MOUNTAIN-PLAINS CONSORTIUM

MPC 23-500 | E. Byron and P. Nelson

PROBABILISTIC MODELING OF LANDSLIDE HAZARDS TO IMPROVE THE RESILIENCE OF TRANSPORTATION INFRASTRUCTURE

MOUNTAIN-PLAINS CONSORTIUM REGION 8

A University Transportation Center sponsored by the U.S. Department of Transportation serving the Mountain-Plains Region. Consortium members:

Colorado State University
North Dakota State University
South Dakota State University
University of Colorado Denver
University of Denver
University of Utah
Utah State University
University of Wyoming

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
MPC-593		

4. Title and Subtitle	5. Report Date
Probabilistic Modeling of Landslide Hazards to Improve the Resilience of Transportation Infrastructure	July 2023
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Elizabeth Byron Peter A. Nelson	MPC 23-500

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Colorado State University 1372 Campus Delivery Fort Collins, CO 80523-1372	11. Contract or Grant No.

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Mountain-Plains Consortium North Dakota State University PO Box 6050, Fargo, ND 58108	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

Supported by a grant from the US DOT, University Transportation Centers Program

16. Abstract

Precipitation-induced landslides pose risks to humans through property damage, disruption of infrastructure, injury, and loss of life. These risks may be altered by climate change, as changes in vegetation cover and associated root cohesion might lead to a change in areas susceptible to landslides. We investigate this possibility through Monte Carlo simulations of slope stability in the Colorado Front Range, where climate change is expected to significantly change vegetation cover across the landscape. Climate change simulations predict an overall increase in the area susceptible to landslides and a shift to more instability on north-facing slopes. Our study suggests that vegetation changes due to climate change could result in major shifts in the people and infrastructure susceptible to landslides.

We also apply a landslide runout model to a large spatial scale to determine whether simplified assumptions using easily accessible data can provide realistic estimates of landslide stopping locations. We consider stopping rules using slope, curvature, and travel distance and find that a combination of a critical angle and a distance the landslide must maintain beneath the critical angle best predicts stopping locations in our study area.

17. Key Word	18. Distribution Statement
debris flows, disaster resilience, floods, forecasting, highways, landslides, mapping, Monte Carlo method, probability	Public distribution

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	69	n/a

Probabilistic Modeling of Landslide Hazards to Improve the Resilience of Transportation Infrastructure

Elizabeth Byron
Peter A. Nelson (PI)

Department of Civil and Environmental Engineering
Colorado State University
Fort Collins, CO 80523

July 2023

Acknowledgments

The authors thank the Mountain-Plains Consortium and Colorado State University (CSU) for providing funds that made this study possible.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ABSTRACT

Precipitation-induced landslides pose risks to humans through property damage, disruption of infrastructure, injury, and loss of life. These risks may be altered by climate change, as changes in vegetation cover and associated root cohesion might lead to a change in areas susceptible to landslides. We investigate this possibility through Monte Carlo simulations of slope stability in the Colorado Front Range, where climate change is expected to significantly change vegetation cover across the landscape. Climate change simulations predict an overall increase in the areas susceptible to landslides and a shift to more instability on north-facing slopes. Our study suggests that vegetation changes due to climate change could result in major shifts in the people and infrastructure susceptible to landslides.

We also apply a landslide runout model to a large spatial scale to determine whether simplified assumptions using easily accessible data can provide realistic estimates of landslide stopping locations. We consider stopping rules using slope, curvature, and travel distance and find that a combination of a critical angle and a distance the landslide must maintain beneath the critical angle best predicts stopping locations in our study area.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE OF NOTATION

Slope stability variables

Variable	Symbol	Units
Air entry matric potential	ψ_s	m of water
Cohesion, soil	C_s	Pa
Cohesion, vegetated	C_v	Pa
Density, soil	ρ_s	kg/m^2
Density, water	ρ_w	kg/m^2
Factor of safety	FS	-
Fractional vegetative cover	V	-
Internal friction angle	φ	°
Pore size distribution	b	-
Residual soil moisture	θ_r	m^3/m^3
Saturated water content	θ_s	-
Slope	S	m/m
Soil depth	h	m
Volumetric water content	θ	m^3/m^3

Soil moisture downscaling variables

Variable	Symbol	Units
Aspect, average	$R_{\#}$	-
Base soil depth, curvature = 0	δ_0	m
Curvature	κ	-
Curvature, minimum where soil depth is present	κ_{min}	-
Deep drainage	G	mm/d
Dimension of input DEM grid cells	c	M
Elevation dependence	τ	1/m
Elevation dependence of PET	ω	1/m
Elevation, average	$Z_{\#}$	m
Evapotranspiration	E	mm/d
Evapotranspiration, aerodynamic limitations due to moisture	β_a	-
Evapotranspiration, radiation limitations due to moisture	β_r	-
Hydraulic conductivity anisotropy	ι	-
Infiltration rate	F	mm/d
Lateral flow	L	mm/d
Orientation dependence	ξ	-
Orientation reference	υ	-
Pore disconnectedness index, horizontal	γ_h	-
Pore disconnectedness index, vertical	γ_v	-
Potential solar radiation index	I_p	-
Precipitation, spatial average	$\bar{P}$	mm/d

Priestley Taylor coefficient minus 1	α	-
Relationship between slope with horizontal hydraulic gradient	ϵ	-
Saturated precipitation rate, vertical	$K_{s,v}$	mm/d
Shading and its impacts	μ	-
Slope, average	$S_{\#}$	-
Soil moisture, aerodynamic evapotranspiration value	θ_A	
Soil moisture, deep drainage value	θ_G	
Soil moisture, lateral flow value	θ_L	
Soil moisture, radiative evapotranspiration value	θ_R	
Soil moisture, weighting for the aerodynamic evapotranspiration component	ω_A	
Soil moisture, weighting for the deep drainage component	ω_G	
Soil moisture, weighting for the lateral flow component	ω_L	
Soil moisture, weighting for the radiative evapotranspiration component	ω_R	
Spatially average potential evapotranspiration	$\bar{\bar{E}}_p$	mm/d
Transpiration in downscaled soil moisture	η	-
Vegetation efficiency (blocking precipitation)	λ	-

Probability model evaluation

Variable	Symbol	Units
False negatives (observed landslides predicted as stable)	*FN*	-
False positives (not observed landslides predicted as unstable)	*FP*	-
Majority class, count	N_{maj}	-
Minority class, count	N_{min}	-
True negatives (not observed landslides predicted as stable)	*TN*	-
True positive (observed landslides predicted as unstable)	*TP*	-

EXECUTIVE SUMMARY

Precipitation-induced landslides pose risks to humans through property damage, disruption of infrastructure, injury, and loss of life. Due to the spatial and temporal heterogeneity of soil moisture and landscape characteristics that impact slope stability and potential impacts of climate change on landslide location, quantifying landslide risk to humans is difficult as uncertainties are not represented in available datasets. Recent developments have improved our ability to probabilistically model landslide initiation, thus allowing for the incorporation of spatial and temporal uncertainty in the prediction of the onset of hillslope failures. The ability to incorporate uncertainty in landslide models is particularly valuable for considering how climate change, which could impact vegetation cover and associated root cohesion, might alter the vulnerability of people and infrastructure to landslides. The aim of this analysis is to probabilistically forecast landslide susceptibility under climate change by incorporating changes in the type and distribution of vegetation while accounting for uncertainties in key properties. Using Landlab, a Python-based toolkit for landscape modeling, we perform Monte Carlo simulations with an infinite slope stability model to make spatially explicit calculations of the probability of landslide initiation. The soil moisture input to the landslide model is from the Equilibrium Moisture from Topography, Vegetation, and Soil (EMT+VS) model, which downscales coarse-resolution soil moisture by incorporating the dependence of soil moisture on topographic, vegetative, and soil characteristics. We evaluate model sensitivity and identify that vegetation, which impacts cohesion and soil depth, has a large impact on the model. We evaluate model performance by simulating landslide susceptibility over a 1333 km^2 area of the Colorado Front Range as there is a large inventory of more than 1,300 landslides from an extreme precipitation event in 2013. One anticipated effect of climate change in the Colorado Front Range is a reduction in the survivability of trees, which we incorporate through applying reductions to vegetative cohesion and vegetation cover. For the 2013 event, the model predicts 79.6% of the mapped landslides and 5.8% of the rest of the study area as being unstable. A deterministic model using mean values from the probability model and assuming $FS \leq 1$ is unstable captures only 42% of observed landslides, supporting the use of the probabilistic model. The probabilities are low ($P(F) < 0.2$) for the majority of predicted failures with a concentration at higher ($P(F) > 0.8$) values, with the latter having higher slopes and lower vegetation. Among nodes with $P(F) > 0$, 66% occur on south facing slopes where trees are less abundant. After incorporating climate change, we see an increase in the areas susceptible to landslides and a shift to more instability on north-facing slopes. Our study suggests that vegetation changes due to climate change could result in major shifts in the people and infrastructure susceptible to landslides in the Colorado Front Range.

In conjunction with landslide initiation, determining landslide runout is important to fully analyze landslide risk. Landslide runout modeling for large areas is difficult due to limited information and the complexity of landslides. The difficulties of physically modeling landslides on large spatial scales have led to the development of empirical methods based on topographic attributes. While empirical models are limited in that they require calibration in new areas and thus can only be applied to areas with landslide inventories, they provide a way to model landslide runout at large spatial scales and identify areas for further, potentially more physically based, analyses. We investigate whether topographic controls can be used to predict landslide termination. We develop a landslide runout model and apply it to a 10-m elevation grid. Our model routes landslides downslope with a d8 flow direction method and uses a critical slope, defined as a minimum slope a landslide must encounter to end, and slope persistence, defined as the distance the landslide must travel under the critical slope, to represent landslide stopping locations. We apply our model to see if it can replicate landslide runout in the Colorado Front Range due to a large landslide inventory from a 2013 precipitation event that induced approximately 1,300 mapped landslides. The calibrated model has a critical slope of 3° and a slope persistence of 20 m and predicts landslide distance in both the calibration and evaluation areas with a Nash-Sutcliffe (NS) value of 0.69 and 0.58, respectively. We compare our calibrated model to an angle of reach approach, an approach that has been applied previously for landslide runout mapping that determines the slope between the start and end of a

landslide, and determine that the best NS value of 0.14 occurs at an angle of 20°. Our results show that within our study area, topographic controls provide plausible initial estimates of runout endpoints and an improvement over similarly simplistic methods such as the angle of reach. The potential of using critical slope combined with slope persistence to capture topographic controls to predict runout endpoints is a promising opportunity for landslide hazard mapping at large spatial extents.

1. INTRODUCTION AND BACKGROUND

1.1 Introduction

Landslides pose risk through economic losses to infrastructure, damage to the environment, and loss of human life. Quantifying risk considers where landslides will occur and where they will travel once initiated to determine what people, infrastructure, and downslope environments are susceptible.

Modeling landslide initiation at large extents requires accurate spatial data for all input variables, including topography, hydrology, and vegetation. While some input data, such as elevation, are readily available for most parts of the world, other data may be unavailable, of low resolution, or deemed highly uncertain. Soil moisture, due to its dependence on elevation, vegetation, soil type, and soil texture, is one variable that is particularly difficult to quantity for precipitation-induced landslides. Soil moisture is often coarse in resolution, requiring downscaling for accurate values if applied to regional models. Peng et al. (2016) discuss approaches to soil moisture downscaling, which includes relating land surface temperature with vegetation and relating vegetative dryness and vegetation temperature controls. Another variable that is difficult to quantify at large spatial extents is soil texture, which impacts infiltration rates and water holding capacity of a soil, both of which impact soil moisture and thus slope stability. Pauly et al. (2020) found that spatially averaged soil texture worked as well as spatially varied soil texture for soil moisture downscaling due to uncertainties. Grieco et al. (2018) identified that using global land use data in soil moisture downscaling did not improve performance due to variable uncertainties and potential misclassifications. These studies demonstrate that while accurate soil data are necessary for soil moisture evaluation, and thus precipitation-induced landslide models, providing accurate soil characteristics for input into models applied over broad spatial extents is challenging. Deterministic predictions for landslide susceptibility are particularly challenging due to these issues. A probabilistic approach can improve upon a deterministic approach by incorporating some of these uncertainties.

Landslide runout can be considered in a variety of ways. While all approaches are aimed at quantifying risk, modeling approaches might consider rules-based, topographic control approaches (Milledge et al., 2019), hazard mapping based on representing the frictional and momentum losses of landslides as they travel (Gibson and Sanchez, 2020; Horton et al., 2013; Quan Luna et al., 2016), and empirical evaluations of landslide inventories (Clerici et al., 2010). Similar to landslide initiation, many modeling approaches have limitations, particularly when applied at large spatial extents. One setback is that the data required for some modeling approaches, such as relating distance traveled to landslide volume, are not readily available in all landslide inventories, leading to the inability to apply models that use that approach over broad extents. Another limitation is that physically characterizing mass movements, particularly debris flows, is a challenge due to the complexity and unsteady spatial extents and temporal aspects (Iverson, 1997), due in part to solid-fluid interactions leading to non-Newtonian behavior (De Blasio, 2011). This leads to challenges in applying runout models to large spatial extents if the physical components of the landslides are fully represented. We aim to develop a simple landslide runout model that evaluates topographic controls to determine whether they represent landslide end points, which could allow for easier and more rapid applications of landslide runout to large spatial extents.

A final component of modeling precipitation-induced landslide risk across large spatial extents is considering how to incorporate the impacts of climate change. Understanding the difference between current and future landslide initiation locations is important to consider but is challenging to quantify. Predicting the impacts of climate change on precipitation-induced landslides requires considering both the changes to hydrologic factors and changes to vegetation as they both can impact landslide susceptibility (Bernardie et al., 2021). Climate models themselves have built-in assumptions and limitations, including often having large spatial extents and thus having similar downscaling problems as soil moisture. In

considering the hydrologic impacts of climate change, downscaling is often needed as the duration and intensity of rainfall can differ between regional-level models and local weather (Mahoney et al., 2013). Hydrologic changes are most often incorporated into landslide models through changes to soil moisture. In considering changes to vegetation, the compounding considerations of changes to vegetation patterns due to temperature, precipitation, and competition is difficult to quantify (Cannone et al., 2007). The impacts of climate change on vegetation are less often incorporated into landslide models, though changes in forest makeup associated with climate change have been considered previously and do impact slope stability (Bernardie et al., 2021). We aim to assess the changes to the risk of landslide initiation due to the impacts of climate change on vegetation in the Colorado Front Range by applying a probability model.

1.2 Case Study

In September 2013, a storm stalled over the Front Range in Colorado. The corresponding rainfall, which lasted almost a week, led to flooding and landslides that resulted in several deaths, thousands of houses being damaged, and millions of dollars in damage to infrastructure. More than 1,300 landslides were initiated by the precipitation event, some traveling as far as 5 km (Coe et al., 2014). The storm was unprecedented in the Front Range due to both the large spatial extent and the duration it remained stationary (Coe et al., 2014). The impacts of the event provide an opportunity to consider regional impacts of precipitation events on landslide susceptibility in the Front Range. The mapped landslides from the event are used to assess model performance for both models applied in this study.

1.3 Structure of This Report

This report is divided into two analyses. The second chapter considers a probabilistic landslide initiation model, which includes consideration of soil moisture downscaling, sensitivity of landslide initiation to landslide model inputs, and a comparison of methods to analyze a probabilistic model. These analyses guide an evaluation of the impacts of climate change on landslide susceptibility in the Front Range. The third chapter presents the development, calibration, and implementation of a landslide runout model. A brief synthesis is provided in the conclusion.

2. LANDSLIDE SUSCEPTIBILITY IN CURRENT AND FUTURE CLIMATES IN THE COLORADO FRONT RANGE

2.1 Introduction

Landslides pose a risk to society in the form of property and infrastructure damage, personal injury, and loss of life. Petley (2012) estimated that 2,600 landslides resulted in more than 32,000 deaths worldwide between 2004 and 2010. With climate change influencing multiple variables that impact slope stability, the distribution of predicted landslide locations is expected to change (Alvioli et al., 2018; Bernardie et al., 2021; Kim et al., 2015) thus impacting the areas, infrastructure, and people that are susceptible to them. Haque et al. (2019) evaluated landslides that resulted in loss of life and determined that between 1995 and 2014, there was an increasing trend of loss of life due to landslides and increased landslides from extreme precipitation events. There is a global trend of increasing high-intensity precipitation events (Wuebbles et al., 2017), and this shift toward extreme precipitation is projected to increase in frequency under climate scenarios (Pendergrass and Knutti, 2018), thus increasing the potential for precipitation-induced landslides. This means that determining which areas of the landscape are susceptible to landslides under current and projected climate scenarios is increasingly important to minimize risk to people and infrastructure.

Quantifying landslide risk requires considerations of the spatial and temporal aspects affecting landslide susceptibility. Landslide risk can be calculated from multiplying frequency, which is how often a landslide will occur, with vulnerability, which is the damage to infrastructure and people once they occur (van Westen et al., 2006; Salvati et al., 2010). Landslide initiation models can also be combined with landslide runout to distinguish the full area susceptible to landslides, leading to the ability to fully capture and represent landslide risk.

Factors that impact slope stability include topography, soil texture, and vegetation. Accurate topographic data are readily available for most areas. Soil texture, which can impact soil moisture and pore water pressures that lead to slope failures, can be difficult to quantify. Pauly et al. (2020) found similar model performance for a soil moisture downscaling model applying spatially averaged soil texture as applying spatially varied texture due to uncertainties. Vegetation includes both vegetation cover and type of vegetation. Vegetation indirectly impacts slope stability by influencing soil moisture, thus affecting the hydrologic component of landslides, and directly impacts slope stability through vegetative cohesion forces. Grieco et al. (2018) found that due to uncertainties and misclassifications in global land use data, model performance for a soil moisture downscaling model did not improve by incorporating land use. These studies demonstrate the difficulty in representing precipitation-induced landslides due to the uncertainties in available datasets. Applying deterministic landslide models can be challenging due to not representing the spatial uncertainties in vegetation and soil characteristics. Probabilistic approaches can improve upon a deterministic approach by incorporating and representing some of these uncertainties.

Particularly for risk analyses for regional applications, climate change is important to consider in landslide modeling to assess how landslide susceptibility and location might be impacted. The impacts of climate change on precipitation-induced landslides are primarily considered in one of two ways: through assessing changes to hydrologic drivers that could lead to landslides (Alvioli et al., 2018; Bernardie et al., 2021; Rosso et al., 2006; Mahoney et al., 2013) and through assessing the impacts of changes in vegetation or land use on landslide susceptibility (Bernardie et al., 2021; Gariano et al., 2018; Vanacker et al., 2003). Bernardie et al. (2021) evaluated the impacts of climate change on landslide susceptibility in the French Pyrenees by considering both hydrological and vegetative factors, incorporating the latter by adjusting cohesion values based on modeled changes in land use. They found that slope stability was impacted by land use types, where areas with decreased slope stability had increased human activity and

thus reduced vegetation and soil water content, which increased in some areas leading to decreased slope stability. The incorporation of climate change is challenging due to spatial and temporal uncertainties, such as when and how vegetation and precipitation will change, but is vital to consider in landslide studies to evaluate site-specific impacts. Due to the large uncertainties in model projections, a probability model can better address and represent changes to the environment.

Assessment of the impacts of climate change on vegetation must consider a range of variables, including changes to temperature, precipitation, and general vegetation dynamics (Cannone et al., 2007). The literature generally shows that vegetation is highly susceptible to climate change, though whether the population of a plant species increases or decreases is dependent on several other factors (Brice et al., 2020; Buma and Wessman, 2013; Wang et al., 2018). Stoddard et al. (2015) evaluated the impacts of climate change on forest management practices for a mixed-conifer forest in the San Juan Mountains in southwestern Colorado. They performed field analyses before treatment and five years after treatment. The field data were used as growth rates for a forest vegetation dynamics model used to evaluate the impacts of climate change on tree survivability. Their model predicted a significant reduction in tree populations under future climate scenarios, a conclusion that has been identified previously in the western U.S. (Tarancón et al., 2014; Charnie et al., 2016). Vegetation cohesion is dependent on species type (Norris et al., 2008), so a reduction in tree abundance or a shift in the distribution of plant species may cause changes in landslide susceptibility.

This study aims to probabilistically model landslide susceptibility to evaluate the impacts of climate change on vegetation distribution and vegetation cover. A probabilistic model is particularly appropriate for climate change analyses as spatial and temporal uncertainties can be directly incorporated into the model. We model landslide probability by combining a soil moisture downscaling model and a probabilistic landslide initiation model, aiming to have consistent variables and assumptions between the two models, and applying the model to the Colorado Front Range. We evaluate the Colorado Front Range because it experienced an extreme precipitation event in 2013 that induced 1,300 mapped landslides, providing an opportunity to evaluate regional landslide susceptibility under current and potential future climates. We model the impacts of climate change on predicted landslide locations based on changes to vegetation to answer the question of what the landslide risk would be for a storm event of similar magnitude to the base scenario under climate scenarios. We do not consider other impacts of climate change, such as fires, changing precipitation patterns both spatially and temporally, and impacts of pests or diseases on vegetation. The goal of the application is to consider regional impacts of changes to vegetation and the consequent relative impacts and uncertainties of landslide initiation.

2.2 Model Design

2.2.1 Soil Moisture Model

A key variable for precipitation-induced landslides is soil moisture. Accurately representing soil moisture for a storm event is difficult as it requires consideration of both antecedent conditions and the timing of rainfall and infiltration to determine when, if ever, a critical loss of slope stability occurs due to saturation (Wicki et al., 2020). Increased soil moisture increases pore water pressure and decreases matric suction, a combination that can lead to slope failure (Wicki et al., 2020). The failure mode caused by increased soil moisture is impacted by soil texture: finer soils are more likely to fail due to the loss of suction while coarser soils are more likely to fail due to positive pore water pressures (Collins and Znidarcic, 2004).

Providing accurate soil moisture values for a landslide model applied to large spatial extents requires the incorporation of the impact of heterogeneity of topography and vegetation on soil moisture. Soil moisture data resolution is often coarse, limiting its accuracy (Wicki et al., 2020; Coleman and Niemann, 2013; Vereecken et al., 2008). The Equilibrium Moisture from Topography, Vegetation, and Soil (EMT+VS) is

a soil moisture downscaling model (Ranney et al., 2015). The model is based on the research of Coleman and Niemann (2013), who studied the relationship between hydrologic processes (evapotranspiration, lateral flow, and deep drainage) and the spatial structure of soil moisture patterns. The EMT+VS model requires topographic, vegetative, soil, and climate variables to determine soil moisture downscaling (Ranney et al., 2015). Additional analyses have considered how best to downscale multiple coarse grid cells (Hoehn et al., 2017) and have incorporated the dependence of soil moisture on orographic precipitation and potential evapotranspiration (Cowley et al., 2017). The model can represent soil moisture in wet and dry states, can represent temporally unstable patterns, and explicitly incorporates topographic dependence of soil moisture (Coleman and Niemann, 2013). The model has been applied to a variety of areas, including the Colorado Front Range, and produced realistic results (Ranney et al., 2015; Timilsina et al., 2021). The use of the EMT+VS model allows for fine-resolution soil moisture to be incorporated into a landslide initiation model and an increased understanding about which variables impact soil moisture downscaling.

Four hydrologic processes are modeled in EMT+VS to downscale soil moisture: infiltration, deep drainage, lateral flows, and evapotranspiration (ET). As an abbreviated model description is provided in Timilsina et al. (2021) and a full model description can be found in Cowley et al. (2017), here we summarize only the main model equations.

The first component is infiltration, which is calculated by considering orographic dependence of precipitation and interception by vegetation cover:

$$F = \bar{P}\frac{[1+\tau(Z_{\#}-\bar{Z}_{\#})]\{1+\xi[S_{\#}\cos(R_{\#}-\upsilon)-\overline{S_{\#}\cos(R_{\#}-\upsilon)}]\}}{1+\tau\xi\left[\overline{Z_{\#}S_{\#}\cos(R_{\#}-\upsilon)}-\bar{Z}_{\#}\,\overline{S_{\#}\cos(R_{\#}-\upsilon)}\right]}(1-\lambda V) \tag{2.1}$$

where F [mm/d] is the infiltration rate, $\bar{P}$ [mm/d] is the spatially averaged precipitation, τ [1/m] specifies the elevation dependence of precipitation, ξ [-] determines orientation dependence of precipitation, $Z_{\#}$ [m], $S_{\#}$, and $R_{\#}$ are average elevation, slope, and aspect around the grid cell at the spatial scale of the orographic effects, υ [-] specifies the direction from which the orientation dependence is calculated, λ [-] specifies the efficiency of vegetation intercepting precipitation, and V [-] is the fractional vegetation cover. This expression neglects surface runoff because Pauly et al. (2020) showed that incorporating surface runoff does not improve the accuracy of the downscaled soil moisture due to associated data limitations.

The second component considers deep drainage by applying Darcy's law under gravity drainage and calculating unsaturated hydraulic conductivity based on Campbell (1974):

$$G = K_{s,v}\left(\frac{\theta}{\theta_s}\right)^{\gamma_v} \tag{2.2}$$

where G [mm/d] is deep drainage, $K_{s,v}$ [mm/d] is the vertical saturated hydraulic conductivity, θ_s is the saturated water content [-], θ is the volumetric water content or soil moisture [m^3/m^3], and γ_v [-] is the vertical pore disconnectedness index.

The third component, lateral flow, applies Darcy's law assuming that the lateral hydraulic gradient is a function of topographic slope. The hydraulic conductivity is again calculated based on Campbell (1974) and soil depth is assumed to be related to topographic curvature, specifically:

$$L = \delta_0\left(\frac{\kappa_{min}-\kappa}{\kappa_{min}}\right)c\imath K_{s,v}\left(\frac{\theta}{\theta_s}\right)^{\gamma_h}S^{\epsilon} \tag{2.3}$$

where L [mm/d] is the lateral flow, δ_0 [m] is the soil depth where topographic curvature is zero, κ_{min} [-] is the minimum curvature where the soil is present, κ [-] is the curvature, c [m] is the dimension of the digital elevation model (DEM) grid cells, ι[-] is the hydraulic conductivity anisotropy, γ_h [-] is the horizontal pore disconnectedness index, and ϵ is the function relating topographic slope with the horizontal hydraulic gradient.

The final soil moisture component is evapotranspiration (ET). Based on the work of Priestley and Taylor (1972), the equation incorporates ET based on the fractional vegetation cover V, shading effects, and water uptake from soil layers below the current soil layer:

$$E = \bar{E}_p[1 + \omega(\bar{Z} - Z)][\eta V + (1 - V)^{\mu}]\left[\frac{I_p}{1+\alpha}\left(\frac{\theta}{\theta_s}\right)^{\beta_r} + \frac{\alpha}{1+\alpha}\left(\frac{\theta}{\theta_s}\right)^{\beta_a}\right] \quad (2.4)$$

where E [mm/d] is the ET, $\bar{E}_p$ [mm/d] is the spatially average of potential ET, ω [1/m] incorporates the elevation dependence into the potential ET, η [-] is the transpiration from the modeled soil, μ [-] is the impact of shading within the cell, α [-] is the Priestley Taylor coefficient minus 1, I_p [-] is the potential solar radiation index (Dingman, 2002), β_r [-] specifies limitations on the radiation component of ET due to moisture, and β_a [-] specifies limitations on the aerodynamic component of ET due to moisture.

The final downscaled soil moisture is calculated as a weighted average:

$$\theta = \frac{\omega_G\theta_G + \omega_L\theta_L + \omega_R\theta_R + \omega_A\theta_A}{\omega_G + \omega_L + \omega_R + \omega_A} \quad (2.5)$$

where $\theta_G, \theta_L, \theta_R, \theta_A$ are soil moisture estimates calculated assuming deep drainage, lateral flow, radiative ET, and aerodynamic ET dominate the water balance and $\omega_G, \omega_L, \omega_R, \omega_A$ are the weights associated with each of the soil moisture estimates. The expressions for the analytical soil moisture estimates and weights can be found in Cowley et al. (2017). We apply a filtering method based on a shifting window approach to avoid unrealistic transitions between the coarse-level soil moisture data; the method and justification for its use is explained in Hoehn et al. (2017), who suggest it is most applicable in broader applications as it requires no additional variables.

2.2.2 Infinite Slope Stability Model

2.2.2.1 Landslide Initiation Model

Landslide modeling approaches vary significantly, though broadly they are sorted into empirically based (Metternicht et al., 2005; Lee et al., 2007; Wu and Sidle, 1995) and process-based analyses (Raia et al., 2014; Strauch et al., 2018; Vandromme et al., 2020). Empirical landslide models establish behaviors using landslide inventories while process-based models physically represent characteristics that lead to slope failures (Bernardie et al., 2021). As such, process-based models can be applied more broadly both spatially, e.g., in areas that do not have landslide inventories, and temporally, which allows the consideration of future climate scenarios (Bernardie et al., 2021). The infinite slope model, which compares the resisting forces to the stabilizing forces on a hillslope (Raia et al., 2014), is one process-based approach for representing landslide susceptibility. The ease of application makes it a widely applied physically based approach for landslide modeling across a range of spatial scales (Baum et al., 2010; Borga et al., 2002; Collins and Znidarcic, 2004; Mergili et al., 2014; Montgomery and Dietrich, 1994; Raia et al., 2014; Strauch et al., 2018; Wu and Sidle, 1995).

As establishing landslide risk requires considerations of the spatial and temporal aspects affecting landslide susceptibility, an important consideration is how to quantify uncertainties. Uncertainties can

arise from spatial heterogeneity, such as differences in vegetation and soil characteristics, and measurement limitations and errors (Strauch et al., 2018). One way to incorporate these uncertainties is to probabilistically model landslide initiation. Strauch et al. (2018) developed a probabilistic Monte Carlo model for landslide initiation using the infinite slope equation. Their model, developed in the Python-based Landlab package (Hobley et al., 2017; Barnhart et al., 2020), is aimed at applications applied to regional extents. The consideration of probability allows their model to incorporate spatial uncertainty within the variables. A probabilistic model can also consider landslide risk more broadly as it can incorporate temporal uncertainty. Borga et al. (2002) discusses how landslide probability models often have higher rates of predicted failures across a landscape than the number of observed landslides from a particular landslide initiation event. These predicted landslides could represent landslide hazard potential through time, thus incorporating an aspect of temporal uncertainty (Borga et al., 2002).

We apply the probabilistic Monte Carlo landslide model developed by Strauch et al. (2018), which is available in Landlab, a Python-based package. The model solves the infinite slope factor of safety (FS) equation based on the equation presented by Pack et al. (1998). The model uses a topographic grid and gridded input data to model slope stability, where an FS value less than 1 is unstable, FS greater than 1 is stable, and FS equal to one is quasi-stable. The FS equation compares stabilizing forces, which include internal friction and cohesion, to destabilizing forces, which include gravity and pore water pressure.

The FS equation is modified to explicitly incorporate volumetric soil moisture and suction stress, based on the work of Timilsina et al. (2021). The final equation applied in the Landlab model is:

$$FS = \frac{C_v + C_s}{h\gamma_s \sin S} + \tan\phi \cot S - \tan\phi \cot S \frac{\rho_w}{\rho_s}\left(\frac{\theta-\theta_r}{\theta_s-\theta_r}\right)\left(\psi_s \left(\frac{\theta_s}{\theta}\right)^b\right)\left(\frac{1}{h\cos S}\right) \quad (2.6)$$

where FS [-] is the factor of safety, C_v is vegetative cohesion [Pa], C_s is soil cohesion [Pa], h is the vertical soil depth [m], γ is the specific weight of soil [kN/m^2], S is the slope [m/m], ϕ is the internal friction angle [∘], ρ_w is the density of water [kg/m^2], ρ_s is the density of soil [kg/m^2], θ_s is the saturated water content [-], θ is the volumetric water content or soil moisture [m^3/m^3], θ_r is the residual soil moisture [m^3/m^3], ψ_s is the air entry matric potential [m of water], and b is the pore size distribution [-].

2.2.2.2 Monte Carlo Simulations

FS has limitations, particularly for risk analyses due to not incorporating uncertainties. A key source of uncertainty in landslide models is spatial variation in the datasets for properties that can impact slope stability. The goal of the probability analysis is to translate FS values into probabilities of failure [P(F)] to consider those uncertainties and allow for increased understanding about the relative P(F) between unstable sites.

To determine P(F), the model is run for a specified number of simulations. Uncertain variables have a specified distribution, and for each model simulation a randomly selected value is used based on the distribution. Upon completion, the number of times FS was recorded as unstable is divided by the total number of simulations to determine the P(F) for each cell in the domain.

The variables randomly sampled in the Monte Carlo simulation in the probability model are chosen due to mean spatial variations not capturing uncertainty in the data. The original model by Strauch et al. (2018) varied C_v, ϕ, h, and soil transmissivity. Due to the approaches we apply to represent FS and the uncertainties therein, we modify the model by removing the uncertainty related to soil transmissivity and h, treat θ as uncertain, and make different assumptions regarding the spatial variability of C_v and ϕ. Our FS equation does not include soil transmissivity, so we instead represent uncertainty in θ. h is not varied due to unknown variations in soil depth.

Cohesion considers both C_v and C_s. We assume C_s=0, an assumption applied previously (Strauch et al., 2018), and we evaluate this by considering soil textures, because soil cohesion is only present in areas with clay. We apply two approaches to modeling C_v based on the land use type. The cohesion for grid cells with land use classified as evergreens or shrubs is varied in the Monte Carlo simulation, assuming a positively skewed, triangular distribution for each vegetation type (Strauch et al., 2018), which requires a minimum, maximum, and mean value. Per the probability model development and literature review by Strauch et al. (2018), the use of a positively skewed distribution is consistent with previous landslide probability applications (e.g., Hammond et al., 1992). The final vegetative cohesion value for each cell is calculated:

$$C_v = (C)(V) \tag{2.7}$$

where C is the cohesion selected from the distribution and V is the fractional vegetative cover (Timilsina et al., 2021).

For areas with all other land use classifications, C is not varied in the Monte Carlo simulation. C for these other land uses is calculated by a linear relationship between C and V (Timilsina et al., 2021).

In considering the ϕ, we make the same assumption as in the original landslide model and assume a positively skewed, triangular distribution (Strauch et al., 2018). Triangular distributions have been applied previously in Monte Carlo analyses (Hammond et al., 1992), and the positive-skew as applied by Strauch et al. (2018) is based on a literature review. We assume a spatially constant value for the mean value. This is different from the original work by Strauch et al. (2018), who related ϕ to the soil texture resulting in a spatial variation. The classification of soil characteristics is difficult as outlined in previous studies. For example, Pauly et al. (2020) showed limited improvements when considering the EMT+VS model in applying spatially varied versus spatially averaged soil characteristics. The uncertainty is meant to incorporate the variability within the soil characteristics.

Due to the uncertainty in θ downscaling, we vary θ based on the uncertainty in the downscaling method. To determine values for θ, we analyze uncertainties incorporated into the EMT+VS downscaling that were in the form of a normal distribution, which we expand on in the methods (2.4.2.3).

Randomly selecting a value for each uncertain variable assumes independence, an assumption applied in the original probability model by Strauch et al. (2018). The downscaled θ does not consider land use type, thus can be considered independent of the C_v, and can be considered independent of ϕ as soil texture is not incorporated into EMT+VS. The independence between ϕ and C_v is based on assuming independence between vegetation type and soil characteristics. Boyce et al. (2005) found that changes in alpine plant community composition in Colorado were strongly correlated with water availability and temperature, but not soil texture.

2.3 Case Study

2.3.1 2013 Event

On September 9-13, 2013, a storm stalled over the Colorado Front Range, leading to precipitation exceeding 10 inches in some areas (Lukas et al., 2013). Woolridge et al. (2020) identified the storm as a midlatitude cyclone, which is a low-pressure storm that occurs over a large spatial scale. The storm was an anomaly for the area as it had a longer duration, lower rainfall intensity, and larger spatial extent than many historic storms (Lukas et al., 2013). The storm resulted in more than 1,100 mapped debris flows and more than 200 earth slides, debris slides, and rockfalls across 3,400 km^2 (Coe et al., 2014). The landslides mapped after the 2013 event are used to evaluate probability model performance, and the storm

conditions are used to better understand the potential changes to shallow landslide initiation under climate scenarios in the Colorado Front Range. Per the analyses of Coe et al. (2014), the landslides typically occurred at lower elevations, below 2,600 m. The landslides were biased toward slopes greater than 25 degrees, and 78% of the debris flows occurred on south facing slopes (SFS). Previous studies (McGuire et al., 2016; Rengers et al., 2016) identified reduced vegetation on SFS in the Front Range; Woolridge et al. (2020) demonstrated that SFS were more prone to saturation during the event, which contributes to the aspect dependence of the landslides.

2.3.2 Application Region

We apply the same study area as Timilsina et al. (2021), who evaluated a deterministic FS model, which allows us to apply many of the same assumptions and determine the difference between their deterministic and our probabilistic landslide model. The study area is located within the Colorado Front Range, which is in the Rocky Mountains in north-central Colorado (Figure 2.1). At 1,333 km^2, the study area contains 836 of the landslides mapped from the 2013 storm event. Within the study area, elevation ranges from 1,555 m to 3,481 m. The land use is predominately evergreen trees (NLCD, 2016). The soil is largely coarse in texture, including sandy loams, loams, and sandy clay loams (Soil Survey Staff, 2020).

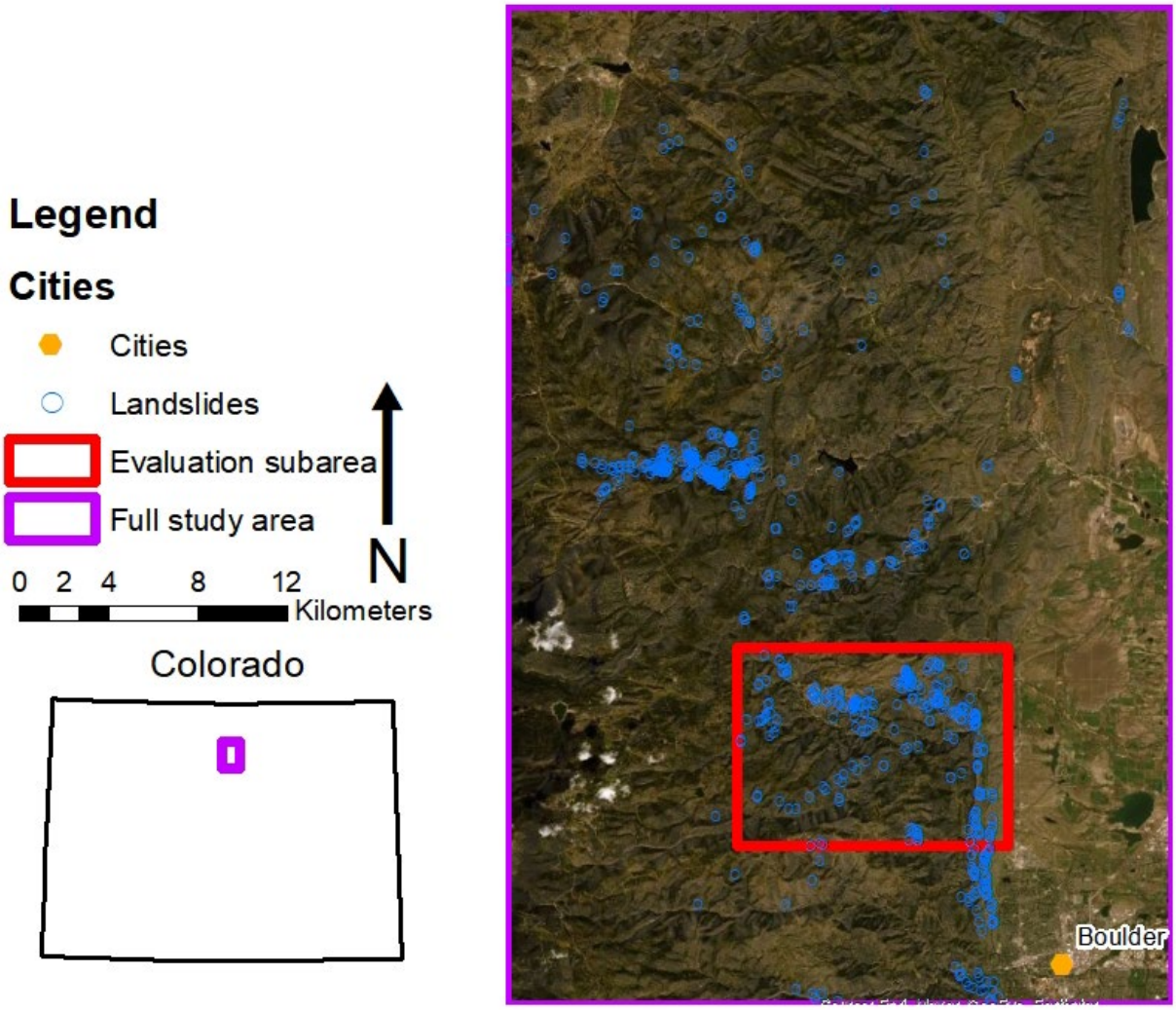

Figure 2.1 The study area located in the Colorado Front Range

The evaluation subarea is used for model sensitivity to consider the impacts of each variable to validate further model evaluations. The area is also used for more in-depth insights into visual behaviors for different modeling considerations. The subarea is contained within the full study area and is 109 km^2, about 10% of the total study area, and contains 298 landslides. The soil and land use characteristics in the evaluation subarea are representative of the full study area.

Within the full study area, there are two types of landslides we consider: debris flows and debris slides. To avoid confusion regarding terminology, debris flows are defined as precipitation-induced soil failures that lead to the fluidization of the surface (Varnes, 1978). Debris slides are slower and lack the fluidization present in debris flows. Coe et al. (2014) suggests that some debris flows are initiated from debris slides. Considering that the mechanisms of initiation between debris flows and debris slides are the same and the difference occurs in their movement after initiation, we assume our landslide initiation model can predict both mass movement events. We henceforth refer to these as landslides.

2.4 Methods

2.4.1 Data Acquisition

For spatially varied data, the probability model requires topographic information, vegetative characteristics, and soil texture data. The spatially varied EMT+VS model variable requirements include the soil moisture to be downscaled, topographic variables, and vegetative characteristics. We use ArcGIS 10.6 for topographic processing, MATLAB R2019a for the EMT+VS model, and Landlab version 1.9.0 in Python 3.7 for all other analyses.

2.4.1.1 Topography

Elevation data include a 1/3 arcsecond resolution DEM resampled in the study area to a 10 m resolution (USGS, 2015). The DEM is from 2015, which could result in changes from landslides from pre-event topography. Using the TauDEM GIS toolbox (TauDEM Version 5), a pit finder tool is applied to the DEM and the d-infinity contributing area and the slope are calculated. Aspect is calculated using ArcGIS tools. The spatially varied topographic variables as used in EMT+VS and Landlab are shown in Figure 2.2.

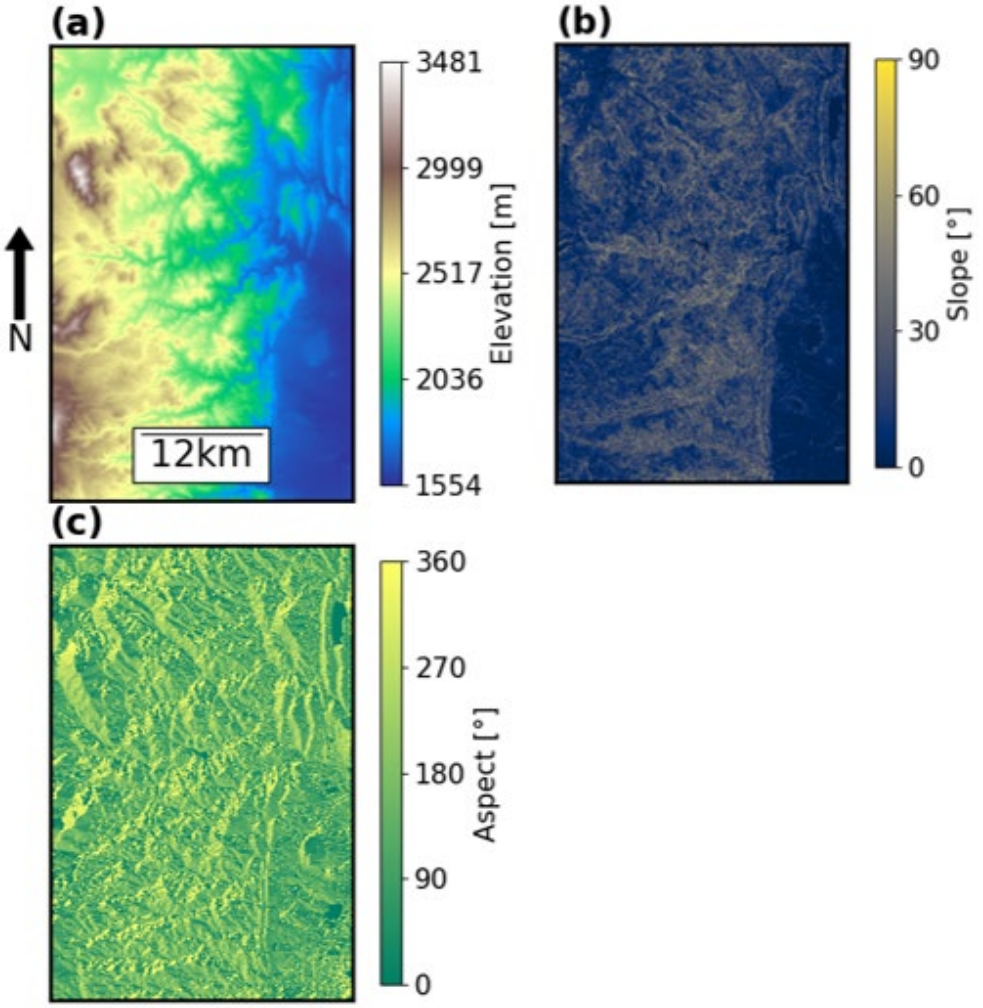

Figure 2.2 The (a) elevation, (b) slope, and (c) aspect for the EMT+VS and the probability models

2.4.1.2 Soil Texture

Soil data for the study area are from the Gridded National Soil Survey Geographic Database for the state of Colorado (Soil Survey Staff, 2020). The percent sand, silt, and clay are applied to the Module Soil Texture Classification in QGIS to determine the USDA soil texture classification (Massei, 2007).

2.4.1.3 Fractional Vegetation Cover

V is calculated using Landsat 5 data from USGS (2016). The data are from September 28, 2011, which was the last September date before the flood event without excess cloud cover, which obscures the land surface and reduces the accuracy of EMT+VS. The Normalized Difference Vegetation Index (NDVI) is calculated from Landsat data using (Rouse et al., 1973):

$$NDVI = \frac{NIR-RED}{NIR+RED} \tag{2.8}$$

where NIR is the near-infrared (Band 4 in Landsat 5) and RED is the visible band (Band 3 in Landsat 5). The NDVI is used to calculate V:

$$V = \frac{NDVI - NDVI_0}{NDVI_\infty + NDVI_0} \quad (2.9)$$

where $NDVI_0$ is the observed NDVI in pixels with no vegetation and $NDVI_\infty$ is the observed NDVI for fully vegetated pixels, with V ranging from 0-1. Based on the work of Timilsina et al. (2021), the minimum vegetation is 0.04 and the maximum vegetation is 0.7 (Figure 2.3).

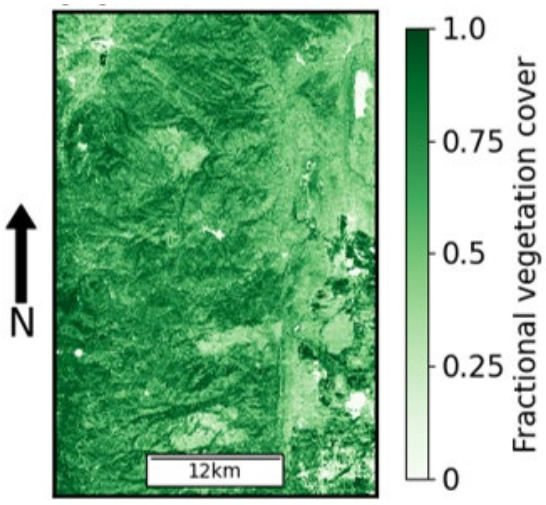

Figure 2.3 V as applied to the probability model

2.4.1.4 Soil Moisture

Large-resolution θ was downloaded from NLDAS Mosaic hourly data (Xia et al., 2012a; Xia et al., 2012b) for the Front Range at 00:00 MDT on September 12, 2013. This was one of the wettest periods during the storm (Timilsina et al., 2021) and a time when multiple landslides are known to have occurred (Coe et al., 2014). The Mosaic model is one of four θ models in NLDAS and has been shown to have better accuracy in Colorado than the other three (Xia et al., 2015). The Mosaic model has hourly soil data for various soil depths, and we apply 0-40 cm as this is closest to observed soil depth values in the Front Range (McGuire et al., 2016; Timilsina et al., 2021). Per comparisons conducted by Timilsina et al. (2021), the NLDAS model is biased during the storm event and as such, before processing in the EMT+VS model, 0.08 m^3/m^3 is subtracted from the values in the study area. The average θ value within our full study area is 0.27 m^3/m^3.

2.4.1.5 Land Use

Land use is obtained from the National Land Cover Database (NLCD) (NLCD, 2016). Initial analyses showed that most of the study area is evergreens and shrubs (Figure 2.4). Because of this, the evergreen and shrub classifications are maintained while all other land use types are combined for the calculations.

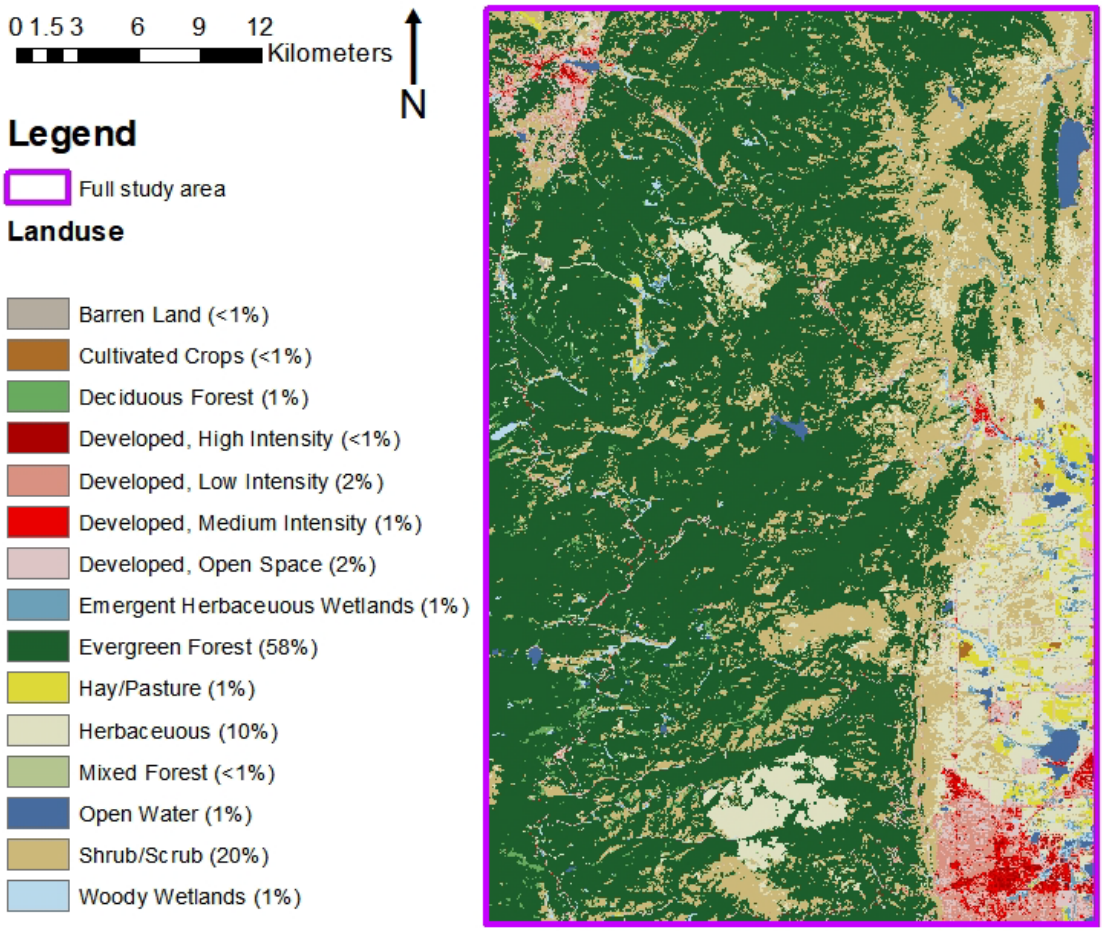

Figure 2.4 The land use types within the study area. NCLD 2016 imagery courtesy of the U.S. Geological Survey

2.4.2 Parameter Estimation

The values for the probability model are found in Table 2.1 and the uncertainty ranges for C_v, ϕ, and θ can be found in Table 2.2. The EMT+VS model variables can be found in Timilsina et al. (2021), where we applied the values from their combined scenario. We do not further discuss EMT+VS here.

2.4.2.1 Soil Characteristics

Soil texture is used to determine ψ_s and b using the values in Cosby et al. (1984). ψ_s is spatially varied based on USDA soil texture classification, with a range of 0.04 to 0.76 m of water (Figure 2.5 Air entry matric potential ($\boldsymbol{\psi_s}$) as applied to the probability modelFigure 2.5). b is a spatially averaged value based on soil texture, which results in a value of 5.14 for the full study area.

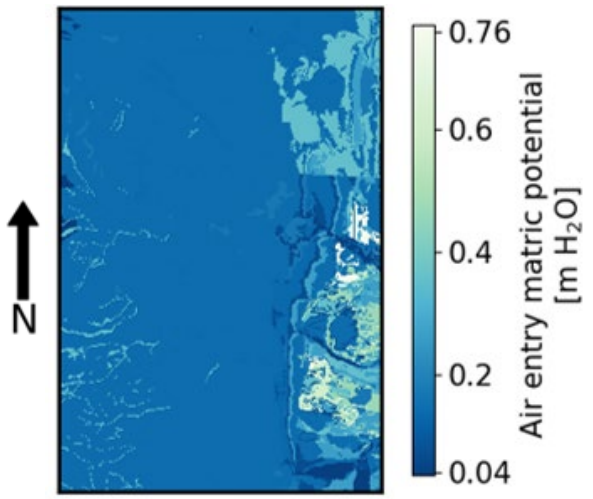

Figure 2.5 Air entry matric potential ($\boldsymbol{\psi_s}$) as applied to the probability model

2.4.2.2 Porosity

θ_s is held constant in the EMT+VS model with a recommended value of 0.409 (Grieco et al., 2018; Timilsina et al., 2021). For consistency between the EMT+VS model and the landslide model, we assume the same θ_s in the probability model.

2.4.2.3 Internal Friction Angle

ϕ is assigned a spatially constant mean value of 30° based on Timilsina et al. (2021). The variability for the Monte Carlo simulation is from Strauch et al. (2018), where the study area is comprised largely of sandy loams, a soil texture to what is in the Front Range. Strauch et al. (2018) generated distributions of ϕ based on the mode and ranges in the Unified Soil Classification System (USCS); we apply their ranges for uncertainty, where calculating the minimum and maximum values uses:

$$\phi_{min} = \phi_{mean} - 0.18 * \phi_{mean} \tag{2.10}$$

$$\phi_{max} = \phi_{mean} + 0.32 * \phi_{mean} \tag{2.11}$$

2.4.2.4 Soil Depth

V is used to calculate h in the probability model. McGuire et al. (2016) determined that south facing slopes in the Colorado Front Range have a mean h of 0.56 m while north facing slopes (NFS) have an h of 0.64 m. Following Timilsina et al. (2021), we represent h as a linear function of V where $h = 0.56 + (V)(0.64)$.

2.4.2.5 Soil Moisture

Soil moisture for the probability model requires specifying θ_r, which is the soil water content under which hydraulic conductivity is assumed to be zero (Pauly et al., 2020). Pauly et al. (2020) showed that incorporating θ_r into the θ downscaling in the EMT+VS model provides no improvements to the model. To maintain consistency between the EMT+VS model and the probability model, the same assumption is applied and the θ_r is assumed to be 0, akin to work by Timilsina et al. (2021).

The EMT+VS model has been calibrated to the Cache la Poudre watershed (Ranney et al., 2015), which is located farther north in the Colorado Front Range than the full study area. We apply the values from Timilsina et al. (2021), who applied the values from the Poudre watershed, estimated some variables based on data, and applied global values from Grieco et al. (2018) when the variables were not well represented by the Poudre dataset.

The probability model incorporates the stochastic variability in θ using a normal distribution, as this is the distribution used to represent θ variability evaluated by Deshon (2018), which requires specifying a mean and a standard deviation. We apply the so-called direct approach developed by Deshon (2018), which considered stochastic variations in θ. Their approach incorporates stochastic variations multiplicatively:

$$\theta_{final} = \theta(1 + a^* + b^*) \tag{2.12}$$

where θ_{final} is the soil moisture [m^3/m^3] applied directly to the probability model, θ is the downscaled soil moisture from the EMT+VS model, a^* is a spatially correlated random variable, and b^* is a spatially uncorrelated random variable.

Both a^* and b^* have means of zero. Deshon (2018) calibrated the standard deviations of a^* and b^* for the Cache la Poudre catchment and obtained $\sigma_a = 0.14$ and $\sigma_b = 0.12$. Deshon (2018) also determined that in the Cache la Poudre catchment, 41% of the variability is a time-stable pattern that is consistent with porosity variations. We assume the remainder (59%) of the stochastic variation is due to θ variability. Our final θ calculation becomes:

$$\theta_{final} = \theta[1 + (1 - 0.41)(a^* + b^*)] \tag{2.13}$$

If this variability results in $\theta > \theta_s$, θ_s is used instead, and negative values of θ are disallowed. We neglect spatial correlation because it is not relevant to the objectives of this study and greatly simplifies the generation of the random variations. We also neglect any large-scale stochastic variations in θ, which would not have been observed by Deshon (2018) within the small Cache la Poudre catchment.

2.4.2.6 Cohesion

Total cohesion is the combination of C_v and C_s. C_s is assumed to be 0 after determining that soils with high clay contents only made up a small portion of the full study area, a modeling approach applied in other probability model applications (Strauch et al., 2018).

C_v is varied based on both land use and V. For land use classifications of "other," the linear relationship applied between C_v and V assumes $C_v = (V)(6200)$ (Timilsina et al., 2021). This is a small percentage of the study area and contains a range of land use types that impact properties such as vegetation, infiltration, and compaction. While Timilsina et al. (2021) predominately justified this assumption for shrubs and evergreens, we only apply it to other land uses.

The areas with land use of evergreen and shrubs have C varied in the Monte Carlo simulations, where the minimum, mean, and maximum cohesion values for fully vegetated grid cells are found in Table 2.2. The values are the same as applied in Strauch et al. (2018). These values are consistent with literature values for shrubs and evergreen tree cover (Bernardie et al., 2021; Norris et al., 2008).

Table 2.1 Values for the landslide probability model

Variable	*Symbol*	*Units*	*Value(s)*
Pore size distribution index	b	-	5.14
Saturated water content	θ_s	m^3/m^3	0.409
Residual soil moisture	θ_r	m^3/m^3	0
Soil cohesion	C_s	Pa	0
Soil density	γ_s	kg/m^3	2,000
Air entry matric potential	ψ_s	m H_2O	0.04 to 0.76
Soil thickness	h	m	0.56 to 0.64
Specific weight of water	γ_w	kg/m^3	1,000

Table 2.2 The values for the uncertain variables

Variable		*Symbol*	*Units*	*Value(s) (min/mean/max)*		
Cohesion	Evergreen Forest	C	Pa	3,000	10,000	20,000
	Shrubs/Scrubs	C	Pa	1,200	4,000	10,000
	Other	C	Pa	0		6,200
Soil internal friction angle		ϕ	°	24.6	30	39.6
Soil moisture		θ	m^3/m^3	0.217 to 0.409		

Note: Soil moisture is the range for the downscaled soil moisture

2.4.3 Climate Change Estimation

We incorporate climate change by changing vegetation in the infinite slope stability model and determining the impacts on landslide susceptibility. We apply the impacts of climate change on vegetation as evaluated by Stoddard et al. (2015). Stoddard et al. used the Central Rockies Variant of Climate-Forest Vegetation Simulator (FVS) model, which models the growth and mortality rates of trees as well as competition within a stand under climate change scenarios (Crookston et al., 2010; Crookston, 2014). The model incorporates climate change by considering changes in site quality, growth, and adaptiveness of trees to climate change (Crookston et al., 2010). The model does not consider the direct impact of increased CO_2 on tree growth (Crookston et al., 2010).

In a field experiment, Stoddard et al. (2015) considered the impacts of various forest management practices (a control, thinning and burn, and a burn) for a mixed-conifer forest in the San Juan Mountains, southwestern Colorado. The management practices were evaluated five years post-treatment to establish growth rates, in which the tree basal area for the control scenario was 26.8 m^2/ha. The growth rates from the field data were applied in FVS. FVS was used to model vegetation dynamics from 2013-2063 for climate scenarios, including an RCP 4.5 and an RCP 8.5 scenario. The climate scenarios are based on the Intergovernmental Panel on Climate Change (IPCC) representative climate pathways (RCPs), which consider the impacts of climate change based on greenhouse gas emissions and land use trajectories and incorporates some socio-economic considerations, including expected technological improvements, to mitigate climate change impacts and anticipated population growth (van Vuuren et al., 2011). RCP 4.5 assumes a maximum radiative forcing of 4.5 W/m^2, or about 650 ppm CO_2 equivalent, by 2,100, while RCP 8.5 assumes a maximum radiative forcing of 8.5 W/m^2, or about 1,370 ppm CO_2 equivalent, and assumes increased energy use, increased population, and minimal technological improvements (van Vuuren et al., 2011). In their evaluation of the FVS model for the RCP 4.5 scenario, tree basal area decreased in their control scenario to 1.9 m^2/ha in 2063, a 93.0% change, and in the RCP 8.5 scenario,

tree basal area decreased to 1.3 m^2/ha in 2063, a 95.3% change (Stoddard et al., 2015). A follow-up field study in the area showed increases in shrubs (Korb et al., 2020).

This suggests that given the similar climate and forest composition between the study site and the Colorado Front Range, a potential effect of climate change is a reduction of the trees currently present and a potential increase in shrubs. Because the exact composition of the understory is unknown and tree species could migrate into the area not represented in the model, the change in vegetation due to climate change is modeled in our study by assuming that trees are replaced by shrubs. We implement this potential change by reducing the mean cohesion in areas with evergreen land use classification. Due to the similarities between the reductions in vegetation for the RCP 4.5 versus the RCP 8.5 scenario, we apply the values from the RCP 8.5. The base scenario uses a mean cohesion of $C = 10{,}000\ Pa$ for areas with evergreens, which becomes a reduction of $C_\Delta = -5{,}715\ Pa$.

Vegetation type can impact vegetation cover and as such, vegetation cover could be impacted by land use changes. In present-day land use, evergreens have greater V than shrubs in our study area (Table 2.3), so a transition from evergreen trees to shrubs in future climates may be accompanied by a reduction in vegetation cover.

Table 2.3 V based on land use and aspect

Variable	Aspect	Shrubs	Evergreens
Mean	NFS	0.51	0.78
	SFS	0.50	0.70
IQR	NFS	0.14 (0.44 - 0.58)	0.17 (0.71 - 0.88)
	SFS	0.13 (0.44 - 0.57)	0.16 (0.62 - 0.78)

Based on these values, we apply a weighting based on the difference between the mean V for shrubs and evergreens. We identify the difference between the mean V in shrublands and the mean V in evergreens on NFS and SFS. For example, based on the values in Table 2.3, the difference between the mean evergreens and mean shrubs on NFS is $0.78 - 0.51 = 0.27$. This is multiplied by the reduction in tree cover (95.3%) to identify the maximum reduction in vegetation cover while results in $V_{\Delta NFS} = 0.27 \times (0.953) = 0.26$. $V_{\Delta NFS}$ is the reduction to vegetation cover for areas on NFS with evergreens. For SFS, the maximum reduction to vegetation cover is $V_{\Delta SFS} = 0.18$.

The adjusted vegetation cover is used to re-calculate θ using EMT+VS. Both EMT+VS and the probability model are modified to use the vegetation from the base scenario to calculate h. This assumes that even if vegetation changes in a shorter period of time, h will take longer to adapt to changes in vegetation.

Due to the uncertainty in timing, vegetation composition during transitioning, and the potential for other species to migrate into the area that are not represented in the model, we evaluate the impacts of combinations of reduced cohesion and reduced vegetation cover on landslide initiation. We apply a full reduction, which assumes 95.3% of evergreens transition to shrubs, and a half reduction. The half reduction in cohesion values corresponds to around the year 2035 for the RCP 8.5 scenario; the RCP 4.5 scenario reaches this same reduction in 2038, while the full reduction corresponds to the year 2063. These reductions are applied by evaluating nine scenarios considering different combinations of vegetation cover and cohesion reductions (

Table 2.4). The values provided in the table include the title for each scenario, the mean value for cohesion, and the reduction for vegetation cover, which is aspect-dependent, for each scenario.

Table 2.4 The title for nine scenarios and the corresponding reduction in mean C and V as applied to areas with evergreen cover

		Cohesion adjustment		
		None (0.0)	Half reduction (0.46)	Full reduction (0.95)
Vegetation cover adjustment	None (0.0)	C(-0.0)_V(-0) C_{Δ}: No reduction $V_{\Delta SFS}$: No reduction $V_{\Delta NFS}$: No reduction	C(-0.46)_V(-0.0) C_{Δ}: -2858 Pa $V_{\Delta SFS}$: No reduction $V_{\Delta NFS}$: No reduction	C(-0.95)_V(-0.0) C_{Δ}: -5715 Pa $V_{\Delta SFS}$: No reduction $V_{\Delta NFS}$: No reduction
	Half reduction (0.46)	C(-0.0)_V(-0.46) C_{Δ}: No reduction $V_{\Delta SFS}$: -0.09 $V_{\Delta NFS}$: -0.13	C(-0.46)_V(-0.46) C_{Δ}: -2858 Pa $V_{\Delta SFS}$: -0.09 $V_{\Delta NFS}$: -0.13	C(-0.95)_V(-0.46) C_{Δ}: -5715 Pa $V_{\Delta SFS}$: -0.09 $V_{\Delta NFS}$: -0.13
	Full reduction (0.95)	C(-0.0)_V(-0.95) C_{Δ}: No reduction $V_{\Delta SFS}$: -0.18 $V_{\Delta NFS}$: -0.26	C(-0.46)_V(-0.95) C_{Δ}: -2858 Pa $V_{\Delta SFS}$: -0.18 $V_{\Delta NFS}$: -0.26	C(-0.95)_V(-0.95) C_{Δ}: -5715 Pa $V_{\Delta SFS}$: -0.18 $V_{\Delta NFS}$: -0.26

In considering the limitations of our approach to climate change, Stoddard et al. (2015) only modeled vegetation for their study site, which is a small area at a set elevation, so the elevation-dependence shift in trees cannot be further quantified from their research. This means that the elevation-dependence and aspect-dependence of vegetation is not incorporated into their modeled reduction in vegetation, and thus is not quantified in our application of changes to vegetation. There might be other large-scale variations on the impacts of climate change on vegetation not captured by Stoddard et al. (2015) in their study area. We do not consider other impacts of climate change, such as fires, changing precipitation patterns both spatially and temporally, and impacts of pests or diseases on vegetation. The goal of the application is to consider regional impacts of changes to vegetation and the relative impacts and uncertainties therein.

2.4.4 Analyses

For the probability analysis, each scenario is iterated 3,000 times, consistent with previous studies (Strauch et al., 2018). Abbaszadeh et al. (2011) demonstrated that more than 1,200 simulations had minimal model improvements for MC simulations of slope stability. In assessing the performance of the probability model, true positives (TP) are mapped landslide locations the model predicts as being unstable, false negatives (FN) are mapped landslide locations the model predicted as being stable, false positives (FP) are non-mapped landslide areas the model predicts as unstable, while true negatives (TN) are non-mapped landslide areas the model predicts as stable.

Due to a lack of information about landslide size in our applied inventory, we identify the eight surrounding cells for point where a landslide occurs. We identify the highest P(F) in the 30 m by 30 m area and assume that the highest P(F) is where the landslide occurs. The nodes applied in the buffer are not considered in the FP and TN fractions. This allows for considering some spatial variability regarding the exact starting location of the landslide.

2.4.4.1 Sensitivity

We perform a sensitivity analysis on the probability model to determine the relative impact of each variable on predicted stability. The sensitivity is applied to the evaluation subarea, which is representative of the full study area but can be evaluated more quickly. A base scenario is run assuming mean values for the uncertain variables, and returning the FS. Each variable is adjusted by 5%, 10%, and 15% of the full range of the variables within the study area, considering both an increase and a decrease. The FS is

identified for the observed landslides within the evaluation subarea. Only the observed landslides are used as these values have lower FS values and thus the adjustment on the variables consider the impacts on the probability model nearer to the stable/unstable transition. The nodes within each landslide buffer are evaluated and the node with the lowest FS is identified for the base scenario. These nodes are then used for the sensitivities. The mean FS value for all the landslides under the base scenario is compared to the mean FS value calculated with the perturbed variable values. By only considering observed landslides with the evaluation subarea, large-scale impacts of adjusting model variables might not be represented.

2.4.4.2 Probability Model Evaluation

To analyze a probability model, there are two key components to consider: the TP/FP and the probabilities for those nodes. In considering the TP/FP nodes, one decision is what constitutes a stable versus an unstable node. One approach that allows for considering a range of probability is calculating the area under the curve (AUC) for receiver operating characteristics curves (ROC-AUC) (Fawcett, 2005) and precision recall (PR-AUC) curves. ROC curves are a non-threshold-based analysis, meaning a range of probabilities can be considered and evaluated. In the case of landslide risk, identifying unstable areas as stable would be non-conservative and lead to potential loss of infrastructure and loss of life, while on the other extreme, identifying areas that are stable as unstable would potentially lead to increased costs for preventative measures (Beguería, 2006). ROC curves can be used to determine a probability threshold to apply in risk analyses and can establish whether a model distinguishes stable from unstable in the landscape (Beguería, 2006). ROC curves compare the true positive fraction (TPF) to the false positive fraction (FPF) (Beguería, 2006; Vakhshoori and Zare, 2018):

$$TPF = \frac{TP}{TP+FN} \tag{2.14}$$

$$FPF = \frac{FP}{FP+TN} \tag{2.15}$$

To develop an ROC curve, TPF is plotted on the vertical axis and FPF is plotted on the horizontal axis. A series of probabilities are chosen, 0, 0.1, etc., as the "threshold" that determines whether an area is stable or unstable and the corresponding TPF and FPF are determined, resulting in a line that extends from the lower left (0,0) to the upper right (1,1). A line of randomly selected values will be a straight line from the bottom left to the upper right. The AUC ranges from 0.5 to 1, where a line closer to the upper left corner, and thus a higher AUC value, is a better performing model that can distinguish between stable and unstable areas (Beguería, 2006).

A PR curve compares recall, same as the TPF, to precision (Davis and Goadrich, 2006):

$$Precision = \frac{TP}{TP+FP} \tag{2.16}$$

The same approach is taken as an ROC curve by identifying a series of probabilities and the corresponding precision (vertical axis) compared with recall (horizontal axis). A better performing model will be toward the upper right corner. A PR curve differs from an ROC curve in that it does not consider the TN fraction. This allows behaviors in highly skewed datasets to be more distinguishable (Davis and Goadrich, 2006).

Considering the number of nodes in the TP/FN/FP/TN fractions applied to the ROC and PR curves is important. The imbalance ratio is calculated for all ROC-AUC and PR-AUC analyses (Zhu et al., 2020):

$$Imbalance\ ratio = \frac{N_{major}}{N_{minor}} \tag{2.17}$$

where $N_{major}$is the number of values in the majority class and N_{minor} is the number of values in the minority class. For the ROC analysis, the N_{major} is the FP plus the TN, while for the PR analysis, N_{major} is the FP. In both analyses, the N_{minor} is the TP plus the FN. The imbalanced ratio is used to consider both model performance and whether the model is being well represented by the analyses. In general, a large imbalance ratio can negatively impact model performance by having decreased ability to correctly classify variables (Zhu et al., 2020). Our analysis uses the imbalanced ratio to evaluate differences in the ROC-AUC and PR-AUC to determine whether the model distinguishes between stable and unstable locations and which selection of nodes should be applied to future analyses.

We compare areas predicted to fail with those predicted to be stable to evaluate whether there are significant differences between stable and unstable locations in the spatially varied variables. We apply the two-sided non-parametric Mann-Whitney statistical test, which considers the null hypothesis P(x > y) ≠ 0.5, where x and y are groups that in our application are the FP and TN values (Helsel and Hirsch, 2002, Fagerland and Sandvik, 2009). The test analyzes inequality to determine whether one group is biased toward having higher values than the other group (Helsel and Hirsch, 2002).

2.5 Results

2.5.1 Model Sensitivities

We perform a sensitivity analysis of the model by evaluating the ratio of the FS of observed landslide nodes using the perturbed values to the FS of landslides nodes using the base scenario (Figure 2.6). The variables shown are calculated or estimated in our analyses. The landslides are located in the evaluation subarea, though percent changes are based on the variable ranges from the full study area.

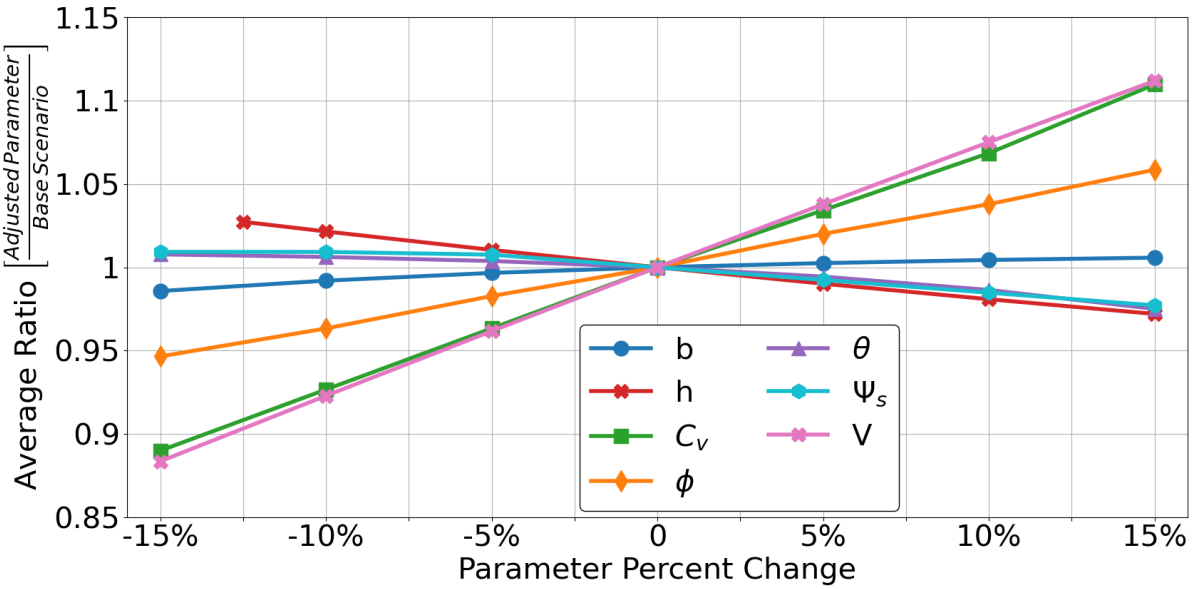

Figure 2.6 The model sensitivity compares the percent change in each variable to the average of the adjusted FS values divided by the base scenario FS values for all landslides in the evaluation subarea

Within the considered ranges, many variables shown in the sensitivity analysis demonstrate relatively linear sensitivity. θ and ψ_s are the exceptions and show some slight non-linearity. V and C_v have the largest impact on the results, and ϕ is also impactful. This implies that probabilistic modeling accounting for uncertainties in C_v and ϕ is critical for assessing the probability of landslide initiation. It also supports considering the impacts of changes to C_v and V under climate change scenarios. While variables demonstrate relative linear responses within the ranges we consider, outside of the chosen percent adjustments the variables could have non-linear responses that would yield different results.

2.5.2 Base Scenario Evaluation

The 2013 storm event is used to evaluate model performance. The downscaled θ provides increased accuracy and incorporates aspect and elevation dependence (Figure 2.7).

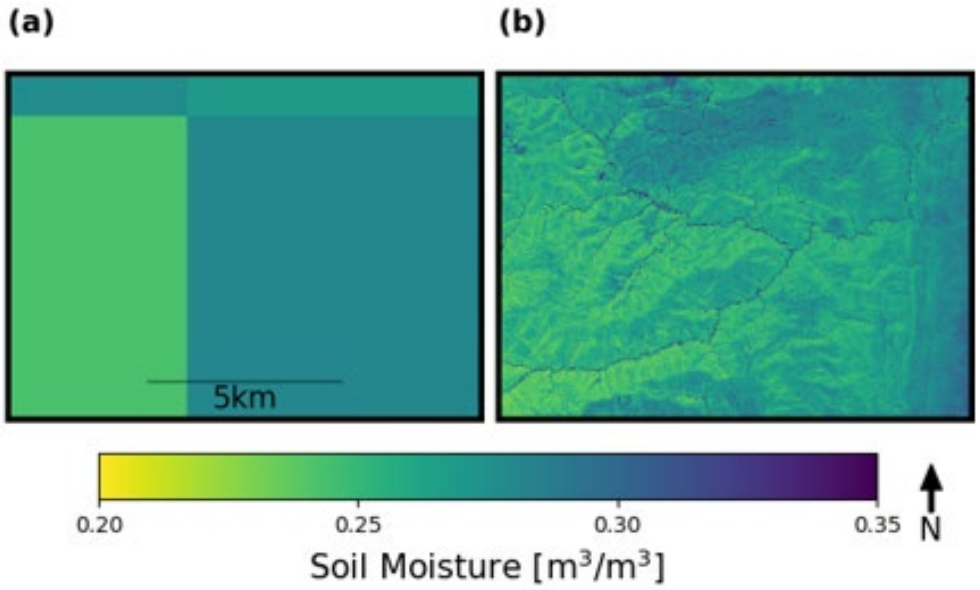

Figure 2.7 The non-downscaled **θ** (a) compared to the downscaled **θ** (b) for September 12, 2013, at 00:00 MDT. The evaluation subarea is shown for better visualization of variations within the downscaled.

To evaluate model performance, we compare the probabilistic model results to a deterministic FS approach, which uses mean values for C_v, θ, and ϕ. We evaluate a range of FS values to determine the percent of both observed landslides and FP values that are predicted as unstable by that threshold (Figure 2.8). We evaluate thresholds for P(F) in a similar manner for comparison.

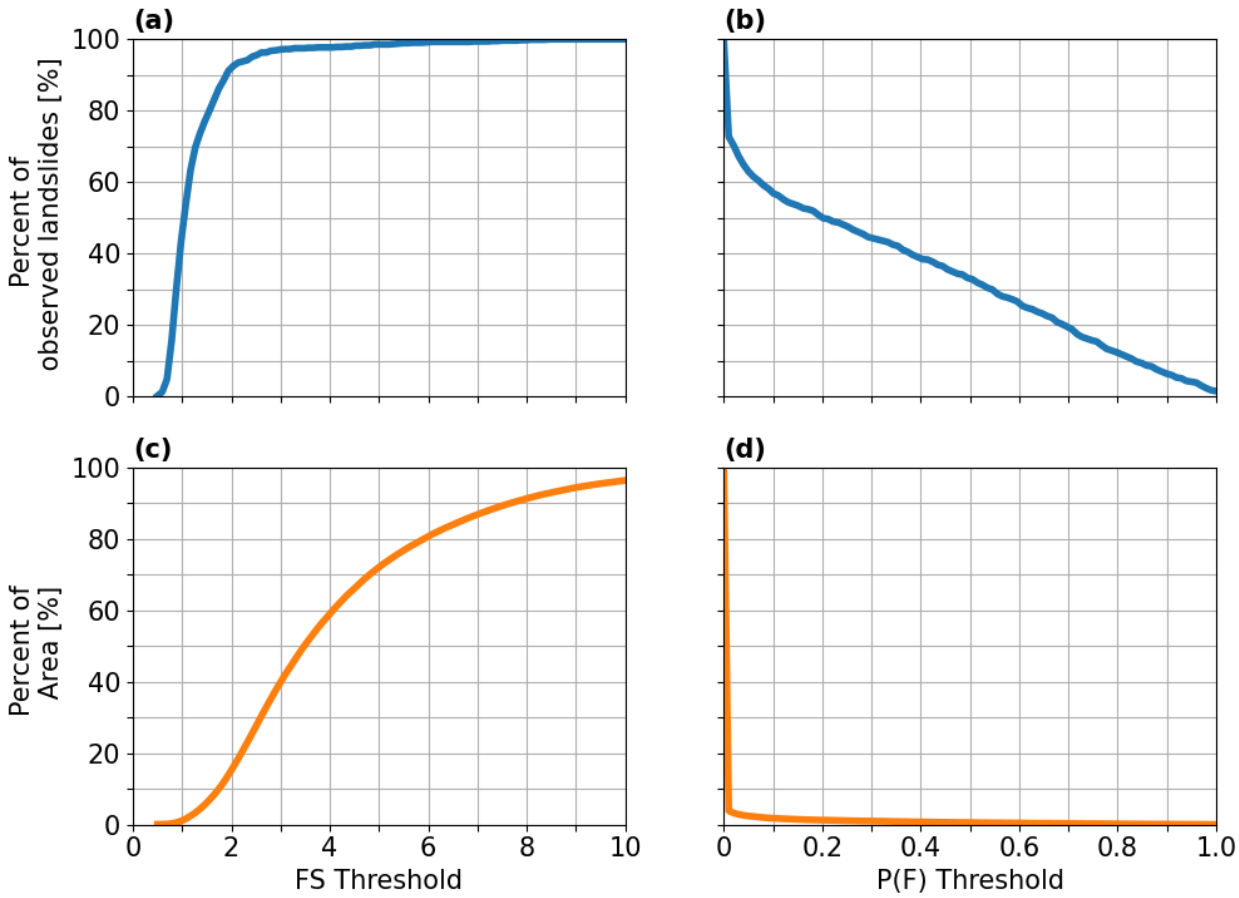

Figure 2.8 The percent of observed landslides predicted by different FS (a) and P(F) (b) thresholds and the FP values predicted by different FS (c) and P(F) (d) thresholds

Figure 2.8a shows that when FS = 1 is the instability threshold, 42% of the observed landslides are predicted. In comparing this to probability thresholds, P(F) > 0 results in 79.6% of the observed landslides being predicted (aboveFigure 2.8b). Using FS = 1 results in less than 1% of the full study area having FP values (Figure 2.8c) while using P(F) > 0 results in 5.8% of the full study area having FP values (Figure 2.8d).

Additionally, we compare the deterministic approach to P(F) through developing an ROC curve (Figure 2.9). The FS ROC curve is determined in the same way as the P(F) curve by choosing different thresholds and determining the TPF and FPF, which considers which nodes are equal to or less than the FS threshold. We randomly select 100,000 nodes for the FPF fraction for FS and the P(F). The ROC curve is similar between the two, with the AUC being slightly higher for the FS at 0.942 as compared with the P(F) value of 0.890.

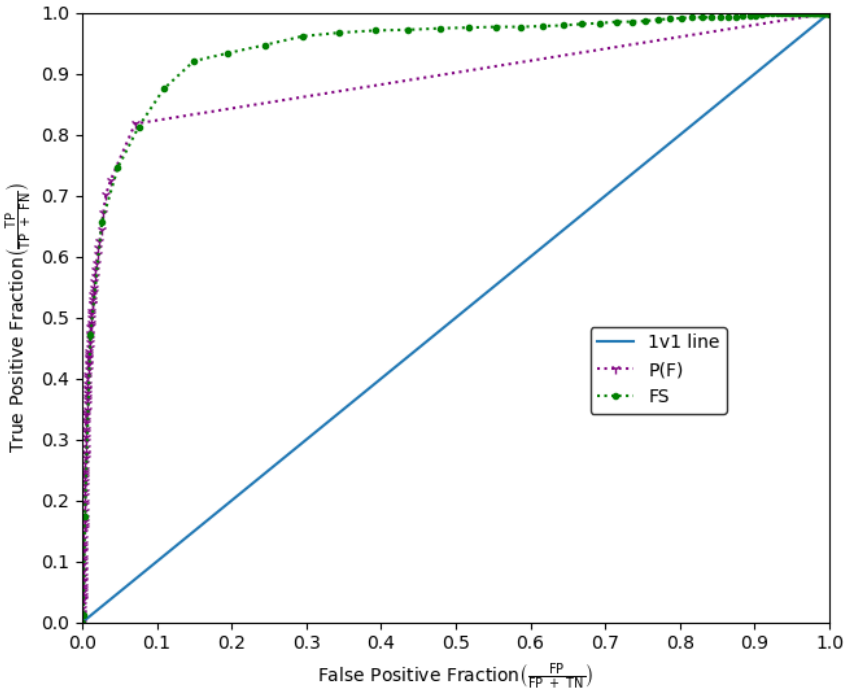

Figure 2.9 The ROC curve comparing FS and P(F)

We further evaluate P(F) by determining the ROC-AUC and the PR-AUC (Table 2.5). In addition to considering all the nodes in the study area, we consider "valid" nodes, which are nodes with a slope greater than 5°. We randomly select nodes (1000, 10,000, and 100,000) greater than 5° to consider the impacts on the ROC-AUC and the PR-AUC with different imbalanced ratios. All ROC-AUC values are greater than 0.5, indicating the model performs better than random. In considering the imbalanced ratios, there is little difference between the ROC-AUC values associated with different imbalanced ratios. The impacts are more obvious for the PR-AUC, which shows that the lower the imbalanced ratio the higher the AUC. The random selection of 10,000 nodes leads to an imbalanced ratio closest to 1. The analysis highlights that quantifying probability results is challenging as the imbalanced ratio impacts how well the model performs within ROC and PR analyses. Determining both the initial imbalanced ratio to understand how different the ratios are between TPF and FPF values and selecting a scenario to consider the imbalanced scenario closest to 1 are performed here to evaluate the differences and be able to move forward with analyses.

Table 2.5 Results for different evaluation approaches

Count	Imbalanced Ratio, ROC	Imbalanced Ratio, PR	ROC-AUC	PR-AUC
1,000	1.20	0.08	0.891	0.904
10,000	11.96	0.82	0.892	0.682
100,000	119.62	8.48	0.890	0.214
Valid Nodes	11940.35	919.45	0.894	0.023
Full Area	15012.78	919.45	0.890	0.023

To further evaluate the probabilistic model, we consider the P(F) associated with the unstable areas (Figure 2.10). In considering the full range of probabilities, the values are skewed toward low values with most failures having a P(F) < 0.2. There is also a higher concentration of values at probabilities near 1.

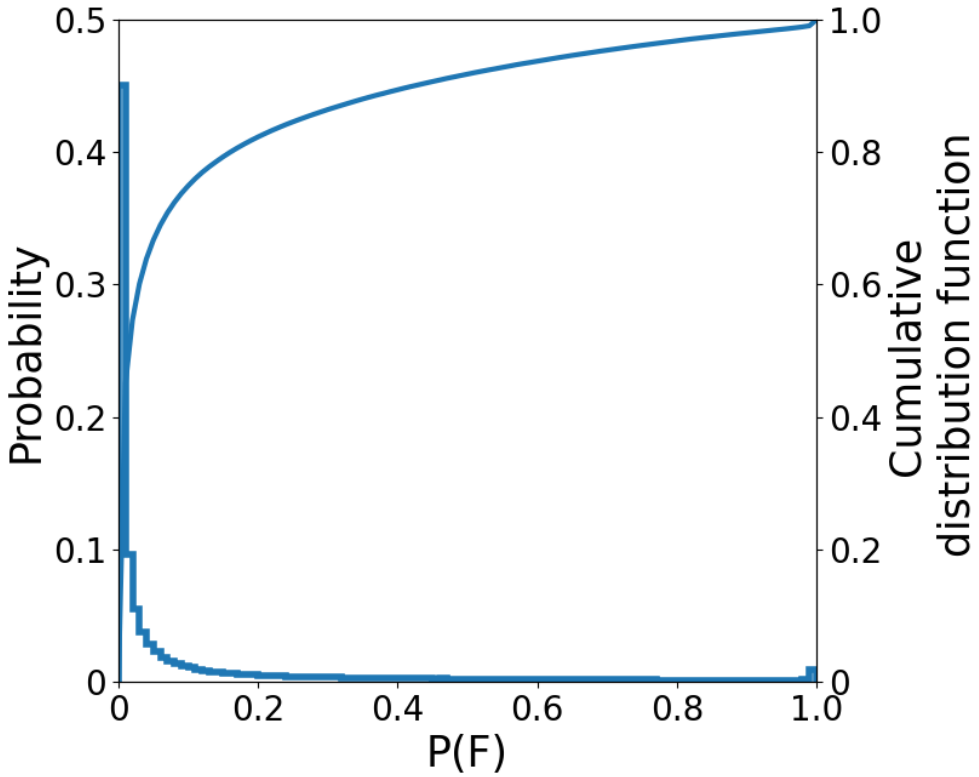

Figure 2.10 A histogram showing the P(F) for the full study area

To identify differences and similarities between distributions, we consider observed landslides, P(F) > 0.8, 0.8 > P(F) > 0, and P(F) = 0. A histogram shows the distribution for elevation, slope, V, θ, and aspect (Figure 2.11). The 0.8 > P(F) > 0 and P(F) = 0 fractions are from a random selection of 10,000 nodes with slopes greater than 5° and the P(F) > 0.8 fraction contains all nodes with that probability of failure (~24,000).

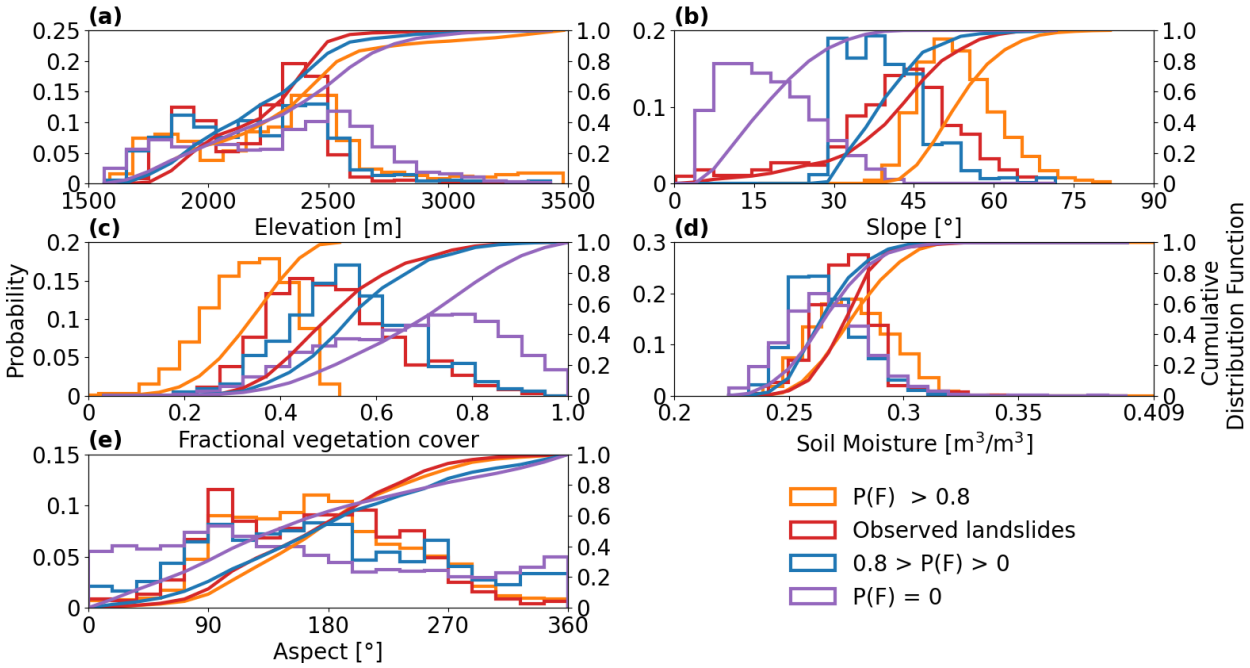

Figure 2.11 Histograms (a-e) and cumulative distribution functions comparing the observed landslides with the FP and TN fraction node types across the spatially varied variables

Based on the histogram, there are three key observations. The first observation is that the 0.8 > P(F) > 0 fraction has similar distributions to the observed landslides, suggesting that the model accurately represents the conditions leading to failure. The second observation is that the 0.8 > P(F) > 0 fraction is unique from the P(F) = 0 fraction by having a higher slope and a lower V. The third observation is that the P(F) > 0.8 fraction is unique from the other three fractions, exhibiting lower V and higher slope.

To consider the differences between 0.8 > P(F) > 0 and P(F) = 0, we apply a Mann-Whitney test to compare the distributions to see whether there are significant differences (Table 2.6). Higher probabilities, P(F) > 0.8, are not considered in the Mann-Whitney test as the histogram analysis suggests more extreme differences than between lower probabilities and P(F) = 0. We apply the same 10,000 nodes shown in the histograms for the statistical test. The Mann-Whitney test indicates statistical significance between the 0.8 > P(F) > 0 and P(F) = 0 fractions for both slope and V.

Table 2.6 Mann-Whitney test comparing 0.8 > P(F) > 0 and P(F) = 0

	Elevation	Slope	Vegetation	Soil moisture	Aspect
Statistic	2535622	6007425	1746586	3023307	3510283
P value	0.00	0.00	0.00	0.35	0.00
Significance	< 0.05	< 0.05	< 0.05	-	< 0.05

Note: The n values for the two variables are: 0.8 > P(F) > 0, n = 664 and P(F) = 0, n = 9308

We determine the percent of each type of land use on NFS versus SFS for the full study area, observed failures, and predicted failures (Figure 2.12). Within the full study area, there are differences between SFS and NFS. SFS have more shrubs while NFS have more evergreens. In predicted failures, there is both an aspect-dependence and more predicted failures in areas with shrubs. In considering observed landslides, there is more prevalence on SFS and in areas with shrubs.

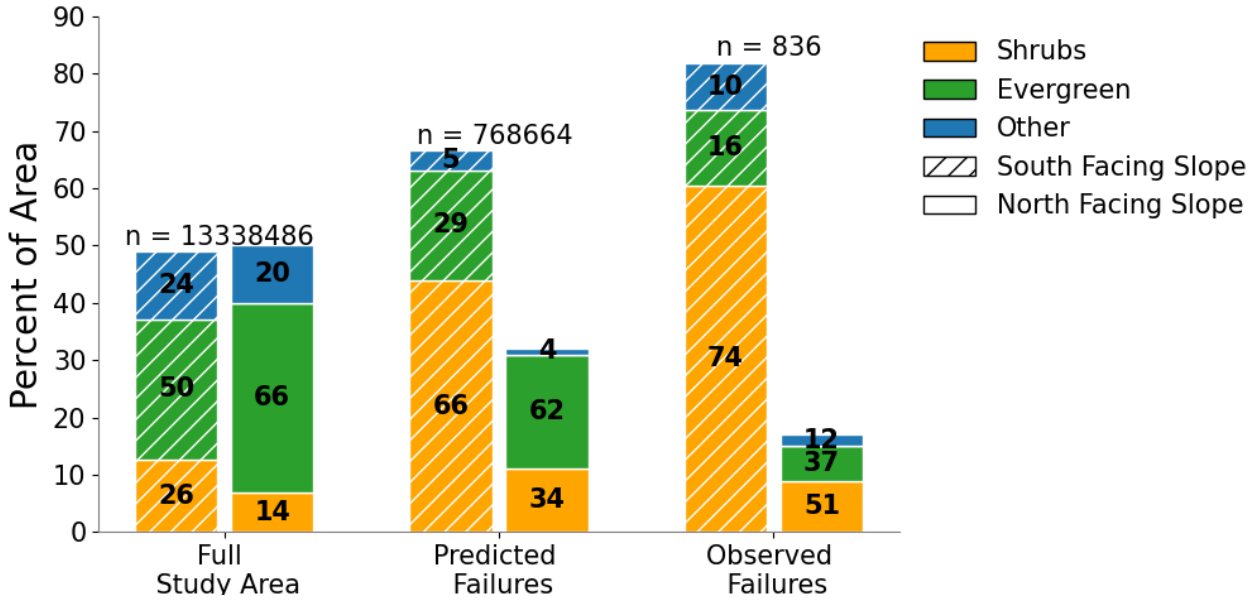

Figure 2.12 The land use types within the general landscape, where the sum of the SFS and NFS for each subset is equal to 1

2.5.3 Climate Change Modeling

Using the sensitivity results and the current understanding of landslide behaviors within the study area to guide the analyses, we consider the impacts of climate change, and the uncertainty therein, on landslide risk. The original, half, and full reduction adjustment applied to V and the corresponding C_v values for the scenarios C(-0.0)_V(-0.0), C(-0.46)_V(-0.46), and C(-0.95)_V(-0.95) are shown in Figure 2.13.

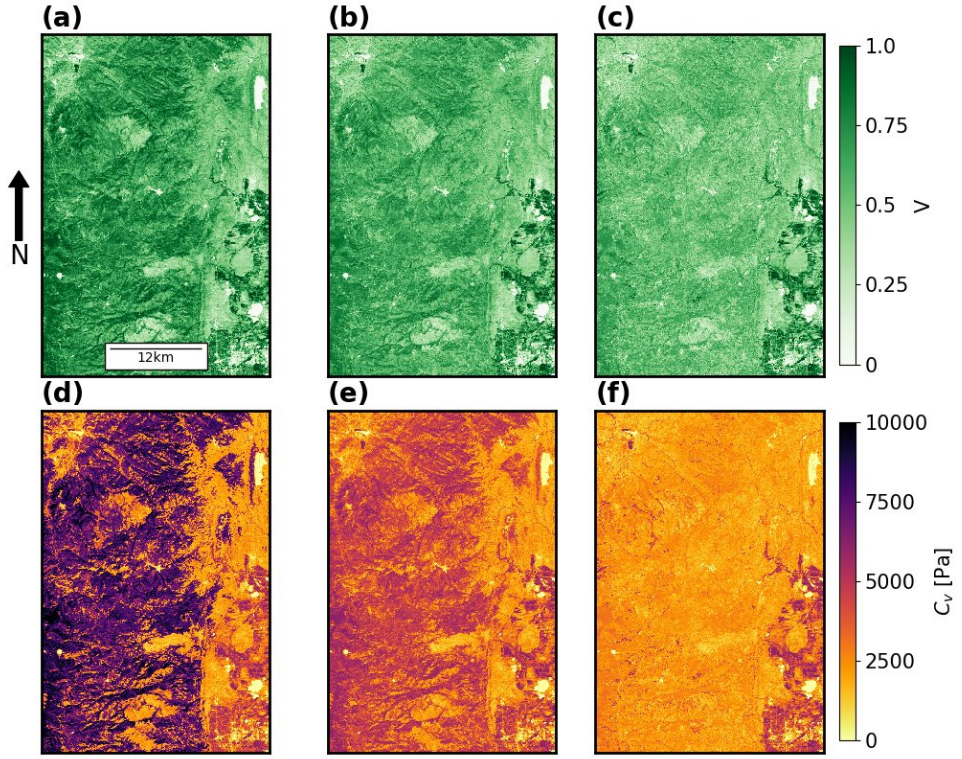

Figure 2.13 The base scenario V (a), a half reduction (b), and a full reduction (c) and C_v values for C(-0.0)_V(-0.0) (d), C(-0.46)_V(-0.46) (e), and C(-0.95)_V(-0.95) (f)

The percent increase in areas with P(F) > 0 between the base scenario and the climate scenarios are identified in Table 2.7. All climate scenarios lead to an increase in areas with P(F) > 0. Reducing cohesion leads to more unstable areas than reducing vegetation cover.

Table 2.7 Percent increase between the scenario and the base scenario for the area with P(F) > *0*

		Cohesion adjustment		
		No Reduction	Half Reduction	Full Reduction
Vegetation cover adjustment	No Reduction	Base scenario	33%	84%
	Half Reduction	17%	50%	102%
	Full Reduction	38%	72%	119%

To evaluate which areas of the landscape transition from stable in the base scenario to unstable in the climate change scenarios, we select 2,000 random nodes with a P(F) > 0 in the climate scenarios and P(F) = 0 in the base scenario and compare the elevation, slope, V, θ, and aspect (Figure 2.14). As the climate scenarios lead to reductions in slope stability, all the areas with P(F) > 0 in the base scenario are still unstable in the climate scenarios. We show the climate scenarios with a full reduction in V, C, or both. The new unstable points have lower slopes than the base scenario. The aspect-dependence of the base scenario being biased toward SFS is no longer visibly present and there is more uniformity across aspects.

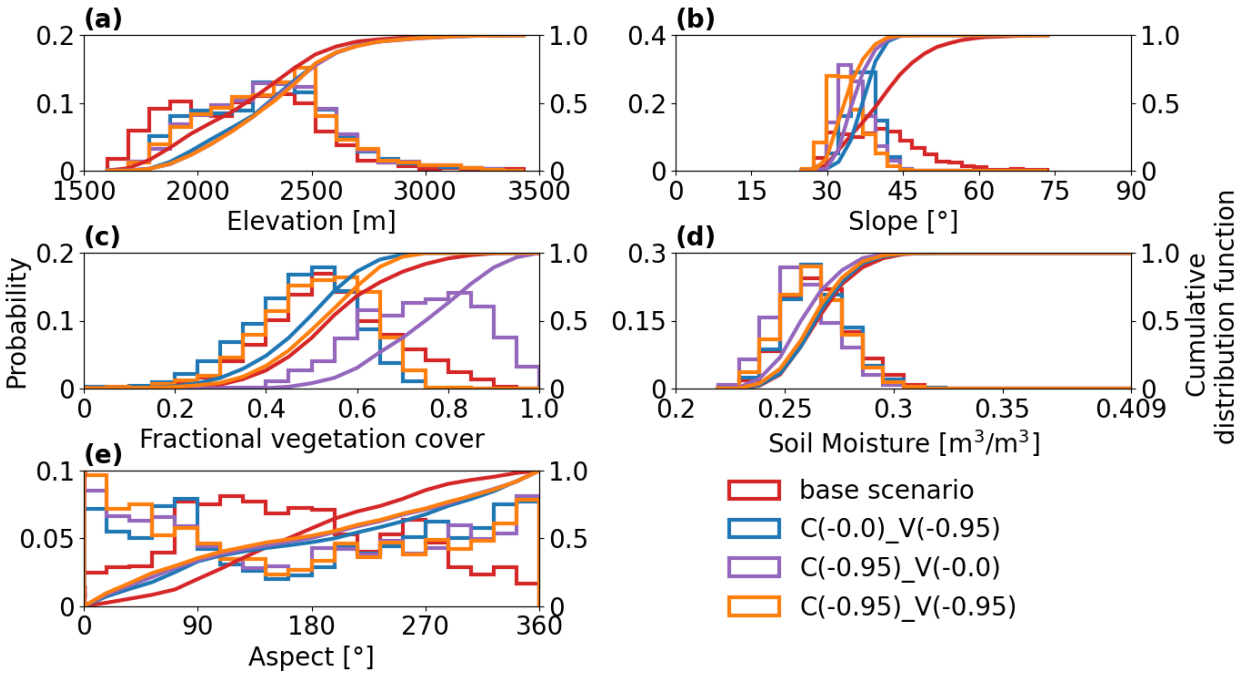

Figure 2.14 Histograms (a-e) and cumulative distribution functions comparing node types between the base scenario and some of the climate scenarios for nodes with P(F) > 0

To quantify the change to aspect-dependence, we determine the percent of new unstable areas in the climate change scenarios that are on SFS (Table 2.8). The climate change analyses correspond to a shift in the aspect-dependence of areas susceptible to landslides. In the base scenario, a minority of failures (33.5%) occur on NFS. In all climate scenarios, there is a shift to increased landslide susceptibility on NFS where greater than 60% of failures occur on NFS.

Table 2.8 Percent of areas with a P(F)>0 on NFS

		Cohesion adjustment		
		No Reduction	Half Reduction	Full Reduction
Vegetation cover adjustment	No Reduction	33.5% (n = 768628)	60.0% (n = 253368)	61.8% (n = 657398)
	Half Reduction	64.1% (n = 132915)	62.5% (n = 398520)	62.7% (n = 786261)
	Full Reduction	65.4% (n = 304602)	64.1% (n = 570586)	63.4% (n = 928671)

Note: n is the total number of nodes with a P(F) > 0 in the climate scenario and P(F) = 0 in the base scenario

We map the climate change scenarios by computing the climate scenario minus the base scenario (Figure 2.15). The shift in location of the climate scenarios is most obvious in Figure 2.15d, which considers a full reduction in both cohesion and vegetation. The least impacted scenario is Figure 2.15c, which assumed no reduction in cohesion and a full reduction in vegetation.

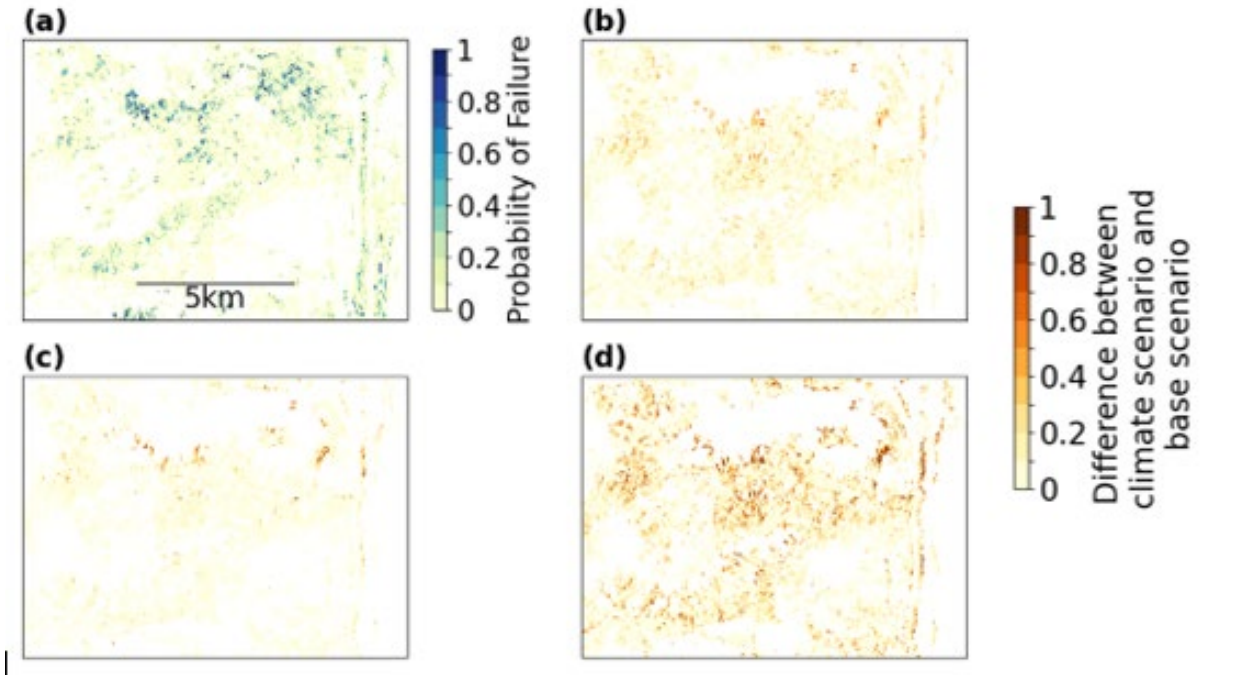

Figure 2.15 The P(F) within the evaluation subarea for the base (a) and the difference between the P(F) for the climate scenario and the base scenario for C(-0.45)_V(-0.0) (b), C(-0.0)_V(-0.95) (c), and C(-0.95)_V(-0.95) (d).

2.6 Discussion

We compare a probabilistic landslide model to a deterministic landslide model. Our results show that for our application, assuming that FS $\leq$ 1, indicates instability is less accurate at predicting observed landslides than the probability model. Because a number of observed landslides are around the FS = 1 threshold, slight uncertainties in the input variables can yield large differences in the fraction of observed landslides that are correctly predicted. This suggests that a probabilistic approach, which can incorporate variable uncertainties, is an improvement over the deterministic approach in terms of being able to better represent observed landslides.

In considering model performance and the predicted probabilities of failure under the base scenario, the probabilities associated with the P(F) > 0 fraction are typically low but with a spike at high probabilities that are associated with lower vegetation and higher slopes. These areas have distinct characteristics from the observed landslides, which suggests that other factors might be impacting the slope stability. The division in probabilities in our base scenario exhibits similarities to the probability model application by Strauch et al. (2018), who attributed the difference to an ecosystem transition. The prominence of very low probabilities is worth analyzing in future model evaluations to determine whether a broader distribution of probabilities is expected, as the probabilities themselves are not calibrated. Differences in our model between predicted and actual probabilities of failure might stem from variables being too high, leading to arbitrarily low probabilities. This is supported by the deterministic FS values for the median values having poor model performance, whereas previous deterministic FS analyses such as the model developed by Timilsina et al. (2021) had higher model performance and accurately predicted 72% of observed landslides in the same study area. Identifying whether P(F) values are genuine, and accurately capturing the return interval in terms of the relative impact of landslides, is challenging due to a lack of information in most landslide inventories about occurrence intervals (van Western et al., 2006). We consider the probabilities in our model relative, where areas with higher probabilities are considered more

unstable but we do not consider them to represent a return period of the landslides due to the lack of data for more than a single landslide initiation event.

In further considering the base scenario, we evaluate the ROC-AUC and the PR-AUC. Our model has an ROC-AUC value of 0.9 for the full study area. While similar ROC values are found in Gudiyangada Nachappa et al. (2019), literature values are typically lower, which highlights an important limitation of ROC analyses. ROC analyses require consideration of both the imbalanced ratio and the ratio of FP to TN. Our model has a high imbalanced ratio and a high ratio between FP and TN values; the combined effect leads to high AUC values in our model, which makes discrimination between model scenarios challenging. The tendency for ROC curves to provide overly optimistic results has been identified previously in other fields (Fu et al., 2019). To avoid some of the shortcomings of the ROC curve, we evaluate the PR-AUC, which is used less often in landslide literature (Bernardie et al., 2021). The differences in PR-AUC are more extreme between different imbalanced ratios, suggesting that under some conditions, PR curves can be used to understand variations between model scenarios; for example, quantifying the differences between calibration model runs. A recent study by Yordanov and Brovelli (2020) compared PR and ROC curves for landslide prediction and identified that if balanced data are used they perform comparably well for model evaluations, but imbalanced datasets led to improved performance based on the PR curves. This suggests that identifying the imbalanced ratio is important to determining the correct landslide probability model evaluation technique. The importance of accurately evaluating landslide models and additional model evaluation approaches has been presented elsewhere (Frattini et al., 2010; Guzzetti et al., 2006; Vakhshoori and Zare, 2018) and while ROC and PR curves provide insight into landslide models, combining them with other evaluation methods might be more appropriate depending on the goal of the analysis. We consider them appropriate in our application as we use them to discuss differences between a deterministic and probabilistic modeling approach, to show that our model can distinguish between stable/unstable locations, and as a comparative consideration to assess how to select FP/TN nodes for further evaluation.

In considering the observed landslides from the 2013 event, previous studies have shown that landslides in the Front Range have a strong aspect dependence (Coe et al., 2014). Recent studies have identified reduced vegetation on SFS in the Front Range (Rengers et al., 2016) and marginally lower soil depths on SFS than NFS (McGuire et al., 2016). Timilsina et al. (2021) found increased soil moisture on SFS, which is captured in the EMT+VS model. Within our study area, 82% of the observed landslides occur on SFS. Figure 2.12NFS have more trees (Coe et al., 2014), which results in higher cohesion and increased vegetation cover. This suggests that under current conditions, NFS are more stable but reductions in trees due to climate change could impact this stability.

Under climate change scenarios, we identified a shift in the areas susceptible to landslides in the Colorado Front Range. All our climate scenarios predict increased landslide susceptibility on NFS. The decreasing prominence of evergreens on NFS is one reason for this, as more areas on NFS are impacted by the adjustments than areas on SFS. The shift to increased instability with decreased vegetation has been modeled previously. Bernardie et al. (2021) modeled land use changes and predicted increased slope stability where forests increased and decreased slope stability in areas where the forest vegetation decreased. There is also increased landslide susceptibility at lower slopes within the climate data. The increase in areas with P(F) > 0 at lower slopes is due to the decrease in C_v and/or vegetation cover, resulting in reduced stability.

We consider the impacts of climate change through analyzing reductions to trees, which is a simplified approach to regional-scale impacts of climate change. We incorporate variation in the values we apply to consider some of the uncertainties about which type of vegetation will be present, the vegetation cover associated with the vegetation type, and the general vegetation dynamics that can impact slope stability. Further, we do not consider aspect, elevation, or spatial heterogeneity that impact changes to vegetation

type in incorporating climate change. In our climate scenarios, we are not predicting the exact composition of vegetation and are not evaluating the results as an absolute. Rather, we evaluate the relative impact of the scenarios to study the effect of vegetation changes to landslide susceptibility in the Colorado Front Range.

2.7 Conclusion

We combine a mechanistic soil model with a probabilistic landslide model to evaluate landslide probability during a storm event in the Colorado Front Range. This provides a baseline for evaluating the impacts of climate change on landslide susceptibility assuming a similarly sized storm event. The main findings of our work are:

- Comparing a deterministic model using mean values to the probabilistic model demonstrates that the probabilistic model is better at predicting observed landslides. More area is predicted as being susceptible to landslides in the probability model as compared with the deterministic model.
- Under current conditions, most observed landslides are predicted as having $P(F) > 0$. For FP values, values $0.8 > P(F) > 0$ exhibit similar vegetative and slope characteristics to the observed landslides, suggesting these unstable areas are correctly being captured by the model. A fraction of landslides also occurs at $P(F) > 0.8$, which have higher slope and lower vegetation than areas with lower probabilities of failure.
- Within our study area, a reduction in tree cover due to climate change leads to increased overall landslide susceptibility. For areas susceptible under the climate scenarios but stable under the base scenarios, there is a shift from predominantly SFS to more uniform risk across all aspects.

We consider our results applicable to considering how to incorporate probabilistic landslide modeling and uncertain climate change projections. Our research can be applied to help improve landslide risk and regional changes to landslide susceptibility in the future. We can better understand future landslide risks by understanding that current practices, such as planting trees that are native and ideal for the environment in current times, might lead to more failures long term due to the species having reduced survivability in the future. These considerations allow us to consider the best ways to model and reduce mass-sliding on a regional scale in the future.

There are several avenues for future work to better quantify landslide risk using this approach. Further evaluating deterministic approaches as compared with probabilistic approaches to be able to show which models are better is important for moving forward with landslide mapping. Quantifying the temporal aspect of the landslide probability model would provide insight into appropriate values for probabilities of failure and provide increased understanding about representing landslide risk through time. Determining ways to incorporate the multi-variable impacts of climate change and incorporate the uncertainty therein can improve landslide risk projections for future climate scenarios. Finally, combining the model with a landslide runout model would allow for a more complete understanding of landslide risk.

3. A LARGE-EXTENT TOPOGRAPHICALLY DRIVEN LANDSLIDE RUNOUT MODEL FOR THE COLORADO FRONT RANGE

3.1 Introduction

Landslides can cause widespread damage to infrastructure and the environment and lead to loss of life. Between 2004 and 2010, more than 2,000 landslides across the world resulted in more than 32,000 deaths (Petley, 2012). Precipitation-induced landslides are likely to increase in frequency due to a trend of increasing extreme precipitation because of climate change (Wuebbles et al., 2017). With the far-reaching factors of land use changes and climate change impacting projected landslide locations (Alvioli et al., 2018; Bernardie et al., 2021; Kim et al., 2015), determining the areas most susceptible to both landslide initiation and landslide runout is vital to building resilient infrastructure and protecting people.

There are three components of representing landslides: identifying the areas of initiation, determining the landslide direction and travel path once initiated, and representing the landslide end point. Landslide runout, which considers both the direction and endpoint of the landslide, is impacted by the physical attributes of both the landslide and the topography. The distance traveled by a landslides can be influenced by the landslide failure type, such as a slide or a flow, and the type of landslide material, such as earth, debris, or rock (Varnes, 1978). The direction of the runout can be impacted by topographic controls, such as encountering bends leading to reduced runout distances (Corominas 1996).

Application of landslide runout models is challenging, particularly at large spatial scales. For one, large spatial scales limit the ability to gather detailed information for landslide mapping (Carrara et al., 2008). The type of landslide, which may not be known from landslide initiation models, can be difficult to represent and capture, particularly for large-spatial scale mapping. Debris flows can be particularly challenging to model due to being temporally and spatially varying (Iverson, 1997) due to solid-fluid interactions that result in debris flows acting as a non-Newtonian fluid (De Blasio, 2011). Another consideration, particularly for large spatial scales, that impacts model performance is data resolution. Huggel et al. (2003) applied a 25 m digital elevation model (DEM) in the Swiss Alps and concluded that the resolution provided adequate results for risk mapping at large spatial extents. Fischer et al. (2012) stated that a 25 m DEM led to appropriate landslide runout distance predictions, with a downside being over-predictions of landslide spreading, and that a 10 m DEM provided increased model accuracy. The complexity of modeling landslide runout through limited data availability, quantifying how to represent different types of landslides, and limitations on data resolution are key elements to consider for large-scale analyses for hazard mapping and identifying subregions that warrant additional, in-depth analyses.

Broadly, runout analyses are categorized into physically based models and empirical models (Peruzzetto et al., 2020). Physically based landslide runout models can be applied more broadly on spatial and temporal scales as they are based on representing the physical characteristics that cause landslides to move. Empirical models are often based on developing relationships from landslide inventories, so they have limited applicability to other areas and are more challenging to use for predictions when conditions change, such as considering the impacts of climate change (Bernardie et al., 2021).

Various landslide and topographic characteristics have been evaluated to estimate landslide runout. Runout distance has been related to landslide volume (Iverson, 1998), kinetic energy (Lari et al., 2014), and the center of mass of the landslide (Feranie et al., 2016; Legros, 2002). One widely applied variable to evaluate runout distance is the angle of reach, defined by Corominas (1996) as the angle between the initiation point of the landslide and the distal point in the landslide deposit, calculated as the inverse tangent of the height divided by the horizontal distance traveled (Lockyear, 2018). The angle of reach is smaller for landslides that are larger in volume as they typically have longer runout (Corominas, 1996).

This relationship has been evaluated in numerous modeling applications as a baseline, initial analysis, or provided as a geometric relationship for landslides inventories (Dai and Lee, 2002; Hunter and Fell, 2003; Leng et al., 2018; Lockyear, 2018). Kappes et al. (2011) applied the angle of reach as a worst case scenario for landslide runout mapping. The angle of reach has been applied elsewhere as an approach for initial hazard mapping (Bathurst et al., 1997; Corominas, 2003; Dahl et al., 2010; Kappes et al., 2011). The simplicity of identifying where landslides will stop using the angle of reach makes it convenient for broad spatial scale applications.

Recent developments in landslide runout models have focused on combining large-scale mapping with accurate landslide runout modeling. AschFlow (Quan Luna et al., 2016) is a 2D model that provides results for landslide runout, spreading, and deposition. AschFlow considers volume, which can be difficult to validate if volume is not part of a landslide inventory. Flow-R (Horton et al., 2013) is an another recently developed model that requires only elevation data and landslide initiation points to determine landslide runout. The Flow-R model, which requires some variable and algorithm selection before running, has been applied to a variety of study areas and landslide types. Both models aim to be computationally simplistic enough to evaluate landslides at a larger spatial scale through representing the physical attributes of landslide runout.

While the goal of some models, like AschFlow and Flow-R, is to physically represent the landslide characteristics that lead to termination, other empirical models are aimed at identifying topographic controls that lead to termination. Benda and Cundy (1990) modeled debris flow termination in the Pacific Northwest based on the runout reaching a critical slope of 3.5° or having a change in direction of greater than 70° on a reach. Fannin and Wise (2000) determine erosion/deposition along the landslide path based on slope. Approaches like these topographic evaluations have more limited applicability due to not always having explicit analyses about the landslide runout paths and they typically require site-specific calibration, but they can be useful in large-scale applications, be applied with ease, and require minimal data as input. While previous models have considered topographic controls, they either did it at larger spatial scales or adjusted landslide runout based on previously encountered topographic biases, such as bends.

This study aims to develop a large spatial scale landslide runout model using only topography as input, allowing for ease of application without computationally intensive demands or large amounts of data. The aim is to evaluate whether a rules-based topographic landslide runout model can perform well enough to provide a middle ground between a simple angle of reach analysis and more complex mechanistic models. We evaluate our model with a landslide inventory from a precipitation-induced event in the Colorado Front Range. The precipitation event, which lasted nearly a week in September 2013, led to more than 1,300 observed landslides across 3,400 km^2 where the combined impact of debris flows and flooding caused hundreds of millions of dollars in damage to roads and infrastructure, damaged thousands of houses, and killed many people (Coe et al., 2014). The study aims to develop an empirical, topographic landslide runout model and in doing so, evaluate whether a simple routing can route landslides to the same topographic control as the observed landslides, evaluate whether our topographically controlled landslide runout model can predict landslide termination, and evaluate how the model performs when compared with an angle of reach approach.

3.2 Model Design

There are two key components to representing landslide runout: the path and the termination. The landslide runout model is developed using tools from the Python package Landlab, version 1.0.9. A pre-processing step uses a built-in tool for flow accumulation, which is based loosely on the work of Braun and Willett (2013). The tool determines which direction flow will travel for each cell in the DEM. We apply this step assuming a d8 routing with steepest line of descent, based on the work of Tarboton (1997),

which finds the steepest slope between each node and the eight surrounding nodes. We also apply a depression finder and router, based on the work of Tucker et al. (2001), to provide routing through depressions in the landscape. The output from this processing step is a flow direction grid.

By assuming a steepest line of descent and applying routing through depressions, there are limitations that we are incorporating into our modeling approach. The steepest line of descent means that the landslide is continuously going to traverse downhill along the steepest gradient, independent of obstructions and other losses in momentum that might impact landslide direction. A depression router assumes that the landslide will continue to travel despite slight discontinuities in slope. In reality, a depression would cause a decrease in the momentum of the landslides and a potential loss in volume. Iverson and Denlinger (2001) discuss the complexity of representing erosion/deposition of debris flows, which can cause significant changes in volume, and conclude that accurate assessment of these patterns requires 3D modeling to capture the interactions between the debris flow and the boundary conditions. This is beyond the scope of this project and is thus neglected, an assumption also applied in previous landslide models (Horton et al., 2013).

The landslide runout model is a modified version of the built-in Landlab trickle down tool. The model requires the output from the flow accumulation processing step and a list of the landslide initiation points. For each initiation point, a path is determined based on the flow direction grid. At each incremental point, the cumulative travel distance, local slope, and local curvature is determined.

Our landslide model predicts termination based exclusively on topography. The model requires specifying a critical local slope in degrees, defined as a minimum slope the landslide must traverse, and slope persistence in meters, which is the distance under the critical slope the landslide must travel before stopping. Landslides will continue traveling until the energy lost surpasses the initial gravitational potential energy (Iverson, 1997). The combination of a critical slope and slope persistence is aimed at assessing this energy reduction indirectly though considering whether the topographic control of a low slope for an extended period predicts where termination occurs.

In addition to the critical slope and slope persistence stopping condition, we apply the angle of reach as a stopping condition. While simplistic, the ease of application for an angle of reach to represent potential landslide stopping locations makes it easy to apply and is thus a widespread initial approach to landslide runout mapping. The angle of reach is applied as a comparison to our model to assess which approach performs better. This involves identifying the distance traveled and noting when, if ever, the landslide reaches the specified angle. For each scenario, all modeled landslides use the same specified angle of reach.

3.3 Modeling Methods

The landslide inventory used in our model calibration is from the 2013 storm event in the Colorado Front Range. Landslide initiation points and runout paths are provided by Coe et al. (2014). The initial evaluation by Coe et al. (2014) provides insight into landslide behaviors within the Colorado Front Range. Of the more than 1,100 debris flows analyzed, they were predominately initiated on open slopes and swales. The landslides initiated as debris slides that then transformed into debris flows, leading to longer runout. The landslide direction largely impacted the travel distance: landslides that entered a channel traveled five times farther than landslides that terminated before reaching a channel (Coe et al., 2014).

Two types of mass movements are debris flows and debris slides. We use debris flows, henceforth termed landslides, in our model and remove all debris slides and other mass movements from our evaluation. A debris slide can be initiated in the same way as a debris flow and as such, we define both terms based on

Varnes (1978). Debris slides are comprised of a translational movement of soil. Debris flows are also made of soil though the water content is higher, leading to more fluid-like characteristics. This formation can lead to increased travel distances when compared with debris slides.

3.3.1 Study Area

The model calibration and evaluation areas are located in the Colorado Front Range (Figure 3.1). The areas are chosen due to the high concentration of landslides. The calibration area is 109 km^2 and contains 258 mapped landslides from the landslide inventory from the 2013 storm event (Coe et al., 2014). The elevation within the calibration area ranges from 1,708 to 2,791 m, the land use is comprised predominately of evergreen trees (NLCD, 2016) with some dispersed shrubs, and soils are predominately sandy loams (Soil Survey Staff, 2020). The model evaluation area is 150 km^2 with 303 landslides from the 2013 event. The evaluation area ranges in elevation from 1,792 to 3,215 m and has soil and vegetative characteristics similar to the calibration area.

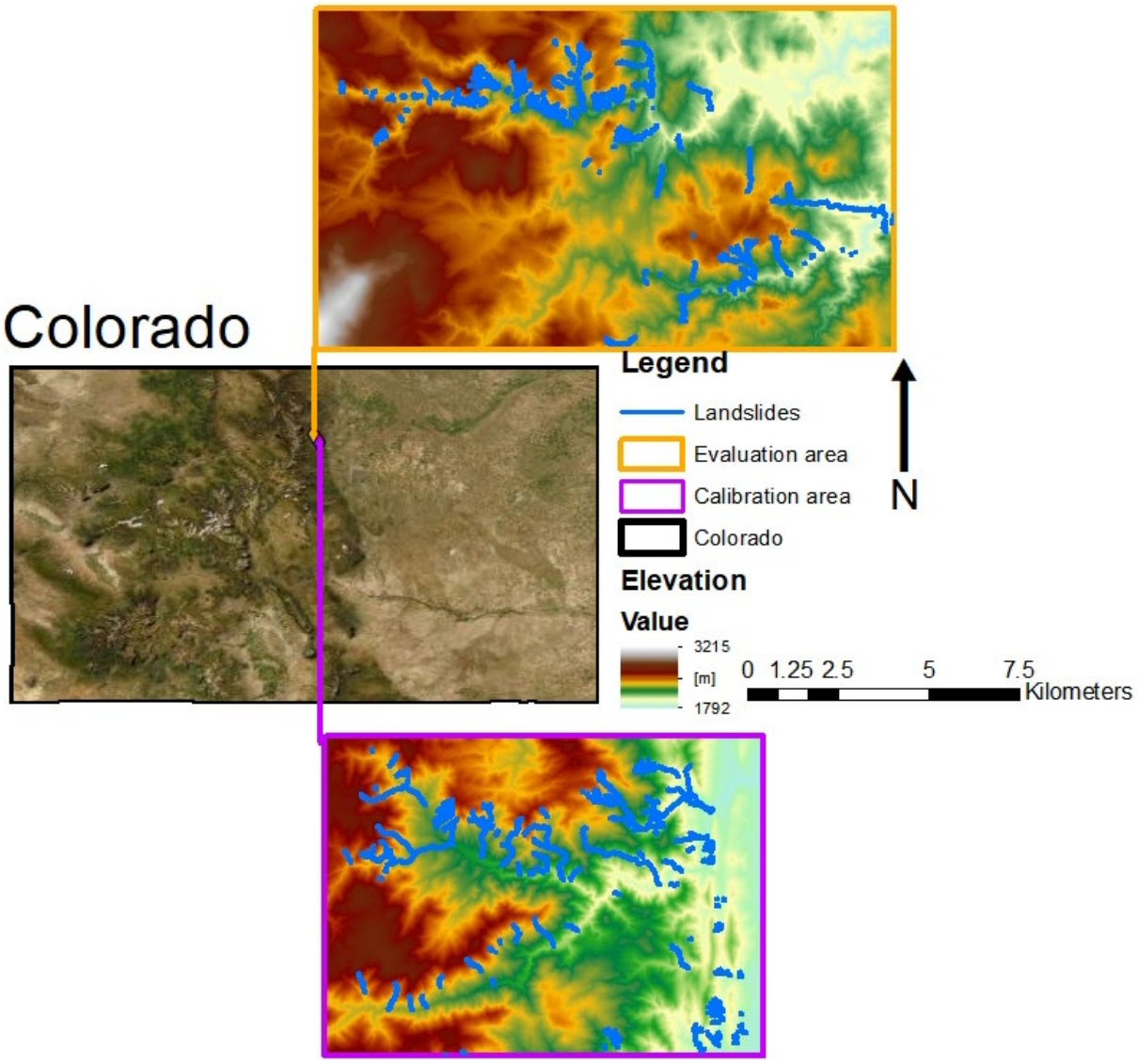

Figure 3.1 The calibration and evaluation areas are located in the Colorado Front Range

3.3.2 Model Calibration

An initial analysis we perform is evaluating the appropriate DEM grid size to use in our study area when applying the runout model by comparing 1 m LiDAR data (USGS, 2018) with 10 m DEM data (USGS, 2015). The LiDAR data, which were collected after the 2013 storm event, are non-continuous in some areas, leading to the inability to map all landslides. The 10 m DEM, also collected post-event (2015), is 1/3 arcsecond resolution and is resampled in the study area to a 10 m resolution. The post-event elevation data are the only elevation data readily available. The mapped landslide tracks are discretized such that the distance between any two points is a maximum of 1 m or 10 m to correspond to the elevation data cell size. The elevations, slopes, and curvature for the landslides are identified to determine differences between the 1 m and the 10 m data. We qualitatively and quantitatively compare the results. Based on this analysis, the results of which we present later, we determine that the 10 m DEM provides necessary model precision and there is not significant loss in accuracy for landslide slope between the start and end of the landslide and as such, apply the 10 m DEM to all following analyses. We use MATLAB version R2019A for this initial data processing.

The landslides are processed in Python 3.7 for subsequent modeling efforts. The landslides are mapped to each DEM grid by determining the grid cells that the landslides pass through. The distance along each landslide is re-calculated using the grid cell distances. The gridded landslides are henceforth termed "observed landslides," while the non-gridded landslides are termed "original landslides." The observed landslide distances are used for comparing with the modeled landslide distances.

Model calibration requires setting the values for the critical slope and for slope persistence distance. For model calibration, two statistics are calculated to compare the modeled with the observed landslide runout distances. The Nash-Sutcliffe (NS) value (Nash and Sutcliffe, 1970; Golmohammadi et al., 2014) is:

$$NS = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}(O_i - M_i)^2}{\frac{1}{n}\sum_{i=1}^{n}(O_i - \bar{O})^2} \tag{3.1}$$

where O_i is the observed distance, M_i is the modeled distance, $\bar{O}$ is the averaged observed distance, and n is the total number of modeled landslides. The maximum NS is a value of 1, which means the modeled values are the same as the observed values, while a value less than zero means the model performs worse than the average observed distance.

The ratio of medians (RM) is calculated as:

$$RM = \frac{M_m}{O_m}, \tag{3.2}$$

where M_m is the median modeled distance and O_m is the median observed distance. Because RM only uses the median value and is not impacted by outliers, it is used to assess whether the model is over- or under-predicting observed landslide distances.

3.4 Results

3.4.1 Model Resolution

The 1 m and 10 m DEM data are compared to assess whether the observed landslides are affected by the increased resolution and precision provided by the 1 m DEM (Figure 3.2).

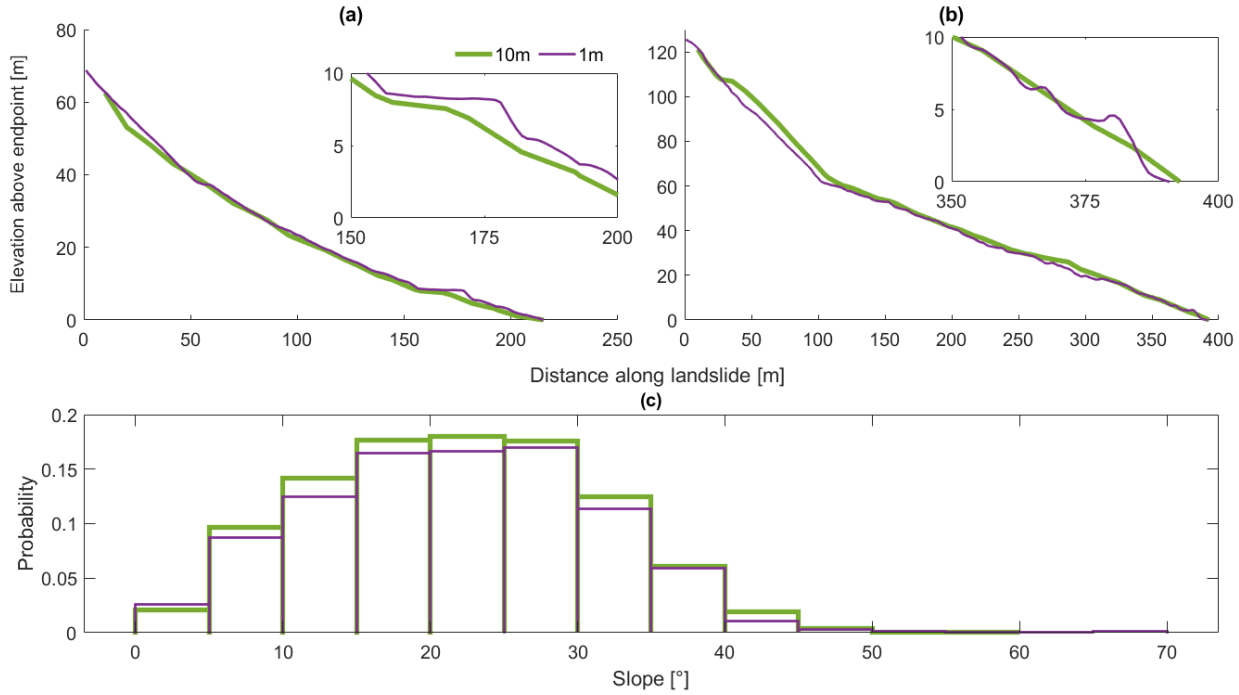

Figure 3.2 Comparisons between the elevation above the endpoint between 10 m and 1 m elevation data for two example landslides and a histogram (c) showing the slope between the first and last points

The slight discontinuities in the 1 m data lead to longer areas with low slope, as shown in the close-up in Figure 3.2a, and at times slight increases in elevation, as shown in the close-up in Figure 3.2b. The general behavior for 1 m versus 10 m data considering overall slope, elevation, and distance along the landslide are similar, despite small-scale differences. The histogram (Figure 3.2c), which considers the slope between the first and last point of the landslide, has no major differences between the 10 m and the 1 m elevation data. The 10 m data have marginally higher occurrences of lower slopes. Due to the runout model assuming a steepest line of descent, the increases in elevation present in the 1 m LiDAR data cannot be captured by our model. Between this and the slope of the landslides being similar, we assume that 10 m data are appropriate for application of our model. The histogram considers all the mass movements in the landslide inventory, the majority of which (>1100) are debris flows. The 10 m data include 1,350 landsides while the 1 m data, which do not cover the entire area containing mapped landslides, include 1,255 landslides.

3.4.2 Landslide Direction

To validate our assumption about d8 routing, we evaluate the landslide path to determine whether our modeled landslide paths encounter the same topographic controls as the observed landslides. For the purpose of comparing the model and the observed landslides, the modeled stopping point is assumed to be the closest node to the observed end point. For the calibration area, 83% of the end points of the modeled

landslides are within 1 node (14 m maximum) of the observed landslides. This suggests that a perfectly calibrated model could stop at the same location as the observed landslide based on our routing assumption. We visually analyze this and compare the original, observed, and modeled landslides (Figure 3.3).

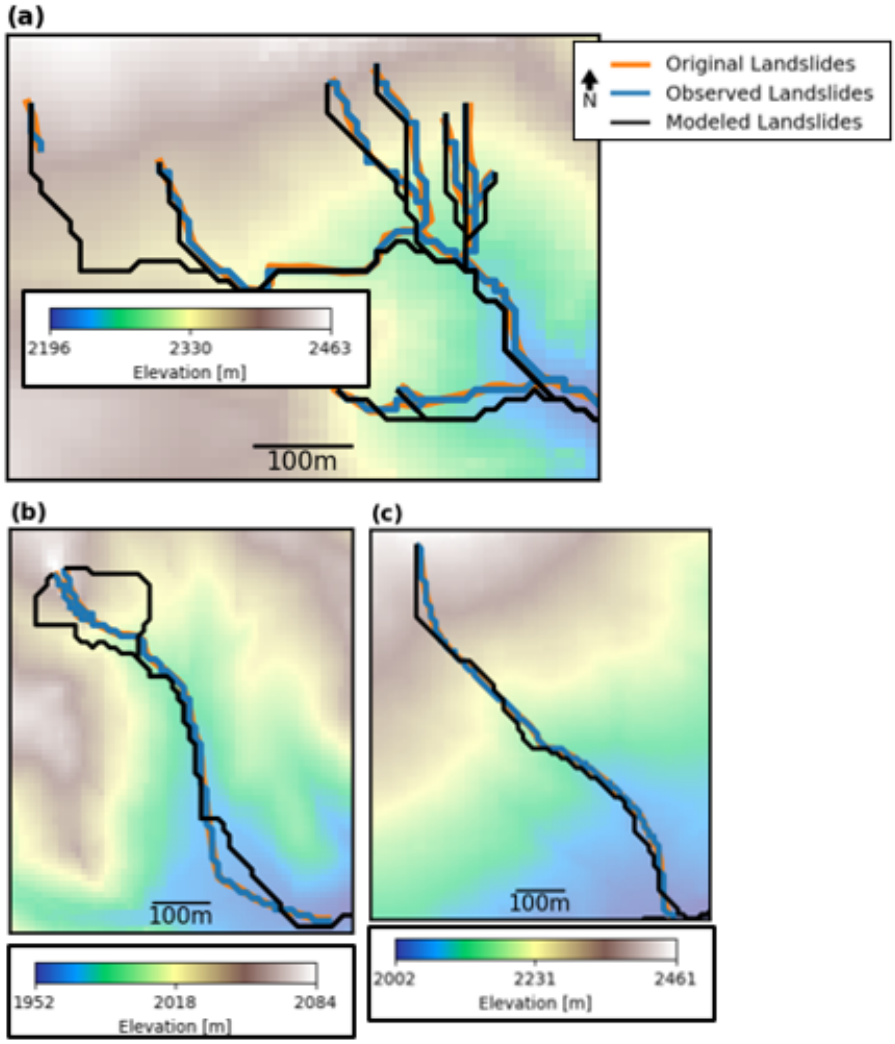

Figure 3.3 Examples of original, observed, and modeled landslides in the calibration area

Figure 3.3 shows that d8 routing generally provides a reasonable initial evaluation of landslide direction. While there are some areas with differences between the modeled and observed landslide paths, particularly closer to the landslide initiation points and in the flatter areas (Figure 3.3b), the model generally reflects the landslide direction.

3.4.3 Model Calibration

Our model calibration is aimed at determining whether modeling a critical slope and slope persistence identifies a topographic control that leads to landslides stopping. Our initial analysis considers critical slope and persistence combinations and the corresponding NS and RM values within the calibration area (Figure 3.4).

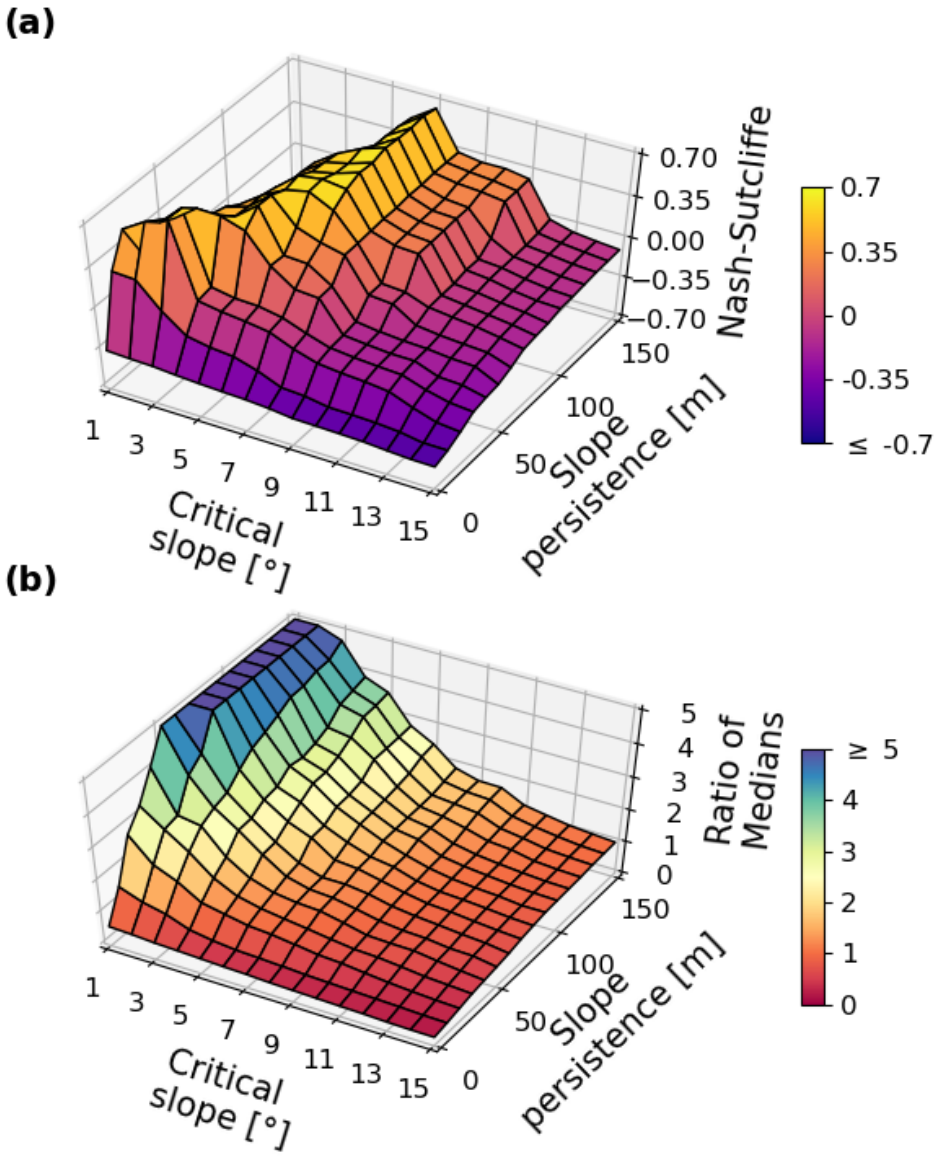

Figure 3.4 Critical slope and the slope persistence combinations compared to the corresponding NS (a) and RM (b) in the calibration area

The analysis shows that the highest performance for NS has two peak values. The first is at a low critical slope and a short slope persistence. The second is a marginally higher slope and longer slope persistence, and this peak does not have a local maximum. More extreme values in both higher and lower critical slopes lead to NS < 0, indicating poor model performance. RM is optimized with both lower critical slope and lower slope persistence or moderate critical slope and high slope persistence, leading to values closest to 1. While there are numerous NS and RM values that are close to the optimized values, for calibration purposes we move forward with the highest NS value.

We perform three comparisons between the calibrated model and characteristics that determine the validity of the calibration. The calibrated model is the critical slope/slope persistence combination that optimizes NS. This scenario is compared with (1) the optimized RM model, (2) a landslide convergence analysis, and (3) the evaluation area (Table 3.1).

Table 3.1 Geometric data corresponding to the model comparisons

Variable	Calibrated Scenario	Comparison 1: Calibration Variable	Comparison 2: Landslide Convergence			Comparison 3: Evaluation Area
		RM	Maximum	Minimum	Random	Evaluation Area
Critical Slope [°]	3	12	3	3	3	3
Slope persistence [m]	20	90	20	20	20	20
NS	0.69	-0.07	-0.43	0.49	0.44	0.58
RM	1.72	1.00	2.49	2.02	2.04	1.54
Mean distance, observed [m]	1264	1251	483	538	547	558
Mean distance, modeled [m]	1259	606	695	618	672	555
Median distance, observed [m]	547	542	212	258	258	279
Median distance, modeled [m]	941	540	527	522	525	429
IQR, observed [distance]	1635	1625	402	567	428	594
IQR, modeled [distance]	1053	503	597	565	564	613
IQR [diff. between observed and modeled distances]	595	1089	480	312	337	92
IQR [diff. between modeled and observed end points]	707	655	475	325	353	143
Unique end points	80	112	80			119
Total Landslides	244	244	244			303

Note: IQR = interquartile range

Comparison 1 contrasts the calibrated scenario, which optimizes NS, to the optimized scenario for RM. NS is optimized at a slope of 3° and a slope persistence of 20 m while RM is optimized at a slope of 12° and a slope persistence of 90 m. The NS is significantly lower (-0.07 compared with 0.69) for the optimized RM. The RM is higher, indicating over-prediction of landslide runout distance, for the optimized NS scenario.

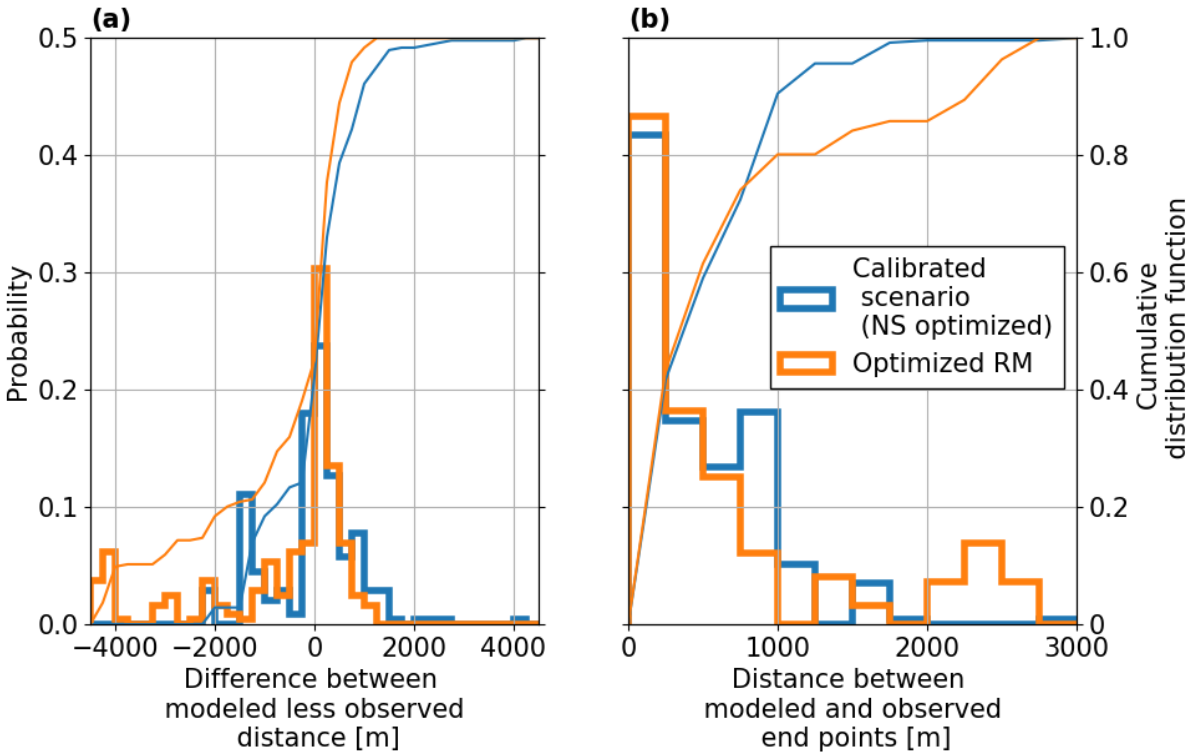

Figure 3.5 The difference between landslide distances (a) and end point locations (b) in the calibration area between the optimization of the two calibration variables

Figure 3.5 presents histograms comparing the difference between the modeled and observed landslide distances (Figure 3.5a) and the distance between the modeled and observed end points (Figure 3.5b). The difference in distances (Figure 3.5a) is shifted toward positive values for both NS and RM values. The RM optimization results in more equal over- and under-prediction of landslide distance, resulting in a number of landslides having distances greater than 2,000 m between the modeled and observed landslide values. This validates using NS as the main calibration variable as all the landslides are considered.

Comparison 2 considers landslide convergence. Within the 290 observed landslides within the calibration area that are at least 10 m (two grid cells) long, there are 172 unique end points. All landslides, both modeled and observed, that converge are stopping due to the same topographic control. As our model evaluates topographic controls, using all the landslides in the calibration where multiple landslides converge might lead to skewed model results as the topographic controls are not unique. For end points where multiple landslides converge, we consider four scenarios: (1) using all landslides that the model predicts as stopping, which is the approach applied in the calibrated NS scenario and the RM scenario in Comparison 1, considering landslides that converge and at each unique end point; (2) choosing the landslide with the minimum absolute distance between the modeled and observed landslide distance; (3) choosing the landslide with the maximum absolute difference between the modeled and observed landslide distance; and (4) randomly selecting a landslide that ends at that location (Figure 3.6). These modeling analyses will henceforth be termed full, minimum, maximum, and random end point analyses.

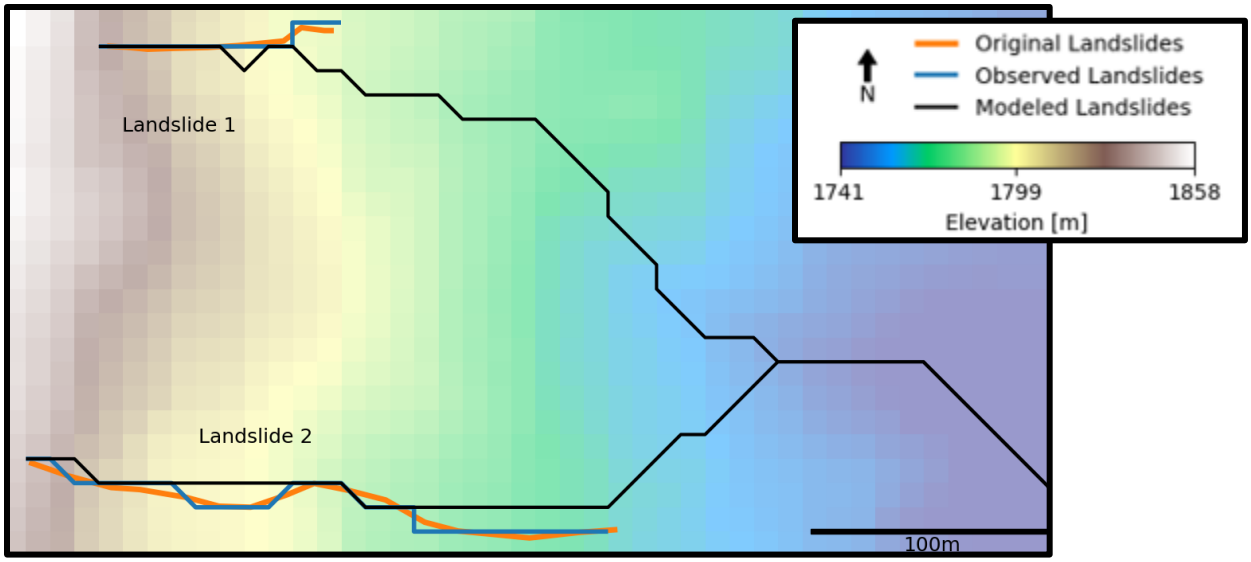

Figure 3.6 A visual example for the four types of end point analyses in which the model predicts both slides terminating: the full analysis uses both landslides 1 and 2; the maximum end point analysis using Landslide 1 as the difference between the observed and the modeled landslide distance is greater; the minimum end point condition using Landslide 2 as the difference between the observed and the modeled landslide distance is smaller; and the random end point analysis randomly selects Landslide 1 or Landslide 2

Comparison 2, which compares the calibrated scenario to the minimum, maximum, and random end point analyses assuming the same critical slope and slope persistence, has a range in predictive abilities. The maximum analysis has an NS value less than 0 and the largest RM value. The minimum analysis has NS of 0.49, which is closer to the calibrated scenario, and an RM value of 2.02. The random analysis has a lower NS than the calibrated scenario and a comparable RM value to the minimum analysis.

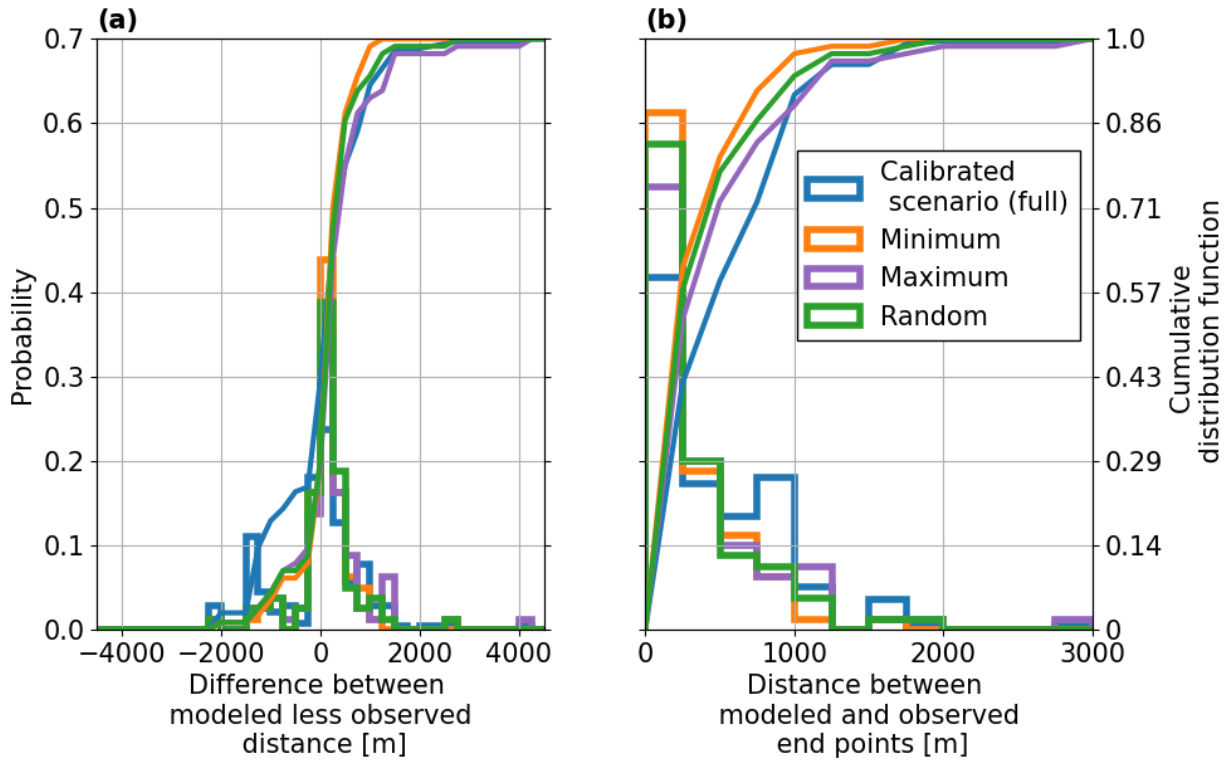

Figure 3.7 The difference between landslide distances (a) and end point locations (b) in the calibration area between end point analyses

In considering the end point analysis histograms (Figure 3.7), there are 244 total landslides in the full dataset that are predicted as stopping and 80 landslides in the minimum, maximum, and random analyses. The maximum is shifted toward increased variability, which means more spread. The minimum analysis minimizes the distance, thus is selecting for the best-fit landslide values. The random selection is variable between model runs. The difference between the landslide stopping analyses suggests landslide convergence is important to consider in topographic models but in our case, maintaining all landslides as in the calibrated scenario is the best way to ensure all landslide lengths and differences between predicted and modeled landslides are considered.

Comparison 3 (Figure 3.8) applies the calibrated model variables to the evaluation area to evaluate model performance. The evaluation area has a marginally lower NS value of 0.58 and an RM value of 1.54, both of which are lower than the calibrated scenario. The landslides in the evaluation area are shorter with a mean of 555 m for modeled versus 558 m for observed landslides as compared with the mean of 1,259 m for modeled and 1,264 m for the observed landslides in the calibration area. The IQR is also significantly less for the difference between modeled and observed landslides in the evaluation area.

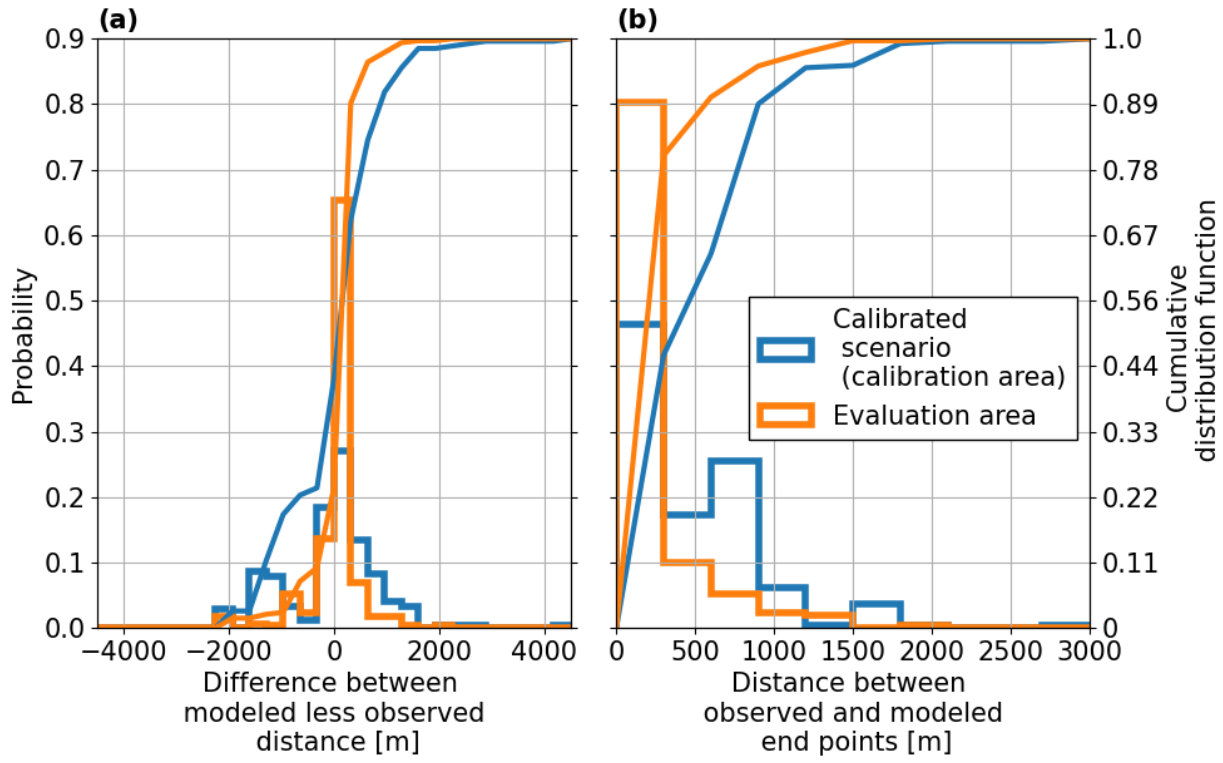

Figure 3.8 The difference between landslide distances (a) and end point locations (b) in the calibration area compared with the evaluation area

The histograms for Comparison 3 (Figure 3.8) show that in the evaluation area, differences between modeled and observed landslide distances and end points are closer than the calibration scenario. The spread in the values for the evaluation area is significantly smaller than in the calibration area.

The results for the evaluation area are mapped in Figure 3.9.

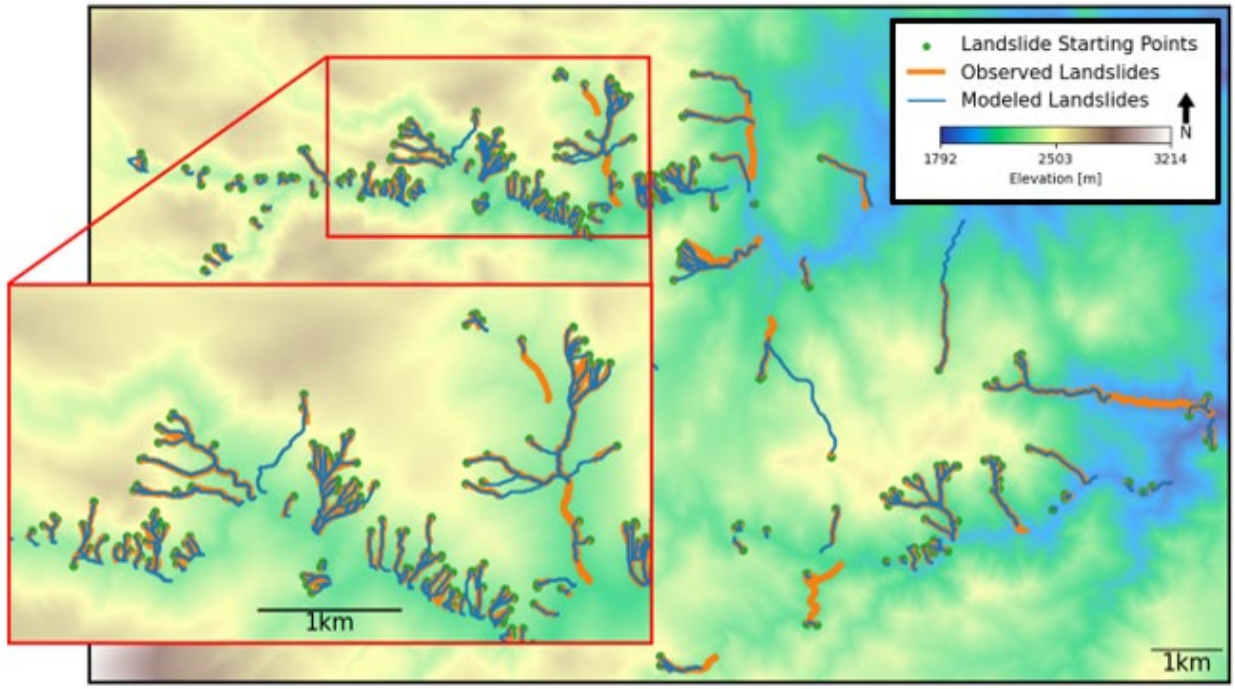

Figure 3.9 *A map from the evaluation area demonstrates the final model performance comparing the observed landslides with the modeled landslides*

The model captures the general behavior in landslide tracks. The zoomed-in area shows that the model tends to predict slightly longer runout paths and that modeled landslides tend to converge more often than the observed landslides.

3.4.4 Comparison to Angle of Reach Modeling

For an initial consideration, we determine the angle of reach for the observed landslides in the evaluation area (Figure 3.10). The observed landslides have an angle of reach ranging from 3.6° to 49.0° with a mean value of 25.1° and a median of 25.5°.

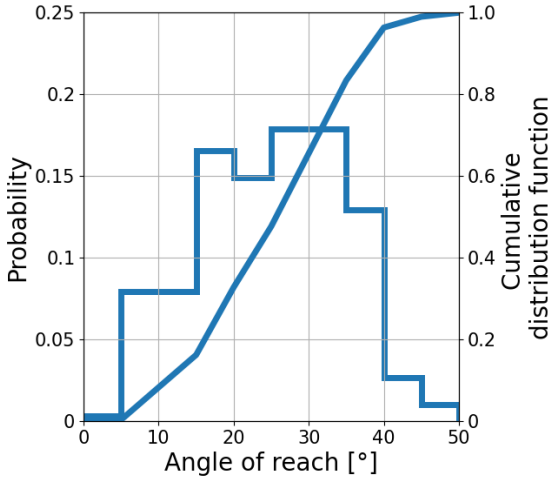

Figure 3.10 The angle of reach for the observed landslides in the evaluation area

To assess whether our model performs better than an angle of reach approach, we apply the angle of reach as a stopping condition by specifying the angle and determining when, if ever, modeled landslides reach that angle. We apply an angle of reach ranging from 3° to 50°, increasing in increments of 1°, as a stopping condition applied to all the landslides in the evaluation area (Figure 3.11). The results have NS > 0 between 18° and 24°. RM is greater than 1 from 3° until 23° and less than 1 at higher slopes. The maximum NS is at 20° with NS = 0.14 and RM = 1.78.

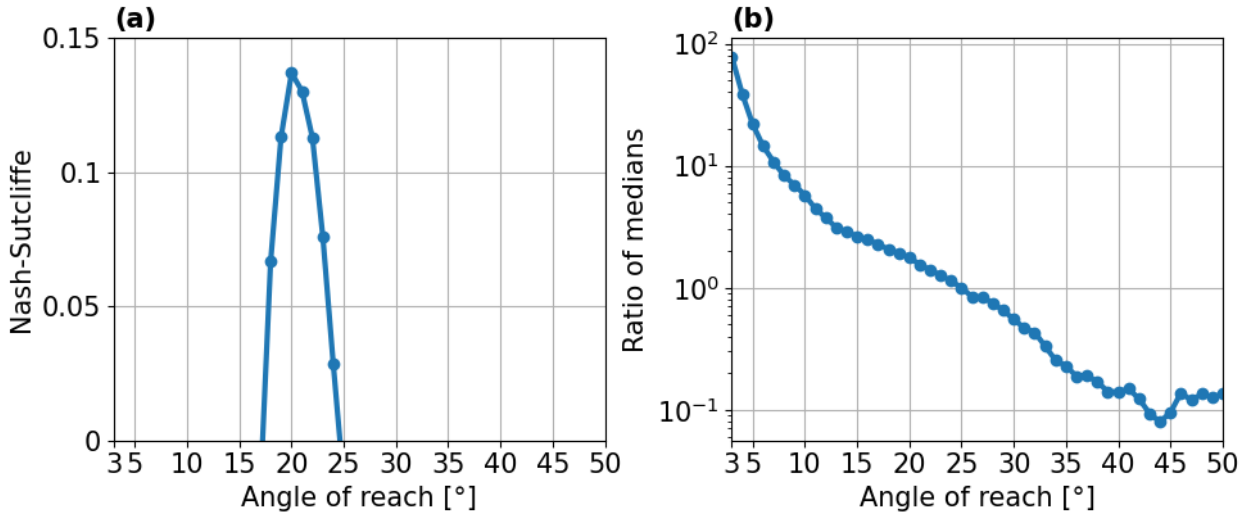

Figure 3.11 The corresponding NS and RM values for a constant angle of reach being applied

3.5 Discussion

Our analyses determine that topographic controls can predict landslide termination in the Colorado Front Range. We compare our model to an angle of reach. Angle of reach analyses are often applied through considering landslide volume, which can lead to increased model performance (Corominas, 1996). Due to a lack of information about landslide volume, we only apply a single angle of reach. Another approach that has been taken to improve topographic models is to consider change in direction. Corominas (1996) evaluated the angle of reach for unobstructed, obstructed, and channelized landslides. Benda and Cundy (1990) applied a landslide stopping condition of whether a junction exceeded 70°. Our model considers only the topographic controls that lead to landslides stopping and does not consider additional upstream conditions, such as changes in direction. This simplifying assumption, along with the lack of information about landslide volume in our landslide inventory, limits the accuracy of the angle of reach in our study. The result is that our model performs better, as it can capture more of the variability in the observed landslides.

Our model evaluates the topographic controls on landslides initiated from a single precipitation event in Colorado. The modeling goal is to evaluate whether our model provides increased accuracy for large-scale runout predictions. To make the model applicable on a broader scale, there are tradeoffs to consider between our simplistic application and being able to represent landslide runout risk more broadly, including considering flow spreading, landslide volume, and landslide type.

In considering simplifying model assumptions, our model does not consider landslide spreading. Huggel et al. (2003) considers glacial outburst floods and found that the d8 routing approach correctly identified the flow path but is more impacted by errors in the DEM and represents fewer realistic deposition patterns than alternative routing approaches. Horton et al. (2013) discusses that, in terms of runout paths, the d8 approach provides unrealistically straight paths and leads to limited landslide spreading. To more

realistically model landslide spreading while maintaining simple model assumptions, Scheidl and Rickenmann (2009) applied a d8 landslide mapping approach combined with a Monte Carlo simulation to represent multiple flow paths. A similar modeling approach, or evaluating whether a simple buffer around the landslide path could realistically represent conditions, could be incorporated to capture landslide spreading more realistically.

Another consideration is the tradeoff between using a single stopping condition and considering landslide volume. Relating landslide stopping location with landslide volume could improve the model by more accurately representing observed landslides. Landslide volume could be considered indirectly by considering landslide runout distance as higher angles of reach correspond to decreased volumes (Corominas, 1996). To maintain simplicity, having unique critical slope and slope persistence combinations dependent on landslide volume could be applied by having a range in the predicted landslide stopping location instead of a single point.

Another tradeoff in our model is that we only consider debris flows though a number of debris slides are present. Debris slides tend to be shorter than debris flows in our study area, so determining if the same critical slope/slope persistence combinations represent the debris slides would need to be evaluated. Further, the d8 routing would need to be checked to assess whether this assumption accurately captures debris slide routing. Identifying the topographic controls that lead to debris slides stopping could contribute to quantifying potential landslide runout distances based on landslide type.

3.6 Conclusion

Our landslide model is an empirical, topographically based framework for landslide runout modeling in the Colorado Front Range. We apply d8 routing, which assumes a steepest line of descent, and a combination of a critical slope and slope persistence to capture landslide end points in the Colorado Front Range. We determine that simplistic d8 routing and steepest line of descent assumptions on a 10 m DEM provide reasonable predictions for landslide runout paths and predicting landslide termination. Finally, we apply an angle of reach to determine landslide termination and determine that the angle of reach performs worse than the calibrated combination of critical slope and slope persistence. The model is simplistic and requires only elevation and landslide initiation points as input, allowing for ease of application.

Future model improvements should evaluate model performance in different study areas to determine whether the combination of a critical slope and slope persistence can act as topographic control for determining where landslides will terminate. Another area for future work is combining the model with a landslide probability model. In integrating the two, a more complete understanding about landslide hazard mapping can be developed.

4. CONCLUSIONS AND FUTURE WORK

The landslide probability model provides an initial evaluation of current and future landslide risk in the Colorado Front Range. The model predicts the landslides initiated by the 2013 storm event and quantifies the potential impacts of climate change. There are areas for further development of the model and analyses. Evaluating whether the probabilities associated with the model are appropriate would allow for better understanding of the temporal aspect of landslide probability modeling. This could include evaluating a broader landslide inventory as all the observed landslides we consider were initiated by a single event. The uncertainties with climate change, relating to both vegetation dynamics and hydrologic changes, should be further analyzed to determine how best to represent uncertainties over large spatial extents and the resulting impacts on landslide susceptibility.

The landslide runout model showed that the topographic controls of a specified distance under a critical slope provide reasonable results for determining landslide end points within our study area. Future model development can focus on applying the model to different areas to determine if the same assumptions yield realistic results. The analyses could also be improved by incorporating landslide volume to better quantify topographic controls. Additional research can also focus on whether there are easily applied variables that could represent landslide spreading. Incorporating landslide spreading would provide additional understanding about areas susceptible to landslides. Finally, considering whether debris slides have topographic controls would benefit model applicability. Because debris flows and debris slides are initiated by the same mechanisms, being able to predict the runout of both types is important for integration with a landslide initiation model.

The landslide probability and landslide runout models provide a basis for mapping landslide risk in the Colorado Front Range, and an area for future work is combining the two models. Areas with a predicted probability of failure from the landslide probability model can be used as the initiation points for the landslide runout model. In the landslide runout model, once a stopping condition is met, a probability for that landslide location can be determined based on the initial landslide probability of failure and the percent of landslides in the study area the model accurately represents with those stopping conditions. This would allow for more complete understanding about current landslide risk and could be combined with climate modeling to provide more complete understanding about areas susceptible to landslides under climate change scenarios.

Our landslide modeling applications contribute to the general understanding about shallow, precipitation-induced landslides, the drivers for accurately capturing landslide probability of failure on a broad spatial extents, the impacts of climate change induced changes to landslide susceptibility in the Colorado Front Range, and the topographic characteristics that make landslides stop. This information can be used by researchers to perform to further analyses of landslide probability and runout, risk managers in considering which practices might impact slope stability, and infrastructure developers who might consider landslide risk in current and future projects. Particularly with climate change potentially increasing the areas susceptible to landslides in the Front Range, determining under what conditions landslides occur and where they will travel is important to guide decision making in the present and in moving forward.

REFERENCES

Abbaszadeh, M., Shahriar, K., Sharifzadeh, M., and Heydari, M. "Uncertainty and Reliability Analysis Applied to Slope Stability: A Case Study From Sungun Copper Mine," *Geotech Geol Eng*, 29, 581–596, https://doi.org/10.1007/s10706-011-9405-1, 2011.

Alvioli, M., Melillo, M., Guzzetti, F., Rossi, M., Palazzi, E., von Hardenberg, J., Brunetti, M. T., and Peruccacci, S. "Implications of climate change on landslide hazard in Central Italy," *Science of the Total Environment*, 630, 1528–1543, https://doi.org/10.1016/j.scitotenv.2018.02.315, 2018.

Barnhart, K. R., Hutton, E. W. H., Tucker, G. E., Gasparini, N. M., Istanbulluoglu, E., Hobley, D. E. J., Lyons, N. J., Mouchene, M., Nudurupati, S. S., Adams, J. M., and Bandaragoda, C. "Short communication: Landlab v2.0: A software package for Earth surface dynamics," *Earth Surf. Dynam*., 8(2), p 379-397, doi:10.5194/esurf-8-379-2020, 2020.

Bathurst, J. C., Burton, A., and Ward, T. J. "Debris flow run-out and landslide sediment delivery model tests," *Journal of Hydraulic Engineering*, 123(5): 410-419, 1997.

Baum, R. L., Godt, J. W., and Savage, W. Z. "Estimating the timing and location of shallow rainfall-induced landslides using a model for transient, unsaturated infiltration," *J. Geophys. Res*., 115, F03013, https://doi.org/10.1029/2009JF001321, 2010.

Beguería, S. "Validation and Evaluation of Predictive Models in Hazard Assessment and Risk Management," *Nat Hazards*, 37, 315–329, https://doi.org/10.1007/s11069-005-5182-6, 2006.

Benda, L. E., and Cundy, T. W. "Predicting deposition of debris flows in mountain channels," *Can. Geotech. J*., 27, 409–417, https://doi.org/10.1139/t90-057, 1990.

Bernardie, S., Vandromme, R., Thiery, Y., Houet, T., Grémont, M., Masson, F., Grandjean, G., and Bouroullec, I. "Modelling landslide hazards under global changes: the case of a Pyrenean valley," *Nat. Hazards Earth Syst. Sci*., 21, 147–169, https://doi.org/10.5194/nhess-21-147-2021, 2021.

Borga, M., Dalla Fontana, G., and Cazorzi, F. "Analysis of topographic and climatic control on rainfall-triggered shallow landsliding using a quasi-dynamic wetness index," *Journal of Hydrology*, 268, 56–71, https://doi.org/10.1016/S0022-1694(02)00118-X, 2002.

Boyce, R. L., Clark, R., and Dawson, C. "Factors Determining Alpine Species Distribution on Goliath Peak, Front Range, Colorado, U.S.A.," *Arctic, Antarctic, and Alpine Research*, 37, 88-96, https://doi.org/10.1657/1523-0430(2005)037[0088:FDASDO]2.0.CO;2, 2005.

Braun, J., and Willett, S. D. "A very efficient O(n), implicit and parallel method to solve the stream power equation governing fluvial incision and landscape evolution," *Geomorphology*, 180–181, 170–179, https://doi.org/10.1016/j.geomorph.2012.10.008, 2013.

Brice, E. M., Miller, B. A., Zhang, H., Goldstein, K., Zimmer, S. N., Grosklos, G. J., Belmont, P., Flint, C. G., Givens, J. E., Adler, P. B., Brunson, M. W., and Smith, J. W. "Impacts of climate change on multiple use management of Bureau of Land Management land in the Intermountain West, USA," *Ecosphere*, 11, https://doi.org/10.1002/ecs2.3286, 2020.

Buma, B., and Wessman, C. A. "Forest resilience, climate change, and opportunities for adaptation: A specific case of a general problem," *Forest Ecology and Management*, 306, 216–225, https://doi.org/10.1016/j.foreco.2013.06.044, 2013.

Campbell, G.S. "A simple method for determining unsaturated conductivity from moisture retention data," *Soil Science*, 117, 311-314, http://dx.doi.org/10.1097/00010694-197406000-00001, 1974.

Cannone, N., Sgorbati, S., and Guglielmin, M. "Unexpected impacts of climate change on alpine vegetation," *Frontiers in Ecology and the Environment*, 5, 360–364, https://doi.org/10.1890/1540-9295(2007)5[360:UIOCCO]2.0.CO;2, 2007.

Carrara, A., Crosta, G., and Frattini, P. "Comparing models of debris-flow susceptibility in the alpine environment," *Geomorphology*, 94, 353–378, https://doi.org/10.1016/j.geomorph.2006.10.033, 2008.

Clerici, A., Perego, S., Tellini, C., and Vescovi, P. "Landslide failure and runout susceptibility in the upper T. Ceno valley (Northern Apennines, Italy)," *Nat Hazards*, 52, 1–29, https://doi.org/10.1007/s11069-009-9349-4, 2010.

Charney, N. D., Babst, F., Poulter, B., Record, S., Trouet, V. M., Frank, D., Enquist, B. J., and Evans, M. E. K. "Observed forest sensitivity to climate implies large changes in 21st century North American forest growth," *Ecol Lett*, 19, 1119–1128, https://doi.org/10.1111/ele.12650, 2016.

Coe, J. A., Kean, J. W., Godt, J. W., Baum, R. L., Jones, E. S., Gochis, D. J., and Anderson, G. S. "New insights into debris-flow hazards from an extraordinary event in the Colorado Front Range," *GSAT*, 24, 4–10, https://doi.org/10.1130/GSATG214A.1, 2014.

Coleman, M. L., and Niemann, J. D. "Controls on topographic dependence and temporal instability in catchment-scale soil moisture patterns," *Water Resources Research*, 49, 1625-1642, https://doi.org/10.1002/wrcr.20159, 2013.

Collins, B. D., and Znidarcic, D. "Stability Analyses of Rainfall Induced Landslides," *J. Geotech. Geoenviron. Eng.,* 130, 362–372, https://doi.org/10.1061/(ASCE)1090-0241(2004)130:4(362), 2004.

Corominas, J. "The angle of reach as a mobility index for small and large landslides," *Canadian Geotechnical Journal*, 33(2), 260-271, https://doi.org/10.1139/t96-005, 1996.

Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R. "A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils," *Water Resources Research*, 20, 682-690, https://doi.org/10.1029/WR020i006p00682, 1984.

Cowley, G. S., Niemann, J. D., Green, T. R., Seyfried, M. S., Jones, A. S., and Grazaitis, P. J. "Impacts of precipitation and potential evapotranspiration patterns on downscaling soil moisture in regions with large topographic relief," *Water Resour. Res*., 53, 1553–1574, https://doi.org/10.1002/2016WR019907, 2017.

Crookston, N. L., Rehfeldt, G. E., Dixon, G. E., and Weiskittel, A. R. "Addressing climate change in the forest vegetation simulator to assess impacts on landscape forest dynamics," *Forest Ecology and Management*, 260, 1198-1211., 2013.

Crookston, N. L. "Climate-FVS Version 2: Content, users guide, applications, and behavior." Gen. Tech. Rep. RMRS-GTR-319. Fort Collins, CO: U. S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, 38, 2014.

Dahl, M.-P. J., Mortensen, L. E., Veihe, A., and Jensen, N. H. "A simple qualitative approach for mapping regional landslide susceptibility in the Faroe Islands," *Nat. Hazards Earth Syst. Sci*., 10, 159–170, https://doi.org/10.5194/nhess-10-159-2010, 2010.

Davis, J., and Goadrich, M. "The relationship between Precision-Recall and ROC curves," in: Proceedings of the 23rd International Conference on Machine Learning - ICML '06, the 23rd international conference, Pittsburgh, Pennsylvania, 233–240, https://doi.org/10.1145/1143844.1143874, 2006.

De Blasio, F. V. “Non-Newtonian Fluids, Mudflows, and Debris Flows: A Rheological Approach,” in: Introduction to the Physics of Landslides: Lecture notes on the dynamics of mass wasting, edited by: de Blasio, F. V., Springer Netherlands, Dordrecht, 89–130, https://doi.org/10.1007/978-94-007-1122-8_4, 2011.

Deshon, J. P. “Stochastic analysis and probabilistic downscaling of soil moisture,” M. S. thesis, Colorado State University, Fort Collins, Colorado, 2018.

Dingman, S. L. *Physical Hydrology, 2nd ed*., 646 pp., Prentice Hall, Upper Saddle River, N. J., 2002.

Escobar-Wolf, R., Sanders, J. D., Vishnu, C. L., Oommen, T., and Sajinkumar, K. S. “A GIS tool for infinite slope stability analysis (GIS-TISSA),” *Geoscience Frontiers*, 12, 756–768, https://doi.org/10.1016/j.gsf.2020.09.008, 2021.

Fagerland, M. W., and Sandvik, L. “Performance of five two-sample location tests for skewed distributions with unequal variances,” *Contemporary Clinical Trials*, 30, 490–496, https://doi.org/10.1016/j.cct.2009.06.007, 2009.

Fannin, R. J., and Wise, M. P. “An empirical-statistical model for debris flow travel distance,” *Can. Geotech. J*., 38, 982–994, https://doi.org/10.1139/t01-030, 2001.

Fawcett, T. “An introduction to ROC analysis,” *Pattern Recognition Letters*, 27, 861-874, https://doi.org/10.1016/j.patrec.2005.10.010, 2005.

Feranie, S., Tohari, A., and Latief, F. D. E. “Prediction of landslide run-out distance based on slope stability analysis and center of mass approach,” IOP Conf. Ser.: Earth Environ. Sci., 29, 012003, https://doi.org/10.1088/1755-1315/29/1/012003, 2016.

Fischer, L., Rubensdotter, L., Sletten, K., Stalsberg, K., Melchiorre, C., Horton, P., and Jaboyedoff, M. “Debris flow modeling for susceptibility mapping at regional to national scale in Norway,” 7, 2012.

Frattini, P., Crosta, G., and Carrara, A. “Techniques for evaluating the performance of landslide susceptibility models,” *Engineering Geology*, 111, 62–72, https://doi.org/10.1016/j.enggeo.2009.12.004, 2010.

Fu, G., Yi, L., and Pan, J. “Tuning model parameters in class-imbalanced learning with precision-recall curve,” *Biom. J*., 61, 652–664, https://doi.org/10.1002/bimj.201800148, 2019.

Gariano, S. L., Petrucci, O., Rianna, G., Santini, M., and Guzzetti, F. “Impacts of past and future land changes on landslides in southern Italy,” *Reg Environ Change*, 18, 437–449, https://doi.org/10.1007/s10113-017-1210-9, 2018.

Gibson, S., and Sanchez, A. *HEC-RAS Mud and Debris Flow Manual, Version 6.0*, US Army Corps of Engineers, 2020.

Golmohammadi, G., Prasher, S., Madani, A., and Rudra, R. “Evaluating three hydrologic distributed watershed models: MIKE-SHE, APEX, SWAT,” *Hydrology*, *1*, 20-39, Doi: 10.3390/hydrology1010020, 2014/

Grieco, N. R., Niemann, J. D., Green, T. R., Jones, A. S., and Grazaitis, P. J. “Hydrologic Downscaling of Soil Moisture Using Global Data Sets without Site-Specific Calibration,” *J. Hydrol. Eng*., 23, 04018048, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001702, 2018.

Gudiyangada Nachappa, T., Tavakkoli Piralilou, S., Ghorbanzadeh, O., Shahabi, H., and Blaschke, T. “Landslide Susceptibility Mapping for Austria Using Geons and Optimization with the Dempster-Shafer Theory,” *Applied Sciences*, 9, 5393, https://doi.org/10.3390/app9245393, 2019.

Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., and Galli, M. “Estimating the quality of landslide susceptibility models,” *Geomorphology, 81*, 166–184, https://doi.org/10.1016/j.geomorph.2006.04.007, 2006.

Hammond, C., Hall, D., Miller, S., and Swetik, P. “Level 1 stability analysis (LISA), documentation for Version 2.0,” USDA, Forest Service, Moscow, ID, Intermountain Res. Sta. Gen. Tech. Rep. INT-285, 1992.

Haque, U., da Silva, P. F., Devoli, G., Pilz, J., Zhao, B., Khaloua, A., Wilopo, W., Andersen, P., Lu, P., Lee, J., Yamamoto, T., Keellings, D., Wu, J.-H., and Glass, G. E. “The human cost of global warming: Deadly landslides and their triggers (1995–2014),” *Science of The Total Environment*, 682, 673–684, https://doi.org/10.1016/j.scitotenv.2019.03.415, 2019.

Helsel, D. R., and Hirsch, R. M. “Chapter A3: Statistical Methods in Water Resources,” in: *Techniques of Water-Resources Investigations of the Unites States Geological Survey, Book 4, Hydrologic Analysis and Interpretation*., U. S. Department of the Interior, 2002.

Hobley, D. E. J., Adams, J. M., Nudurupati, S. S., Hutton, E. W. H., Gasparini, N. M., Istanbulluoglu, E. and Tucker, G. E. “Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics,” *Earth Surface Dynamics*, 5(1), p 21-46, 10.5194/esurf-5-21-2017, 2017.

Hoehn, D. C., Niemann, J. D., Green, T. R., Jones, A. S., and Grazaitis, P. J. “Downscaling soil moisture over regions that include multiple coarse-resolution grid cells,” *Remote Sensing of Environment*, 199, 187–200, https://doi.org/10.1016/j.rse.2017.07.021, 2017.

Horton, P., Jaboyedoff, M., Rudaz, B., and Zimmermann, M. “Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale,” *Nat. Hazards Earth Syst. Sci.,* 13, 869–885, https://doi.org/10.5194/nhess-13-869-2013, 2013.

Huggel, C., Kääb, A., Haeberli, W., and Krummenacher, B. “Regional-scale GIS-models for assessment of hazards from glacier lake outbursts: evaluation and application in the Swiss Alps,” *Nat. Hazards Earth Syst. Sci*., 3, 647–662, https://doi.org/10.5194/nhess-3-647-2003, 2003.

Hunter, G., and Fell, R. “Travel distance angle for “rapid” landslides in constructed and natural soil slopes,” *Can. Geotech. J*., 40, 1123–1141, https://doi.org/10.1139/t03-061, 2003.

Iverson, R. M. “Objective delineation of lahar-inundation hazard zones,” 13, 1998.

Iverson, R.M. “The physics of debris flows,” *Rev. Geophys*., 35, 245–296, https://doi.org/10.1029/97RG00426, 1997.

Iverson, R. M., and Denlinger, R. P. “Mechanics of debris flows and debris-laden flash floods,” Seventh Federal Interagency Sedimentation Conference, Reno, Nevada, USA, 2001.

Kim, H. G., Lee, D. K., Park, C., Kil, S., Son, Y., and Park, J. H. “Evaluating landslide hazards using RCP 4.5 and 8.5 scenarios,” *Environ Earth Sci*, 73, 1385–1400, https://doi.org/10.1007/s12665-014-3775-7, 2015.

Korb, J. E., Stoddard, M. T., and Huffman, D. W. “Effectiveness of Restoration Treatments for Reducing Fuels and Increasing Understory Diversity in Shrubby Mixed-Conifer Forests of the Southern Rocky Mountains,” USA, 20, 2020.

Lari, S., Frattini, P., and Crosta, G. B. “A probabilistic approach for landslide hazard analysis,” *Engineering Geology*, 182, 3–14, https://doi.org/10.1016/j.enggeo.2014.07.015, 2014.

Lee, S., Ryu, J.-H., and Kim, I.-S. “Landslide susceptibility analysis and its verification using likelihood ratio, logistic regression, and artificial neural network models: case study of Youngin, Korea,” *Landslides*, 4, 327–338, https://doi.org/10.1007/s10346-007-0088-x, 2007.

Legros, F. “The mobility of long-runout landslides,” *Engineering Geology*, 63, 301–331, https://doi.org/10.1016/S0013-7952(01)00090-4, 2002.

Leng, Y., Peng, J., Wang, Q., Meng, Z., and Huang, W. “A fluidized landslide occurred in the Loess Plateau: A study on loess landslide in South Jingyang tableland,” *Engineering Geology*, 236, 129–136, https://doi.org/10.1016/j.enggeo.2017.05.006, 2018.

Lockyear, R. A. “Identification of parameters for predicting long-runout landslides in the western United States,” M. S. Thesis, Golden, Colorado, Colorado School of Mines, 2018.

Lukas, J., Wolter, K., Mahoney, K., Barsugli, J., Doesken, N., Ryan, W., Rangwala, I., Livneh, B., Gordon, E., Hoerling, M., Kiladis, G., and Nacu-Schmidt, A. “Severe flooding on the Colorado Front Range,” September 2013: A preliminary assessment from the CIRES Western Water Assessment at the University of Colorado, NOAA ESRL Physical Science Division, and the CSU Colorado Climate Center, 2013.

Mahoney, K., Alexander, M., Scott, J. D., and Barsugli, J. “High-Resolution Downscaled Simulations of Warm-Season Extreme Precipitation Events in the Colorado Front Range under Past and Future Climates,” *Journal of Climate*, 26, 8671–8689, https://doi.org/10.1175/JCLI-D-12-00744.1, 2013.

Massei, G. “Module Soil Texture Classification,” SAGA-GIS Module Library Documentation v2.2.1 [code], 2007.

Matsuo, M., and Kuroda, K. “Probabilistic approach to design of embankments,” *Soils and Foundations*, 14, 2, pg. 1-17, https://doi.org/10.3208/sandf1972.14.2_1.

McGuire, L. A., Rengers, F. K., Kean, J. W., Coe, J. A., Mirus, B. B., Baum, R. L., and Godt, J. W. “Elucidating the role of vegetation in the initiation of rainfall-induced shallow landslides: Insights from an extreme rainfall event in the Colorado Front Range: SHALLOW LANDSLIDES,” *Geophys. Res. Lett*., 43, 9084–9092, https://doi.org/10.1002/2016GL070741, 2016.

Mergili, M., Marchesini, I., Rossi, M., Guzzetti, F., and Fellin, W. “Spatially distributed three-dimensional slope stability modelling in a raster GIS,” *Geomorphology*, 206, 178–195, https://doi.org/10.1016/j.geomorph.2013.10.008, 2014.

Metternicht, G., Hurni, L., and Gogu, R. “Remote sensing of landslides: An analysis of the potential contribution to geo-spatial systems for hazard assessment in mountainous environments,” *Remote Sensing of Environment*, 98, 284–303, https://doi.org/10.1016/j.rse.2005.08.004, 2005.

Milledge, D. G., Densmore, A. L., Bellugi, D., Rosser, N. J., Watt, J., Li, G., and Oven, K. J. “Simple rules to minimise exposure to coseismic landslide hazard,” *Nat. Hazards Earth Syst. Sci*., 19, 837–856, https://doi.org/10.5194/nhess-19-837-2019, 2019.

Montgomery, D. R., and Dietrich, W. E. “A physically based model for the topographic control on shallow landsliding,” *Water Resour. Res*., 30, 1153–1171, https://doi.org/10.1029/93WR02979, 1994.

Nash, J. E., and Sutcliffe, J. V. “River flow forecasting through conceptual models Part I –A discussion of principles,” *Journal of Hydrology*, 10, 282-290, North-Holland Publishing Co., Amsterdam, 1970.

National Land Cover Database (NLCD). CONUS [data set], www.mrlc.gov, 2016.

Norris, J. E., Stokes, A., Mickovski, S. B., Cammeraat, E., van Beek, R., Nicoll, B. C., and Achim, A. (Eds.) “Slope Stability and Erosion Control: Ecotechnological Solutions,” Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-1-4020-6676-4, 2008.

Pack, R. T., Tarboton, D. G., and Goodwin, C. N. “The SINMAP approach to terrain stability mapping,” in: Proceedings of the 8th international congress of the international association of engineering

geology and the environment, Vancouver, British Columbia, Canada, 21–25 September, vol. 2, AA Balkema, Rotterdam,1157–1165, 1998.

Pauly, M. J., Niemann, J. D., Scalia, J., Green, T. R., Erskine, R. H., Jones, A. S., and Grazaitis, P. J. "Enhanced hydrologic simulation may not improve downscaled soil moisture patterns without improved soil characterization," *Soil Sci. Soc. Am. J.*, 84, 672–689, https://doi.org/10.1002/saj2.20052, 2020.

Pendergrass, A. G., and Knutti, R. "The Uneven Nature of Daily Precipitation and Its Change," *Geophys. Res. Lett.*, 45, 11,980-11,988, https://doi.org/10.1029/2018GL080298, 2018.

Peng, J., Loew, A., Zhang, S., Wang, J., and Niesel, J. "Spatial Downscaling of Satellite Soil Moisture Data Using a Vegetation Temperature Condition Index," *IEEE Trans. Geosci. Remote Sensing*, 54, 558–566, https://doi.org/10.1109/TGRS.2015.2462074, 2016.

Petley, D. "Global patterns of loss of life from landslides," *Geology*, 40, 927–930, https://doi.org/10.1130/G33217.1, 2012.

Peruzzetto, M., Mangeney, A., Grandjean, G., Levy, C., Thiery, Y., Rohmer, J., and Lucas, A. "Operational Estimation of Landslide Runout: Comparison of Empirical and Numerical Methods," *Geosciences*, 10, 424, https://doi.org/10.3390/geosciences10110424, 2020.

Priestley, C., and Taylor, R. "On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters," *Monthly Weather Review*, 100, 81-92, 1972.

Quan Luna, B., Blahut, J., van Asch, T., van Westen, C., and Kappes, M. "ASCHFLOW - A dynamic landslide run-out model for medium scale hazard analysis," *Geoenviron Disasters*, 3, 29, https://doi.org/10.1186/s40677-016-0064-7, 2016.

Raia, S., Alvioli, M., Rossi, M., Baum, R. L., Godt, J. W., and Guzzetti, F. "Improving predictive power of physically based rainfall-induced shallow landslide models: a probabilistic approach," *Geosci. Model Dev.*, 7, 495–514, https://doi.org/10.5194/gmd-7-495-2014, 2014.

Ranney, K. J., Niemann, J. D., Lehman, B. M., Green, T. R., and Jones, A. S. "A method to downscale soil moisture to fine resolutions using topographic, vegetation, and soil data," *Advances in Water Resources*, 76, 81–96, https://doi.org/10.1016/j.advwatres.2014.12.003, 2015.

Rengers, F. K., McGuire, L. A., Coe, J. A., Kean, J. W., Baum, R. L., Staley, D. M., and Godt, J. W. "The influence of vegetation on debris-flow initiation during extreme rainfall in the northern Colorado Front Range," *Geology*, 44, 823–826, https://doi.org/10.1130/G38096.1, 2016.

Rosso, R., Rulli, M. C., and Vannucchi, G. "A physically based model for the hydrologic control on shallow landsliding: HYDROLOGIC CONTROL ON SHALLOW LANDSLIDING," *Water Resour. Res.*, 42, https://doi.org/10.1029/2005WR004369, 2006.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., Freden, S. C., Mercanti, E. P., and Becker, M. A. (Eds.) "Monitoring vegetation systems in the great plains with ERTS," Third Earth Resources Technology Satellite-1 Symposium-Volume I: Technical Presentations, NASA, Washington, D.C., pp. 309-317, 1973.

Salvati, P., Bianchi, C., Rossi, M., and Guzzetti, F. "Societal landslide and flood risk in Italy," *Nat. Hazards Earth Syst. Sci.*, 10, 465–483, https://doi.org/10.5194/nhess-10-465-2010, 2010.

Scheidl, C., and Rickenmann, D. "Empirical prediction of debris-flow mobility and deposition on fans," *Earth Surf. Process. Landforms*, n/a-n/a, https://doi.org/10.1002/esp.1897, 2009.

Soil Survey Staff. Gridded National Soil Survey Geographic (gNATSGO) Database for Colorado, United States Department of Agriculture Natural Resources Conservation Service, [dataset], https://nrcs.app.box.com/v/soils, 2020.

Stoddard, M. T., Sánchez Meador, A. J., Fulé, P. Z., and Korb, J. E. "Five-year post-restoration conditions and simulated climate-change trajectories in a warm/dry mixed-conifer forest, southwestern Colorado, USA," *Forest Ecology and Management*, 356, 253–261, https://doi.org/10.1016/j.foreco.2015.07.007, 2015.

Strauch, R., Istanbulluoglu, E., Nudurupati, S. S., Bandaragoda, C., Gasparini, N. M., and Tucker, G. E. "A hydroclimatological approach to predicting regional landslide probability using Landlab," *Earth Surf. Dynam*., 6, 49–75, https://doi.org/10.5194/esurf-6-49-2018, 2018.

Tarancón, A. A., Fulé, P. Z., Shive, K. L., Sieg, C. H., Meador, A. S., and Strom, B. "Simulating post-wildfire forest trajectories under alternative climate and management scenarios," *Ecological Applications*, 24, 1626–1637, https://doi.org/10.1890/13-1787.1, 2014.

Tarboton, D.G. "A New Method for the Determination of Flow Directions and Contributing Areas in Grid Digital Elevation Models," *Water Resources Research*, 33, 309–319, 1997.

TauDEM Version 5. https://hydrology.usu.edu/taudem/taudem5/, last access: 14 June 2021.

Timilsina, S., Niemann, J. D., Rathburn, S. L., Rengers, F. K., and Nelson, P. A. "Modeling hydrologic processes associated with soil saturation and debris flow initiation during the September 2013 storm, Colorado Front Range," *Landslides*, https://doi.org/10.1007/s10346-020-01582-5, 2021.

Tucker, G., Lancaster, S., Gasparini, N., and Bras, R. "The Channel-Hillslope Integrated Landscape Development Model (CHILD)," in: *Landscape Erosion and Evolution Modeling*, edited by: Harmon, R. S. and Doe, W. W., Springer US, Boston, MA, 349–388, https://doi.org/10.1007/978-1-4615-0575-4_12, 2001.

U.S. Geological Survey (USGS). 1/3rd arc-second Digital Elevation Models (DEMs), USGS [dataset], https://nationalmap.gov/elevation.html, 2015.

U. S. Geological Survey (USGS). Landsat 5, USGS 20110928 from 20160830 [dataset], doi: /10.5066/F7N015TQ, 2016.

Vakhshoori, V., and Zare, M. "Is the ROC curve a reliable tool to compare the validity of landslide susceptibility maps?" *Geomatics, Natural Hazards and Risk*, 9, 249–266, https://doi.org/10.1080/19475705.2018.1424043, 2018.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K. "The representative concentration pathways: an overview," *Climatic Change*, 27, 2011.

van Westen, C. J., van Asch, T. W. J., and Soeters, R. "Landslide hazard and risk zonation—why is it still so difficult?" *Bull Eng Geol Environ*, 65, 167–184, https://doi.org/10.1007/s10064-005-0023-0, 2006.

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

Vandromme, R., Thiery, Y., Bernardie, S., and Sedan, O. "ALICE (Assessment of Landslides Induced by Climatic Events): A single tool to integrate shallow and deep landslides for susceptibility and hazard assessment," *Geomorphology*, 367, 107307, https://doi.org/10.1016/j.geomorph.2020.107307, 2020.

Varnes, D.J. “Slope movement types and processes,” in: Schuster RL, Krizek RJ (eds) Landslides, analysis and control, special report 176: Transportation research board, National Academy of Sciences, Washington, DC., 11–33, 1978.

Vereecken, H., Huisman, J. A., Bogena, H., Vanderborght, J., Vrugt, J. A., and Hopmans, J. W. “On the value of soil moisture measurements in vadose zone hydrology: A review: SOIL MOISTURE AND HYDROLOGY,” *Water Resour. Res*., 44, https://doi.org/10.1029/2008WR006829, 2008.

Wang, W. J., He, H. S., Thompson, F. R., Spetich, M. A., and Fraser, J. S. “Effects of species biological traits and environmental heterogeneity on simulated tree species distribution shifts under climate change,” *Science of The Total Environment*, 634, 1214–1221, https://doi.org/10.1016/j.scitotenv.2018.03.353, 2018.

Woolridge, D. D., Niemann, J. D., Perry, M. A., Bauer, K. E., and McCormick, W. T. “Identifying Runoff Production Mechanisms for Dam Safety Applications in the Colorado Front Range,” *J. Hydrol. Eng*., 25, 05020016, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001958, 2020.

Wicki, A., Lehmann, P., Hauck, C., Seneviratne, S. I., Waldner, P., and Stähli, M. “Assessing the potential of soil moisture measurements for regional landslide early warning,” *Landslides*, 17, 1881–1896, https://doi.org/10.1007/s10346-020-01400-y, 2020.

Wu, W., and Sidle, R. C. “A Distributed Slope Stability Model for Steep Forested Basins,” *Water Resour. Res.*, 31, 2097–2110, https://doi.org/10.1029/95WR01136, 1995.

Wuebbles, D.J., Easterling, D.R., Hayhoe, K., Knutson, T., Kopp, R.E., Kossin, J.P., Kunkel, K.E., LeGrande, A.N., Mears, C., Sweet, W.V., Taylor, P.C., Vose, R.S., and Wehner, M.F. “Chapter 1: Our globally changing climate,” in: Climate Science Special Report: Fourth National Climate Assessment, Volume I, edited by: Wuebbles, D.J., Fahey, D.W., Hibbard, K.A., Dokken, D.J., Stewart, B.C., and Maycock, T.K., U.S. Global Change Research Program, 35-72, doi:10.7930/J08S4N35, 2017.

Xia Y., Ek M.B., Wu Y., Ford T., and Quiring, S.M. “Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture Part I: comparison and analysis,” *Journal of Hydrometeorology*, 16(5), 1962–1980, https://doi.org/10.1175/JHM-D-14-0096.1, 2015.

Xia Y., Mitchell K., Ek M., Sheffield J., Cosgrove B., Wood E., Luo L., Alonge C., Wei H., Meng J., Livneh B., Lettenmaier D., Koren V., Duan Q., Mo K., Fan Y., Mocko D. (a). “Continental scale water and energy flux analysis and validation for the North American land data assimilation system project phase 2 (NLDAS-2): 1. Intercomparison and application of model products,” *Journal of Geophysical Research Atmospheres*, 117(3):1–27, https://doi.org/10.1029/2011JD016048, 2012.

Xia, Y., et al. (b). NCEP/EMC (2009), “NLDAS Mosaic Land Surface Model L4 Hourly 0.125 x 0.125 degree V002,” Edited by David Mocko, NASA/GSFC/HSL, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), [data set], 10.5067/EN4MBWTCENE5, 2012.

Yordanov, V., and Brovelli, M. A. “Comparing Model Performance Metrics for Landslide Susceptibility Mapping,” *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLIII-B3-2020, 1277–1284, https://doi.org/10.5194/isprs-archives-XLIII-B3-2020-1277-2020, 2020.

Zhu, R., Guo, Y., and Xue, J.-H. “Adjusting the imbalance ratio by the dimensionality of imbalanced data, Pattern Recognition Letters,” 133, 217–223, https://doi.org/10.1016/j.patrec.2020.03.004, 2020.