accessed July 20, 2015, on <http://nldr.library.ucar.edu/repository/assets/soars/SOARS-000-000-000-027.pdf>

Eeckhaut, M. van den, Poesen, J., Verstraeten, G., Vanacker, V., Nyssen, J., Moeyersons, J., Beek, L.P.H., & Vandekerckhove, L. (2007). Use of LiDAR-derived images for mapping old landslides under forest. *Earth Surface Processes and Landforms*(32), 754–769.

- Environment Agency (2001). *Estimation of open water evaporation*. Accessed June 8, 2015, on [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/290578/sw6-043-tr-e-e.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/290578/sw6-043-tr-e-e.pdf)
- Foster, C., Jenkins, G.O, & Gibson, A.D. (2007). *Landslides and mass movement processes and their distribution in the York District (Sheet 63)*. Accessed January 20, 2015, on [nora.nerc.ac.uk/3728/1/YORK\\_FINAL](http://nora.nerc.ac.uk/3728/1/YORK_FINAL)
- Garcia, C., & Fearnly, C.J (2012). Evaluating critical links in early warning systems for natural hazards. *Environmental Hazards*(11), 123-137.
- Geoscience (sine anno). *What is continental rifting?* Retrieved December 1, 2014, from <http://www.geosci.usyd.edu.au/users/prey/Teaching/Geol-3101/Rifting02/rift.html>
- Gomez-Hernández, J.J., & Journel, A.G. (1993). Joint Sequential Simulation of Multigaussian Fields. *Geostatistics Troia*(92), 85-94.
- Gorsevski, P.V., Brown, K.M., Panter, K., Onasch, C.M., Simic, A., & Snyder, J. (2015). Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: a case study in the Cuyahoga Valley National Park, Ohio. *Landslides*.
- Guimarães, R.F., Montgomery, D.R, Greenberg, H.M., Fernandes, N.F., Gomes, R.A.T., & Carvalho Júnior, O.A. d. (2003). Parameterization of soil properties for a model of topographic controls on shallow landsliding: application to Rio de Janeiro. *Engineering Geology*(69), 99-108.
- Gunn, D., Chambers, J., Hobbs, P., Ford, J., Wilkinson, P., Jenkins, G., & Merrit, A. (2013). Rapid observations to guide the design of systems for long-term monitoring of a complex landslide in the Upper Lias days of North Yorkshire, UK. *Quarterly Journal of Engineering Geology and Hydrogeology*.
- Guzetti, F., Peruccacci, S., Rossi, M., & Stark, C. (2008). The rainfall-intensity duration control of shallow landslides and debris flows: an update. *Landslides*(5), 3-17.
- Hobbs, P.R.N, Entwisle, D.C., Northmore, K.J., Sumbler, M.G., Jones, L.D, Kemp, S., . . . Meakin, J.L.(2012). *Engineering Geology of British Rocks and Soils - Lias Group*. Accessed December 4, 2014, on <http://nora.nerc.ac.uk/17270/1/OR12032.pdf>
- Haby, J. (sine anno). *Applying tropospheric moisture to forecasting*. Accessed June 8, 2015, on <http://www.theweatherprediction.com/habyhints/114/>
- Hagemann, S., Botzet, M., Dümenil, L., & Machenhauer, B. (1999). *DERIVATION OF GLOBAL GCM BOUNDARY CONDITIONS FROM 1 KM LAND USE SATELLITE DATA*. Hamburg: Max Planck Institute for Meteorology. Accessed on July 29, 2014, van [http://www.mpimet.mpg.de/fileadmin/publikationen/Reports/max\\_screp\\_289.pdf](http://www.mpimet.mpg.de/fileadmin/publikationen/Reports/max_screp_289.pdf)
- Hungr, O. (2004). Landslide Hazards in British Columbia. Achieving Balance in Risk Assessments. *Innovation*, 12-16.
- Hwang, T., Band, L.E., Hales, T.C., Miniati, C.F., Vose, J.M., Bolstad, P.V., Miles, B., & Price, K. (2015). Simulating vegetation controls on hurricane-induced shallow landslides with a distributed ecohydrological model. *Journal of Geophysical Research: Biogeosciences*(120), 361-378.
- Intrieri, E., Gigli, G., Casagli, N. & Nadim, F. (2013). "Landslide Early Warning System: toolbox and general concepts". *Natural Hazards and Earth System Sciences*(13), 85-90.

Jakob, M., Owen, T., & Simpson, T. (2011). A regional real-time debris-flow warning system for the District of North Vancouver, Canada. *Landslides*(9), 165-178.

Krzeminska, D.M. (2012). *THE INFLUENCE OF FISSURES ON LANDSLIDE HYDROLOGY*. PHD Thesis. Available on: [http://repository.tudelft.nl/assets/uuid:a6539f84-f607-49d4-9bbc-f61c5b70f998/Krzeminska\\_Thesis.pdf](http://repository.tudelft.nl/assets/uuid:a6539f84-f607-49d4-9bbc-f61c5b70f998/Krzeminska_Thesis.pdf)

Krzeminska, D.M., Boogaard, T.A., Asch, T. W. v., & Beek, L.P.H. v. (2012). A conceptual model of the hydrological influence of fissures on landslide activity. *Hydrology and Earth System Science*(16), 1561-1567.

Kuriakose, S. (2010). *Physically-based dynamic modelling of the effect of land use changes on shallow landslide initiation in the Western Ghats of Kerala, India*. Enschede, Netherlands: ITC Printing Department.

Kurnik, B., Kafjež-Bogataj, L., & Ceglar, A. (2012). Correcting mean and extremes in monthly precipitation from 8 regional climate models over Europe. *Climatic Past Discussions*(8), 953-986.

Kustas, W.P., & Goodrich, D.C., (1994). Preface to the special section on Monsoon 90. *Water Resources Research* 30 (5), 1211–1225.

Lanni, C., Borga, M., Rigon, R., & Tarolli, P. (2012). Modeling shallow landslide susceptibility by means of a subsurface flow path connectivity index and estimates of soil depth spatial distribution. *Hydrology and Earth System Sciences*(16), 3959-3971.

Le, H.L, Miyagi, T., Shinro, A., Hamasaki, E., & Dinh, V.T. (2014). In: Sassa, K., Canuti, P., & Yin, Y. (eds) *Landslides Science for a Safer Geoenvironment*.

Liu, X. (2008). Airborne LiDAR for DEM generation: some critical issues. *Progress in Physical Geography*(32), 31–49.

Medina-Cetina, Z. & Nadim, F. (2008). Stochastic design of an early warning system. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*(2), 223-236.

Mercogliano, P., Segoni, S., Rossi, G., Sikorsky, B., Tofani, V., Schiano, P., . . . Casagli, N. (2013). A prototype forecasting chain for rainfall induced shallow landslides. *Natural Hazards and Earth System Sciences*(13), 771-777.

Merz, R., Parajka, J., & Blöschl, G. (2009). Scale effects in conceptual hydrological modeling. *Water Resources Research* (45).

Meteorologisk institutt (2014). *Weather statistics for Yorkshire, England (United Kingdom)*. Accessed June 11, 2015, on [www.yr.no/place/United\\_Kingdom/England/Yorkshire/statistics.html](http://www.yr.no/place/United_Kingdom/England/Yorkshire/statistics.html)

Met Office (2015). *North East England: Climate*. Accessed July 8, 2015, on <http://www.metoffice.gov.uk/dimate/uk/regional-dimates/ne>

Met Office (2014). *UK 3-hourly site-specific forecast detailed documentation*. Accessed July 19, 2015, on <http://www.metoffice.gov.uk/datapoint/product/uk-3hourly-site-specific-forecast/detailed-documentation>

Michaud, J., & Sorooshian, S. (1994). Comparison of simple versus complex distributed runoff models on a mid-sized semiarid watershed. *Water Resources Research* 30(3), 593–606.

Mulungu, D.M.M., & Munishi, S.E. (2007). Simiyu River catchment parameterization using SWAT model. *Physics and Chemistry of the Earth*(32), 1032-1039.

Nash, J.E., Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*(10), 282-290.

NERC(2004). *Evaporation and soil moisture deficits*. Accessed June 8, 2015 on <http://www.nerc-wallingford.ac.uk/ih/nrfa/yb/yb2004/evaporation.html>

NERC (2014). *Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current)*. Accessed July 20, 2014, on [http://badc.nerc.ac.uk/view/badc.nerc.ac.uk\\_\\_ATOM\\_\\_dataent\\_ukmo-midas](http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_ukmo-midas)

Nicótina, L., Tarboton, D.G., Tesfa, T.K. & Rinaldo, A. (2011). Hydrologic controls on equilibrium soil depths. *Water Resources Research*(47).

Opadeyi, J., Ali, S., & Chin, E. (2003). *Status of Hazard Maps Vulnerability Assessment and Digital Maps in the Caribbean. Final Report*. Accessed June 26, 2015, on <http://reliefweb.int/sites/reliefweb.int/files/resources/35C55AFA44647D97C1256FAF00326F42-cdera-disred-18feb.pdf>

Pack, R.T., Tarboton, D.G., Goodwin, C.G., & Prasad, A. (2005). *SINMAP 2: A STABILITY INDEX APPROACH TO TERRAIN STABILITY HAZARD MAPPING*. Accessed July 20, 2015, on <http://hydrology.usu.edu/sinmap2/sinmap2.PDF>

Pack, R.T., Tarboton, D.G., & Goodwin, C.N. (1998). *The SINMAP Approach to Terrain Stability Mapping*. Accessed July 18, 2015, on <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.41.5799&rep=rep1&type=pdf>

Pechlivanidis, I.G., Jackson, B.M., McIntyre, N.R., & Wheeler, H.S.(2011). Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST Journal*(13), 193-214.

Penížek, V., & Boruvka, L. (2006). Soil depth prediction supported by primary terrain attributes: a comparison of methods. *Plant Soil Environ*(9), 424-430.

Powell, J.H. (2010) *Jurassic sedimentation in the Cleveland Basin: a review*. Accessed November 29, 2014, on <http://core.ac.uk/download/pdf/56592.pdf>

Pritchard, O., Farewell, T., & Hallett, S. (2013). *Coastal Landslides within the UK Context*. Cranfield University, UK, NSRI.

Ramesh, M. (2014). Design, development, and deployment of a wireless sensor network for detection of landslides. *Ad Hoc Networks*(13), 2-18.

Riddell, J. (2007). *Bedrock Geology*. Accessed May 22, 2015, on <http://www.centrialsaanich.ca/Assets/Central+Saanch/Publications/CS+Resource+Atlas/Bedrock+Geology.pdf?method=1>

Samani, Z. (sine anno). *Estimating Solar Radiation and Evapotranspiration Using Minimum Climatological Data (Hargreaves-Samani equation)*. New Mexico State University, Civil Engineering Department. Accessed on July 28, 2014, on [http://www.zohrabsamani.com/research\\_material/files/Hargreaves-samani.pdf](http://www.zohrabsamani.com/research_material/files/Hargreaves-samani.pdf)

- Schmugge, T., Jackson, T.J., Kustas, W.P., Roberts, R., Parry, R., Goodrich, D.C., Amer, S.A., & Wertz, M.A. (1994). Push broom microwave radiometer observations of surface soil moisture in Monsoon '90. *Water Resources Research* 30(5), 1321–1328.
- Schulz, W.H. (2007). Landslides susceptibility revealed by LiDAR imagery and historical records, Seattle, Washington. *Engineering Geology* (89),67–87.
- Shahidian, S., Serralheiro, R., Serrano, J., Teixeira, J., & Haie, N. (2012). *Hargreaves and Other Reduced-Set Methods for Calculating Evapotranspiration*. InTech.
- Sidle, R., & Dhakal, A. (2002). Potential effects of environmental change on landslide hazards in forest environments. In R. Sidle, *Environmental changes and geomorphic hazards in forests*. (pp. 123-167). Cabi Publishing.
- Sirangelo, B., & Braca, G. (2004). Identification of hazard conditions for mudflow occurrence by hydrological model: Application of FLAIR model to Sarno warning system. *Engineering Geology*(73), 267-276.
- Stephenson, G.R., & Freeze, R.A (1974). Mathematical simulation of subsurface flow contributions to snowmelt runoff. *Water Resources Research*(10), 284-294.
- Tesfa, T.K., Tarboton, D.G., Chandler, D.G., & McNamara, J.P. (2009). Modeling soil depth from topographic and land cover attributes. *Water Resources Research*(45).
- Texas A&M University (sine anno). *SWAT Input Data: Overview*. Accessed July 17, 2015 on [http://swat.tamu.edu/media/69302/ch01\\_overview.pdf](http://swat.tamu.edu/media/69302/ch01_overview.pdf)
- The Geological Society of London. (2001). *The where, when and how of landslides*. The Geological Society of London. London: The Geological Society of London.
- Tsai, T.-L., & Wang, J.-K. (2011). Examination of influences of rainfall patterns on shallow landslides due to dissipation of matric suction. *Environmental Earth Sciences*(63), 65-75.
- Ulmer, M., Molnar, P., & Purves, R. (2009). *Influence of DEM and Soil Property Uncertainty on an Infinite Slope Stability Model*. Accessed July 21, 2015, on <http://geomorphometry.org/system/files/ulmer2009geomorphometry.pdf>
- USGS. (2012). *Horst and graben diagram*. Retrieved December 1, 2014 from <http://earthquake.usgs.gov/learn/glossary/?term=horst>
- USGS. (2008). *Global Land Cover Characteristics Data Base Version 1.2*. Accessed June 20, 2014 on [http://edc2.usgs.gov/gfcc/globdoc1\\_2.php](http://edc2.usgs.gov/gfcc/globdoc1_2.php)
- Utrecht University (2014). *Iddcreate*. Retrieved July 20, 2014, from [http://pcraster.geo.uu.nl/pcraster/4.0.0/doc/manual/op\\_iddcreate.html](http://pcraster.geo.uu.nl/pcraster/4.0.0/doc/manual/op_iddcreate.html)
- Varnes, D.J. (1978). Slope movements: type and processes. In: Schuster, R.L., & Krizek, R.J. (eds) *Landslides Analysis and Control. Special Report*(176), 11-33.
- Wienhöfer, J., Lindenmaier, F., & Zehe, E. (2011). Challenges in Understanding the Hydrologic Controls on the Mobility of Slow-Moving Landslides. *Vadose Zone Journal*(9), 496-511.
- Wong, H.N., Ho, K.K.S., & Chan, Y.C. (1997). Assessment of consequences of landslides. In: Cruden, R., & Fell, R. (eds) *Landslide Risk Assessment*, 111-149.

Yu, Z. (2002). *Hydrology: Model and Prediction*. University of Nevada, Department of Geoscience. Elsevier Science Ltd. Retrieved March 10, 2014, from <http://hydro.nevada.edu/pub/books/book7.pdf>