Empirical evaluation of shallow landslide susceptibility, frequency, and runout by landform

Unstable Slope Criteria Project Team: Julie Dieu, Jeff Keck, Dan Miller, Ted Turner, (Greg Stewart, former member)

Executive Summary

This study design entails two of the four studies that comprise the Unstable Slope Criteria Project. These four studies are intended to evaluate and potentially improve the current criteria for identifying areas where public resources may be impacted or public safety threatened by landslides. The current criteria resulted from years of diligent effort by experienced scientists working with the Watershed Analysis and Landslide Hazard Zonation projects in Washington. They are based on well-established methods for the assessment of landslide susceptibility. Two events motivated a re-evaluation of these criteria: 1) the December 2007 storm in southwest Washington, which triggered many landslides outside of areas recognized as susceptible to landslides, and 2) the March 2014 Oso landslide, in which the death of 43 individuals highlighted the lack of effective criteria for evaluating areas susceptible to impacts from upslope landslides[1]. Traditional methods and data sources are unlikely to effectively evaluate the current criteria, as it is those methods and data sources on which the current criteria are already based. Neither will traditional methods and data sources have much to offer for improved assessment of susceptibility to landslide runout, or that would have already been done. However, the last two decades, particularly the last several years, have seen the development of expansive new data sources and of computational methods for using those data effectively. As a result, we now have the resources to re-evaluate the unstable slopes criteria. These new resources require iterative development and testing of new methods for assessing shallow landslide susceptibility, frequency, and runout. This document lays out a study design for accomplishing that.

We have three data sources that did not exist when the current criteria were developed: 1) aerial lidar point clouds, 2) contiguous land cover maps based on lidar point cloud and multiband imagery analyses, and 3) contiguous precipitation time series based on interpolation from gauge records, weather radar, and analysis of satellite imagery. We also have three data analysis techniques that were not previously available: 1) measurement of elevation changes between lidar acquisitions to provide measures of landslide location, area, and volume precisely aligned with lidar-derived measures of topography, 2) analysis of large datasets and determining the confidence to place in the predictions based on those datasets using machine learning algorithms, and 3) implementation of those algorithms using open-source software (python, R). Standard protocols for using these data and computational resources to analyze landslide susceptibility have not been developed; however, robust statistical and machine-learning protocols for data analysis are well developed. This study design applies these well-developed techniques to the analysis of shallow landslide susceptibility.

Current unstable-slope criteria in Washington have no explicit protocol for estimating susceptibility of downslope resources to impacts and threats from landslides that originate upslope. A key study design component is explicitly linking landslide initiation and runout. Such a linkage is crucial for determining

---

[1] While Oso was a deep-seated landslide and the focus of this project is shallow, distributed landslides, it highlighted the fact that we have no standard or systematic way to quantitatively assess potential runout for landslides of any type and was a motivating factor for putting together a team to examine the criteria used in forest practices to identify areas where public safety may be threatened by landslides.

the impacts to public resources and threats to public safety posed by landslides because these impacts and threats primarily exist downslope, sometimes far downslope, from the sites of landslide initiation.

This document describes how landslide susceptibility can be measured, lays out the challenges posed in obtaining those measurements, and finally shows how new data and analysis resources can be applied to meet those challenges.

Table of Contents

1 Introduction

1.1 Overview of Unstable Slope Criteria Project

The Unstable Slope Criteria Project is part of the Cooperative Monitoring, Evaluation, and Research (CMER) Committee's Mass Wasting Effectiveness Monitoring Program. This project addresses the Forests & Fish Report[2] Schedule L-1 research topic "**Test the accuracy and lack of bias of the criteria for identifying unstable landforms in predicting areas with a high risk of instability**" and will answer the Unstable Slopes Rule Group critical question posed by CMER **"Are unstable landforms being correctly and uniformly identified and evaluated for potential hazard?"[3]**

In February 2017, the Technical Writing and Implementation Group (TWIG) for this project submitted to CMER the document "Unstable Slope Criteria Project – Research Alternatives," hereafter called the Alternatives document. It articulated research objectives, reviewed current best-available science for identifying unstable slopes, and proposed a set of five research projects. Subsequently, in April 2017, the Timber/Fish/Wildlife (TFW) Policy Committee (hereafter called Policy) approved the research projects recommended by the TWIG in the Alternatives document.

In response to initial Independent Scientific Peer Review (ISPR) comments of that document, a final set of four projects were identified:

1. Automated Object-Based Landform Mapping with High-Resolution Topography;
2. Empirical Evaluation of Shallow Landslide Susceptibility and Frequency by Landform;
3. Empirical Evaluation of Shallow Landslide Runout; and
4. Models to Identify Landscapes/Landslides Most Susceptible to Management.

The primary goal of the Forest Practices unstable slopes rules is to avoid areas where management-induced landslides could impact public resources or threaten public safety. The research objectives are to reduce errors associated with the unstable-slope and landslide-delivery-assessment criteria. Those errors include 1) misidentification of rule-identified landforms (RIL), 2) exclusion of unstable slopes that do not meet RIL criteria (i.e., not identifying unstable slopes – false negatives), 3) inclusion of stable slopes that meet RIL criteria (i.e., identifying stable slopes as unstable – false positives), 4) incorrect determination of the potential for a landslide to deliver material to a downslope public resource or to threaten downslope public safety[4].

This set of four projects seeks to capitalize on newly available data and analysis tools to evaluate and refine RIL definitions to thereby reduce errors associated with the unstable-slope criteria. Project 1 serves as a proof of concept for automated and objective mapping of landforms as a baseline for subsequent analyses; that study is underway. The study design[5] for Project 1 was approved in March 2020; CMER staffing shortages have delayed the completion of the project, but a report is in preparation. Projects 2 and 3 (this design) will provide accurate statistics about where and how many landslides occur on Washington's timberlands. During the development of the study designs, it was

---

[2] https://www.dnr.wa.gov/publications/fp_rules_forestsandfish.pdf
[3] See the CMER workplan: https://www.dnr.wa.gov/publications/fp_cmer_2023_2025_wrkplan.pdf
[4] Only the first three objectives were listed in the Alternatives document; the fourth is added here for completeness and consistency.
[5] https://www.dnr.wa.gov/publications/bc_tfw_unstable_study_20200305.pdf

decided to combine Projects 2 and 3 because they will rely on the same data and types of analyses. Hence, this document provides a study design that addresses shallow landslide susceptibility and runout. This study utilizes object-based mapping techniques developed with Project 1, but the development of this project design is not dependent on the details of those techniques. Therefore, to maintain timely progress on the sequence of projects, this study design was produced prior to the completion of a report for Project 1. This study, Projects 2 and 3, examines the statistics of landslide occurrence: where they occur, their size, runout, and (potentially) frequency. It does not, however, look at why they occur or how forest practices might alter those statistics; that is for Project 4. A study design for Project 4 will be developed after the susceptibility and runout project results are available.

The critical question posed in the Alternatives document for the Unstable Slope Criteria Project is:

> "What modifications to the unstable slopes criteria and delivery-assessment methods would result in more accurate and consistent identification of *1) unstable slopes and landforms*, 2) unstable slopes and landforms sensitive to forest-practices-related changes in landslide processes, and 3) *unstable-slope and landform conditions where landslide runout would likely have an adverse impact to public resources or a threat to public safety is possible*?"

Evaluation of shallow landslide susceptibility and runout addresses the first and third items in the critical question above: accurate and consistent identification of 1) unstable slopes and landforms and 2) locations where landslide runout will likely harm public resources or pose a threat to public safety[6]. The second depends on the first because to identify locations susceptible to impacts from upslope landslides, we must first identify upslope locations susceptible to landslide initiation. The potential for downslope impacts depends on both the potential for upslope landslide initiation and the potential for landslide runout to the downslope point. For our purposes, the downslope points of interest are determined by the presence of a public resource, any typed water of the state, and any human infrastructure, including public roads, buildings, pipelines, transmission lines.[7]

The critical question does not mention "potential," but it is an important component of this task because the WA forest practices rules use potential for landslide initiation and runout to define unstable slopes and landforms (WAC 222-16-050(1)(d)). The potential that a hillslope site will fail in any year varies from point to point. The potential for impacts to any downslope point depends on both the potential for failure and the potential for runout to that point. Because shallow landslides can run out long distances, a downslope point may be susceptible to impacts from many upslope initiation sites, each with a different potential for landslide initiation and a different potential for runout to that location. Assessing susceptibility to impacts from upslope landslides requires the determination of the probability that any of those upslope initiation sites could fail and run out to that point. This probability depends on the number of upslope initiation sites, the probability of initiation and landslide size, and the runout distance to each one. This assessment is worthwhile because a site susceptible to landslide

---

[6] Note, however that effects of forest cover on landslide occurrence and runout will be examined empirically with these projects.

[7] Washington State has not yet provided a clear definition of what poses a threat to public safety. We therefore need methods that can work for any specified point of interest.

impacts once every 10,000 years might reasonably be viewed differently than a site susceptible to landslides every ten years.

It is also worth considering the nature and magnitude of the impacts. A landslide that clears ten thousand square meters of hillslope and deposits five thousand cubic meters of material in a fish-bearing stream has a different impact than a ten-square-meter landslide that dumps five cubic meters into that stream. A landform prone to the first poses different impacts to public resources and threats to public safety than a landform prone to the second, even if there are hundreds of small landslides for every one of the larger ones. Not all impacts and threats are created equal. To identify locations where, as required by the critical question, "landslide runout would likely have an adverse impact to public resources," we need to look at both the potential for impacts and the nature of those impacts – are they *adverse* impacts? Likewise, to identify locations where "a threat to public safety is possible," we need some way to gauge the magnitude of that threat. Does "is possible" include an event occurring once every 500 years? Oso made this more than just an academic issue.

The goal of this project is a quantitative assessment of shallow landslide initiation and runout susceptibility. The above paragraphs outline the scope of this task: we need quantitative measures of susceptibility for failure and runout extent for all hillslope locations, and we need the capability to integrate those measures over multiple failure sites to obtain quantitative measures of downslope susceptibility to impacts and threats from landslides originating upslope. In addition, quantitative measures for downslope impacts and threats require quantitative measures of landslide size. Traditional methods to develop current RIL-criteria are not up to these tasks. However, newly available data products and analysis techniques could be leveraged. To accomplish this, we identify seven objectives:

1. Use of lidar differencing to build landslide inventories.

2. Association in terms of a spatial probability for landslide number, location, surface area, volume, and runout extent as functions of measurable terrain elements. These elements include:

   a. Topographic attributes, such as surface gradient, curvature, and contributing area.

   b. Mapped geologic unit and soil type.

   c. Precipitation intensity averaged over specified periods (e.g., 3 hours, daily, three days).

   d. Forest stand characteristics, such as stand height and age.

   e. All upslope initiation sites and intervening flow paths are included in determining runout probability.

   f. Calculated factor-of-safety from a process-based model of slope instability that includes spatially variable estimates of soil depth and upslope contributing area associated with storms of variable duration.

3. Identification of those terrain elements most relevant for the determination of probability.

4. Translation of probabilities to relative landslide density as a function of the terrain elements identified in step 3.

5. Association of absolute landslide density with measures of storm magnitude. Translation of this relationship to landslide rate as a function of the frequency distribution of storm magnitudes (intensity times duration) or storm return interval.

6. Integrating the resulting spatially distributed density values over delineated landforms gives the proportion of landslide numbers, areas, volumes, and runout track length originating from each landform type.

7. Use and development of well-documented and accessible software tools and public-domain data so that the methods developed with this study can be used and replicated by anyone with the requisite computer hardware, software, and expertise.

Research tasks are divided into four main groups:

1. Create landslide inventory from lidar differencing.
2. Compile terrain-element datasets (e.g., precipitation, forest cover, topographic attributes).
3. Develop statistical relationships between landslide location and size with terrain elements.
4. Use these results to evaluate the RIL and other potential criteria for identifying and delineating unstable slopes and landforms.

These tasks will provide the following deliverables:

1. Workflow and protocols for building landslide inventories using lidar differencing.
2. Computer code and scripts for the derivation of elevation derivatives.
3. Python and R scripts for statistical analysis of inventory data.
4. Evaluation of current RILs in terms of the landslide proportions and density. Identification of potentially useful subdivisions or ranking of RILs. Identification of potential additions or modifications to current RILs.
5. A report describing the developed methodology and results of the study.
6. Digital maps of model outputs, including landslide inventories and modeled initiation and delivery probability.

The efforts outlined under this work plan will advance our ability to assess landslide susceptibility, frequency, and runout by landform with which to evaluate the effectiveness of current RIL criteria. This research is a formidable challenge that is possible now due to recent advances in relevant technologies. The scope of work includes using new methods employed on new dataset sources. This Research is an iterative process, and the deliverables will reflect this. This study design sets ambitious targets for each task. The deliverables will report on the progress made for each target in terms of the progress made, the challenges discovered, and the direction of future research.

The remainder of Section 1 describes the conceptual foundation for the analysis tasks proposed. Section 2 then describes the limitations of traditional methods and data sources for assessing susceptibility, focusing on landslide inventories. Section 3 describes how newly available data and analysis techniques can overcome these limitations. Section 4 then lays out the analysis tasks for this study. Section 5 elaborates on methods to be used for calculating susceptibility, including measures of uncertainty. Section 6 provides an explanation and elaboration of why this particular set of methods has been chosen. Section 7 provides a task list, and Section 8 provides a set of deliverables for the proposed study. An appendix provides additional details for methods that may be applied for the study.

1.2 Definitions

This document uses terms specific to the field of geomorphology and analysis of spatial data that may not be familiar to all readers. Likewise, although we seek consistency, our use of terms may differ from how they are used and interpreted elsewhere. Therefore, here is a glossary of potentially unfamiliar or ambiguous terms, most of which are elaborated on in the text. These are listed in thematic, rather than alphabetic, order.

Geomorphology. The study of Earth's surface and of the processes that act to create and modify that surface.

Landscape. In a geomorphic context, "landscape" refers to some contiguous portion of Earth's surface and includes all the features that compose that surface. In a broad sense, the term "landscape" refers to the features and the physical processes and process interactions that create and modify those features. Therefore, this document also uses the term "terrain" as a synonym for "landscape," that infers a quantified representation of landscape features and processes.

Landform. In the broadest sense, a landform is a delineated portion of the landscape having some homogenous set of attributes. These attributes may include numerous measures of the topography; soil textures, types, and ages; and substrate lithology and structure. The definition of a landform may include interpretations of surface-process types and rates (e.g., flood plains, terraces, fans, hillslopes). The criteria for delineating a landform, both on the ground and on a map, will vary with the purpose for delineating landforms. For a given set of criteria, the delineated landforms will vary in size and extent depending on who does the delineation and the tools and data they use.

Rule-identified landform (RIL). In Washington, placement of unstable-slope buffers is determined through 1) screening of remote data (e.g., topographic maps, historic air photos, geologic data) and 2) field identification and delineation of certain well-defined "Rule-Identified Landforms" (RILs, WAC 222-16-050[8]) assessed in combination with site-specific indicators of slope instability. Criteria for identifying RILs rely on a field-based interpretation of topographic form (e.g., inner gorges, convergent headwalls, bedrock hollows) and landscape position (e.g., the foot of a deep-seated landslide, the outside of a meander bend) and measurements of the surface gradient. These attributes can be estimated on maps and with computer-based analyses of digital elevation models (DEMs) so that estimated RIL locations and boundaries can be drawn as polygons on maps. The final determination is field based. Observations of landslides over many years across Washington state indicate five RILs (from WAC 222-16-050):

(A) Inner gorges, convergent headwalls, or bedrock hollows with slopes steeper than thirty-five degrees (seventy percent);

(B) Toes of deep-seated landslides, with slopes steeper than thirty-three degrees (sixty-five percent);

(C) Groundwater recharge areas for glacial deep-seated landslides;

(D) Outer edges of meander bends along valley walls or high terraces of an unconfined meandering stream; or

(E) Any areas containing features indicating the presence of potential slope instability, which cumulatively indicate the presence of unstable slopes.

---

[8] https://app.leg.wa.gov/wac/default.aspx?cite=222-16-050

Landform type. Criteria for delineating landforms typically seek to differentiate different landform types. The number and diversity of landform types vary with different sets of criteria. For example, RILs are intended to delineate potentially unstable slopes.

Landform polygon. A polygon drawn on a map may be on paper or within a computer geographic information system (GIS) to delineate an individual landform.

Pixel. A digital image, such as a digital photo taken with a smartphone, comprises a grid of small (typically square) elements called pixels, each of a uniform color. When viewed from a distance, that grid forms an image – a picture –but the individual pixels are visible up close. Digital imagery collected by satellites and digital aerial photographs also consists of grids of pixels.

Digital elevation model (DEM). A grid of ground-surface elevation values. Each grid point defines a grid cell. A DEM cell is similar to an image pixel, representing a value over some spatial extent. Derivatives calculated from the elevation values, such as surface gradient, are also stored and displayed over the same grid points and cells. In an image, each pixel has a single color or intensity value. With a DEM, it is typically assumed that the elevation of the ground surface and elevation derivatives vary continuously between the grid points and across the cells.

Lidar DEM. A DEM created from data collected by laser altimetry (Lidar[9]). Lidar-derived DEMs typically have a horizontal resolution of 1 meter (3 feet). These are a great improvement over the previously available DEMs in Washington state, created by interpolating elevations from contour lines on 1:24,000-scale topographic maps over a 10-m horizontal spacing. Lidar DEMS can resolve topographic features over several-meter length scales; however, the actual resolution depends on the spatial density (number per unit area) of the detected laser signals that penetrate the vegetation layer and reflect from the ground surface (ground returns) and the algorithm and cell-size chosen to create the DEM from the lidar point cloud. The ground-return density is lower in areas with dense forest canopy. In addition, the signal density (total number of laser pulses per unit area) is generally higher with more recent lidar acquisitions, so newer lidar DEMs can typically resolve greater detail than older lidar DEMs.

Geographic Information System (GIS). Software for analysis of digital data. ArcGIS[10] by ESRI is the GIS used by state agencies in Washington.

Raster data[11]. A grid-based representation of spatial data used in GIS. A binary image, composed of regular arrays of pixels, and a DEM, composed of a regular grid of elevation point values, are both raster representations of values over some spatial extent. A raster may represent continuously varying values, as with a DEM, or zones of homogenous values, as with mapped soil units, lithologic units, or land-cover type.

Vector data[12]. A representation of spatial features using points, lines, and polygons in GIS. Vector data work well for features with distinct locations and boundaries, such as a polygon outlining a landslide scar, a point indicating the midpoint of the landslide scar, and a line indicating the center line of the runout path of a landslide.

---

[9] https://www.dnr.wa.gov/lidar
[10] https://www.esri.com/en-us/arcgis/about-arcgis/overview
[11] https://desktop.arcgis.com/en/arcmap/latest/manage-data/raster-and-images/what-is-raster-data.htm
[12] https://docs.qgis.org/2.8/en/docs/gentle_gis_introduction/vector_data.html

Pixel-based image analysis. Raster data can be categorized and analyzed using pixel or grid-point values. The calculation returns a value for each pixel or grid point. This type of analysis can be done on any grid-based (raster) set of values. For example, a surface gradient can be calculated for every grid point in a DEM based on the elevation and distance to neighboring points to produce a new grid of values. Pixel-based analyses produce raster outputs. Because modeled values can vary from pixel to pixel, the outputs can be challenging to interpret regarding what they represent on the ground. This challenge is where object-based image analysis comes in.

Object-based image analysis (OBIA)[13]. Object-based image analysis seeks to translate a raster representation of some spatially variable attributes to a vector polygon representation of individual features or objects, where each polygon represents a distinct feature or an area with homogeneous attributes. Originally developed to aid in the interpretation of medical imagery, the techniques are now extensively applied for analysis of satellite imagery and in GIS for delineation and analysis of features on the ground, such as roads, rivers, and landforms. Geographic object-based image analysis (GEOBIA) is the application of OBIA techniques for geographic features, as done in this set of studies for the delineation of landforms and other automated analyses of geospatial data; it is the method used to delineate RIL and non-RIL landforms in Project #1 and for this project. OBIA uses image segmentation techniques, where pixels are segmented into homogenous groups, with "homogeneity" determined by the types of identified features. Segmentation may utilize a variety of types of information, including pixel (grid point) values, the spatial distribution of values, and the size and shape of the segmented objects. Segmentation may also use information from multiple rasters. For example, landform delineation may depend on surface gradient and curvature rasters. The delineated objects are then classified into a distinct set of object types. Here, these types are different landform types.

eCognition[14]. Software used for GEOBIA.

Terrain element. This term refers to some quantifiable set of attributes of a point or area on the ground. This set includes continuously variable attributes, such as surface gradient, and thematic features, such as mapped rock type or land-cover class. We seek to identify terrain-element attributes associated with landslide initiation and runout and to develop quantitative models that relate terrain-element values to the probability of landslide initiation and runout extent. Because these include ground-surface topography attributes calculated from raster elevation data (a DEM), terrain elements are initially represented using a raster data model where each pixel (grid point) has multiple attributes. Models based on analysis of terrain elements and observed landslides produce raster outputs for the probability of landslide initiation and runout to some specified features, such as stream channels. Raster representations of the terrain elements may be analyzed using OBIA to delineate terrain-element objects or landforms. Every landform has an associated set of terrain element values. Therefore, this study's delineated terrain-element objects – landforms – will directly relate to modeled landslide initiation and runout potential.

Terrain-element-based analysis. An analysis seeking functional relationships between terrain-element values and landslide probability. Because terrain-element values can vary continuously with position and are represented using raster data, the output of a terrain-element-based analysis is also a raster. The

[13] https://gisgeography.com/obia-object-based-image-analysis-geobia/
[14] https://geospatial.trimble.com/products-and-solutions/trimble-ecognition

horizontal DEM grid spacing sets the minimum pixel size of the analysis, although many of the terrain-element-attribute values may be calculated over length scales spanning many DEM cells.

Landform-based analysis. An analysis seeking functional relationships between landform type and landslide probability. In this case, the minimum pixel size of the analysis is the size of the individual, delineated landforms.

Landslide. This document addresses shallow landslides that fail suddenly with the rapid downslope movement of failed material. Types of shallow landslides are described in Table 1.

Initiation. The location of initial soil failure for a shallow landslide. The zone of initial failure is identified by the scar created when a landslide occurs. The scar may be larger than the initial zone of failure since the initial loss of lateral buttressing may trigger progressive soil failures along the margin of the initial scar. However, we only have the final scar extent to identify the failure zone. The landslide scar may be mapped as a polygon or a point feature.

Runout. The downslope extent of travel for mobilized landslide debris.

Landslide density. Landslide density is typically considered regarding the number of landslides per unit area. For this study, we expand the context for landslide density to look also at the area of landslide scars per unit area, the volume of landslide-mobilized debris per unit area, and the landslide (or debris-flow) runout-track length per unit area. Obtaining a measure of landslide density by number requires counting the number of observed landslides within a delineated study area and dividing the number observed by the area examined. Measuring landslide density by landslide area requires measuring the area of observed landslide scars within some delineated study area and dividing the total measured landslide-scar area by the total area examined. Measuring landslide density by volume requires measuring the mobilized volume of all observed landslides within the delineated study area and dividing the total mobilized volume by the total area examined. The density of runout tracks is obtained by measuring the runout track length of all observed landslides within a delineated study area and dividing the cumulative runout-track length by the total area examined. Landslide density varies with physical controls on landslide occurrence, so higher densities occur where physical characteristics (topography, soils, geology, land cover) are associated with landsliding. We use spatial variations in landslide density to identify and rank those physical characteristics – the terrain elements - associated with landsliding. Observable landslide occurrence is strongly influenced by rainfall intensity, so the landslide densities at any time also reflect the preceding sequence of storms and the duration of observations.

Landslide proportion. Integration of landslide density over some specified area gives the number (area, volume, track length) of landslides within that area. We can thus compare terrain elements and landforms in terms of the proportion of all landslides within a study area (e.g., a watershed administrative unit) originating from each. Landslide densities vary with storm history, so densities measured at two different study sites are not directly comparable. Proportional values vary between zero and one. Hence, the use of proportions normalizes measures of landslide density and provides a more robust comparison of landslide occurrences across different study sites.

Landslide susceptibility. Susceptibility analysis seeks to delineate areas on the ground that may be affected by landsliding. This study seeks to expand traditional susceptibility measures, typically reported as "high, medium, low" with no means of validation, to provide a quantitative measure that can be

replicated and validated. We seek to characterize two types of susceptibility: 1) susceptibility to landslide initiation and 2) susceptibility to traversal or deposition of landslide debris originating from an upslope initiation site. Landslide density provides a quantitative measure of landslide susceptibility. We do not use density directly because measured values depend on the recent sequence of landslide-triggering events – spatial and temporal variations in rainfall intensity confound observed density. Rather, we seek functional relationships between terrain elements and observed landslide density and integrate those relationships over a study area to get the total landslide number (or area, volume, or runout-track length). The number, area, volume, or track length of landslides for any combination of terrain elements is then divided by the total number, area, volume, or track length to give the proportion of landslides (or landslide area or volume or runout-track length) associated with that combination of terrain elements. The raster of proportion-by-pixel can be overlain on any sub-area of the study area to predict the proportion of landslides (or area or volume or runout-track length) originating within that subarea. Using a single terrain element as an example, surface gradient, we would seek a functional relationship between observed landslide densities and gradient. With that relationship, we can then go to any delineated area – a harvest unit, for example – and integrate the landslide density associated with each increment of gradient over all hillslopes within that area to get the number (area, volume, track length) of landslides expected within that delineated area. We can also compare how many landslides (or how much landslide area, volume, or runout track length) occur on slopes of some particular increment of the gradient. We could, for example, calculate the proportion of landslides that were observed or are expected on slopes of 60-70% gradient and compare to the number observed or expected on slopes of 70-80% gradient. We propose identifying these relationships for single terrain elements and the sets of terrain elements that influence landslide density. This approach provides a quantitative measure of susceptibility regarding landslide proportions so that terrain elements and delineated landforms can be ranked in terms of the proportion of landslides (by number, area, volume, or track length) they produce. Additionally, this is a measure for which confidence intervals can be calculated and predictions tested against observations.

Landslide probability. Density can be translated to the probability of encountering a landslide scar or runout track at any point on the ground, so the probability is also used as a quantitative measure of susceptibility. Probability can likewise be integrated to give proportions, so these measures are synonymous.

Topographic attribute. This document defines a topographic attribute as a terrain-element value calculated from an elevation raster (DEM). e.g., Maxwell and Shobe (2022) provide a useful review of the attributes used in geomorphic analyses. These include surface gradient, curvature, contributing area, and relief for landslide susceptibility analysis.

Dependent and independent variables. In an empirical model, one set of dependent variables is described as functions of another independent set of variables. The values of the independent variables determine the value of the dependent variable(s). Dependent variables are also referred to as response variables, and the independent variables as explanatory or predictor variables. In this study, landslide density and proportion will be the dependent (response) variables, and the terrain element values will be the independent (explanatory) variables.

Statistical analysis. Various modeling techniques are used to quantify relationships between dependent and independent variables and the uncertainty in those relationships. These techniques include machine

learning, in which relationships are found and evaluated by building and testing many models using subsets of the entire dataset. Machine-learning protocols can provide estimates of certainty (e.g., confidence intervals) for complex relationships involving many independent variables.

R. An open-source computer language used for statistical analyses.

Rule-Identified Landforms (RIL). From WAC 222-16-050(1)(d)(i), They are listed in the rule from (A) to (E) as follows:

A. Inner gorges, convergent headwalls, or bedrock hollows with slopes steeper than 35 degrees (>70%);

B. Toes of deep-seated landslides with slopes steeper than 33 degrees (>65%); C. Groundwater recharge areas for glacial deep-seated landslides;

D. Outer edges of meander bends along valley walls or high terraces of an unconfined meandering stream; or

E. Any areas containing features indicating the presence of potential slope instability, which cumulatively indicate the presence of unstable slopes.

1.3 Landslide Type

The focus here is on *shallow* landslides. These landslides typically involve the failure of a layer of soil overlying a more competent substrate and the failure of soils undercut along stream banks. These include debris slides, flows, and avalanches (Table 1) in the updated Varnes classification (Hungr, Leroueil and Picarelli, 2014). In Washington, these are referred to as shallow landslides, as the depth of soils that fail tends to be a couple of meters or less. This shallow failure depth contrasts deep-seated landslides, which tend to be larger in areal extent and extend to deeper depths. Deep-seated landslides may move slowly or intermittently, although they may also fail catastrophically (e.g., Oso, Keaton et al., 2014). Deep-seated landslides form distinct features in the landscape that can persist for thousands of years, whereas evidence of shallow landslides may persist for only a few decades or less (Reid and Dunne, 1996).

Table 1. Shallow landslide classification, taken from Table 5-3 in Stewart et al., 2013.

Landslide Type	Definition
Debris Slide	Aggregations of coarse soil, rock, and vegetation that lack significant water move at speeds ranging from very slow to rapid down slope by sliding or rolling forward. Debris slides typically travel short distances and form hummocky, poorly sorted deposits.
Debris Flow	The rapid flow of slurries composed of sediment, water, vegetation, and other debris. Debris flows typically initiate as shallow landslides on steep, saturated slopes and travel down convergent channelized pathways.
Debris Avalanche	Partially or fully saturated, rapid landslides similar in process and material to debris flows but not channelized over most of their length. They tend to behave morphologically similar to snow avalanches that splay across the slope.

1.4 Landslide Proportions[15]

In previous work (UPSAG, 2006), landform susceptibility was measured regarding landslide density as the number of mapped landslides per unit area. This density was translated to an estimate of rate, number per unit area per unit time, by dividing by the period of observation, generally, the period for which aerial photography was available. This study takes a similar approach, but to improve the identification of unstable slopes and areas susceptible to impacts from upslope landslides, we seek to expand our measure of susceptibility to focus more specifically on proportions. We will look at proportions from four perspectives:

1. What proportion of shallow landslides originates from different parts of the landscape? These "parts" may be particular delineated landforms or areas associated with some combination of terrain elements.

2. What proportion of the total landscape-scar area is associated with each of the landslides?

3. What proportion of the volume of mobilized material is associated with each of the landslides?

4. And what proportion of downslope impacts are associated with each of the landslides?[16]

Previously, available data were generally inadequate to attempt such comparisons, particularly for items 3 and 4. That situation is changing, however, with new lidar data sources and analysis techniques for using those data (e.g., Scheip and Wegmann, 2022).

The shift in focus to proportions does not ignore landslide density and rate. Landslide numbers (and total area and volume) are obtained by integrating density over area (or rate over area and time). Using density and proportions provides a quantitative and testable measure of susceptibility. If we can identify functions that relate landslide density to terrain attributes, we can predict the proportion of landslides within any portion of a basin. We can then rank different parts of the landscape – different landforms – by the proportion of all landslides in each part. We can also monitor with periodic landslide inventories to see if those predictions hold up and then modify our methods if they do not. Highlighting proportions from different parts of the landscape improves our ability to assess aggregate impacts. Here is a hypothetical example using landforms identified in Washington State Board Manual Section 16.[17]

Consider a 500 square-kilometer (193 square miles) basin. Landslides and landforms are mapped using aerial photos, topographic maps, and field surveys. Three landslide-prone landform types are mapped: steep, convergent slopes (bedrock hollows), inner gorges, and steep, planar slopes. One hundred thirty-five landslides were identified, 45 carrying material to a stream channel. These are distributed as follows:

---

[15] This and subsequent sections focus on where landslides occur. The relative influence of forest practices on landslide proportions will be examined with Study Project 4.

[16] These impacts might be measured a variety of ways, e.g., in terms of the proportion of those landslides that runout to a channel or the proportion of material deposited in channels originating from each landform type

[17] https://www.dnr.wa.gov/publications/bc_fpb_manual_section16.pdf

*Table 2. Landslides by landform, hypothetical example*

Landform	Landform Area (km²)	Landslides	Total Volume (m³)	Density (#/km²)	Delivery to stream (%)	Delivered Density (#/km²)	Delivered Volume (m³)
Hollow	10	15	22,500	1.50	33%	0.50	7,425
Inner gorge	12	40	2,000	3.33	90%	3.00	1,800
Planar slope	145	80	2,400	0.55	5%	0.03	120
Entire Basin	500	135	26,900	0.27	33%	0.09	9,345
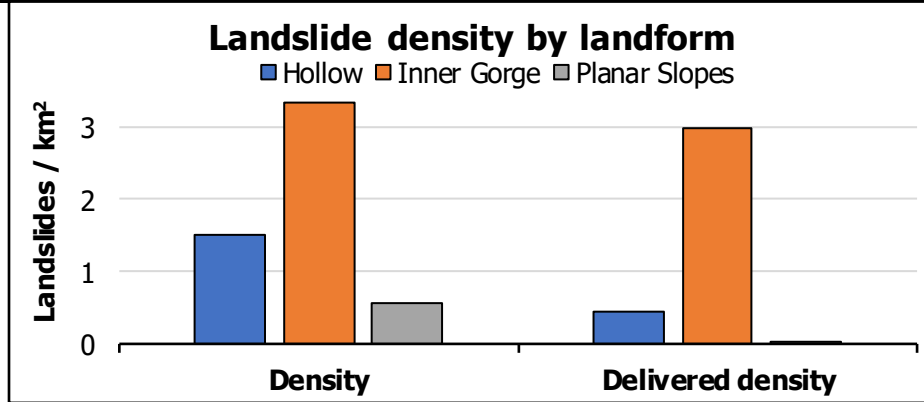
All three landforms produced landslides but at different densities. Based on density in this hypothetical example, inner gorges are the most prone to failure, and steep, planar slopes are the least failure-prone. When we look at the spatial density of landslides that deliver material to stream channels, these relationships remain unchanged, but the relative magnitudes change. Now a look at proportions.

Table 3. Landslide Proportion by Landform

Landform	Proportion of all landslides	Proportion of delivering landslides	Proportion of all volume	Proportion of delivered volume
Hollow	11%	11%	84%	80%
Inner Gorge	30%	80%	7%	19%
Planar slopes	59%	9%	9%	1%
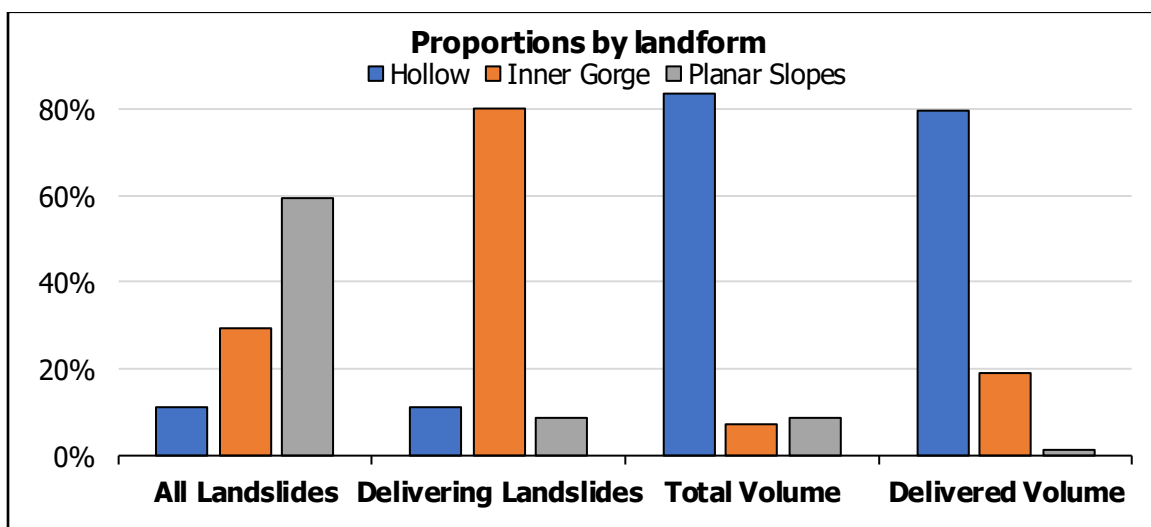
These provide a different perspective. Planar slopes, the most stable of the three evaluated landforms, produce the most landslides overall but, not surprisingly, the inner gorges produce the most landslides that put material directly into stream channels. This story changes again with a look at volume. Hollows and the long-runout debris flows they produce mobilize and deliver the largest volume of material, while planar slopes deliver the least, despite producing the most landslides. Although these numbers are made up to illustrate a point, they are consistent with actual landslide inventories. The observed ratios would vary with location and the timing of the study. The point is that the evaluation of adverse impacts might evolve depending on how landslides are measured and counted. Each perspective provides potentially useful information for determining which parts of the landscape threaten public resources or public safety.

The proportions and associated landslide densities are typically determined by overlaying inventories of mapped landslides on measured, mapped, and modeled terrain elements and counting or measuring the landslide numbers, areas, and volumes associated with each (Guzzetti et al., 2012; Wieczorek, 1984). With many observed landslides, statistics can be calculated by associating landslide potential and proportions as functions of terrain-element attributes. This empirical approach has been tested and refined with hundreds of studies across the globe (Reichenbach et al., 2018). This range of experience with empirically evaluating landslide susceptibility also informs us about this study's challenges. The following sections list potential issues, with strategies for dealing with them.

This overlay approach, however, has been primarily applied for assessing susceptibility to landslide initiation. Landslide runout models tend to focus on estimating runout extent for single events. There are fewer examples in the literature for building statistical models of runout with which to estimate

runout probability for any arbitrary initiation site. However, landslide volume and runout extent can be more difficult to discern accurately, so many photo-based inventories do not include volume or runout, limiting the runout data available for building statistical models. The following sections describe the options available.

1.5 Landforms in the Context of Terrain Elements

Current RIL definitions are based on remote and field-based observations of the landform types and topographic attributes where landslides occur (e.g., Benda et al., 1998). These definitions have been vetted through Watershed Analyses (Washington Forest Practices Board, 2011) and the Landslide Hazard Zonation Project (UPSAG, 2006). Division of a landscape into distinct landforms has proven a useful framework for identifying areas where different geomorphic processes are active. The strategies and methods used to delineate landforms vary with the questions posed (Otto and Smith, 2013). Some landforms are the direct products of the mobilization of landslide debris, so landform characteristics can indicate where landslides may most often occur. Focusing on processes of sediment storage, transport, and deposition can therefore aid in identifying areas susceptible to landslide impacts (Marden et al., 2015; Theler et al., 2010). Hence, we use landforms – RILs – to identify potentially unstable terrain. Similarly, British Columbia also developed a mapping protocol to identify landslide-prone terrain (Schwab and Geertsema, 2010).

With the advent of high-resolution lidar DEMs, expanded landslide inventories, newly available precipitation and land-cover data, and new methods for integrating all that information into quantitative measures of landslide potential, we now have an opportunity to look anew at landslide susceptibility. First, Lidar DEMs can obtain accurate spatially continuous measures of surface gradient and curvature measured over length scales consistent with the size of landslide initiation zones and runout tracks. We seek to parameterize and quantify those factors that affect the balance of forces acting to move soil downslope to those acting to hold it in place. Primary among those factors are soil depth and depth of soil saturation. We have no way of measuring these quantities directly, but we can measure hillsurface gradient and curvature, both of which influence soil depth and saturation. Hence, these topographic elements, derived from lidar DEMs, are important terrain elements for inclusion in susceptibility analyses. Antecedent soil moisture and precipitation intensity and duration during storms are important landslide-triggering factors; spatially contiguous estimates of daily precipitation amounts now allow us to look for relationships between the sequence of precipitation events and landslide density associated with specific storms. Finally, the ability to derive spatially and temporally continuous measures of forest stand characteristics using lidar data, and satellite imagery provides the ability to assess the influence of forest cover on landslide density[18]. These information sources were not available for earlier analyses of landslide susceptibility done with Washington Watershed Analyses and the Landslide Hazard Zonation Project.

Initially, observed landslide locations guided the identification of landforms where landslides occur. Those observations were translated into narrative descriptions of specific landforms and hillsurface gradients associated with landslides, forming the bases for the regulated landforms under WA forest

---

[18] Other physical attributes also affect the potential for landslide initiation, such as soil geotechnical properties, soil pipes, and bedrock fracture density. These attributes cannot be measured over the regional extents for which these analyses will be conducted and spatial variability in these attributes will produce uncertainty in measures of landslide susceptibility.

practices rules: the RILs. Observed landslide locations can be used to build statistical models relating terrain-element attributes to landslide density. We can then use those relationships to identify locations in the landscape – the landforms – where landslides initiate and where debris from upslope landslides travels and deposits. Moreover, we can rank these locations in terms of the proportion of total landslide numbers, area, and volume and of the proportion of total landslide runout-track length. This capability will allow evaluation of the efficacy of current RIL definitions. Once functional relationships between terrain element attributes and landslide density are determined, histograms like those shown in Section 1.4 can be constructed anywhere RILs or other landforms are delineated. If these analyses indicate that current RIL definitions are missing some important proportion of landslides, this capability will also guide modifications to RIL definitions in Project 4.

Previous efforts relied on field observations coupled with analyses of aerial photos, paper maps, and DEMs derived from those maps. Translation of observed landslides, landform boundaries, and measured topographic attributes across these data sources introduced considerable uncertainty. In contrast, this study will obtain landslide locations, landform boundaries, and topographic attributes from the same data source, lidar DEMs, with an order-of-magnitude greater spatial precision. Additionally, we now have digital data for storm intensity and land cover to incorporate into our analyses. These new data and analysis methods will improve our ability to estimate the physical controls on landslide initiation and runout at landscape scales in ways that were not practical with prior remote sensing and field methods, so we expect that they will improve our ability to assess landslide susceptibility more accurately. Maps created by models developed with this project will help guide field identification of unstable slopes and the layout of associated buffers. However, these new data and analysis methods cannot identify or resolve all process controls on landslide initiation and runout. Field observations will still be essential for verifying attribute values measured from digital data (e.g., landslide-scar boundaries) and looking for indicators of instability not seen in the data or recognized with the models. Likewise, we will look for ways to translate model outputs to field-observable and measurable attributes to help on-the-ground delineation of unstable-slope buffers. This translation will depend on our interpretations of model results, that is, on which terrain-element attributes show significant relationships with landslide density.

2 Landslide Inventories

The landslide inventory is our measure of reality. It is used to calibrate landslide susceptibility models and test and validate model results. However, imprecision, error, and bias in the inventory translate to imprecision, error, and bias in resulting estimates of landslide proportions and susceptibility (Steger et al., 2016; Steger et al., 2017). It is important, therefore, to acknowledge limits on precision and identify sources of error and bias.

2.1 Precision and Accuracy of Mapped Landslide Locations.

Landslide mapping is done within a geographic information system (GIS). All mapping and subsequent computer analyses are based on digitized landslide locations. Our initial focus is on landslide initiation. The location where a landslide initiates may be digitized as a point or a polygon. For example, with a field-based inventory, the initiation point may be based on a GPS reading taken from a location the surveyor interprets as within the likely initiation zone. With an aerial-photo-based inventory, the landslide scar may be digitized as a polygon. Based on the photo analyst's interpretation, a point within the polygon may also be flagged as the initiation point.

The precision of current hand-held GPS receivers in rough terrain with forest cover is about 10 meters[19] (although much higher precision options are becoming available [20]). Mapping based on GPS coordinates from older receivers will have lower precision. For mapping from aerial photographs, 10 meters on the ground translates to less than a millimeter in a 1:12,000-scale photo. Ten meters is probably the best precision achievable for field and photo-based inventories. Actual precision and accuracy will be less (more than ten meters) because of other errors in mapping and interpreting what is observed on the ground and in photographs.

For this project, we want to associate landslide locations with topographic attributes derived directly from the digital elevation data. These attributes include spatial variations in gradient, curvature, and contributing area. With lidar DEMs, these attributes can be measured over length scales of several meters. Therefore, we need to maintain this level of resolution in our analyses, which requires a degree of precision considerably finer than the ten-meter-or-more level typically available with existing landslide inventories.

One approach to deal with this lack of precision is to snap landslide initiation points to the inferred lowest-stability point, based on the topographic attributes derived from the higher-resolution DEM, within a specified radius (e.g., 30 meters, Miller and Burnett, 2007) or, for landslides mapped as polygons, to choose the lowest-stability-point within the polygon (Dietrich, Bellugi and de Asua, 2001). Unfortunately, these methods bias the identified landslide initiation sites to those locations the analyst considers less stable. The influence of this type of bias can be estimated by looking at the sensitivity of the results to point placement.

A better approach is to identify landslide locations using the same elevation data from which topographic attributes are derived. As described later in this document, this strategy is now feasible using sequences of overlapping lidar surveys. Landslides between surveys are evident as changes in ground-surface elevation, with elevation losses in zones of failure and scour and elevation gains in deposition zones (Coe et al., 2021; Scheip and Wegmann, 2022). The pre-failure topography can be measured from the earlier lidar survey, and landslide initiation and runout locations can be precisely associated with the topographic attributes derived from that same data.

2.2 Sampling Bias

Landslide inventories also suffer from intrinsic biases resulting from incomplete sampling of the population of landslide events. Associations found between environmental factors and landslide location, area, and volume will depend explicitly on the set of landslide scars examined. If this set does not provide a complete sample of the locations, areas, and volumes for landslides that can occur in some regions, our identification of the factors associated with these landslides will also be incomplete and potentially biased.

Two factors hinder the collection of a representative sample of landslide events:

1. Inability to see some subset of landslides because of their size or location, and

---

[19] https://www.garmin.com/en-US/AboutGPS/, https://www.gps.gov/systems/gps/performance/accuracy/
[20] https://www.gpsworld.com/commentary-high-precision-positioning-is-going-mainstream/

2. Absence of some subset of landslides to observe because an event capable of triggering those landslides has not occurred within the period over which the resulting landslide scars would be visible.

The first of these is referred to as detection bias. In mapping from aerial photographs, there is a size limit below which landslide scars are not visible on the photos. Landslide sizes exhibit a distinctive frequency distribution (Stark and Hovius, 2001), with many small and fewer large landslides. Consequently, most landslides may not be visible in aerial photographs (Brardinoni, Slaymaker and Hassan, 2003; Miller and Burnett, 2007; Turner et al., 2010), and resulting inventories are biased toward larger landslides. Estimates of landslide density and proportions based on those inventories are also biased toward larger landslides. Any variation in the distribution of landslide sizes across landform types will then create errors in comparisons like those illustrated in Section 1.1. Because landslides exhibit distinct size distributions (Malamud et al., 2004), the degree of detection bias can be estimated by extrapolating the observed distribution to smaller sizes (Guzzetti et al., 2002; Miller and Burnett, 2007).

The size limit of landslide scars detectable in aerial photos increases for landslide scars under forest canopy; hence, the degree of bias varies with land cover. This observation is a crucial confounding factor when comparing landslide densities across different forest ages, as when assessing the impact of timber harvest on landslide susceptibility (Pyles and Froehlich, 1987).

Field-based inventories can potentially map all landslide scars detectible on the ground and overcome the size bias inherent in photo-based inventories (Brardinoni, Slaymaker and Hassan, 2003). However, because of the large labor and time investments required for field surveys, they cover limited spatial extents and may include only a subset of more readily encountered landslides, focusing, for example, on those that runout to a stream channel (Robison et al., 1999; Stewart et al., 2013).

The second factor potentially biasing the sample of landslides in an inventory arises from the limited time over which a shallow landslide scar is detectable. Field evidence of small landslides may persist for only a decade or two (Reid and Dunne, 1996). This duration is not an issue for field inventories that focus on landslides associated with specific recent extreme storm events (Robison et al., 1999; Stewart et al., 2013), but those inventories are then sampling only the subset of landslides associated with those events. In aerial photos, landslide scars are primarily discernable because of the disruption they produce in the vegetation. As the scar revegetates, it becomes more difficult to see and interpret. Thus, a photo-based inventory provides the subset (biased by size) of landslides that have occurred over the time the scars are visible in the photos, estimated by Brardinoni, Slaymaker and Hassan (2003) as a maximum of about 30 years. A series of photos may span several decades, providing a longer-term record of landslide events. It is still constrained, however, to the set of landslide-triggering events that occurred over that period.

2.3 Forest Cover

Trees not only limit the resolution of landslide scars visible in air photos, but many studies have also found systematic differences in landslide density associated with differences in forest cover (e.g., Brardinoni, Hassan and Slaymaker, 2002; Goetz, Guthrie and Brenning, 2015; Miller and Burnett, 2007; Montgomery et al., 2000; Turner et al., 2010). Increased landslide density following changes in forest cover from wildfire or timber harvest is primarily associated with the loss of effective soil cohesion provided by the tree roots, which die and decay away after the loss of the trees (Schmidt et al., 2001).

As trees grow back, root strength can recover, so there is a decade or more over which the potential for soil failure increases (Imaizumi, Sidle and Kamei, 2008; Sidle, 1991). Tree canopy may also modulate spikes in soil pore pressures associated with intense rainfall (Keim and Skaugset, 2003), although Dhakal and Sullivan (2014) found no such effect.

Variability in forest cover, therefore, presents a confounding factor for comparisons of landform density across landform types. Therefore, the spatial distribution of forest ages or stand types must be determined for the study area for the time of each landslide in an inventory (Imaizumi, Sidle and Kamei, 2008; Robison et al., 1999).

2.4 Substrates: Soil, Lithology, Geology

Properties of the soil – its origin (colluvial or residual), composition, texture, depth – and underlying bedrock – lithology, bedding, fracture density, and orientation – profoundly influence the hydrologic and geotechnical properties that determine landslide potential. Over the spatial extent of entire landforms, these properties cannot be measured directly but manifest in topography characteristics: e.g., the relief, the frequency distribution of hillsurface gradient, and variations of the gradient with aspect. To some degree, spatial variability in soil and substrate properties on landslide potential can be inferred from spatial variability in topography attributes. Mapping of soil and rock types and geologic structure can aid in delineating these topographic zones; hence such maps are typically included as explanatory variables in empirical landslide susceptibility analyses (e.g., Amato et al., 2019).

Variations in soil and substrate properties over shorter length scales, e.g., variations in soil depth across a bedrock hollow or a meter-thick fracture zone in the bedrock, also influence landslide locations. This smaller-scale variability cannot be resolved from topographic analysis or soil and geologic mapping. This scale of variability, therefore, introduces an unavoidable degree of uncertainty in comparisons of landslide susceptibility across terrain elements and landforms.

2.5 Storm Characteristics

As described previously, the Washington Administrative Code identifies certain landforms as potentially unstable slopes[21]. The definitions for these RILs were based on extensive photo-based landslide inventories and field reconnaissance as part of the Watershed Analysis (Washington Forest Practices Board, 2011) and Landslide Hazard Zonation projects (UPSAG, 2006). These projects used air-photo records that typically spanned 40 years or more. It was assumed that this provided a representative sample of landslide types and locations from which the range of potentially unstable landforms could be identified. In December 2007, southwest Washington experienced an extreme storm that triggered many landslides (Stewart et al., 2013; Turner et al., 2010). Many landslides occurred outside areas interpreted as rule-identified landforms (Murphy, Sarikhan and Slaughter, 2013; Stewart et al., 2013). Such events had not been recorded in the air-photo record or experienced over the lifetimes of geologists working in the area. The inventories and experience on which the RIL definitions are based may have been biased by the absence of such events. The importance of this bias in terms of the proportion of landslide numbers, area, and volume is currently unquantified. We do not know if the nontypical landslide locations associated with such infrequent storms are major players or a minor blip

---

[21] WAC 222-16-050

in the landslide regime and sediment budget of Washington watersheds. This lack of knowledge is a major motivating factor for this study.

The number of landslide scars available for inventorying is also dependent on the characteristics of the storms preceding data collection. For example, in an analysis of the 2007 storm, Turner et al. (2010) found a highly nonlinear relationship between rainfall intensity and landslide density (number per unit area). This observation suggests extreme storms have a disproportionate influence in landslide sediment budgets (Lee, 2017). This result is also an issue to be explored in this study.

Antecedent conditions and storm duration may also play a role in determining where in the landscape landslides occur (Jones et al., 2021). Wieczorek (1987), for example, found that landslides in deeper soils (1-3m) located in concave areas at mid to low hillslope positions were associated with longer-duration storms. In contrast, failure of shallower soils on more planar slopes occurred on mid to upper hillslope locations in association with higher-intensity, shorter-duration storms. Snow melt, particularly associated with rain-on-snow events, may also be an important driver of landslide initiation (Guthrie et al., 2010b).

Establishing these storm and landslide relationships requires that statistics on landslide location and density be determined as functions of storm characteristics and return intervals (Reid and Page, 2002), requiring the association of each landslide event with a specific storm.

2.6 Landslide History

A recent multi-temporal landslide study in Italy showed that earlier landslides impact a substantial fraction of landslides, and those landslides following earlier landslides may differ from those that do not (Temme et al., 2020). The effects of past landslides on current landslide susceptibility span a range of spatial and temporal scales.

Deep-seated landslide alterations of surface topography and sub-surface properties can locally increase the potential for shallow landslide initiation and alter runout potential. For example, steepened surface gradients at deep-seated landslides' toes and head scarps may increase the potential for shallow landslide initiation in those zones. Hence the inclusion of steep deep-seated landslide toes in the set of RILs, and the topographic bench typically formed by the body of a deep-seated landslide below the head scarp can limit the runout extent of landslides initiated on the head scarp. These observations suggest that local effects on topography, subsurface hydrology, and geotechnical properties, either caused by or associated with deep-seated landslides, may also alter shallow landslide susceptibility under certain storm conditions. Furthermore, deep-seated landslides' surface and sub-surface effects may persist for millennia.

Shallow landslide events also locally alter future landslide susceptibility. Loss of the soil layer in the zone of initial failure and along scoured zones of debris-flow runout tracks acts to reduce the potential for landslide initiation where the soil layer is removed (D'Odorico and Fagherazzi, 2003; Dunne, 1991) and to increase the potential for landslide initiation along the cut banks and scarps created by the event (Murphy, Sarikhan and Slaughter, 2013). These effects persist until the soil layer is re-established, which may take decades to several centuries (May and Gresswell, 2003). Likewise, loss of roughness elements, such as large logs and boulders, along a debris-flow track may increase the potential runout length of future debris-flow events traversing that path. The loss of forest cover caused by landsliding alters local hydrology with potential impacts to downslope landslide potential (Mirus, Smith and Baum, 2017). A

time series of landslide events could be used to examine these influences on landslide susceptibility, but data for constructing such time series extend over only a small portion of the time span these effects persist. Hence, the unknown history of shallow landslide events creates a source of unresolvable variability in estimates of landslide susceptibility.

2.7 Landslide Area and Volume

Aerial-photo-based landslide inventories with landslide scars digitized as polygons can be used to measure landslide surface area (e.g., Malamud et al., 2004). These areas can be multiplied by average failure depths obtained from field measurements to obtain estimates of landslide volume (e.g., Brardinoni, Slaymaker and Hassan, 2003; Martin et al., 2002). Landslide volume also exhibits a power-law relationship to landslide area (e.g., Larsen, Montgomery and Korup, 2010), with which area measurements can infer volume. Channelized debris flows exhibit a power-law relationship between scoured volume and runout length (May, 2002). These relationships can be used to estimate landslide volume from remotely mapped landslides but with relatively low confidence. Field measurements of landslide geometry (length, width, and depth) provide volume estimates with greater confidence (e.g., Miskovic and Powell, 2009) but for smaller sample sets than obtainable from air-photo mapping.

2.8 Landslide Runout

Runout extent is observed to vary with multiple factors: mobilized volume (Griswold and Iverson, 2008), gradient and confinement along the runout path (Guthrie et al., 2010a), changes in the flow direction (Benda and Cundy, 1990), and the abundance and size of standing trees and down wood along the runout path (Booth et al., 2020). These relationships are based primarily on field observations. With landslide inventories that include sufficient such observations, statistical models for predicting runout extent can be developed (Fannin and Wise, 2001). By relating field-measured values to map-based values or by direct mapping from aerial photos onto base maps, map-based (or GIS-based) relationships can also be developed for predicting runout extent (Benda and Cundy, 1990; Berti and Simoni, 2014; Griswold and Iverson, 2008; Guthrie et al., 2010a; Horton et al., 2013; Miller and Burnett, 2008; Reid, Coe and Brien, 2016).

Statistical models for runout are, as with statistical models for landslide initiation, constrained by the completeness, precision, and accuracy of the landslide inventory data from which they are built. With high-resolution lidar DEMs, topographic attributes along observed runout paths can be well-constrained, leaving two primary sources of uncertainty. The first is the determination of the full extent of runout. For photo-based inventories, the downslope extent of deposition may be hidden under the forest canopy. For both photo- and field-based inventories, the full extent of runout may be indeterminate because all or some portion of the deposit has been removed by subsequent fluvial erosion. The second primary source of uncertainty is in estimates of mobilized volume. On the ground, this is a challenging and time-consuming measurement to make, with variability in measured values between surveyors (Miskovic and Powell, 2009). For photo-based inventories, volume estimates are based on averages of field-based relationships between area and volume. Landslides that evolve into channelized debris flows can also entrain material along the runout path, thus growing in volume downslope. The tendency for a debris flow to scour or deposit material varies with the gradient and degree of topographic confinement (Benda and Cundy, 1990; Fannin and Wise, 2001). Through zones of scour, the volume entrained can be estimated from the average volume of colluvium stored in steep, low-order channels. This estimate requires measuring that average volume through field surveys (Benda and Cundy, 1990; May, 2002) or

by comparing high-resolution DEMs obtained before and after the debris flow occurred (Reid, Coe and Brien, 2016; Scheip and Wegmann, 2022). It also requires mapping zones of scour, transition (no net scour or deposition), and deposition to determine the gradient and degree of confinement associated with each zone (Miller and Burnett, 2008; Scheip and Wegmann, 2022). These zones are not easily distinguished from aerial photography, so inventories that include such information are typically field-based (Robison et al., 1999).

2.9 Available Inventories

The Washington Department of Natural Resources (DNR) has compiled landslide inventories across the state[22] (Figure 1). These are primarily from aerial-photo-based mapping and include digitized landslide polygons with limited locational precision and variable accuracy (Slaughter, 2015). The "post-mortem" project (Stewart et al., 2013) has field-surveyed landslide initiation sites digitized as points with field-estimated length, width, and mean and maximum depths. This inventory provides a census of landslides with runouts to stream channels and those initiating at forest roads within the study areas. Landslide locations were recorded as points using handheld GPS receivers. The precision of the points (in addition to the approximately 10-m precision of the GPS units) as a reference to the initiation location primarily depended on the interpretation of the field analyst. This study focused on landslide densities within different harvest treatments; a high precision level for landslide location was not required.

These inventories can potentially serve as a valuable resource for this project. However, they have been collected by numerous different analysts for a variety of purposes using a variety of protocols and data resources. In addition, they exhibit variable error levels (e.g., features visible on aerial photographs incorrectly interpreted as landslides) and precision (Slaughter, 2015). Therefore, their use will require considerable vetting and involve uncertainty in landslide location on the order of tens of meters.

The DNR landslide hazard mapping program was reactivated in 2016 to produce accurate, precise, and consistent inventories. Current protocols for mapping landslides (Slaughter et al., 2017) focus on deep-seated landslide features observed in shaded-relief images created from high-resolution lidar-derived DEMs. Because of their typically smaller size, shallow-rapid landslide features are more difficult to discern in a lidar-shaded-relief image. Likewise, shallow-rapid landslides often occur under a forest canopy, where the low density of ground returns in the lidar point cloud limits the resolution of lidar-derived topography. Hence, shallow-rapid landslides cannot be reliably mapped based on lidar-shaded relief imagery interpretation. The published inventories[23] focus primarily on larger deep-seated landslide features mappable from lidar-shaded relief imagery but also include point locations for small and shallow landslides observed in aerial imagery and field observations. These point map features provide an additional source of inventory data. Likewise, because of the potential association of deep-seated landslides with shallow-landslide susceptibility, the deep-seated landslide in these inventories should be included as an input variable in the analyses for this project.

---

[22] https://data-wadnr.opendata.arcgis.com/documents/landslide-compilation-pre-2017-analysis/about
[23] https://fortress.wa.gov/dnr/geologydata/publications/data_download/ger_portal_landslide_inventory.zip
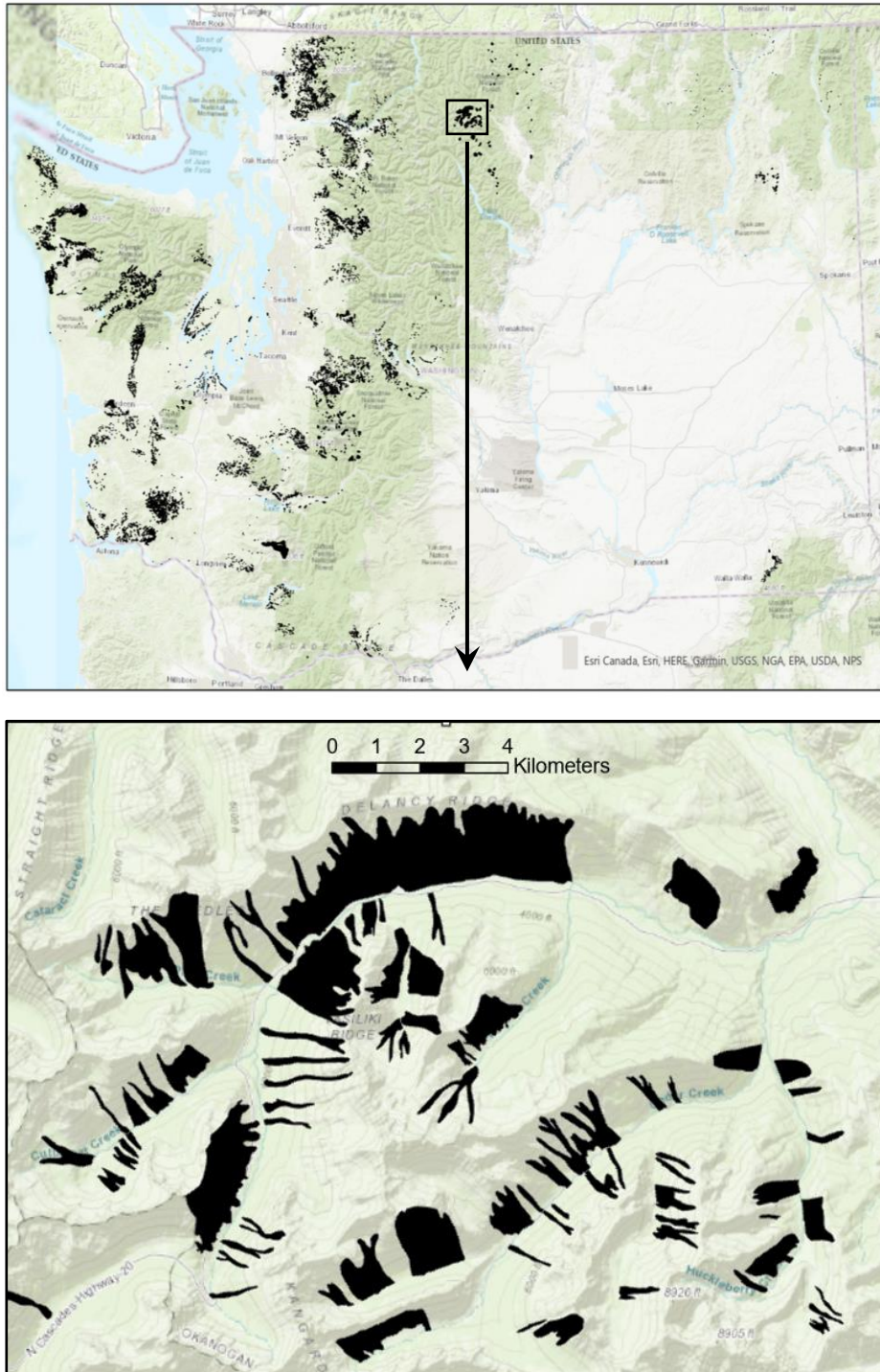
Figure 1. DNR Landslide Compilation. Black polygons show debris flows, debris slides, and debris avalanches recorded in the compilation. The upper panel shows the distribution of available inventory polygons across the state. The lower panel shows that, in some cases, landslide zones were mapped, rather than individual landslide scars. The black polygons are digitized as single entities, not individual landslides.

3 New Opportunities

We have the option of collecting new inventories for this project. This option would enable using newly available data and methods to address many of the abovementioned issues. Washington DNR has been actively collecting lidar datasets across the state[24] (Figure 2), with extensive survey areas planned for 2022-23. Repeat lidar data collections provide a new resource for landslide mapping. DEM or point-cloud differencing, which measures elevation changes occurring during the period between two lidar acquisitions (Figure 3), can be used to map individual landslides (Bull et al., 2010; Burns et al., 2010; Cavalli et al., 2017; Scheidl, Rickenmann and Chiari, 2008). Lidar differencing can potentially provide precision near the 1-m resolution of the lidar DEMs and consistent measures of landslide area and volume (Bernard, Lague and Steer, 2021).

The date of landslide occurrence is constrained by the dates of the lidar collections and can be further constrained using aerial photography collected during that interval (Fernández et al., 2021). In addition, landslide dates are required for correlation with forest-stand type and characteristics of the landslide-triggering storms.

Land cover can be inferred using the supervised classification of satellite imagery. Several gridded land cover data sets based on analysis of Landsat imagery are available for Washington (e.g., Ohmann and Gregory, 2002). In conjunction with multiband NAIP and satellite imagery, Lidar data can be used to map
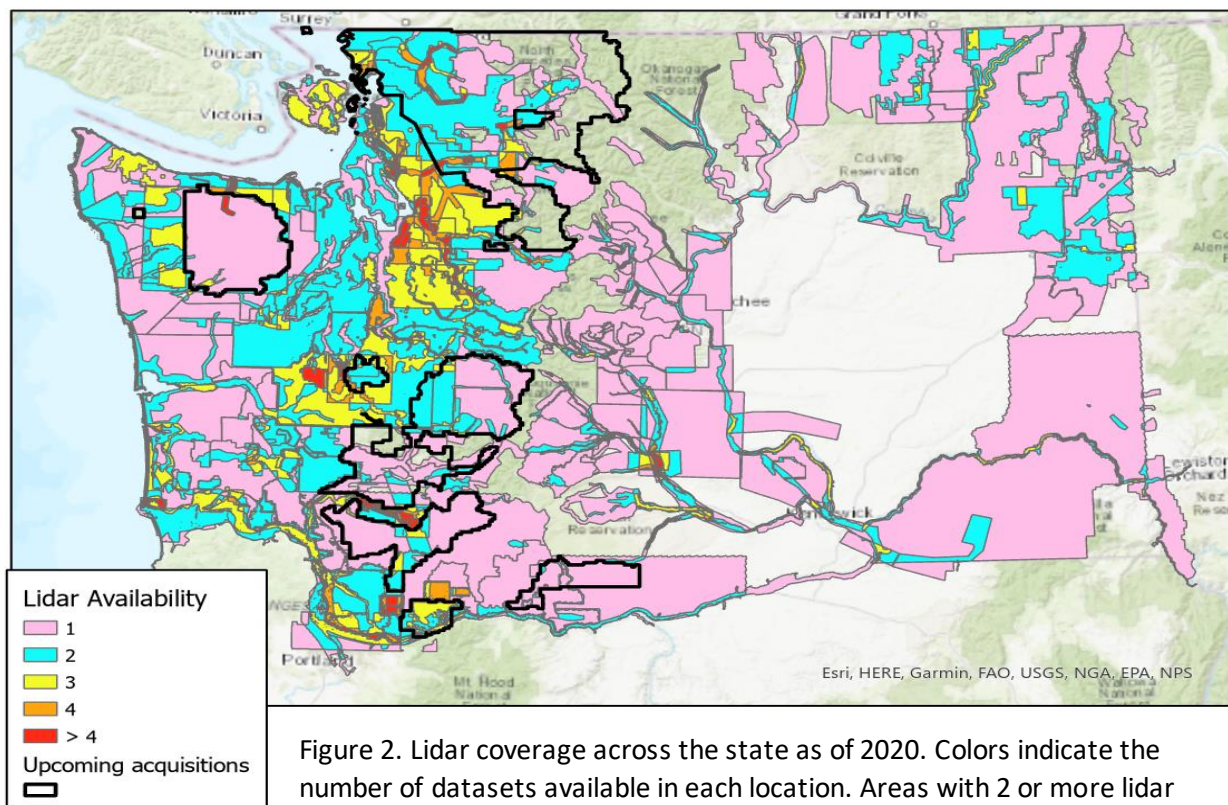

Figure 2. Lidar coverage across the state as of 2020. Colors indicate the number of datasets available in each location. Areas with 2 or more lidar datasets can be used for lidar differencing. Upcoming acquisitions scheduled for 2022-2023 target areas with recent landslide activity and will expand the area available for lidar differencing.

---

[24] https://www.dnr.wa.gov/lidar

forest stand characteristics to differentiate different stand types and ages with a high spatial resolution (e.g., Matasci et al., 2018; Su et al., 2015). Washington DNR has used canopy height models obtained from lidar point clouds and structure-from-motion analysis of stereo imagery, referred to as PhoDAR, with height-to-age relationships to map forest stand age across the state. With such an analysis, stand type and age can be used as a variable in statistical analyses of landslide density. Such an analysis would also aid in evaluating the influence of forest harvest on landslide susceptibility.

Various gridded precipitation data might be used to characterize antecedent moisture and storm characteristics associated with landslide occurrences (Stanley et al., 2020). There are several sources for these data[25] (Prat and Nelson, 2015): rain gauge records (Buban, Lee and Baker, 2020), terrestrial weather radar (Sokol et al., 2021), and satellite data (Kidd and Levizzani, 2022). Data are available online with temporal resolutions from three hours to daily, spatial resolutions from one to ten kilometers, and temporal extent from 1980 to the present. Each data product has different strengths and weaknesses and is, therefore, best applied to specific types of analyses (Hu et al., 2019; Li et al., 2020; Molter, Collins and Risser, 2021; Ombadi et al., 2021; Prat and Nelson, 2015; Sun et al., 2018; Thornton et al., 2021; Timmermans et al., 2019). Our interest is primarily in extreme events because it is these that trigger the most landslides, for which several data products may be best suited (Molter, Collins and Risser, 2021; Pierce et al., 2021; Prat and Nelson, 2015; Rajulapati et al., 2020; Thornton et al., 2021; Timmermans et al., 2019; Werner et al., 2019). For any particular day, all of the different data products indicate slightly different precipitation amounts. There may not be a single ideal dataset for this project; the available options will need to be evaluated.

Repeat lidar point clouds, NAIP and Landsat multiband imagery, and gridded land cover and precipitation data provide a suite of publicly available data that this project can capitalize on to build landslide inventories that potentially overcome many of the constraints discussed above. In addition, in areas where multiple lidar datasets overlap existing landslide inventories, the performance of the new and old inventories can be compared.

---

[25] https://climatedataguide.ucar.edu/climate-data/precipitation-data-sets-overview-comparison-table
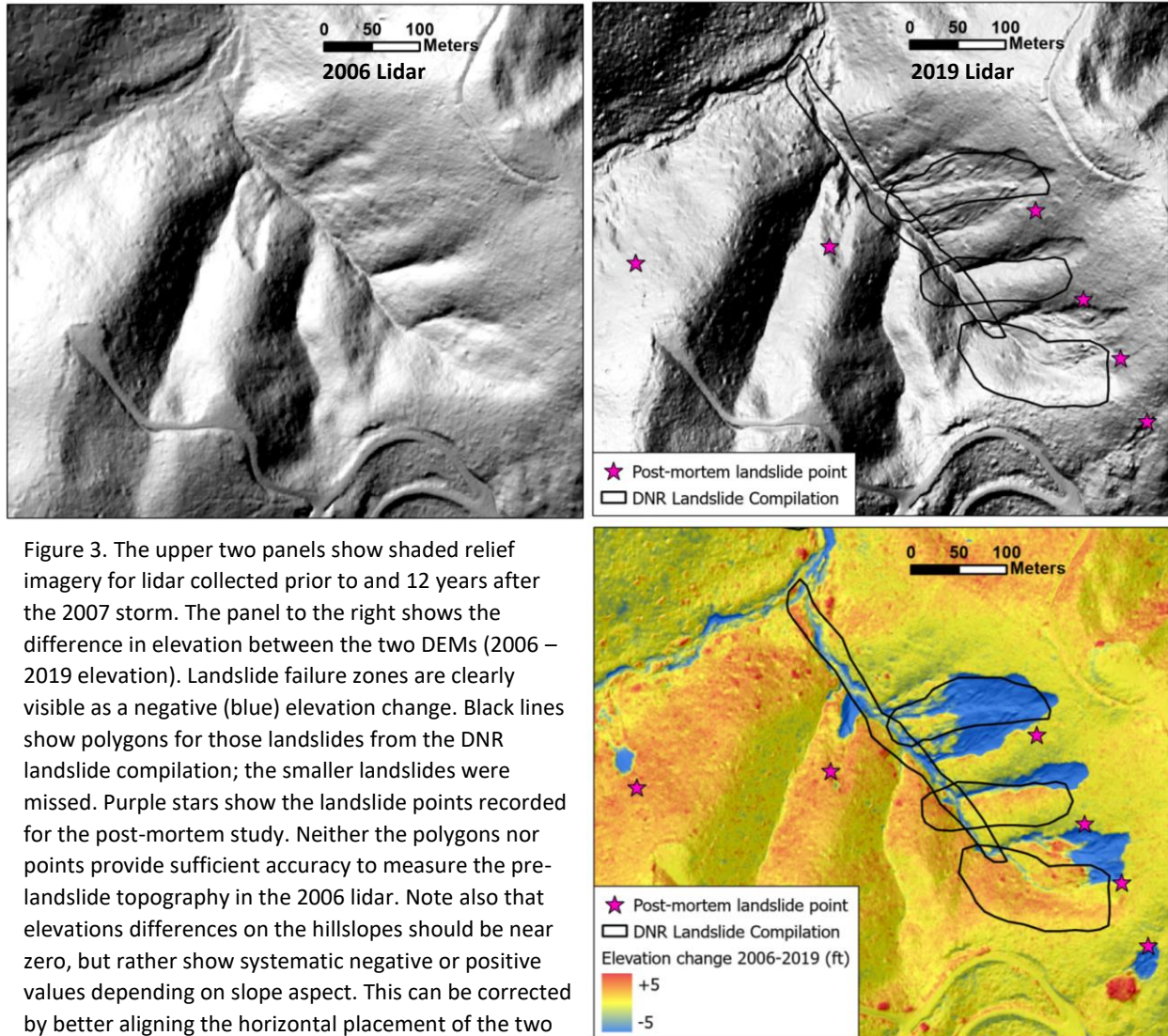
Figure 3. The upper two panels show shaded relief imagery for lidar collected prior to and 12 years after the 2007 storm. The panel to the right shows the difference in elevation between the two DEMs (2006 – 2019 elevation). Landslide failure zones are clearly visible as a negative (blue) elevation change. Black lines show polygons for those landslides from the DNR landslide compilation; the smaller landslides were missed. Purple stars show the landslide points recorded for the post-mortem study. Neither the polygons nor points provide sufficient accuracy to measure the pre-landslide topography in the 2006 lidar. Note also that elevations differences on the hillslopes should be near zero, but rather show systematic negative or positive values depending on slope aspect. This can be corrected by better aligning the horizontal placement of the two lidar datasets.

3.1 Lidar Differencing

The strategy is to compare elevations from point-cloud ground returns or the derived DEMs for lidar datasets collected at different times. Landslides occurring in the period spanning the datasets will produce a negative change in elevation in the failure zone and along scoured portions of the runout path and an increase in elevation through the area of deposition. Other processes can also create changes in elevation, such as road construction and maintenance or fluvial erosion and deposition along river courses. Other factors, such as surface gradient and position in the landscape, must also be considered in using elevation differences to delineate landslide features.

There are several issues to address when comparing two lidar datasets to infer elevation changes. On steep slopes, a small difference in the horizontal registration (Cheng et al., 2018; the latitude-longitude reference for each point) between data sets can produce a substantial difference in the reported elevation. This issue is evident in Figure 3, where systematic differences in elevation with slope aspect occur between the 2006 and 2019 datasets. Differences in flight lines relative to slope aspect can also

result in slight local differences in point registration between datasets. Various co-registration methods can correct misalignment (Cucchiaro et al., 2020). Open-source software available at https://www.cloudcompare.org/ provides resources for this task.

Older lidar datasets tend to have lower point densities and correspondingly lower ground-return densities. Ground returns under a dense canopy may be widely spaced, hindering ground surface resolution and elevation changes between lidar datasets. Ground-return point density and data accuracy (e.g., the number of incorrectly classified ground returns) may therefore differ for the datasets. Measures of elevation change should include estimates of confidence based on ground-return point density and local surface roughness (Bernard, Lague and Steer, 2021). The precision to which elevation changes can be resolved with available lidar data remains to be determined, but one approach would be to eliminate from analysis all values from the DEM of difference raster in the domain of the mean ± 1 standard deviation (e.g., Scheip and Wegmann, 2022). Elevation differences between datasets may be obtained by subtracting one DEM from another, as shown in Figure 3 (note that registration differences were not corrected in the figure). An alternative is to compare ground returns in the point clouds directly. The point positions will not correspond between datasets. Various methods have been devised for interpolating between points to estimate elevation differences (Okyay et al., 2019). The current favorite is the multiscale model-to-model cloud comparison (M3C2) algorithm (Bernard, Lague and Steer, 2021; Lague, Brodu and Leroux, 2013; Winiwarter, Anders and Höfle, 2021). It essentially involves fitting a surface to a local neighborhood of points in each dataset and comparing the surfaces. It also provides confidence measures in the differences based on ground-return point density and local surface roughness. The algorithm can be applied over a regular mesh of points to provide a grid of elevation differences (Bernard, Lague and Steer, 2021). Open-source software for implementing M3C2 is available at https://www.cloudcompare.org/.

The grid of elevation differences obtained with DEM or point-cloud differencing provides a raster image that can be used with image segmentation techniques to delineate landslide scars. Our initial focus is on elevation loss to identify the failure zone. However, segmentation of the elevation-difference raster may employ other rasters, such as lidar ground point density, roads, surface gradient, and landscape position to better delineate landslide scars from elevation losses created by other processes. Techniques must also be explored for differentiating actual landslide scars from noise in the DEM. These may be as simple as thresholds in the surface area and depth of delineated elevation-loss zones, which may vary with the DEMs' ground-return point spacing and local surface roughness.

We are fortunate to have three lidar datasets that overlap portions of the post-mortem study sites, one of which was collected in 2006, predating the 2007 storm. The post-mortem field-based inventory provides a census of landslides that reached stream channels or originated at forest roads within the study sites. The proportion of the post-mortem-study landslides identified with lidar differencing provides a test of this new method. The post-mortem data include landslide size and depth measures, with which we can look for thresholds in the size and depth of landslide scars discernable with lidar differencing, at least with the available point clouds[26]. These datasets (the post-mortem landslide inventory, the Lewis 2006, SWWA Foothills 2017, and Southwest WA OPSW 2019 lidar point clouds) provide the resources to develop and test a workflow for automated landslide mapping using lidar

---

[26] Older lidar datasets have lower ground-return-point density, which may be the limiting factor for the size of landslide scars that can be identified.

differencing. With available precipitation data, landslide densities can then be compared to estimated rainfall totals for the December storm, and these results, parsed by surface gradient and forest-stand class, compared to those found by Stewart et al. (2013) and Turner et al. (2010). These data resources provide the means to develop and test a new method for collecting landslide inventories to provide high locational precision and consistent failure area and volume measures. Once a workflow is established, and if it proves useful and feasible, it can be extended to other areas of the state where overlapping lidar datasets exist (Figure 2).

The datasets above provide an example of landslide occurrences associated with a major storm event. Other available datasets will provide information about landslide responses to different storm events. The gridded precipitation data might allow us to characterize the storms associated with those landslides (Reid, 1998; Reid and Page, 2002), constrained at least to the potential storms between lidar acquisitions.

In areas where repeat lidar data overlap with inventories in the DNR Landslide Compilation, the landslides mapped with lidar differencing can be compared to the set in the compilation based on the landslide date recorded in the compilation. This exercise can provide both a check on the completeness of each inventory, what proportions of mapped polygons have corresponding failure zones delineated in the lidar differencing data, and vice versa, and the confidence to place on landslide sizes recorded by the polygons in the compilation. Depending on the success of the above-listed exercises, it may also be deemed useful during the study to collect additional photo-based (with field checking) inventories.

Lidar data can also be used to measure attributes for landslide runouts. Many failure zones indicated by elevation loss between two lidar acquisitions have corresponding depositional zones downslope indicated by an increase in elevation (Bernard, Lague and Steer, 2021). In some cases, the deposit is distinct and well-defined. In others, the deposit is indistinct, spread out over a considerable portion of the downslope area, or missing altogether. The extent of channel scour may also be discerned for channelized debris flows as a loss of elevation (Reid, Coe and Brien, 2016) (Figure 4). Those sites where the downslope extent of the deposit or the downslope extent of channel scour can be readily distinguished provide useful information for seeking correlations between runout extent, landslide volume, and topographic attributes (e.g., gradient, curvature, and contributing area; Scheip and Wegmann, 2002). Those sites where an intersection with a stream truncated the downslope extent of runout may also provide useful information along the portion of the runout track traversed because we know runout extended beyond those points.
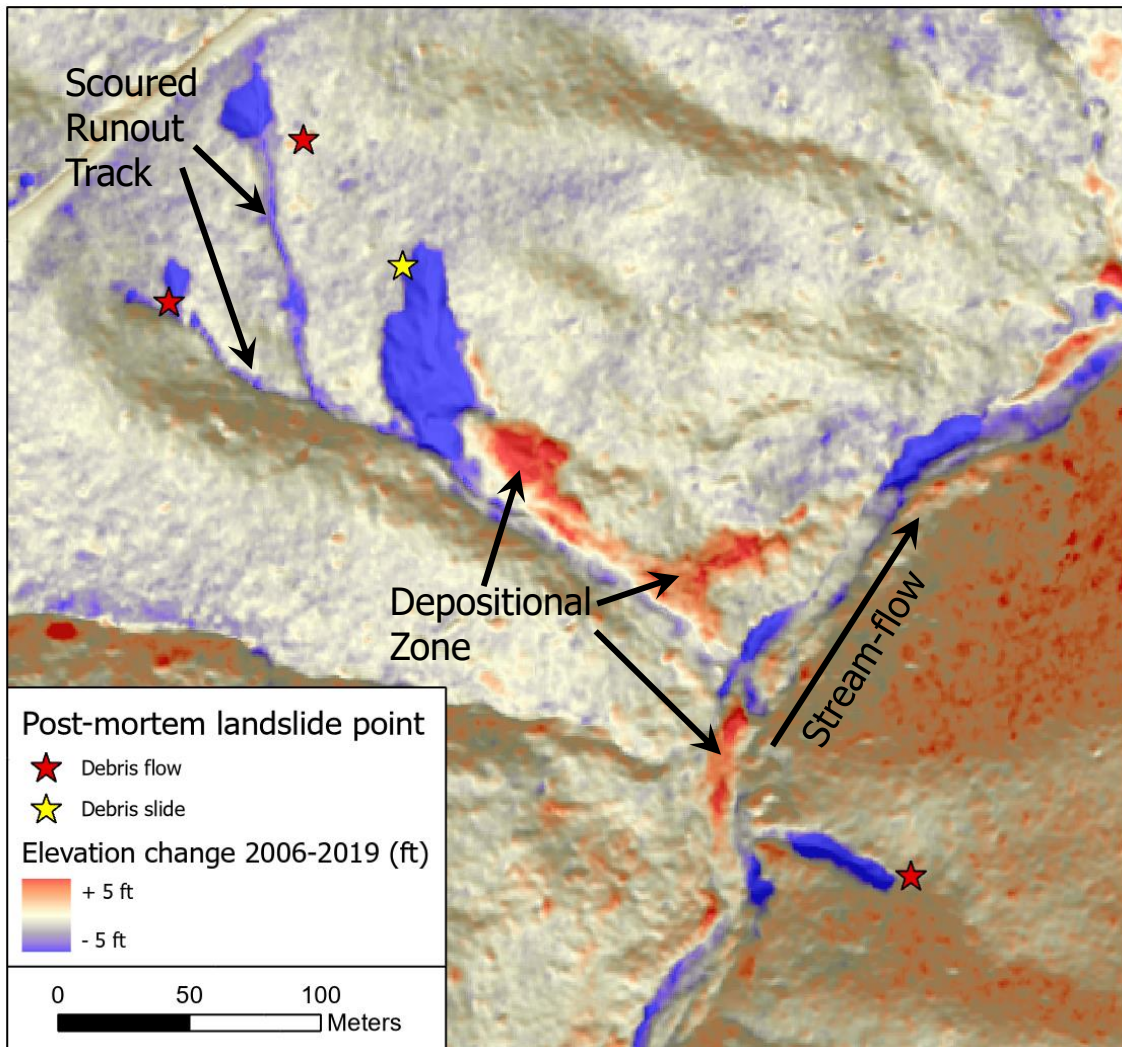
Figure 4. Lidar DEM differencing showing landslide initiation zones, runout tracks, and depositional zones. For the two debris flows in the upper left, scoured runout tracks can be followed to the channel at the base of the hill, with no clear depositional zones. For the debris slide (yellow star), a large depositional zone is evident down slope. For the debris flow in the lower right, the scoured runout track and a zone of deposition along the valley floor are visible.

The discussion has dealt with laser reflections interpreted as ground returns in the lidar point clouds. Vegetation returns may also provide useful information. A landslide can knock down and displace trees through the failure and runout zones. Loss or reduction of tree cover corresponding to zones with elevation changes indicative of landsliding may provide additional evidence for delineating landslide locations with lidar differencing.

4 Analyses

Locations on the landscape will be characterized in terms of terrain-element attributes measured or interpolated to the spatial scale of DEM cells. These proportions are measured over the area examined (the area spanned by overlapping lidar datasets used to create a landslide inventory using lidar differencing) and are defined in four ways:

1. The proportion of all shallow landslide events,
2. the proportion of the total failure surface area,
3. the proportion of the total mobilized volume, and
4. the proportion of the total volume delivered to the channel network or some portion of the network, such as fish-bearing streams.

The landslide inventories collected with lidar differencing will be used to measure these proportions and parse them by terrain element and landform, but only for the areas and periods spanned by the available data. The measured proportions will always depend on the area and period examined. Given the large control that storms have on shallow landslide density, types, and locations, we would like to have landslide inventories span a large range of storm characteristics. One could then look at how the proportions listed above vary with storm events and convolve those proportions with the probability of storm occurrence (return interval) to obtain long-term rates for each landform type (e.g., Reid, 1998; Reid and Page, 2002). This relationship is needed to characterize landslide potential by landform fully, but results are constrained by the periods spanned by sequential lidar acquisitions and by the ability of the available precipitation data to characterize storm events accurately.

These proportions will also vary with other factors – all the terrain elements that might influence landslide initiation and runout. Statistical analyses seek to identify the degree to which each terrain element examined is associated with landslide locations, but the ability to identify relationships will depend on the amount and quality of available data.

4.1 Landslide Initiation

The measure of susceptibility to landslides will be relative density. Statistical models will be used to determine density as a function of a set of terrain elements. These densities will be normalized – divided by the largest density value – to span a range from zero to one. Relative density: it allows a more direct comparison of spatial variability between sites where absolute density differs, e.g., for inventories collected following storms of different magnitudes. As described in the introduction, density can be integrated over an area to provide proportions. The calculated proportions are not affected by having first normalized the density values.

Density will be determined as a function of terrain-element values. For example, an initial set of candidate terrain elements to examine might include surface gradient, mean curvature (Minár, Evans and Jenčo, 2020), forest-stand age, some grouping of underlying geology (e.g., consolidated or unconsolidated in regions with extensive glacial deposits), upslope contributing area associated with the triggering-storm duration, and the daily precipitation associated with the landslide-triggering storm. These values will be calculated, interpolated, or sub-sampled to provide a set of values for each DEM cell. The set of terrain elements found to reproduce observed landslide densities best will likely evolve throughout the study based on the results of statistical analyses aimed at identifying those attributes most strongly associated with landslide locations.

The probability of encountering a landslide in the inventory is then determined as a function of those terrain attribute values, and a raster of probability values is computed. Integration of probability over an area gives the number of expected landslides; dividing that number by the area gives density. With this approach, each landslide provides a training (or testing) data sample point. Regression (e.g., generalized linear and additive models, Wood, 2017), frequency ratio (e.g., Miller and Burnett, 2007), and machine learning techniques (e.g., convoluted neural networks, random forest, Dao et al., 2020) can be used to relate terrain attributes to the probability of encountering a landslide.

The concept of density discussed above concerns the number, area, and mobilized volume of landslides per unit area. This concept can be extended to incorporate landslide runout to describe source areas for landslides in terms of the number of landslides that deliver sediment and debris to public resources per unit area and the volume of material carried by landslides to public resources, as illustrated in the introduction. Both measures apply to the source areas where landslides originate.

4.2 What to Examine: Landslide Initiation

The discussions above highlighted factors known to affect landslide densities. These provide obvious candidates for evaluation: topography, forest cover, storms, substrate, and climate. The details of how these are quantified and over what scales depend on available data and processing tools.

4.2.1. Topography

We are limiting analyses to areas with available high-quality lidar data. Lidar offers a vast range of topography characterization over various spatial scales. Empirical studies have examined a large range of topographic attributes. These include elevation derivatives such as surface gradient, curvature, contributing area, topographic wetness index, surface roughness, and stream power (e.g., Mahalingam, Olsen and O'Banion, 2016). Other topography-related factors include distances to roads, faults, or streams (e.g., Gaidzik and Ramirez-Herrera, 2021). Conceptual models of shallow soil stability guide this study's choice of topographic attributes. For example, surface gradient determines the magnitude of the gravitational force acting to move material downslope; contributing area and curvature constrain the depth of soil (Patton et al., 2018) and soil saturation (e.g., Montgomery and Dietrich, 1994).

In rough terrain, measures of gradient and curvature vary with the length over which they are measured. If you are standing on the mound of a tree-fall pit and measure the gradient over 3 meters, you will likely get a different value than if you measured over 30 meters. Lidar DEMs can potentially resolve tree-fall pits and mounds, but we are not particularly interested in these. Likewise, the noise in a lidar DEM caused by reflections from vegetation interpreted as ground returns may be of the same length scale. Gradient and curvature should be measured over length scales similar to the dimensions of a landslide-initiating failure (Sîrbu et al., 2019), on the order of 30 meters (based on field-measured landslide-initiation widths reported in Stewart et al., 2013). Topographic attributes along potential debris-flow tracks may require a shorter length scale commensurate with the average width of a debris flow. Robison et al. (1999) found an average width of 7 meters in field surveys following the 1996 Oregon storms.

The DEM-estimated contributing area requires the determination of flow directions for each DEM cell. Various flow-direction algorithms have been devised (Wilson et al., 2008). Those that include downslope dispersion are found to perform best in comparisons with field observations (Orlandini et al., 2012). DEM roughness also affects flow routing. We are interested in patterns of shallow subsurface flow; the

effects of tree-fall pits and mounds are probably not too important. Hence, it is appropriate to calculate flow accumulation using a smoothed DEM. However, the appropriate length scale for smoothing is not readily apparent (Erdbrügger et al., 2021) because flow lengths and drainage divide locations can be very sensitive to the degree of smoothing. Using terrain-preserving smoothing algorithms (Lindsay et al., 2019) may somewhat alleviate this. Nevertheless, some experimentation will be required to find a satisfactory smoothing length and algorithm.

The discussion above has primarily addressed topographic attributes associated with soil failures triggered by the pressure of water in-filling soil pore spaces. However, inner-gorge and channel-adjacent landslides may be triggered by a different mechanism, undercutting of stream banks, which requires a different set of topographic attributes. For channel-adjacent zones, these attributes might include channel width, valley-floor width, surface gradient, and relief to a slope break of the adjacent hillslope.

4.2.2. Forest Cover

Determination of forest-cover influences on landslide density requires the spatial distribution of cover types over the study area at the time of each landslide in an inventory (Imaizumi, Sidle and Kamei, 2008). The USGS maintains the National Land Cover Database[27], updated every five years, and the Land Change Monitoring, Assessment, and Projection[28] database, updated annually. These data sources provide land cover information inferred from the classification of Landsat imagery at a spatial resolution of 30 meters. In addition, the LEMMA project at Oregon State University has published gridded data providing a variety of forest-stand characteristics inferred using gradient nearest neighbor imputation relating Landsat imagery with forest inventory plots[29] (Ohmann and Gregory, 2002). These data are also at a spatial resolution of 30 meters, with the last update based on imagery from 2017[30].

Forest-stand characteristics, such as stand height, can be measured using lidar point clouds and structure-from-motion analysis (PhoDAR) of stereo photo pairs. Washington DNR has published gridded date-of-origin for forest stands across the state using canopy height models applied to height-age relationships[31]. Grid-cell size is 10 feet (~3m). There is a high degree of variability from cell to cell, but this data provides an estimate of stand age to apply as a terrain element for this project.

Differences in stand height during lidar acquisitions can be mapped at high spatial resolution. These maps will precisely delineate clear-cut harvested areas and unharvested zones, including unstable-slope buffers. Next, landslide density, proportions, and modeled susceptibility can be compared across harvested and non-harvested sites. Likewise, computer-delineated landforms can be compared directly to unharvested zones to examine the coincidence of inferred landform boundaries and buffered zones.

4.2.3. Storm Attributes

Ideally, all landslide-triggering storms could be characterized in terms of some measures of intensity and magnitude. A variety of gridded daily precipitation data products are currently available. The options most applicable for this application are shown in the table below.

---

[27] https://www.usgs.gov/centers/eros/science/national-land-cover-database
[28] https://www.usgs.gov/special-topics/lcmap
[29] https://lemma.forestry.oregonstate.edu/methods/home
[30] https://lemma.forestry.oregonstate.edu/data
[31] See the "Forest Inventory" link at https://www.dnr.wa.gov/opendata

Data source	Type of data	Grid size	Available dates
PRISM Climate Group, Oregon State University	Rain Gauge interpolated	~4km	1981-2021
DAYMET (NASA, DOE)	Rain Gauge interpolated	1 km	1980 to present
National Weather Service	Terrestrial weather radar calibrated to rain gauge records	~4 km	2005 to present
NASA IMERGE	Multiple satellites	~10 km	2000 to present
PERSIANN-CCS-CDR	Multiple satellites	~4 km	1983 to present

It is unclear which data product is best suited for characterizing landslide-generating storms. There are other available data products, though at coarser spatial resolution and lacking current records (e.g., NCA-LDAS). The table above focuses on higher-resolution data that are continuously updated. Data for characterizing rain-on-snow events may be challenging to find. Stanley et al. (2020) used the NCA-LDAS data to estimate snow-melt-event occurrences; Washington DNR has published a rain-on-snow zone GIS layer[32].

4.2.4. Substrate

Geologic mapping is available for the entire state at a 1:100,000 scale[33]. In addition, the Soil Survey Geographic Database (SSURGO[34]) at 1:24,000 scale are available for most of the state. Lithology and soils can be grouped into several types based on geomorphic and hydrologic criteria.

4.2.5. Climate

The PRISM website[35] offers a variety of climate-related data, such as mean annual precipitation, that can be used to characterize variations in climate across the state. Storm frequency-magnitude estimates are also available for 2-hr and 24-hour storm durations (Wallis et al., 2007). It is unlikely that this project can collect a sufficient number of landslide inventories spanning a sufficient range of climatic factors to resolve climatic influences on the distribution of landslides across landform types. Nevertheless, it will be worth seeking study sites that span this climatic range. If landslide densities and storm magnitude relationships can be discerned, these can be used with the frequency-magnitude maps to extrapolate landslide rates to other areas. Such extrapolations serve primarily as a hypothesis to guide other investigations.

4.2.6. Process-Based Modeling

This study uses empirical methods to seek relationships between landslide density and measurable terrain and environmental attributes. Process (or physically) based models may, however, also prove useful in this search. These models are based on concepts of how physical processes and environmental

---

[32] https://data-wadnr.opendata.arcgis.com/datasets/4a8339bfe8ca46b8a0a674195827e6d3_6/about

[33] https://www.dnr.wa.gov/programs-and-services/geology/publications-and-data/gis-data-and-databases

[34] https://catalog.data.gov/dataset/soil-survey-geographic-database-ssurgo

[35] https://prism.oregonstate.edu/

controls of those processes interact to trigger shallow landslides. We have so far used them to guide the selection of pertinent terrain attributes. Here we look at how they might be used in combining functional relationships between these attributes in informative ways. The output of the process-based model is not used directly but rather as an input (i.e., a terrain element) to an empirical model. We are looking to see if a process-based model can provide clues about landslide locations that might be difficult to discern using the input variables alone.

The infinite slope model (Skempton and deLory, 1957) provides a simple abstraction of these processes that have proven remarkably useful for anticipating shallow soil failures. The downslope component of soil weight drives a column of soil downslope. Soil cohesion and friction at the base of the soil column act to hold it in place. Friction is reduced by water pressure in saturated pores of the soil column, which is approximately proportional to the depth of saturation in the column. For a given steady-state rainfall intensity, this depth is proportional to the contributing area to the soil column (e.g., Montgomery and Dietrich, 1994). For the more realistic finite-duration storm, this contributing area is a function of slope steepness and soil transmissivity, which together with rainfall intensity determine the rate at which water flows through saturated portions of the soil. This simple conceptual model identifies soil depth and the contributing drainage area for the triggering storm duration, which varies with storm intensity, as key factors governing slope stability (Iida, 1999).

The soil depth varies systematically with the steepness and curvature of a slope and with positions relative to the top and bottom of a slope (Catani, Segoni and Falorni, 2010; Liu et al., 2013; Patton et al., 2018). These attributes can be used for simple models of soil depth in slope stability analyses (Dietrich et al., 1995). The US Geological Survey has recently published Fortran source code for a model to provide spatially distributed estimates of soil depth[36]. These models rely primarily on the assumption of slope-dependent soil creep rates. Landslides act to remove accumulated soils, so in landslide-prone terrain, soil depth is also a function of the time since the last landslide (D'Odorico and Fagherazzi, 2003; Dunne, 1991; Reneau et al., 1989).
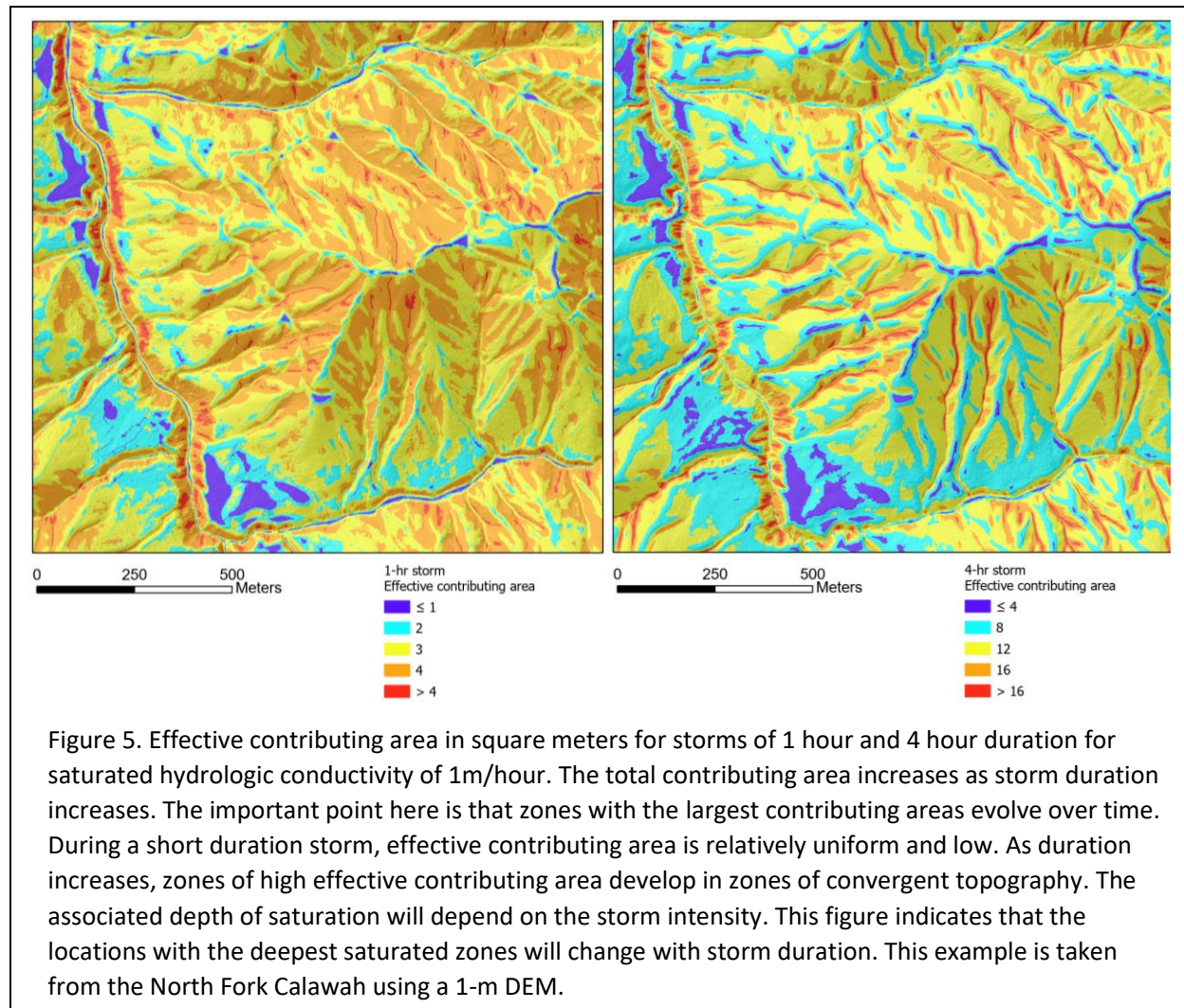
Contributing area for a specified duration storm can be roughly estimated using the Darcy velocity, $K\sin\theta$ (K=hydraulic conductivity, θ = gradient) for saturated flow with a DEM-based flow-routing algorithm. For each DEM cell, the transit time for subsurface flow from all contributing upslope cells can be calculated, and identification of those cells with a transit time less than or equal to the specified duration. Examples are shown in Figure 5.

Linking topographically based models for relative soil depth and contributing area for a specified storm duration with the infinite slope stability model provides calculated factors of safety (FoS), the ratio of forces tending to hold soil in place to those acting to move it downslope, for all DEM cells. Uniform values for rainfall intensity, soil bulk density, saturated hydraulic conductivity, friction angle, and cohesion can be used because the intent is not to predict failure but to look at how local topographic attributes might interact to create spatial variability in relative stability. The calculated values are not predictive but diagnostic and can serve as inputs to the empirical analyses described above. This combination of simple process-based models provides a means for anticipating the influence of different duration storms and therefore is worth exploring with this project.

---

[36] https://www.usgs.gov/software/regolith-fortran-95-program-estimating-soil-mantle-thickness-digital-landscape-landslide

A process-based model can also provide an index of relative stability for channel-adjacent landslides. A plain-strain (2-dimensional) factor-of-safety analysis using either a wedge or method-of-slices (Miller, 1995) can be applied at regular intervals (e.g., the DEM grid spacing) along both sides of all channels. Again, uniform soil parameters can be used because the focus here is on topographic controls. The analysis can be run twice, first with the existing DEM-derived topography and again with the base of the channel-adjacent hillslopes cut in a specified amount to represent bank erosion or channel incision. The difference provides an estimate of the sensitivity of channel-adjacent slopes to undercutting.

Note again that these process-based models are not used directly for measures of landslide susceptibility but as inputs to empirical models. The combined soil depth, partial contributing area, and infinite slope models described above use surface gradient, curvature, and flow direction as inputs, all derived from the DEM. These could be used directly as inputs to an empirical model, but perhaps the process-based model will use them to better resolve topographic controls on landslide locations. The same applies to the channel-adjacent model. Distance from the channel, surface gradient, and relief to a slope break could be used as inputs to an empirical model, but perhaps the sensitivity analysis described above will better resolve topographic controls on such landslides.

Figure 5. Effective contributing area in square meters for storms of 1 hour and 4 hour duration for saturated hydrologic conductivity of 1m/hour. The total contributing area increases as storm duration increases. The important point here is that zones with the largest contributing areas evolve over time. During a short duration storm, effective contributing area is relatively uniform and low. As duration increases, zones of high effective contributing area develop in zones of convergent topography. The associated depth of saturation will depend on the storm intensity. This figure indicates that the locations with the deepest saturated zones will change with storm duration. This example is taken from the North Fork Calawah using a 1-m DEM.

4.3 Landslide Runout

In Washington, the downslope progression of shallow landslides can be divided into two end-member types of movement: slides and flows. With a slide, the failed material slides or tumbles downslope, disintegrating as it goes, until frictional resistance within the failed material and across the surface it traverses brings it to rest. With a flow, as the failed material tumbles downslope, it becomes fluidized, and the soil grains, rocks, and other debris composing that material form a high-density slurry that suspends a wide range of grain sizes. For the three shallow landslide types in Table 1, debris slides exhibit slide-type movement, and debris flows exhibit flow-type movement. Debris avalanches can exhibit both slide- and flow-type movement. Debris flows involve fluidized debris constrained in its downslope travel by a topographically defined channel, whereas debris avalanches traverse unchannelized topography or are much larger than any encountered channels can contain. A flow can travel further than a slide, and a flow constrained by a channel – a debris flow – tends to travel further than a flow traversing unchannelized terrain. Debris flows can entrain material as they go, thus gaining volume, so it is important to distinguish the conditions that favor debris-flow initiation. The environmental factors that influence runout extent are similar for all cases: landslide volume, surface gradient, topographic confinement, slope length, and forest cover, although debris flows tend to form when landslide debris enters a steep, confined hillslope corridor. To characterize the runout length of observed landslides to assess the potential for downslope impacts on public resources and public safety, we seek relationships between runout length and these environmental factors. We focus on empirical methods for finding these relationships because physical models are very sensitive to soil parameters we cannot constrain over the spatial scales of interest. To further simplify this task, we may limit our search to one-dimensional models that only characterize runout distance, not inundation areas.

The Oregon Department of Forestry has developed a protocol for estimating susceptibility to impacts from upslope landslides[37]. These include guidelines for field-based estimation of probable runout extent. These are a good start, but they cannot combine assessments over multiple upslope landslide sites. For example, do two upslope sites with moderate initiation and runout potential pose a greater-than-moderate hazard to downslope sites? What if there are 20 upslope sites?

A variety of empirical equations and methods have been proposed to describe runout extent in terms of the initiating volume and the angle and length of the hillslope traversed by the failed material (Hungr, Corominas and Eherhardt, 2005; Hunter and Fell, 2003; Legros, 2002; Rickenmann, 2005)[38]. However, these have large uncertainties, partly because they do not address other important controls on runout extent. Mass-balance models have been developed to account for the entrainment of material along the debris flow runout path to track the potential for erosion and deposition as functions of topography and forest cover along runout paths (Cannon, 1993; Fannin and Wise, 2001; Guthrie et al., 2010a; Miller and Burnett, 2008); runout extent is determined as that point where the volume eroded equals the volume deposited. The delineation of zones for erosion or deposition in these models was calibrated to field surveys. Miller & Burnett (2008) and Guthrie et al. (2010) then used the field-surveyed zones to calibrate the models to DEM-derived attributes to be used within a GIS framework. These models incorporate

---

[37] https://www.oregon.gov/ODF/Documents/WorkingForests/HighLandslideHazardLocationsTechNote2.pdf and https://www.oregon.gov/odf/Documents/workingforests/LandslideImpactRatingTechNote6.pdf

[38] These are reviewed in the Washington Board Manual, Section 16: https://www.dnr.wa.gov/publications/bc_fpb_manual_section16.pdf

more site-specific information than those based solely on volume and slope angle, but few published examples exist.

Dependence on field data may explain the paucity of such models in the literature[39]. Reid, Coe and Brien (2016) used DEM-differencing to differentiate these zones for debris flows associated with the 1996 storms in Oregon. Their DEMs were constructed from aerial photo stereo pairs using structure from motion. Lidar differencing now opens the potential for building regional mass-balance models using remotely sensed lidar altimetry data (Scheip and Wegmann, 2022). Eroded and deposited volumes can be measured, and erosion and deposition zones along runout paths can be mapped directly (Bernard, Lague and Steer, 2021). A workflow will be developed to delineate initiation zones and search for eroded runout tracks and depositional zones downslope.

Lidar differencing will not work for all cases. Some identified landslide initiation sites will not have well-defined runout tracks and depositional zones. Some will have erosional zones evident from elevation loss along runout tracks but no deposit, particularly for those that run out to stream channels where the deposit is subsequently removed by fluvial erosion (e.g., Scheip and Wegmann, 2022). The inventory of landslide runout will thus be less complete than initiation zones and will include truncated runout tracks lacking depositional zones. We describe a strategy for dealing with that in the next section.

5 Susceptibility

Recall from the introduction:

> "We need quantitative measures of susceptibility for failure and runout extent for all hillslope locations, and we need the capability to integrate those measures over multiple failure sites to obtain quantitative measures of downslope susceptibility to impacts from landslides originating upslope. We also need quantitative measures of landslide magnitude to assess the threat posed by upslope landslides. These measures can then be parsed across landforms and integrated to provide quantitative susceptibility measures by individual landform and landform type."

Lidar-differencing-based inventories of landslide initiation sites and runout zones partitioned into erosion, transport, and deposition with volume measures provide the data needed to accomplish these tasks, or at least some portion of these tasks. To do this, we need two measures of probability:

1. The probability of landslide initiation for all hillslope points and,

2. the probability that a landslide or debris flow stops (or continues through) each increment of travel downslope.

These probabilities will be integrated along DEM-traced flow paths to calculate the probability of landslide impacts for all points along those tracks and then repeated for all potential runout tracks

---

[39] However, the Oregon Department of Geology and Mineral Industries has recently published Special Paper 53 (https://www.oregongeology.org/pubs/sp/SP-53/p-SP-53.htm) in which they describe mapping of zones of scour and deposition along runout tracks based on interpretation of aerial photographs and lidar-DEM shaded relief imagery

(Miller and Burnett, 2008). This procedure provides four maps showing the modeled spatial distribution for:

1. the probability of initiation,

2. the probability of delivery to a public resource downstream, e.g., a stream channel or a fish-bearing channel, or poses potential threats to public safety,

3. the probability that any point is traversed by a landslide that initiated upslope and continues to a public resource or infrastructure downslope, and

4. the probability of being in the runout or depositional path of an upslope landslide, even if it does not reach a public resource.

These calculated probabilities are based on landslide initiation sites and runout extents included in the inventory. They can be interpreted as the probability that any location was 1) within an inventoried initiation site, 2) that an initiated landslide traveled to a public resource (e.g., a stream channel) or posed a threat to public safety, 3) that any point was included in a mapped runout zone in the inventory for a landslide that continued to a public resource or posed a threat to public safety, and 4) that any point was within an initiation point, or runout zone, or within the depositional zone of any inventoried landslide, whether or not it continued to a public resource or posed a threat to public safety. As the introduction describes, these modeled results will reflect errors, uncertainties, and biases in the landslide inventory.

These probabilities can also be calculated by landform type. For a given landslide inventory over a delineated analysis area, the probability that a landslide is found within any landform type is simply the proportion of landslides located within that landform type. For example, if a landslide inventory contains 100 landslides and 80 are located in bedrock hollows, then the probability that a landslide is located within a hollow is $80/100 = 0.8$. If there are 1000 bedrock hollows in the inventory area, the probability that any individual hollow contains an inventoried landslide is $80/1000 = 0.08$. This method is the approach taken with watershed analysis and LHZ protocols. Likewise, the probability that any individual landform (e.g., an individual hollow) contains a mapped landslide is given by the product of the proportion of all landslides occurring within that landform type and the proportion of that landform-type area encompassed by the individual landform. Terrain elements can resolve spatial variation in landslide density over a finer spatial scale than a typical landform. Note that terrain-element values are recorded for each DEM grid point, but the attributes they represent can be measured over any scale. This ability allows us to look at how gradient, for example, varies across individual bedrock hollows and then, if landslide density varies with gradient, how susceptibility varies within a hollow or across an inner gorge. This approach provides a way to compare susceptibility across a population of landforms within a single landform type. For example, we can see if susceptibility varies across some populations of bedrock hollows based on the frequency distribution of gradients within a hollow. This test is a key requirement in seeking to nuance the definition of potentially unstable landforms. We can use existing data sets to show how terrain elements can be used to characterize variations in susceptibility within and across landform types and highlight current data limitations.

5.1 Initiation Probability

Ultimately, we need a quantitative measure of susceptibility calculated as the probability that any point on the ground could be, at some time, directly impacted by a shallow landslide. We include impacts within the initiation zone, along the runout track, and depositional zones. This probability can be estimated from landslide density. Given a landslide inventory for a delineated area, landslide density is the number of landslides divided by the area. As described previously, we will also use measures of density based on landslide area and volume. The probability that a landslide was found within any portion of that delineated area is proportional to the density and, if the location of any landslide is independent of the locations of all other landslides, is given by the Poisson Binomial Distribution (PBD). That defines the probability only for that inventory. This PBD approach provides a means of anticipating where future landslides will occur, but only to the extent that the observed landslides indicate future landslides.

With watershed analysis and the LHZ project, landforms were identified based on observed associations between certain topographic attributes, such as surface gradient and landslide locations. It was found that landslide density varied consistently across landform types, which led to the development of Rule-Identified Landforms. We illustrate this approach using an existing landslide inventory and show how landslide density, our measure of susceptibility, varies with the rule set used to delineate landforms and how these variations in density are associated with differences in surface gradient. The gradient is just one of the terrain elements found to correlate with landslide density and illustrates how such correlations can be used to refine our rules for identifying potentially unstable slopes. We focus here on landslide initiation. We separate analyses of initiation and runout because the attributes associated with initiation will not necessarily be associated similarly with runout extent.

Consider an example from the Calawah basin on the Olympic Peninsula (Figure 6). We have a DEM for the basin, derived from lidar collected in 2018, and a landslide inventory based on aerial photo interpretation and field mapping (not lidar differencing) following the Washington Watershed Analysis and LHZ project protocols. Using the DEM, we delineated a set of basic landforms within the boundaries of the study area, shown in Figure 6.

These landforms were defined using "geomorphons," a topographic classification based on pattern recognition (Jasiewicz and Stepinski, 2013) that divides the landscape into valleys, ridges, and convergent, divergent, and planar slopes. We subdivided the slope classes based on gradient, with low-gradient areas as those less than 20%, moderate-gradient slopes including those between 20% and 60%, and high-gradient slopes as those greater than 60%. We expect that lower-gradient, divergent landforms will have few to no landslides and that higher-gradient, convergent landforms will have more. The landslide inventory was recorded in GIS as a set of polygons. An initiation zone was estimated for each polygon as the area at the upslope end of the polygon extending downslope as far as the polygon is wide. We then compiled the primary landform type overlain by each initiation zone to count the number of landslide initiation sites for each landform type. Results are shown in Figure 6 and Table 4.
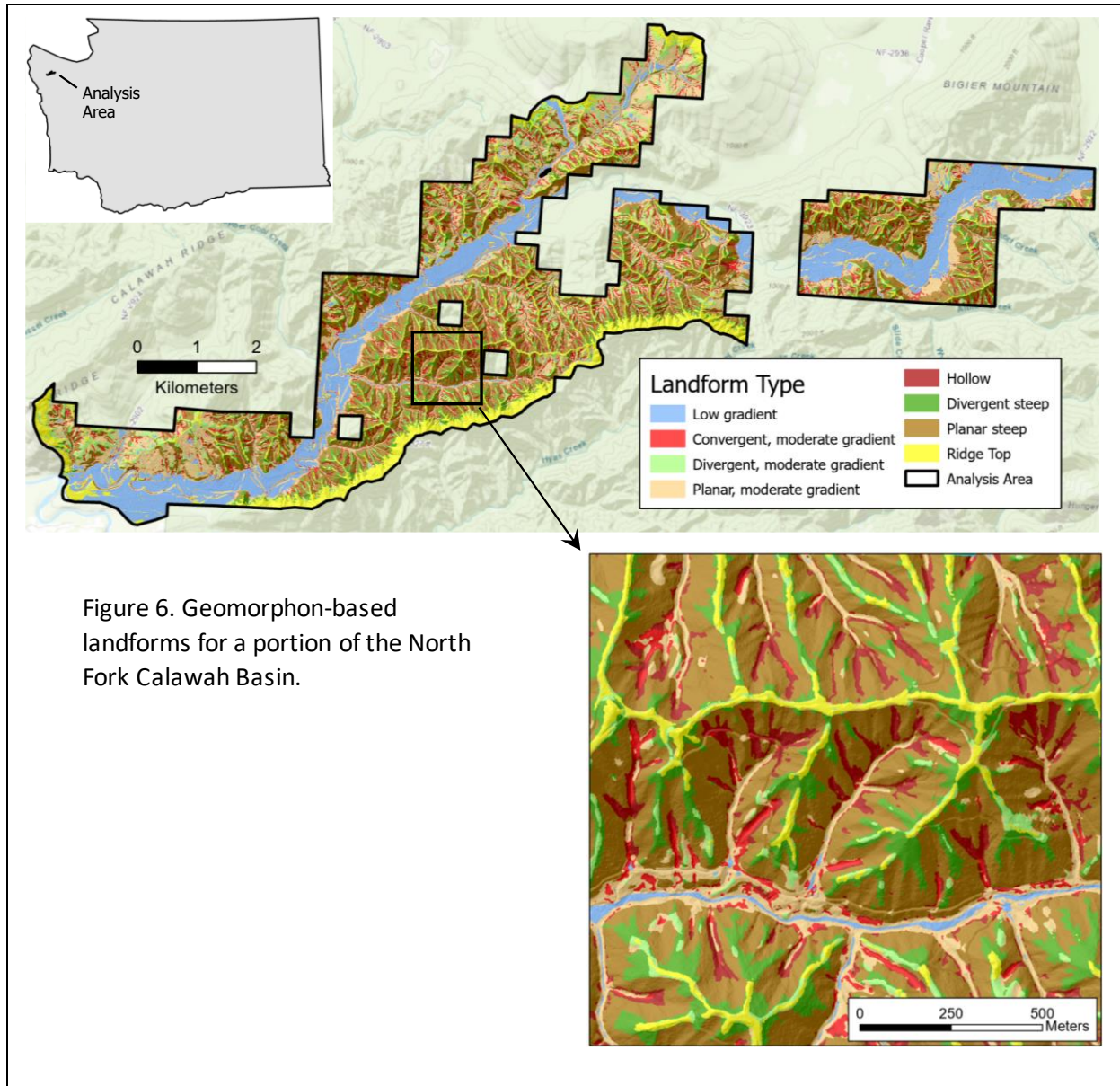
Figure 6. Geomorphon-based landforms for a portion of the North Fork Calawah Basin.

Table 4. Landslide Density by Geomorphons for the North Fork Calawah basin study area.

Landform Type	Area (km²)	Number of Landslides	Landslide Density (#/km²)
Low Gradient	7.082	0	0.000
Convergent, Moderate Gradient	1.564	0	0.000
Divergent, Moderate Gradient	2.401	1	0.416
Planar, Moderate Gradient	8.591	3	0.349
Convergent, Steep	1.365	6	4.395
Divergent, Steep	2.771	9	3.248
Planar, Steep	11.105	37	3.332
Ridge top	4.007	13	3.244
Total	38.887	69	
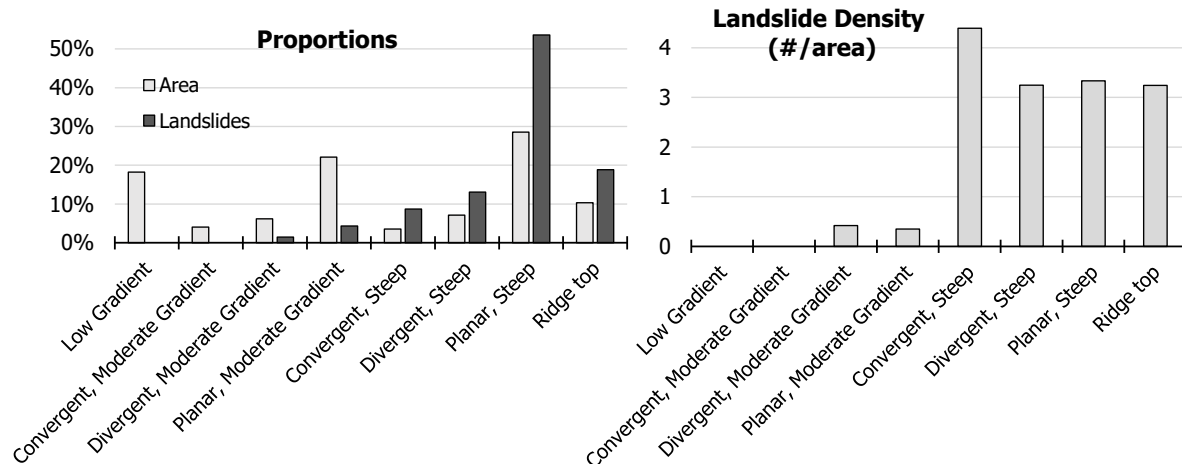
A more detailed analysis would seek to estimate the uncertainty in these results, which will be discussed shortly; for now, this example illustrates that we can define landform types with expected relative variation in landslide potential. What happens with landform definitions more specific to potentially unstable slopes? Figure 7 shows mass-wasting map units focused on current rule-identified landforms based on the landslide inventory, aerial photograph analysis, and extensive field verification (Dieu, 2015). Figure 7 and Table 5 show statistics for these landforms.

The following definitions for the landforms mapped within this analysis area are extracted from Dieu (2015): (A). Inner Gorges. "Inner gorges in the North Fork Calawah WAU are formed by two distinctly different erosional processes. Along low gradient, fish-bearing channels, such as Devil's Creek, the inner gorge edges are formed by fluvial undercutting and downcutting. The inner gorge surface is >70% and can exceed 100% in places." "The second type of inner gorge is incised into the steep hillslopes in the watershed and carries high gradient streams, >20%, to the valley floor. These are formed by repeated debris flows and were called "debris flow tracks" in the original watershed analysis; ... Where channels

within the debris flow tracks are larger, near the valley floor, or flow across debris fans and glacial terraces on the valley floor, fluvial processes play more of a role and there is a transitional zone between the two types of inner gorge. With rare exception, hillslope inner gorges have one to many bedrock hollows within their headwater drainage."

(B). Bedrock Hollows. "Bedrock hollows are steep (>70% along one or more of the axes), spoon-shaped or elongate areas of convergent topography, typically with concave profiles."

(C). Convergent Headwall. "Convergent headwalls are steep (>70% and often much greater than 70%) teardrop-shaped landforms that are concave in both the vertical and horizontal planes. A convergent headwall is comprised of many bedrock hollows and hillslope inner gorges that converge into a single channel in a midslope position."


Figure 7. RILs for the analysis area.
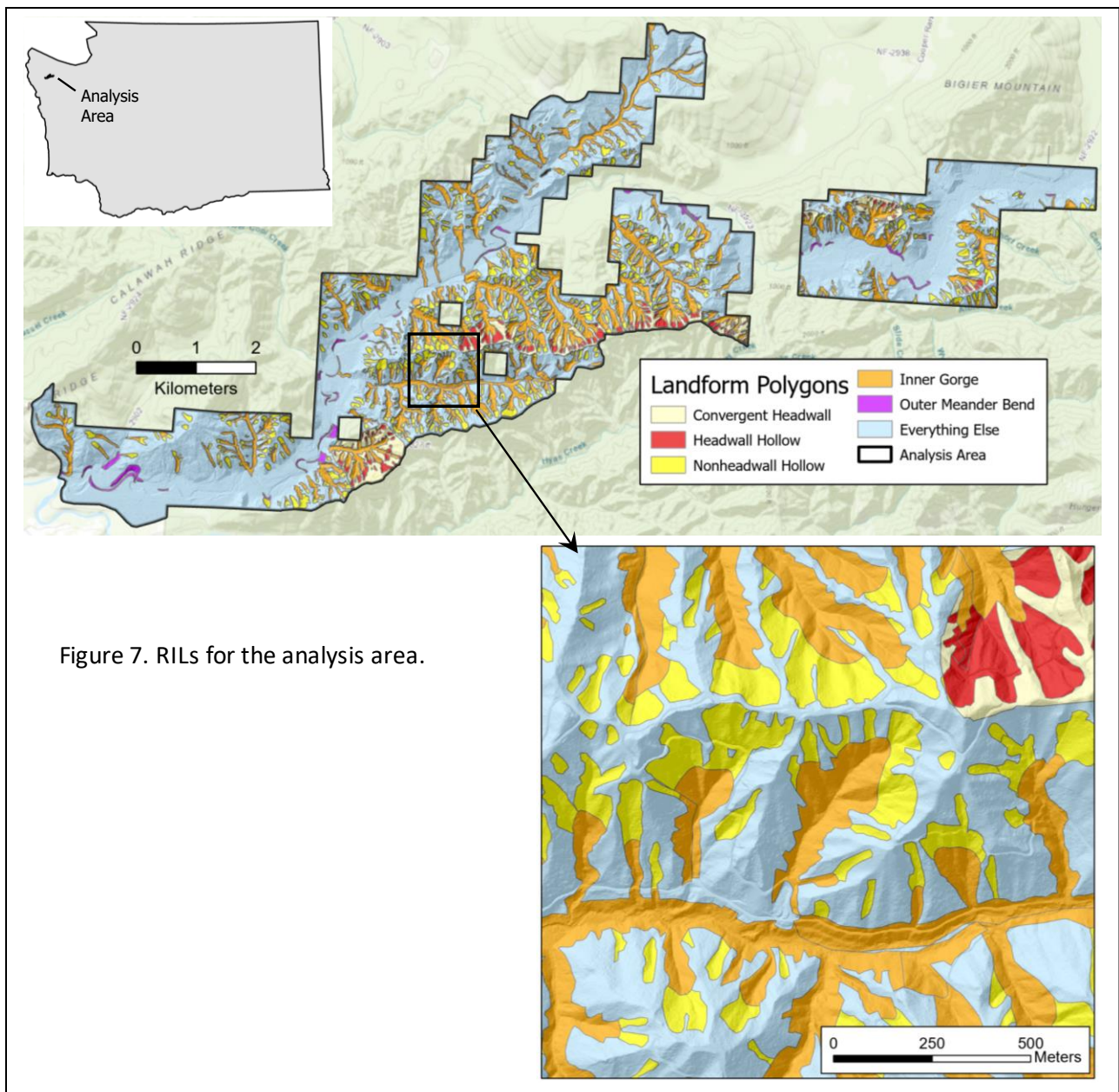
(D). Outer Meander Bends. "Rule-identified outer edges of meander bends occur where the North Fork Calawah River is experiencing channel migration that causes a meander bend to erode into a glacial terrace or the base of the hillslope for a period of years or decades."

Results here are reported for the entire area (mapped as convergent headwalls, hollows within headwalls, and hollows outside of headwalls) and for the headwall area outside the mapped hollows. The table and corresponding bar charts below show that convergent headwalls are susceptible to landslide initiation and that the hollows found within headwalls are significantly more prone to landslide initiation than those found outside of mapped headwalls.

Table 5. Landslide Density for Mass-Wasting Map Units

Rule Identified Landforms	Landslides (#)	Analysis area (km²)	Landslide Density (#/km²)
Headwall, Including Hollows	16	1.854	8.629
Headwall, Excluding Hollows	4	1.266	3.159
Headwall Hollows	12	0.588	20.402
Nonheadwall Hollows	13	3.261	3.986
Inner Gorge	22	6.632	3.317
Outside Meander	3	0.386	7.774
Outside RIL	15	26.852	0.559
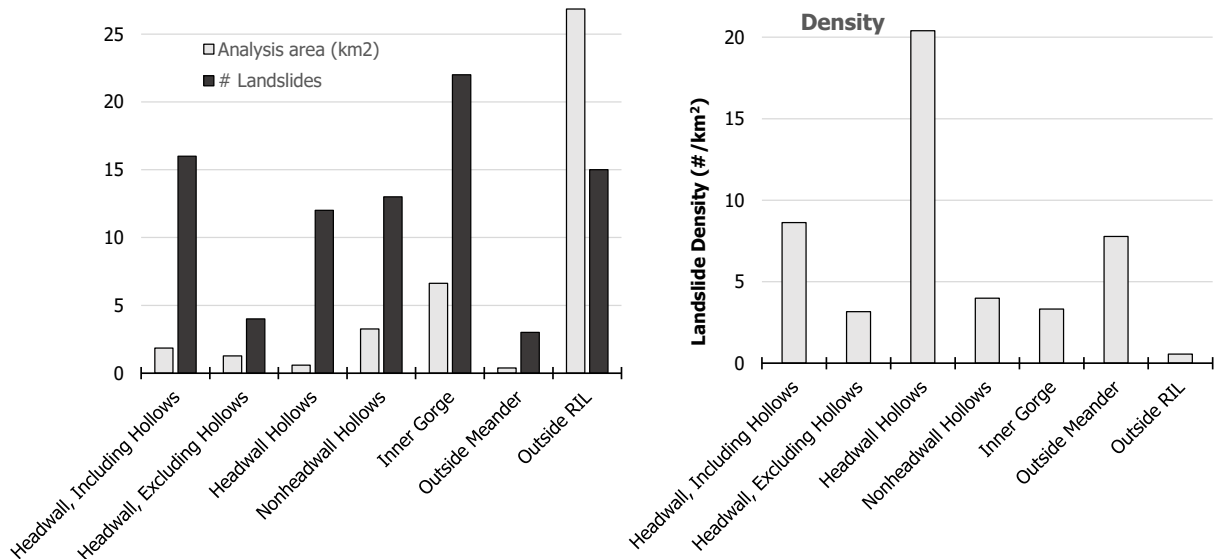
Landslide densities for these RIL-focused landforms are significantly greater than those measured for the more generic landforms in the previous example, showing that we can adjust landform definitions to better target potentially unstable slopes. This project seeks to evaluate our current definitions and refine landform definitions to improve detection probabilities potentially. As mentioned above, there is uncertainty in these results, which arises from the various sources discussed in previous sections, such as the precision of landslide mapping. For example, Figure 8 shows two mapped landslide polygons relative to the mass-wasting map units. The landslide on the left originated within a zone mapped as a

Landform Polygons
- ▢ Convergent Headwall
- ▮ Headwall Hollow
- ▮ Nonheadwall Hollow
- ▮ Inner Gorge
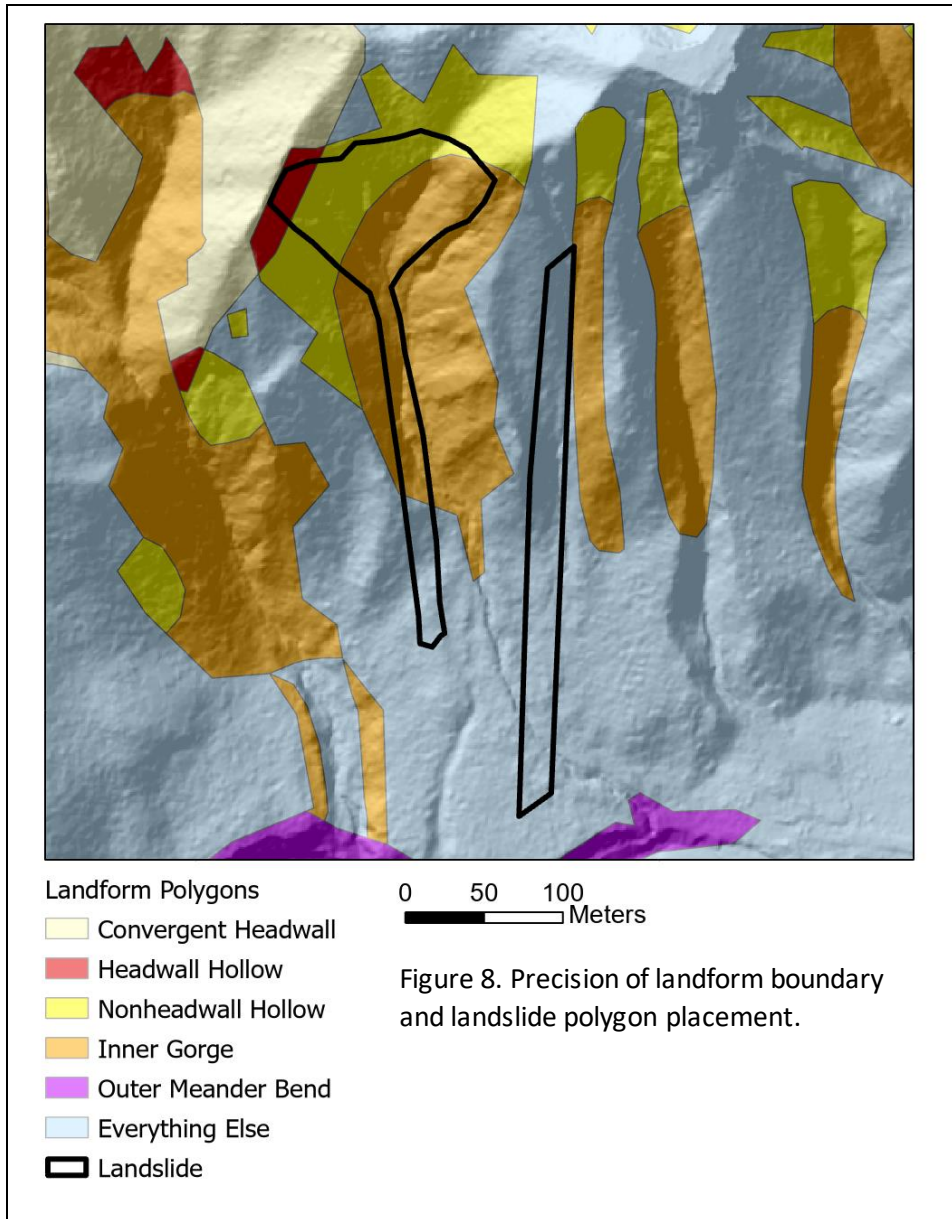- ▮ Outer Meander Bend
- ▮ Everything Else
- ▭ Landslide

Figure 8. Precision of landform boundary and landslide polygon placement.

bedrock hollow. The one on the right is shown entirely outside any RIL but close to a mapped hollow and inner gorge. The difference from the center of the landslide polygon to the center of the adjacent RIL is about 15m, at the limits of precision for mapping from field and aerial photograph observations onto USGS topographic quadrangles in the pre-GPS and pre-lidar era. If the left landslide fell within the mapped RIL, the results in table 5 above would change. Lidar differencing will eliminate uncertainty about landslide placement relative to topography.

Incorporating terrain elements into this analysis should also improve our ability to resolve potential controls on landslide locations. We illustrate below using only one terrain element, surface gradient. Figure 9 shows the gradient at each DEM pixel measured over 30 meters. These were parsed into 10% gradient bins, and each bin's area was measured. In addition, the modal value of the gradient within each landslide initiation zone was measured and parsed into 10% gradient bins. Results are reported in Table 6 below.
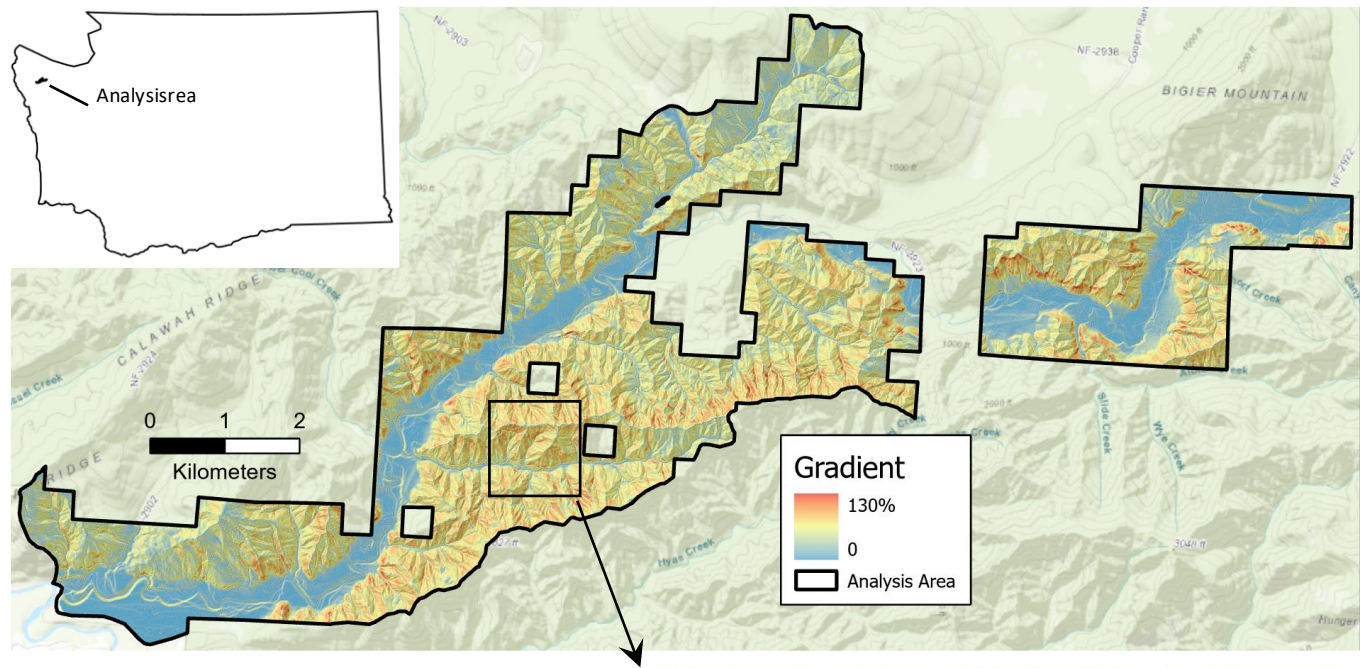
Figure 9. Gradient measured over a length of 30 meters at every DEM grid point.
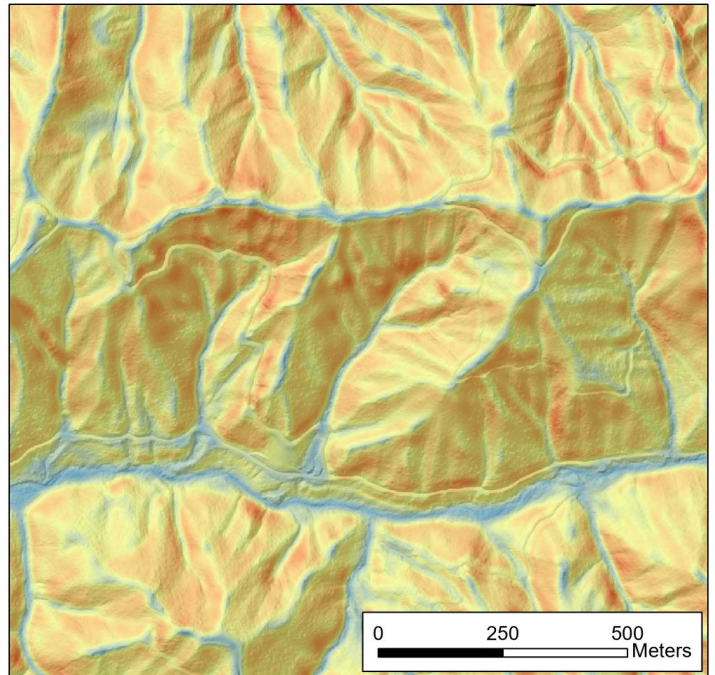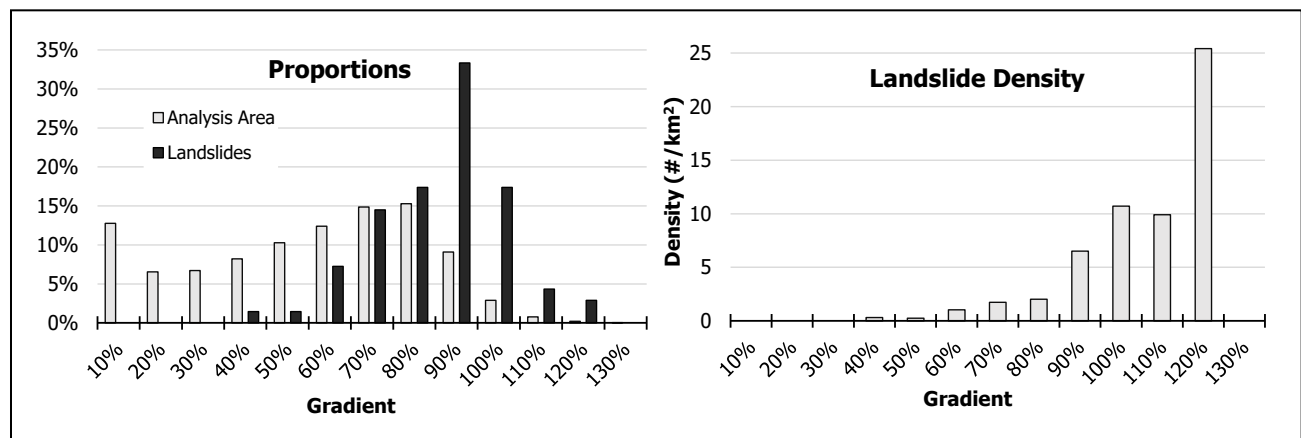
Table 6. Landslide Density by Gradient

Gradient Bin	Analysis Area (km²)	Landslides (#)	Landslide density (#/km²)
0%			
10%	4.966	0	0
20%	2.540	0	0
30%	2.607	0	0
40%	3.186	1	0.314
50%	3.995	1	0.250
60%	4.818	5	1.038
70%	5.770	10	1.733
80%	5.940	12	2.020
90%	3.530	23	6.516
100%	1.121	12	10.705
110%	0.303	3	9.914
120%	0.079	2	25.425
130%	0.002	0	0
> 130%	0.001		
Total	38.857	69	

Landslide density varies dramatically with gradient, up to a point beyond which there are no landslides, perhaps because little soil can accumulate on slopes steeper than 120%. Substantial landslides were found only on surface gradients greater than 60%, with the highest densities on slopes with gradients greater than 80%, consistent with the definitions cited above for the mass-wasting map units. Moreover, we can track gradient variations over length scales smaller than most delineated landforms with a high-resolution DEM. Gradient varies within any delineated landform type, and since landslide density appears to vary with gradient, landslide susceptibility must also vary within any delineated landform type. Look at Figure 10 below.

Figure 10. Cumulative distribution of slope gradient across four mass-wasting map unit landforms.
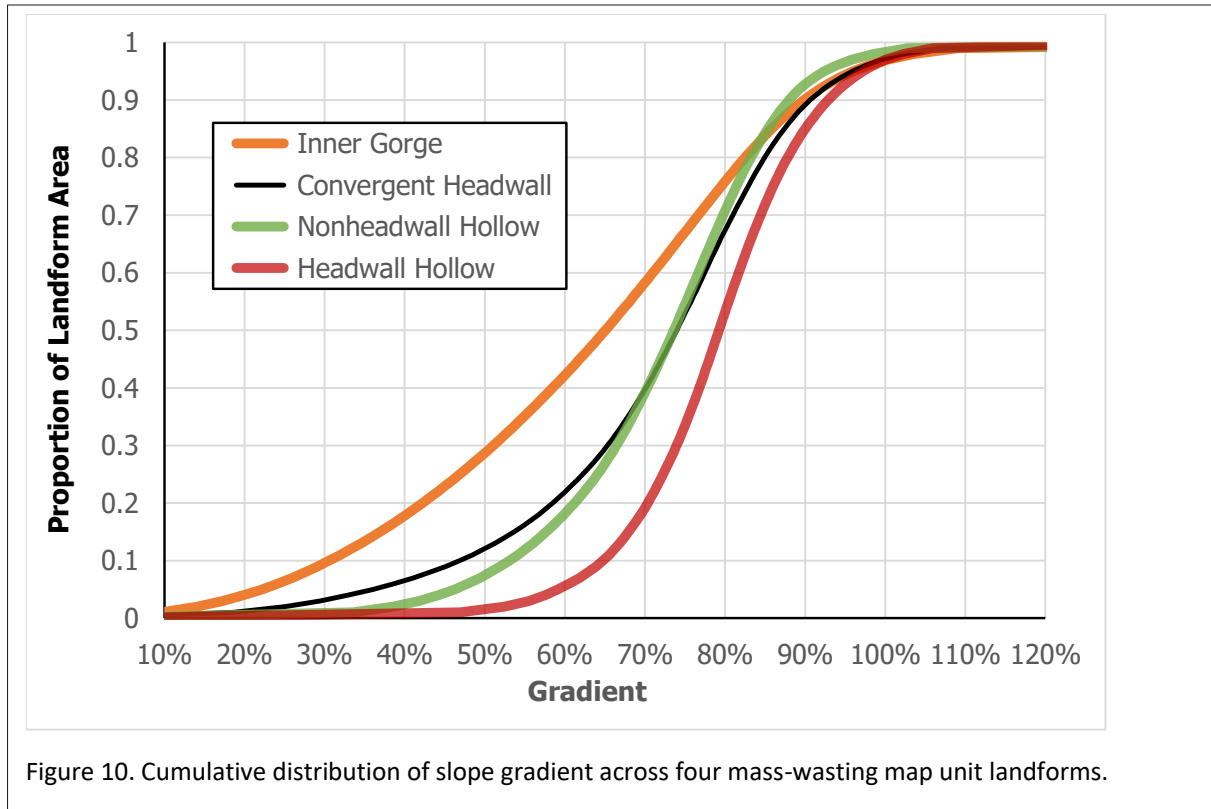
The graph shows cumulative surface gradient distributions across four mass-wasting map unit landforms. These curves provide several useful types of information.

Inner gorges, headwall hollows, and non-headwall hollows exhibit distinctly different distributions. Note that the landslide densities correlate with the proportion of landform area in steeper gradients (Table 6). Headwall hollows, with 80% of their area exceeding 70% gradient, exhibit the highest landslide density; inner gorges, with only about 40% of their area exceeding 70% gradient, exhibit the lowest density. Non-headwall hollows, with about 60% of their area greater than 70%, exhibit a density lower than headwall hollows and higher than inner gorges. Convergent headwalls have a distribution similar to non-headwall hollows but overall exhibit a higher density because of the headwall hollows they contain. Variations in surface gradients can be used to delineate and rank landforms in terms of landslide density. In this example, bedrock hollows fall into two distinct populations, one with a higher proportion of area in steeper gradients and a correspondingly higher landslide density. Each increment of gradient corresponds to a specific landslide density. The results shown in Table 6 and the chart in Figure 10 can be translated to a map similar to that shown in Figure 9, but instead of gradient, showing how landslide density, measured as a function of gradient from Table 6, varies across the landscape and within each landform type. This map would show that some hollows are more susceptible to landslide initiation, based on landslide density than others. Integrating density over an area gives the number of landslides, so this result predicts how landslides will be distributed across and within delineated RILs, a prediction we can test against additional landslide inventories. With a larger set of terrain elements and lidar-based landslide inventories, we hope to expand our ability to associate landslide locations with terrain attributes and enhance our ability to differentiate stable and potentially unstable terrain, with which to both evaluate current RIL criteria and, if warranted, refine them.

However, why do the mapped landforms in this example have any portion of their area falling on gradients less than 70%, given the definitions cited earlier? With the watershed analysis and LHZ protocols, it was expected that small areas of lower-gradient ground exist within any given landform because the mapping was done at a 1:24,000-scale, which cannot resolve small features such as the floor of hillslope inner gorges, and because the mapping effort was conservative, in the sense that anything that might meet the criteria was included with the understanding that field review could then eliminate some of the features because they were less than 70%.

This example used only one terrain element. We intend to use multiple terrain elements, which should improve our ability to identify the different combinations of factors associated with landslide location. However, we require more sophisticated methods for estimating the landslide density across the different combinations of terrain-element values found within the study domain. Estimating landslide density can be done using classification methods like logistic regression and random forests. For a given spatial domain with a given set of terrain-element values and landslide locations, these methods estimate the probability that a landslide site was observed within each combination of terrain-element values. There is a catch, however. These methods do not generally use the entire spatial domain as input. Rather, a subset of the domain must be sampled. The strategy for how that subset is obtained affects the model outcome (Conoscenti et al., 2016; Dornik et al., 2022; Gupta, 2020; Hussin et al., 2016; Kong et al., 2020; Menardi and Torelli, 2012; Oommen, Baise and Vogel, 2010; Shao et al., 2020; Titti et al., 2021; van den Goorbergh et al., 2022). This effect is an important consideration in sampling strategies. The sampling strategy will also depend on the extent of the spatial domain, the number of landslides in the inventory, and the number and range of values found for the terrain elements examined. Hence, a sampling protocol cannot be specified a priori but will depend on what is found during the analysis. An example is provided in the appendix.

Determination of probable landslide locations is a classification problem involving a yes or no answer: any location is or is not within an observed landslide initiation zone. Landslide size, measured in terms of area and volume, varies over a large range of values. Relating landslide size to terrain-element values is a regression problem for which general linear and additive models are well suited.

5.2 Runout Probability

We want to calculate the probability that the downstream edge of debris mobilized by a landslide will stop in any increment of travel. If this probability is uniform, then the distribution of runout lengths for a population of landslides will be exponential. If the probability gradually increases or decreases downslope, then the distribution of runout lengths will follow a Weibull distribution. More generally, the distribution of runout lengths provides information about the probability for the leading edge of a landslide to progress through any increment of travel. We expect this probability to vary with several factors, including the landslide volume, topography, and vegetation along the travel path. This problem is similar to that faced in other disciplines. What is the probability that a machine part will fail over the next week? What is the probability that an individual will die of a heart attack in the next year? Instead of an increment of time, we are interested in an increment of length: what is the probability that a landslide will stop in the next 10 meters? In characterizing the downslope movement of soil particles in processes of soil creep and diffusion, Furbish and coworkers employ a similar conceptual framework: once a particle is in motion, what is the probability that it will stop in the next increment of travel (Furbish and Roering, 2013; Furbish et al., 2021; Williams and Furbish, 2021). Survival analysis, a well-

developed field of statistics, addresses these questions directly. Survival analysis has developed primarily in response to medical and engineering issues; many textbooks describe the methods (e.g., Kleinbaum and Klein, 2012), and an ongoing output of research papers expands its applicability.

None of these, however, include an analysis of landslide runout. Neither can we find any examples in the landslide literature citing survival analysis. For example, the methods described by Miller and Burnett (2008) used a survival-analysis approach but were not cited as such (at that time, we were unaware that such a statistical method existed). Now, however, we have a wealth of resources, specifically, a variety of R[40] packages (e.g., Cekic et al., 2021; Rizopoulos, 2016; Therneau, 2022).

To estimate the probability that a debris flow will travel to any point downslope, we need to define the probability of stopping in an increment of travel distance. This location is called the hazard function (or rate) in survival analysis. The hazard rate may depend on conditions at the landslide initiation point – distance zero (equivalent to time). Such conditions include the initiating volume or local attributes indicative of volume. The hazard function may also vary with distance traveled (equivalent to time elapsed). This variation may be a function of topographic and forest-cover attributes along the travel path. In Miller and Burnett's (2008) implementation of this approach, the hazard function varied with the modeled ratio of deposited to scoured volume. These volumes were estimated proportionally based on the probability of scour or deposition integrated along the runout track. Where this ratio is small, the hazard function (the rate) is small: the probability of stopping is low. When the volume deposited is close to the volume scoured, the ratio and probability of stopping are high. The probability of scour or deposition was based on logistic regression of field-surveyed zones of scour and deposition to DEM-based measures of channel gradient and confinement. The same or a similar approach will be applied using runout tracks mapped with lidar differencing, providing volume estimates more accurately and at a much finer spatial scale than previous efforts (e.g., Scheip and Wegmann, 2022). However, existing software tools (R packages cited above) exist for the analyses. The analysis must accommodate variables that change over time (distance), such as that described in Rizopoulos (2012) and illustrated in the appendix.

An advantage to using survival analysis techniques is the ability to incorporate time series, or in this case, debris flow and landslide runout tracks, that are truncated; that is, the time of death (downslope location of the landslide deposit) is not known. Locations where runout occurred, provide information even if the endpoint is unknown.

These techniques can be applied for both slide- and flow-type landslides. Lidar differencing can delineate three zones: initiation sites, erosional portions of runout tracks, and depositional zones. Based on other studies (Benda and Cundy, 1990; Fannin and Wise, 2001; Miller and Burnett, 2008), we expect the tendency for erosion or deposition to vary with surface gradient, topographic confinement, geology, and forest cover. These mapped zones can be used to calibrate those relationships, e.g., using logistic regression. If topography is a primary control on whether a landslide evolves into a debris flow, then runout length should remain a function of these variables.

5.3 Uncertainty and Validation

Certain landslide initiation and runout controls cannot be identified or resolved with remotely sensed data, even with extremely detailed lidar DEMs. Hence, a certain degree of uncertainty is unavoidable,

---

[40] https://www.r-project.org/

even with abundant, high-quality data. Field observations can identify potential sources of uncertainty, such as spatial variation in soil depth or shallow bedrock fracture density. To apply these observations to a GIS-based model, they need to be quantified as functions of mapped soil and geologic units. That is beyond the scope of this project. However, the distribution of landslide density values associated with terrain elements will be calculated. This distribution is partly determined by the spatial variation in physical attributes like soil depth and fracture density, so these sources of uncertainty are captured within our estimates of uncertainty – we cannot, however, point to the source of that uncertainty.

Likewise, the quantity of data available, particularly the number and geographic range of landslides found with lidar differencing, will further determine the confidence to place in measured and modeled landslide density and susceptibility measures. Statistical protocols for evaluating the uncertainty associated with any particular dataset are well established, and analysis tools for fitting and evaluating empirical models and estimating uncertainty in model results are readily available (e.g., as R packages). Models should be tuned and validated using nested K-fold cross-validation. Measured and modeled quantities, including observed landslide densities and maps of initiation and runout probability, should include confidence intervals and measures of the model's ability to predict observed landslide occurrences.

Field observations need to be incorporated as integral components for two parts of this project:

1. To verify the interpretation of elevation differences observed with lidar differencing. Using relatively recent sequences of lidar datasets, for which landslide scars will not be more than a few years old, a set of field-accessible landslide sites based on lidar differencing should be located, and the interpretation and lidar-measured landslide area, volume, and runout extent compared to field observations. Field volume measures can be based on standard surveying techniques or, preferably, terrestrial lidar. A range of landslide sizes should be sampled – but this may depend on what recent landslide sites are available. The field-based inventory from the post-mortem study (Stewart et al., 2013) provides an initial comparison of field- and lidar-based measures of landslide size, but for this study, more detailed measures of landslide area and volume, particularly of debris-flow width and scoured volume along runout tracks, would be very useful. This field check should occur as early in the project as possible.

2. To identify sources of uncertainty in model results. Identifying uncertain sources is done by finding areas where the model performs poorly, where predicted and observed densities of landslide initiation sites and runout tracks do not match. Field visits of landslide sites within these areas should then seek to identify possible causes for the model mismatch. These may include such factors as substrate conditions inconsistent with the mapped geology and soils, topographic attributes not captured by the topographic attributes applied in the model or anomalous accumulations of boulders and large wood within runout tracks. The surface gradient measured from the lidar DEM can also be compared to ground-based measures. Identifying potential reasons for locally poor model performance will aid in identifying site-specific factors that will alter the interpretation of model results. Identification of these factors could also guide improvements in model design.

The number of field observations appropriate for each will depend on the results obtained as the project progresses.

6 Discussion

6.1 Why Build Empirical Models?

With a landslide inventory, landslide densities and proportions can be calculated directly. So why bother with the modeling? There are several reasons:

1. Multiple factors influence landslide susceptibility. The distribution of these factors across terrain elements and landform types is not uniform. Direct comparison of observed densities and proportions is hindered by the bias resulting from these non-uniform distributions. For example, areas recognized on the ground as potentially unstable slopes (as a RIL) have no harvest buffers; thus, these areas will have older forest stands relative to others. The modeling can separate the relative influence of different factors to compare terrain elements and landforms better. In this example, if the modeling indicates that forest-cover type influences landslide density, then the densities can be recalculated as though all areas were under a uniform forest stand, thus removing (or minimizing) that confounding factor from our comparisons. Regression models (e.g., generalized linear and generalized additive models (Wood, 2017)) are best suited for this type of use. However, machine-learning models could also be used by running the trained model with a uniform forest cover, in this example, and comparing the resulting probability raster to the original.

2. Extrapolation to areas with no inventories. If we can use empirical models to identify the influence of the primary factors affecting landslide occurrence, then we can use the model to predict relative landslide densities in locations without inventories. This use comes with the caveat that empirical models cannot be applied with high confidence for conditions outside those they were built with, but it does not preclude their use. Current RIL criteria are applied in areas where we have no landslide inventories. The methods proposed here will provide estimates of susceptibility in quantified terms with quantified estimates of uncertainty. That is considerably more than we can do now with mapped or field-delineated RILs. Model predictions also indicate how landslide susceptibility varies regionally. Consider a scenario: contributing area is a primary factor explaining landslide density for steep, convergent slopes (hollows) in a watershed. This finding suggests that hollow size is important. If we go to a neighboring watershed where hollows tend to be smaller, the model would suggest a lower landslide density for that landform. The empirical model allows us to account for differences in topographic characteristics found in different populations of the same landform type and predict how those differences will manifest in landslide densities and proportions. These are testable predictions.

3. Statistical modeling provides confidence measures through subsampling and testing of model predictions. How representative of the population of landslides that can occur in a region is the sample of events contained in a landslide inventory? How confident can we be that the densities and proportions measured from one study would be the same as those from another study in the same or different regions? That confidence will depend on the size of the sample, both spatially and temporally, and the range of conditions that produced it. Statistical models use that information to provide quantitative estimates of confidence. We are not after "high, medium, and low" but quantified measures of the proportion of landslides and mobilized volume that impact public

resources and threaten public safety. Likewise, we are after quantified estimates of how far off those measures might be based on confidence intervals. These measures allow decision-makers to quantitatively compare different management options, such as variations in the extent of unstable-slope buffers, and to access the degree of confidence in those comparisons.

6.2 Expectations

There is a well-developed methodology for empirical estimation of landslide-initiation susceptibility documented in many hundreds of publications in the peer-reviewed literature. These involve the techniques of mapping landslide locations, overlay of those locations on other mapped attributes, and statistical analyses seeking relationships with these attributes. The current understanding of shallow landslide initiation and the resulting RIL definitions in Washington are based on exactly that type of analysis. So what can the analyses proposed here tell us that we do not already know?

1. Use of lidar differencing will provide precise measures of landslide location, area, and volume, and perhaps of the extent of landslide runout. Such precise measurements have not been available for previous studies in Washington. Improved precision will result in improved resolution and identification of the physical factors associated with landsliding and will allow comparison of the proportion of mobilized volume across different terrain-element combinations and landforms. As illustrated with the example in the introduction, analysis of mobilized volume in terms of proportions might alter our understanding of the relative importance of different landscape positions and landforms in the context of sediment budgets and impacts to stream-channel systems.

2. Inclusion of spatially and temporally distributed precipitation data in these analyses might identify relationships of storm magnitude with landslide density (by number, area, and volume) and landslide location. Including spatially distributed precipitation data could reduce the confounding influence of spatial variations in antecedent moisture conditions and storm intensity on modeled landslide susceptibility. It could also provide the ability to estimate landslide rates regarding storm intensity and duration frequency distributions. This capacity would reduce the bias of incomplete landslide inventories and allow insights into the consequences of evolving storm distributions associated with a changing climate. It will also provide a means of evaluating the importance of extreme events that trigger landslides in unusual locations – outside of current RILs – as occurred during the 2007 storm in southwest Washington.

3. Inclusion of spatially and temporally distributed stand-type data will allow an analysis of how forest conditions (e.g., stand age inferred from tree heights) influence landslide susceptibility.

4. Initiation probability is linked explicitly with runout potential to define susceptibility to landslide impacts. Susceptibility to landslide impacts depends on both the potential for upslope landslide initiation and the potential that a landslide from any potential initiation site will travel to the point of interest.

5. Inclusion of terrain-element analyses of landslide density, proportions, and probability, which include the previous four listed items, into GEOBIA-based landform delineations for landslide susceptibility. These inclusions will characterize landforms in terms of landslide-driven fluxes of material (sediment, wood) across the landscape and to channel systems and downslope human infrastructure. Upslope landforms will be linked directly to downslope consequences. Upslope landforms where landslides initiate can be evaluated and ranked in terms of their impacts and threats to landforms and human infrastructure downslope. Downslope landforms (stream channels, flood plains) and human

infrastructure can be evaluated and ranked in terms of the potential impacts of upslope landslides. This quantification provides a measure of susceptibility directly tied to potential impacts on public resources and threats to public safety. These quantitative susceptibility measures for both initiation and runout will be characterized independently of any landform delineation and at the resolution of the DEM so that they can be integrated to provide direct susceptibility measures for any set of delineated landforms. This approach will provide a means of evaluating both current RIL criteria and any alternatives regarding susceptibility.

Progress on these five items will result in more accurate and consistent identification of unstable slopes and landforms and of unstable-slope and landform conditions where landslide runout would likely harm public resources or pose a threat to public safety, the first and third items in the critical question for the Unstable Slope Criteria Project.

Accomplishing these expectations will involve using new data sources and analysis methods. Therefore, a large part of the project will involve assessing data sources and developing and testing new analysis methods. The degree to which these five analysis goals can be met is uncertain. Nevertheless, this is a worthwhile effort. Even partial success on any of the five can provide substantial progress in understanding the ecological and geomorphic role of shallow landsliding and our ability to assess forest-management influences on landslide susceptibility. This effort must be taken one step at a time, hoping to do a better, more detailed job while advancing some of what we ultimately hope to understand regarding unstable slope delineation across managed forest lands.

7 Task List

Tasks can be divided into four primary groups. These are:

1. Creation of landslide inventories using lidar differencing.
2. Compilation and evaluation of terrain-element data sets.
3. Identification of relationships between landslide locations (and size and volume) and runout extents with terrain elements. These relationships will be used as inputs in the production of landslide initiation and probability of delivery rasters.
4. Use of modeled probabilities to evaluate current RIL criteria and to explore possible modifications or additions to these criteria that may improve the identification and delineation of unstable slopes and landforms.

Each group has a set of subtasks.

1. Lidar differencing to produce landslide inventories.
   This task will be done initially for the post-mortem study area using 2006, 2017, and 2019 lidar acquisitions, allowing for comparing landslide locations, numbers, and sizes to those identified with field surveys. Subtasks include:
   a. GEOBIA delineation of road prisms. These may be used for co-registration and to help distinguish road-related landslides.
   b. Co-registration of lidar point clouds.
   c. Creation of elevation difference rasters, including point-cloud-based analysis of local roughness and confidence in elevation differences (Bernard, Lague and Steer, 2021).

    d. OBIA to delineate candidate landslide sites using rasters of elevation difference, gradient, and other data that may prove useful in distinguishing landslide scars from other sources of surface change.

    e. Use of available imagery to validate delineated landslide sites and further constrain the dates of landslide occurrence.

    f. Landslide surface area and mobilized volume from elevation differencing.

    g. Delineation of downslope runout paths, differentiating scour zones, transport with no net scour or deposition, and deposition.

    h. Comparison of landslide sites mapped from lidar differencing with the post-mortem landslide inventory. Determine ability to identify field-surveyed sites; compare length, width, average depth, and maximum depth from field measures to those from lidar. This step determines the feasibility of lidar differencing for mapping landslide sites (if lidar differencing proves untenable, we will have to rely on the field- and photo-based inventories).

       i. What proportion of the field-surveyed sites are identified?

      ii. Is there a correlation between confidence in the lidar-determined elevation difference and detection (or not) of the field-mapped site?

     iii. Is there a threshold of landslide size that can be detected? If so, how does this threshold depend on confidence in the elevation difference?

    i. The above steps will determine the confidence to place in the number of landslides identified with lidar differencing using a comparison with a field-based inventory for the same region. To better constrain the confidence to place in delineated landslide initiation polygons and runout tracks, landslide sites mapped from lidar differencing using more recent lidar datasets (so that landslide scars within a few years old are identified) should be compared with field observations of the same sites. This step will aid in interpreting elevation differences in differenced datasets and provide additional estimates of the confidence to place in lidar-based landslide size and runout extent measures. It may be that some landslides will be identified in the overlapping area of the 2017 and 2019 lidar datasets for the post-mortem area. These could provide field sites for comparison. If not, other more recent lidar datasets must be examined to find field sites.

    j. The above steps will establish a workflow for creating landslide inventories using lidar differencing. This workflow will be automated as much as possible. The next step is to then work systematically through available overlapping lidar datasets, starting with those where the most known landslides have occurred[41].

2. <u>Compilation of terrain-element data sets.</u> This compilation can occur concurrently with Task 1.

    a. Evaluation of gridded precipitation data to obtain antecedent rainfall and storm-specific precipitation measures for each landslide.

    b. Forest-cover classification over the study area(s) for each age class of landslide occurrences. The spatial distribution of forest types over the entire study area is needed for each landslide date. The origin-age rasters available from DNR provide a data source for this purpose. LEMMA stand-structure data and the US Geological Survey LCMAP and NLCD data provide additional resources.

---

[41] https://www.dnr.wa.gov/programs-and-services/geology/geologic-hazards/landslides#recently-reported-landslides

    c. Computation of surface gradient, mean curvature, and, as needed, other topographic-attribute rasters.

    d. Computation of effective contributing areas for a range of storm durations and intensities.

    e. Computation of modeled soil depth rasters.

    f. Computation of factor-of-safety rasters as described in the text.

    g. Computation of channel-adjacent-slope sensitivity to channel incision and bank erosion, as described in the text, and applied specifically for inner-gorge landslides.

3. <u>Identification of relationships between landslide locations and runout extents with terrain elements</u>.

    a. Empirical models to obtain terrain-element functional relationships with landslide density. This analysis requires subsampling of non-landslide sites. Use appropriate sampling techniques for training, validation, and testing; cross-validation for confidence intervals. Frequency-ratio techniques may also be used.

    b. Linkage of process-based and empirical techniques. Models for soil depth, saturation depth, and factor of safety. Use modeled factor-of-safety as input to empirical models for landslide density. Here we seek evidence of variations in landslide location and type with measures of storm magnitude. These measures include the storm duration and intensity applied in the hydrologic model and the precipitation amounts obtained for each landslide site from gridded precipitation data. If such relationships are modeled and observed, this method provides a means of convolving modeled landslide density over the entire storm distribution for a region to estimate the long-term proportion of landslides originating from different landscape locations. It also provides a means of extrapolating model results to areas or future times with different storm characteristics.

    c. Calibration of models for probability of scour and deposition along runout paths. Explanatory variables may include pre-failure surface gradient, topographic confinement (plan curvature), and forest cover along the runout path.

    d. Testing of models for initiating volume. If the initial volume of a landslide can be constrained in terms of measurable quantities, it can then be used as an input variable for survival analysis. Potential candidate variables for initial volume include topographic attributes (contributing area, gradient, curvature), topographic position, and forest cover.

    e. Calibration of the hazard rate as a function of conditions at the initiation site and along the runout path using survival analysis methods. A mass-balance approach can potentially be used. For example, integrating the probability of scour along a runout track provides an index of potential accumulated volume. Integration of the probability of deposition provides an index of potential deposited volume. Their ratio measures the volume deposited relative to the volume accumulated along the runout track. This ratio can provide an input variable that changes as a function of runout length and incorporates the influences of topography and vegetation.

    f. Identification of areas where the model performs poorly. This identification is done by comparing modeled to observed landslide densities over landform types or across areas defined by a specified length scale. Use field visits to identify and document potential causes for poor model performance.

4. <u>Use of modeled probabilities to evaluate current RIL definitions and, if needed, to show how those definitions could be modified or expanded in response to the critical question for this project</u>. There are at least two ways to approach this.

   a. Use landform polygons representing current RIL definitions delineated with GEOBIA. These will not precisely match RIL boundaries that would be delineated on the ground, but they provide a consistent means of delineating RIL polygons across different regions. Statistics for each landform type can then be computed by integrating modeled initiation and delivery probabilities over landform areas. These statistics can be compared to observed landslide proportions for areas with landslide inventories. This comparison provides an evaluation of current RIL definitions.

   b. In areas subject to Washington forest practice rules, unharvested areas include no-harvest buffers delineated on the ground. Unharvested zones can be precisely mapped using canopy heights from lidar point clouds. Modeled landslide densities and proportions can then be compared for harvested and unharvested zones. Additionally, GEOBIA-delineated RIL polygons can be overlain on the harvest/no-harvest zones to gain an estimate of the no-harvest area consisting of unstable-slope buffers. Comparison of harvested zones with modeled landslide initiation and delivery proportions can be used to identify harvested areas that lie within areas having the potential for landslides that can impact public resources or threaten public safety. These zones provide another opportunity for field checking of model outcomes.

8 Study Deliverables

Next is a list of targets. Final deliverables will depend in part on what is discovered in terms of the capacity to use available data throughout the study.

1. Workflow and protocols for building landslide inventories using lidar differencing and auxiliary data sets (multiband imagery, precipitation). Objectives for each inventory include:
   a. Surface area and mobilized volume of each landslide initiation site.
   b. Runout distance for all initiation sites where this can be determined, divided into zones of scour, transport, and deposition.
   c. Spatial distribution of stand types over the inventoried area at the time of each landslide.
   d. Storm attributes associated with each landslide.
   e. Aggradation of mapped lithology (1:100,000 scale) and soils (SSURGO, 1:24,000 scale) into geomorphic and hydrologically homogenous groups, the spatial distribution of each group, and distribution of landslides across each group.

2. Generation of scripts for deriving elevation derivatives: gradient, curvature, total contributing area, and estimated effective contributing area for a range of storm durations. These will be derived at appropriate length scales with appropriate DEM smoothing and flow-routing algorithms.

3. Python and R scripts for statistical analysis of inventory data. These will include options to subsample datasets for training, validation, and testing and options for regression (GLM, GAM) and machine learning (random forest, CNN) modeling techniques.

4. A report describing the results of the study, including the applicability of the different datasets for the intended tasks (e.g., suitability of available precipitation data for characterizing storm intensity

and duration), confidence in the number, size, and volume of landslides mapped with lidar differencing, and relationships (or not) of landslide density with forest-stand types, topographic attributes, lithology, soils, and precipitation.

5. Evaluation of current RILs regarding landslide proportions and density, including the proportion of landslides that occur outside RILs. Identification of potentially useful subdivisions or ranking of RILs, such as slope thresholds that vary with lithology. Identification of potential additions or modifications to current RILs. These analyses will be based solely on susceptibility; sensitivity to forest practices will be evaluated with Project 4.

6. Digital maps of model outputs, including landslide inventories, modeled initiation, delivery probability, and landforms delineated using GEOBIA. These data may be distributed as interactive online maps, but such online distribution is not a specified deliverable for this project.

All data analyses will be implemented using Python and R computer language scripts. These scripts may call executable libraries developed with other languages and available through open-source GIS (e.g., GDAL[42]). The scripts and all analysis tools should be licensed as open source and available on a public repository, such as GitHub. Ideally, data used for all analyses would be archived and available so anyone could use the scripts to replicate all analyses. However, these data sets may be prohibitively large, so an illustrative subset of the data sufficient to complete the analyses for some smaller areas will be provided via a public link (e.g., via Box[43]). Data analysis workflow will include outputs showing the results of validation and testing.

---

[42] https://gdal.org/
[43] https://www.box.com/

Appendix: Analysis Details

1 Lidar differencing

We cited differencing of overlapping aerial lidar datasets as a promising method for creating a landslide inventory to be used for an empirical analysis of landslide susceptibility. The primary benefit over traditional mapping from aerial photographs is that the topographic attributes associated with landslide initiation and runout are derived from the same lidar datasets used to map landslide locations. Identification of landslide locations and runout extents directly from the same data used to characterize the topography would greatly reduce the uncertainties in associating terrain elements with landslide occurrence that plague analyses based on aerial photo mapping and field surveys alone. Likewise, lidar differencing might allow determination of landslide surface area and mobilized volume.

This is a relatively new technology for which standardized methods have not been established, but an increasing number of studies provide examples of how sequential aerial lidar datasets can be used for detecting geomorphic change (Eitel et al., 2016; Okyay et al., 2019; Qin, Tian and Reinartz, 2016). These include analysis of valley-floor changes following floods (Croke et al., 2013; Goodwell et al., 2014; Lallias-Tacon, Liébault and Piégay, 2014; Lane, Westaway and Murray Hicks, 2003; Piovan et al., 2023); post-fire erosional processes (DiBiase and Lamb, 2019; Guilinger et al., 2023; Morell et al., 2021; Orem and Pelletier, 2015; Pelletier and Orem, 2014; Rengers et al., 2021); characterization of landslide-size distributions (Ju, Zhang and Xiao, 2023; Massey et al., 2020); measuring surface displacements across faults (Nissen et al., 2012; Scott, DeLong and Arrowsmith, 2020), over large deep-seated landslides and earthflows (Ghuffar et al., 2013; Paulin et al., 2022; Prokešová et al., 2014; Rechberger, Fey and Zangerl, 2021), and of rock glaciers (Fleischer et al., 2021); analysis of regional erosional processes and delineation of geomorphic regimes (Cavalli et al., 2017; Heckmann and Vericat, 2018; Müller et al., 2022; Sailer et al., 2012; Scorpio et al., 2022; Scott et al., 2022); analysis of debris-flow initiation and bulking rates (Coe et al., 2021; Morino et al., 2019; Scheidl, Rickenmann and Chiari, 2008; Scheip and Wegmann, 2022); and creation of landslide inventories (Bernard, Lague and Steer, 2021; Burns et al., 2010; DeLong et al., 2022; Fernández et al., 2021).

Remarkably, every cited study employs a unique workflow for lidar differencing, using a variety of strategies, algorithms, and software for pre- and post-processing of the lidar data and measuring of surface changes occurring during the period between data acquisitions. There are efforts to establish standard protocols (Bailey et al., 2022; Scott et al., 2021) and numerous studies examine error sources and approaches for characterizing and reducing uncertainty (Anderson, 2019; Cheng et al., 2018; Cucchiaro et al., 2020; Glennie et al., 2014; Kamp et al., 2023; Kharroubi et al., 2022; Kusari et al., 2019; Lague, Brodu and Leroux, 2013; Passalacqua et al., 2015; Qin, Tian and Reinartz, 2016; Schaffrath, Belmont and Wheaton, 2015; Viedma, 2022; Wheaton et al., 2010; Zhang, Glennie and Kusari, 2015), so there is a solid conceptual basis with which to develop a feasible workflow.

Our goal is to develop a workflow for change detection using the DTM (Digital Terrain Model)[44], DSM (Digital Surface Model), and point cloud data available on the Washington Lidar Portal. Resulting data products will include a DEM of Difference (DoD) raster, spatially distributed confidence intervals for the

---

[44] We have referred to gridded bare-earth elevations as a DEM (Digital Elevation Model). The Lidar Portal refers to these datasets as DTMs (Digital Terrain Models). We use DTM here for consistency with the Lidar Portal.

DoD, and a landslide inventory based on analysis of the DoD, DTM, DSM, and potentially other available data products (e.g., NAIP imagery).

The simplest approach for creation of a DoD is to subtract one existing DTM from another with no pre- or post-processing, as shown in Figures 3 and 4 in the main document of this study design. Several studies take this approach (Burns et al., 2010; Coe et al., 2021; Heckmann and Vericat, 2018; Ju, Zhang and Xiao, 2023; Scheidl, Rickenmann and Chiari, 2008). However, errors in the spatial registration in the DTMs translate to errors in calculated elevation changes. This is particularly pertinent for landslide detection because the potential error increases with increasing surface gradient. Several approaches are used to improve spatial alignment between lidar datasets. If stable zones where no change has occurred can be confidently delineated, then one dataset can be shifted relative to the other to minimize differences in those stable zones. These shifts may be performed on the point clouds themselves prior to interpolation of the ground returns to a DTM and again to the resulting DTMs (Cucchiaro et al., 2020). Most of the cited studies employed co-registration of the point clouds prior to creation of the DTMs. If biases persist through stable zones, the DoD can then be warped to minimize those differences (DeLong et al., 2022; Morell et al., 2021).

Co-registration of point clouds is now done almost exclusively using the Iterative Closest Point (ICP) algorithm (Nissen et al., 2012), which shifts one point cloud relative to the other to minimize the distance between points in the two point clouds. These shifts are determined over subsets of the point clouds, centered over known stable points or across a regular mesh spanning stable areas, so that the correction can vary spatially. Several variants of the algorithm have been developed; Scott et al. (2021) evaluate the performance of these variants and provide guidelines for setting the window size needed as a function of point density.

Change detection can be measured from the point clouds directly, without first interpolating elevations to DTM rasters. This may offer some benefits because changes can be measured orthogonal to the ground surface, which can reduce uncertainties in change magnitude on very steep slopes (Bernard, Lague and Steer, 2021). However, change measured in a vertical direction over a regular grid, that is, a DoD, is better suited for subsequent identification of landslides using image segmentation and for analysis of landslide areas and volumes.

Uncertainties in elevation differences recorded in a DoD constrain the minimum change that can be confidently detected. It is appropriate, therefore, to determine that level of detection and limit mapped changes only to those greater than that limit (Lane, Westaway and Murray Hicks, 2003). Several factors influence the degree of confidence to place in measured elevation changes. Some noise arises from the inherent precision of the laser scanners themselves, which is typically measured by the data vendors at control points of known elevation and is generally small relative to other potential sources of error. Errors arise from disparities in the spatial registration of the two DTMs, as discussed above, which can be minimized through co-registration of the point clouds and subsequent warping of the resultant DoD to minimize bias through stable zones.

In rough terrain, such as that where landslides occur, the largest source of error in DTM elevation is the incomplete sampling of the ground surface by the lidar laser pulses (Aguilar et al., 2005; Guo et al., 2010). This becomes particularly acute under dense vegetation where few pulses might actually reach the ground surface (Cățeanu and Ciubotaru, 2021; Hyyppä et al., 2005; Viedma, 2022). Uncertainty in DTM elevations thus varies in response to spatial variability in forest stand characteristics and terrain

steepness and roughness. Uncertainty associated with spatial variability in surface elevation can be estimated from a measure of terrain roughness, as described by Lague, Brodu and Leroux (2013). Roughness at a point is estimated as the deviation of ground-return elevations from a plane fit to all points within a specified radius. If the elevation deviations are normally distributed, then confidence intervals for elevation difference is estimated as

$$LOD_{x\%} = t(x\%)\left(\sqrt{\frac{\sigma_1(d)^2}{n_1} + \frac{\sigma_2(d)^2}{n_2}} + \varepsilon\right)$$

1.

where $LOD_{x\%}$ is the level of detection at $x\%$ confidence, $\sigma_1(d)$ and $\sigma_2(d)$ are the variance of elevation deviations for DTM$_1$ and DTM$_2$ within diameter d, n_1 and n_2 are the number of points within diameter d, ε is the registration error remaining after co-registration, assumed spatially uniform, and $t(x\%)$ is the two-sided t statistic at a confidence level of $x\%$ and a degree of freedom given by

$$DF = \frac{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}{\left(\frac{\sigma_1^4}{n_1^2(n_1-1)} + \frac{\sigma_2^4}{n_2^2(n_2-1)}\right)}$$

2.

The diameter d of the planar surface fit to the ground-return points determines the length scale over which surface roughness is measured. It must be large enough that a sufficient number of points are included to first fit the plane (at least 3 points) and then determine the variance. Lague et al. (2013) suggest a value of d equal to 10 times the average point spacing with at least 5 points, determined for the DTM with lowest point density. Hence, the point density determines the length over which roughness can be measured and will vary across a DTM (Petras et al., 2023).

A tentative workflow might thus include the following steps:

1. Create a DoD with the existing DTMs available on the Lidar Portal. Use these to delineate stable areas where the elevation difference is spatially consistent and small.
2. Filter point clouds to identify ground returns, the appropriate filtering algorithm will need to be determined (e.g., Meng, Currit and Zhao, 2010); identify and remove outliers. This will provide a consistent method for creation of all DTMs.
3. Use ICP over a mesh defined on the stable zones to co-register the two point clouds, with mesh size determined by the ground-return density (Scott et al., 2021) and using the highest quality of the two lidar acquisition (generally the newer one) as the reference.
4. Interpolate each point cloud to a DTM. The appropriate interpolation algorithm will need to be determined (e.g., Habib et al., 2020; Szypuła, 2017).
5. Create a DoD from the two DTMs, or alternatively using the M3C2 algorithm (Lague, Brodu and Leroux, 2013) but with a vertical measurement direction (e.g., Guilinger et al., 2023) with control points over a regular grid.
6. Evaluate the DoD on a piece-wise basis to remove bias over the delineated stable zones (e.g., DeLong et al., 2022).
7. Calculate LOD at 95% confidence using Equations 1 and 2 at each grid point over the DoD, with ε estimated from the remaining DoD values over the delineated stable zones. This gives a spatially variable level of detection.
8. Use image segmentation of the DoD to delineate zones of elevation loss and gain. Filter these by topographic attributes (from the earlier, pre-landslide DTM) appropriate for landslide sites (e.g.,

slopes > 40%), by the juxtaposition of upslope elevation loss and downslope elevation gain (e.g., Burns et al., 2010), and potentially by vegetation changes apparent from the two DSMs and from NAIP or other imagery (e.g., Scheip and Wegmann, 2021) aligned with the lidar acquisition dates. This provides a starting point for object-based image analysis to delineate landslide sites and runout zones (e.g., DeLong et al., 2022).

Open-source projects offer a variety of tools and code libraries for lidar point-cloud analysis (lastools.github.io, portal.opentopography.org/tools/listTools, pdal.io/en/latest/, r-lidar.github.io/lidRbook, laspy.readthedocs.io/en/latest), including co-registration (pointclouds.org, www.cloudcompare.org, www.cvlibs.net/software/libicp) and lidar differencing for change detection (gcd.riverscapes.net, Las2DoD, www.whiteboxgeo.com/manual/wbt_book/available_tools/lidar_tools.html). We anticipate that available software resources will not meet all our needs and that we will develop new code libraries and packages using R, python, Fortran, and C as needed. These will also be licensed as open source and publicly available, e.g., via github.

We fully expect this workflow to evolve as the project progresses. There may be opportunities for co-registration by flight line (e.g., Piovan et al., 2023), depending on the parameters included in the LAS files available from the Lidar Portal; given the poor ground-return density of the older lidar datasets, we may need to experiment with a variety of filtering and interpolation schemes (e.g., Cai et al., 2023; Chen et al., 2020; Hui et al., 2021; Meng, Currit and Zhao, 2010; Montealegre, Lamelas and de la Riva, 2015; Silva et al., 2018); and given the rate at which new studies are being published, the available algorithms will evolve.

2 Susceptibility

When we wrote the study design, we pointed to specific types of analysis methods, but provided no discrete examples. The literature on landslide susceptibility presents a multitude of options: in a recent literature review on empirical methods for landslide susceptibility, Lima et al. (2022) examined 2585 publications. Yet, no standard procedure has emerged from all that effort. This lack of standardization likely reflects the diversity of data types available and the diversity of objectives for susceptibility analyses. Despite this wealth of studies, we are not aware of any published examples that closely match the type of analysis required here (of course, we have not reviewed everything that has been published). To restate what was said in Section 1.1, what we need is a methodology with the following two capabilities:

1. Calculate the proportion of all landslides with potential to impact downslope public resources or threaten public safety that originate within some specified portion of a defined study area. This capability would enable ranking of delineated landforms in terms of the proportion of all "delivering" landslides that originate within each landform type. Here, "delivering" refers to those that impact public resources, e.g., stream channels. Note that to calculate proportions across diverse landform types, we must first quantify spatial variability in landslide density. This also enables ranking of landforms in terms of landslide density averaged over landform area. This is a more typical measure of susceptibility, but we expand on that here to look at the spatial density of *delivering* landslides.

2. Identify those public resources potentially impacted by upslope landslides and rank them by probability of occurrence. This provides a means of explicitly identifying those public resources

potentially at risk. This ranking can then be associated with the upslope source areas where the impacting landslides could originate.

To meet these capabilities requires, for each potential landslide initiation site, calculation of both probability of landslide initiation and probability of runout to all points downslope. Because multiple initiation sites may generate landslides that all cross common points downslope, these probabilities need to be combined to give the probability that any upslope initiation site will generate a landslide that impacts a public resource.

Work by Miller and Burnett offers a template for such an analysis (Burnett and Miller, 2007; Miller and Burnett, 2007, 2008). That work is dated however, data sources and analysis techniques have evolved significantly in the last 15 years. . Following is an updated example using work done for the Oregon Private Forest Accord (PFA) steep-slopes analysis described in Chapter 3 and Appendix B of the PFA Report (Oregon Department of Forestry 2022). We will use those analyses here as examples of how available data can be used to derive terrain-attribute values and apply statistical methods to generate susceptibility maps. The PFA steep-slopes analysis used the landslide inventory published by the Oregon Department of Geology and Mineral Industries (DOGAMI) in Special Paper 53 (Burns, Franczyk and Calhoun, 2022). This inventory included both initiation points and debris-flow runout tracks.

In the study design, we refer to variables that can serve as predictors of landslide density and runout extent as terrain elements. These are measurable attributes of the terrain that are used as inputs to build an empirical model. Although our focus is on landform objects, the analysis methods proposed are pixel based. This is for several reasons:

- Pixel-based analyses do not rely on an *a-priori* decision of what comprises a landform.

- Delineation of a landform object is done through segmentation of pixel-based information, so results of a pixel-based analysis can be integrated within existing landform objects or used in a segmentation scheme to delineate new landform objects.

- Pixel-based analyses can accommodate a great range of predictors over a large range of spatial scales. Landform objects are constrained by the length scales and attributes by which they are identified. For example, runout analyses require point-by-point or cell-by-cell determination of gradient and topographic confinement along a travel path, available with a pixel-based analysis.

The following sections thus address pixel-based analyses. The regular grid of elevation point values from a DEM provides the spatial frame. Rather than "pixel", we will refer to DEM cells. Each square cell is centered over a grid point, or in some GIS implementations such as ArcGIS, a cell spans the area between four grid points.

2.1 Landslide Susceptibility

Susceptibility is quantified in terms of probabilities and the proportions calculated from those probabilities. We can calculate probability from two perspectives:

- Spatial probability; the probability that any DEM cell contains a known landslide initiation site or runout track. This probability can be expressed as landslide density.

- Temporal and spatial probability; the probability that any cell did or will experience a landslide initiation or traversal over some specified interval of time. This probability can be expressed as landslide rate.

Ability to make these calculations depends on the data available. A landslide inventory enables calculation of spatial probability. If the inventory includes all landslides in the study area over a known period of time, then the landslide density (number per unit area) can be divided by that time span to estimate rate (e.g., UPSAG, 2006). This is a poorly constrained estimate, because rate is a function of the time span involved, or more specifically, of the sequence of storms experienced over that time. If, however, landslide density is related to some measurable characteristic of storms for a site (e.g., Marc et al., 2019; Reid and Page, 2002), then landslide rate can be inferred by convolving the storm-dependent rate with the probability distribution of the storm attributes associated with landslide density (e.g., an attribute such as rainfall intensity, Turner et al., 2010). The primary objective of this study is to better constrain spatial probability as a function of measurable landscape attributes, i.e., the terrain elements discussed in the study design. This is what is needed for an assessment of susceptibility and will be examined in the examples to follow. However, recognizing the confounding effects of spatial and temporal variability in storm characteristics on landslide density, we also want to see if we can resolve relationships between some (as now unspecified but to be explored) storm characteristics and landslide density. The availability of gridded precipitation data offers that opportunity and there are examples in the literature to guide initial efforts (e.g., Marc et al., 2019; Thomas et al., 2023; Turner et al., 2010).

2.1.1. Spatial Probability of Initiation

2.1.1.1. Terrain Elements as Predictors of Landslide Density

Landslide density derived from a given inventory of landslide initiation sites translates directly to the probability that any DEM cell within the study area contains or is within a mapped initiation site. We seek to resolve spatial variability in landslide density and then to define that density as a function of some set of spatially distributed terrain elements. There is a large array of possibilities for the choice of terrain elements: Lima et al. (2022) counted 116 different predictors used in the 2585 studies they reviewed. We seek a parsimonious set. Inclusion of predictors unrelated to landslide density will introduce noise, potential bias, and increased danger of overfitting a model. We also do not want to exclude any potentially useful predictors. The choice of terrain elements to serve as predictors can be guided by theoretical understanding of the processes of soil failure. That choice is also constrained by the data available.

The physics of soil failure is complex, yet simple physically based models prove remarkably successful for explaining and anticipating conditions for failure. The infinite slope approximation (Skempton and deLory, 1957) forms the basis for models such as STALSTAB (Montgomery and Dietrich, 1994) and SINMAP (Pack, Tarboton and Goodwin, 1998). These models identify hillslope gradient, weight of soil (bulk density integrated over depth of soil), saturation depth, and soil strength (friction angle and cohesion, including the apparent

cohesion associated with the network of plant roots) as primary controls on soil stability. Which of these attributes can be measured or inferred from remotely sensed data?

Soil weight varies with soil depth; field studies find that soil depth varies systematically with hillslope gradient (e.g., Dietrich et al., 1982) and curvature (e.g., Patton et al., 2018). Depth of saturation varies with upslope contributing area and rainfall intensity. Contributing area increases over time as water infiltrating the soil flows downslope. The rate of water flux through saturated soil varies with hillslope gradient and saturated hydraulic conductivity, so contributing area can be estimated from the upslope distribution of hillslope gradient and aspect. These models thus suggest three topographic attributes as potentially useful predictors of landslide density: hillslope gradient, curvature, and contributing area. These are all measurable from a DEM.

Another potentially important control on soil depth is the frequency of landsliding. Shallow landslides typically expose the underlying bedrock. If the landslide evolves into a debris flow, soil may be scoured over portions of the downslope debris-flow path. Subsequently, soil and organic debris accumulate in the landslide scar and along the debris-flow corridor (Dietrich et al., 1982; May and Gresswell, 2003). Soil depth in these locations is thus a function of the time since the last landslide or debris flow. Landslides can occur at variable locations up or downslope; an upslope-positioned landslide can scour soil along portions of its travel path, thus reducing landslide potential through that downslope zone (Dunne, 1991). Any downslope zone may have many potential upslope landslide sources; the frequency with which it is scoured by debris flows is thus a function of the number of upslope sources. Without having run a susceptibility model, we have no measure of upslope landslide potential, but in landslide-prone terrain, the number of potential upslope sources will increase with increasing total drainage area. Thus, total drainage area, measured to the drainage divide, offers another DEM-derived terrain-element to include as a candidate predictor in building an empirical model for landslide initiation.

Physical models to calculate a factor-of-safety for identifying zones prone to landslide initiation require a determination of pore pressures exerted by water flowing through the soil layer. A variety of approximations and assumptions are employed to develop mathematical descriptions of soil water flux. Montgomery and Dietrich (1994), for example, assume steady-state rainfall with water flow through the soil parallel to the ground surface, which essentially assumes an infinite-duration rainstorm. Wu and Sidle (1995), Iida (1999), and Borga, Fontana and Cazorzi (2002) used a kinematic-wave (quasi-steady-state) approximation for the flux of infiltrating rainwater through the soil layer, which accounted for the increasing upslope area contributing shallow groundwater flow to a hillslope location during a rainstorm. Iverson (2000) critiqued these approaches and developed a transient pressure response based on approximations of the Richards equation that removed constraints on flow direction. Each additional detail can improve the degree to which a model represents reality, but also adds physical attributes required to apply the model. We lack information to constrain these attributes (soil depth and transmissivity) directly, but we can use a kinematic-wave approach to look at how topography can influence contributing area over time, which serves as a proxy representing saturation depth and pore pressure. To delineate contributing area to a single DEM cell for a specified

time duration, we can trace flow upslope using the Darcy velocity $v = K\sin\theta$, where K is saturated hydraulic conductivity and θ is angle of the ground surface, assuming surface parallel flow. We assume uniform K and, for the examples below, set it to a value of one meter per hour. The value is arbitrary, but one m/hr is representative of many soils (Gupta et al., 2021) so that the specified duration has some physical meaning. Flow directions are calculated using D-infinity (Tarboton, 1997). Flow is traced DEM grid point to grid point until the transit time, determined from the Darcy velocity, equals the specified duration. Spatial variation in the calculated contributing area may correlate in some way with the spatial variation in soil pore pressures and, consequently, with variation in the spatial density of landslide initiation points. If so, the calculated contributing area might provide an informative predictor for an empirical model. The predicted spatial pattern of contributing-area size varies with the specified duration. Short durations produce a more uniform pattern; as duration increases, the highest values are concentrated in convergent topography, i.e., the axis of hollows, as illustrated in Figure 5 of the study design.

We have no data with which to measure soil strength directly over regional scales. However, friction angle and cohesion vary with soil texture and mineralogy. These attributes vary depending on the rock types from which a soil originates and with mapped soil types. Landslide density may, therefore, vary with mapped lithology and soil types (e.g., Swanson and Dyrness, 1975). Ability to resolve any associations of rock or soil type with landslide density will depend on the rock and soil types included in areas where landslide inventories are made. Geologic mapping covering the entire state is available at a scale of 1:100,000 (https://www.dnr.wa.gov/programs-and-services/geology/publications-and-data/gis-data-and-databases). This database is the compilation of many separate geologic mapping studies and aggregates lithologic types across those studies into consistent categories. There are still 184 different lithologic groupings. This project cannot collect landslide inventories with a sufficient number and spatial distribution of landslides to include all of these, so these groupings need to be aggregated to a considerably greater degree. For the PFA analysis, I used five groupings: sedimentary rocks, volcanic rocks, volcaniclastic rocks, igneous+metamorphic rocks, and unconsolidated deposits. What is appropriate for this study will depend on where inventories are collected. SSURGO provides soil type mapping at a scale of 1:24,000. The database includes soil properties that may correlate with landslide potential. Strauch et al. (2018), for example, used reported grain-size distributions to estimate the range of friction angles to associate with different soil types. Effective cohesion from roots likely varies with the age, size, species, stem density, and health of the trees in a forest stand (e.g., Schmidt et al., 2001). Stand age provides a potential, although incomplete, proxy that has been found to correlate with landslide density (e.g., Miller and Burnett 2007; Turner et al. 2010). Other broad stand characteristics, such as "sparse, open, semi-open, and closed" (Goetz, Guthrie and Brenning, 2015) have also been correlated to landslide density. The LEMMA project and DNR forest inventory datasets provide GIS data on stand structure and age that can be used for this project. We have not listed specific geologic, soils, or forest-stand attributes to examine. Those choices depend on the range of rock, soil, and forest-stand types available across the areas covered by the inventories collected for the project. We expect that these choices will evolve as different possibilities are explored with the data analyses.

2.1.1.2.   Landslide Density

Given a set of potential predictors, we need a method to relate landslide density to the predictor values. As listed in the main document, there are several options. I used logistic regression for the PFA analysis. As with our analysis here, those results are intended to guide field operations. The data analysis should lead directly to improved understanding of how landform attributes are associated with landslide susceptibility. The results of logistic regression are relatively easy to interpret and understand, so logistic regression serves as a useful starting point for this project. However, the analysis should include at least one other modeling approach. Classification schemes with which probability of occurrence can be estimated use a variety of ways to look at how the proportion of landslide and nonlandslide locations in the inventory are distributed across the data space. For example, logistic regression characterizes that distribution using a linear equation for the odds; decision-tree-based analyses parse the data space into variably-sized chunks. It is unlikely that any method can characterize that distribution totally accurately, so it is worthwhile to see how the performance of different methods compares. A useful starting point is to look at how landslide density changes across the range of individual predictors.

The cumulative area and the cumulative number of landslides can be plotted as a function of a single predictor, as shown below using results of the PFA analysis for gradient, with gradient here equal to the tangent of the hillslope angle (i.e., rise/run).
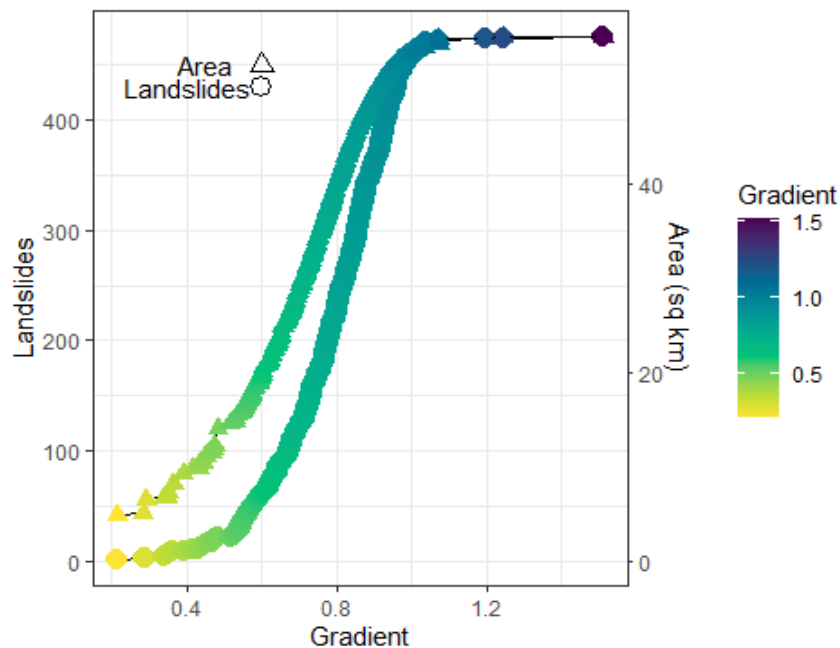
Figure 1: Cumulative area and cumulative number of landslides vs gradient from the PFA analysis.

Each point plotted along the curve corresponds to one landslide in the inventory. Taking the area- and landslide-gradient values at each point, another curve showing the cumulative number of landslides versus area can be plotted:
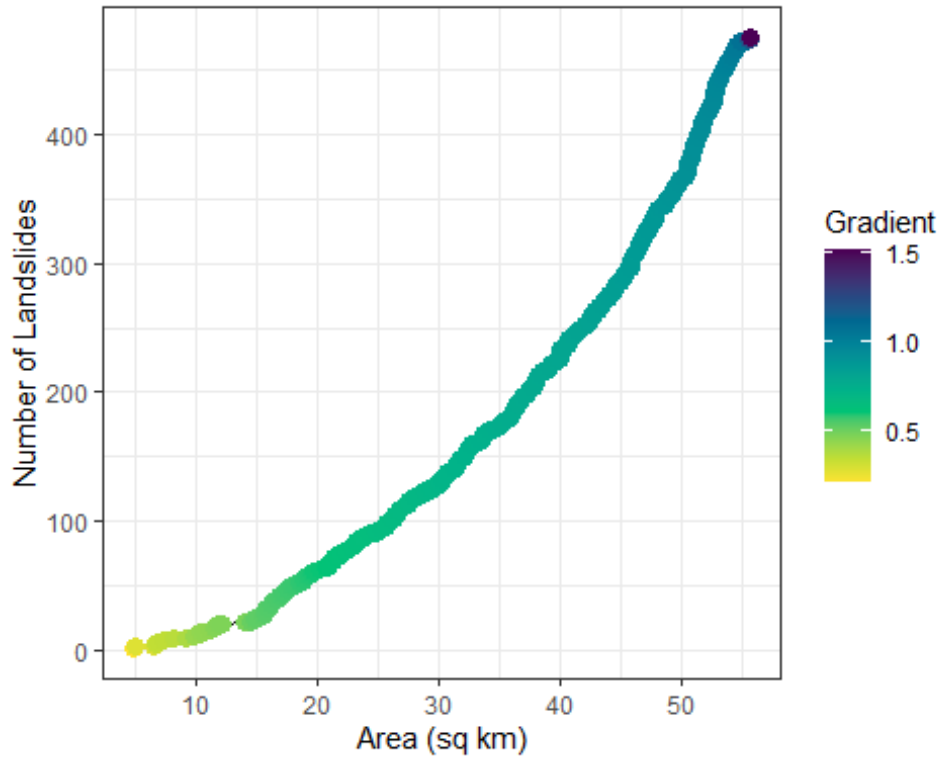

Figure 2: Cumulative number of landslides vs cumulative area; ordered by increasing gradient.

Landslide density is defined as $\Delta\ landslides/\Delta\ area$, so the slope of this curve gives landslide density. Each location along that curve corresponds to a value of gradient, so these two plots together can be translated to landslide density as a function of gradient.
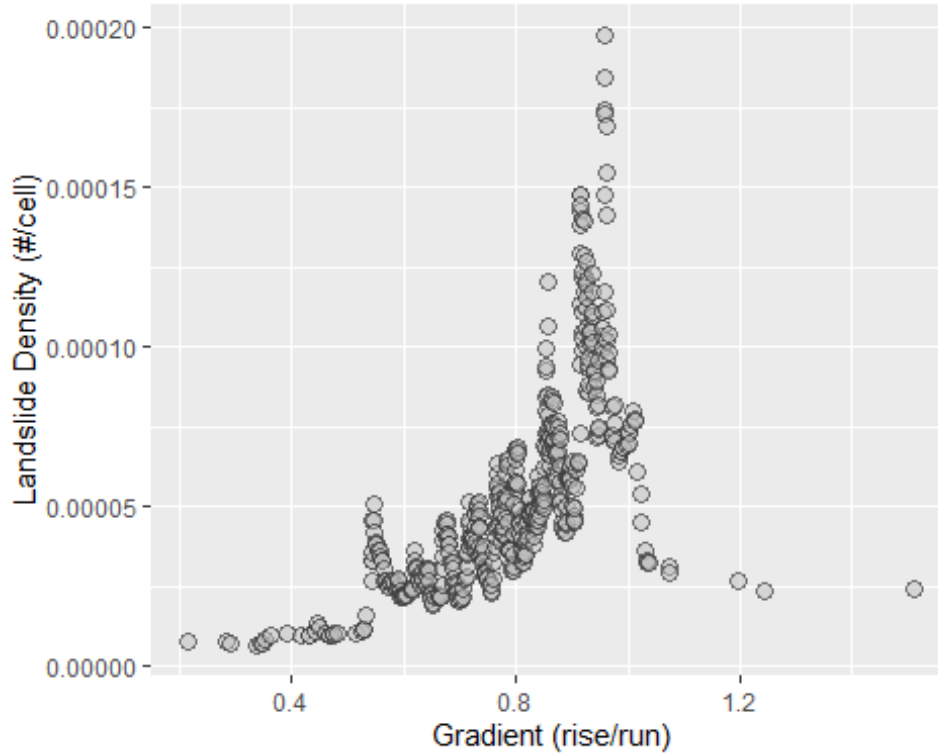
Figure 3: Landslide density vs gradient.

Density in [Figure 3](#) was calculated over a window centered at each landslide point and extending five points up and down the curve. There is considerable scatter in the density values, but also a clear trend indicating very low density at gradients below 55%, increasing density to a peak near 100%, and a very rapid reduction to low values above 100%. Note that density was plotted as number of landslides per DEM cell (this analysis used a 2-meter DEM grid spacing, so four-square-meter cells). This gives the empirical probability of encountering a mapped landslide initiation point in any DEM cell. We want a mathematical expression that will mimic this trend.

We use logistic regression here to illustrate this concept. For a set of predictor values x for some DEM cell, logistic regression expresses the probability that the cell contains a mapped landslide initiation point $p(x)$ as

$$p(x) = \frac{1}{1 + e^{\beta x}} \qquad (1)$$

where β is a vector of empirical coefficients. The ratio of the probability that the cell contains an initiation point and the probability that it does not gives the odds:

$$\frac{p(x)}{1 - p(x)} = e^{\beta x} = odds \qquad (2)$$

The logarithm of the odds is a linear equation in x:

$$log\left(\frac{p(x)}{1 - p(x)}\right) = \beta x = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n = log(odds) \qquad (3)$$

For a single predictor, $log(odds) = \beta_0 + \beta_1 x$. The density shown in Figure 3 gives probability; the logarithm of the odds is shown below:
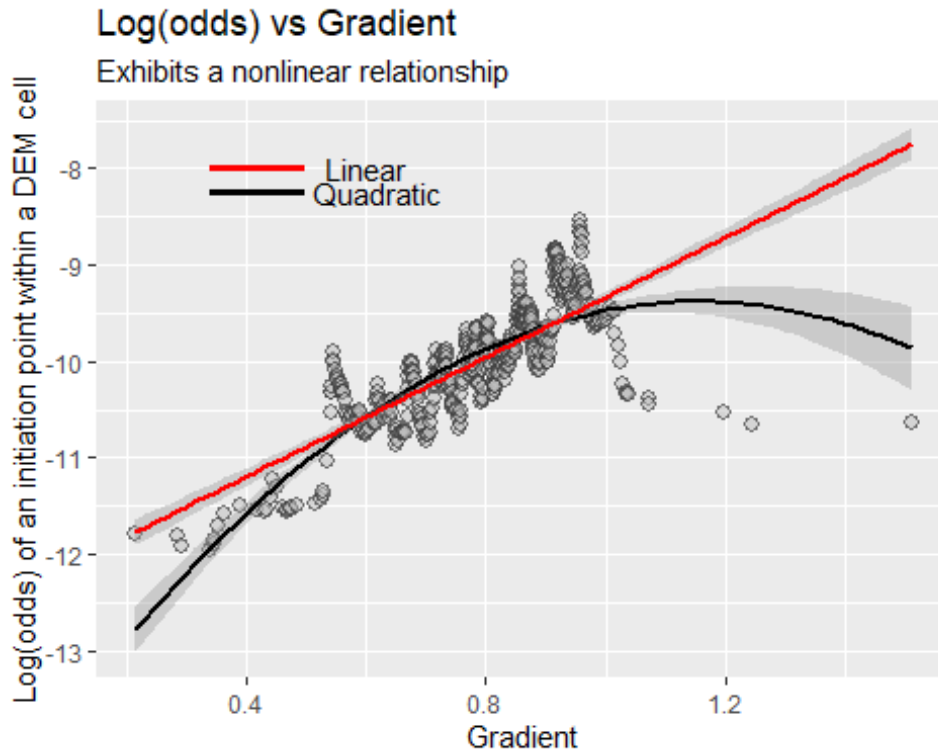

Figure 4: Log(odds) of landslide density versus gradient

The red line shows a linear fit to these values. In this example, a logistic regression model trained on gradient alone would match observed landslide densities fairly well up to a gradient of 100% and over-predict density at higher gradients. Other types of models might better mimic the observed behavior. For logistic regression, we can remedy this lack-of-fit somewhat by adding an x^2 term: $log(odds) = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2$, as shown with the black line in Figure 4.

In this dataset, curvature and contributing area exhibited similar patterns:
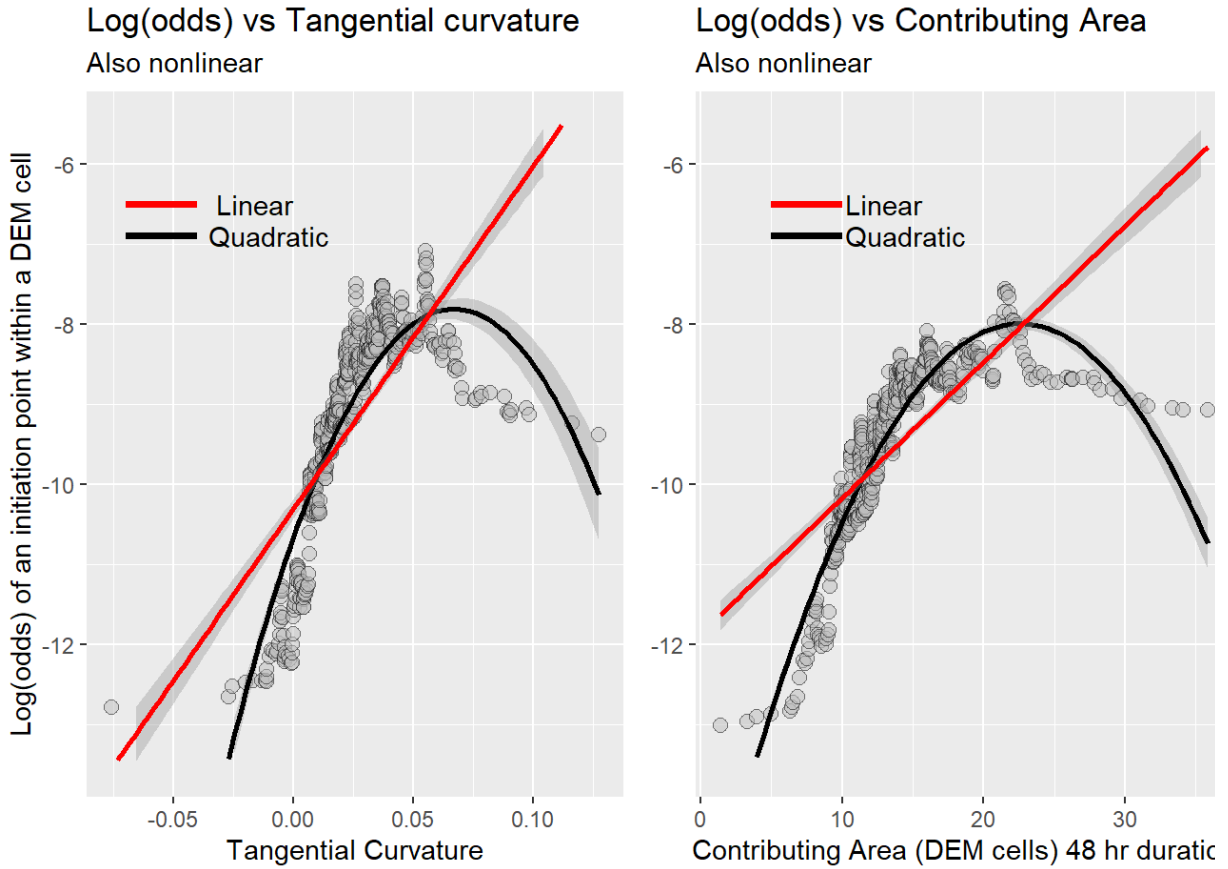
Figure 5: Log(odds) of curvature and contributing area

In the examples above, we looked at how landslide density varies over the range of single predictors. At any predictor value (e.g., at gradient = 0.80), the landslide density indicates the proportion of the area, (i.e., the proportion of DEM cells), with gradients within a small increment of that value that include initiation points. We can look at those proportions directly using density plots. These show the proportion of all initiation points and of all non-initiation-point cells within our sample as a function of predictor value.

Distribution of Predictor Values
for mapped initiation sites and all other areas
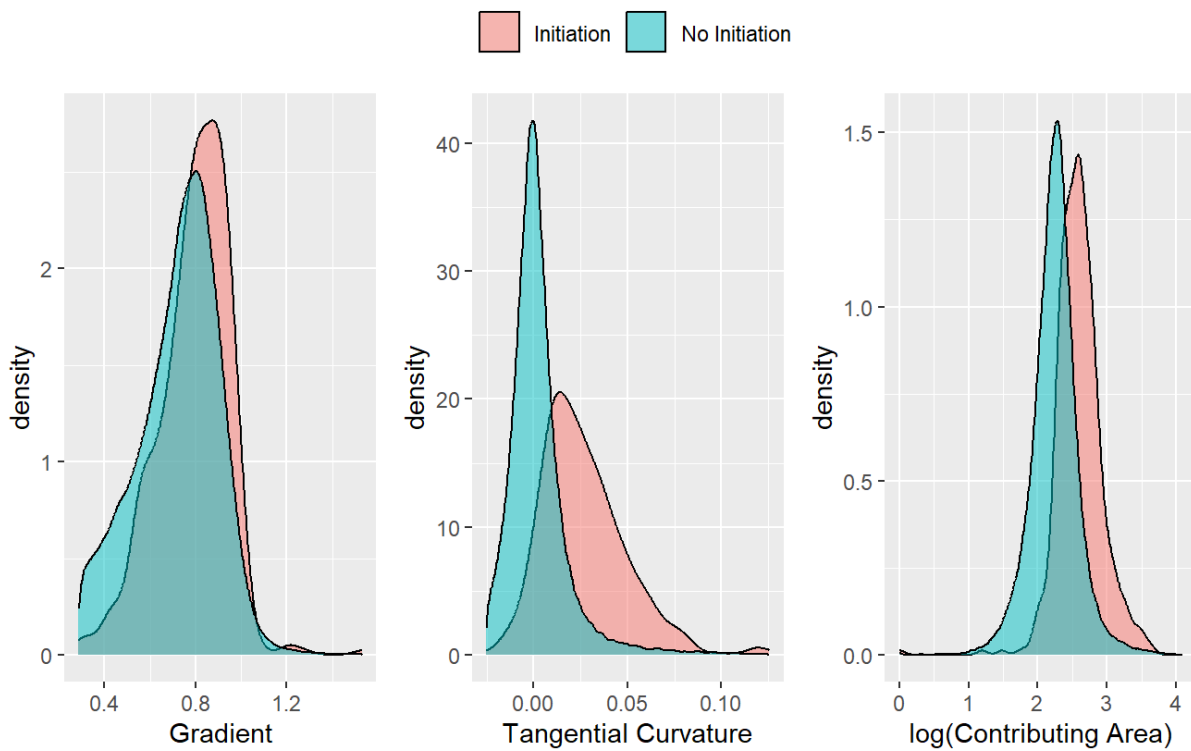
Figure 6: Density plots show how the distribution of topographic attributes varies between landslide sites and the remainder of the study area where landslides were not observed, but are assumed could occur.

For these examples, the portion of the study area examined was constrained to include only predictor values within the range observed for mapped initiation points. Any model should produce a probability of zero for values outside the range of observed values, so we want to focus on those areas where landslides could occur. Hence, there is overlap between locations with and without mapped initiation points across the entire range of all predictor values. The degree to which the two distributions differ determine the degree to which conditions associated with initiation sites differ from conditions without initiation sites. The degree of these differences will influence the standard errors and associated z and p values calculated for model coefficients. Standard errors with z and p values are often used as measures of model performance, but here these values indicate the degree to which a predictor helps to distinguish spatial variation in landslide density. Large z and p values are not a reason to eliminate predictors here, as is typically done in stepwise feature (predictor) selection. We want to include all predictors that provide some information.

So far, we have looked at how landslide density is distributed when all data is projected onto one dimension of the predictor data space. It is challenging to visualize this distribution in the multidimensional data space, but a projection onto two dimensions is informative. Here is the

distribution of mapped initiation points and all the remaining DEM cells over gradient, contributing area, and curvature.
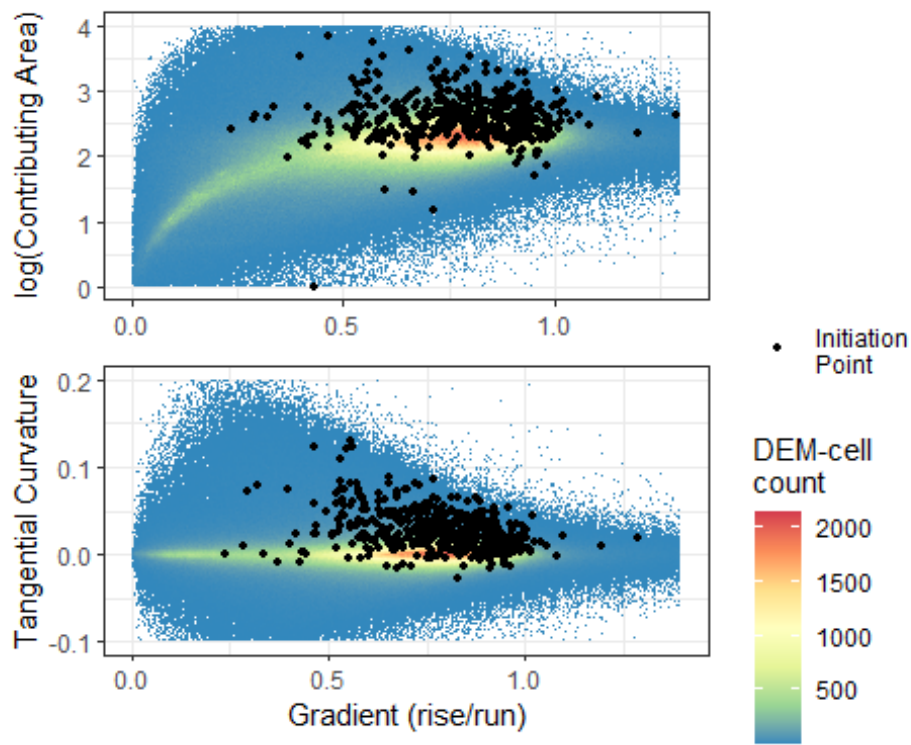

Figure 7: Distribution of DEM-cell values across two dimensions of the predictor data space. Colors indicate the spatial density of non-initiation DEM-cell values, points indicate mapped landslide values.

In multiple dimensions, the distributions of initiation and non-initiation sites may be better distinguished than with the one-dimensional view previously. Different classification algorithms use different tactics to map out these distributions and estimate how the proportions of each (the landslide density in this case) are distributed at all points in the data space. In two dimensions, a logistic regression model without quadratic terms fits the density as a plane. This is clearly a poor representation of that surface; a quadratic surface would be a better approximation. That requires both quadratic (squared) terms and interaction terms. The point of all this is to demonstrate the importance of looking at the data and determining how different modeling strategies will deal with it.

The mapped initiation points from a landslide inventory typically involve a very small portion of the entire study area. In the examples above, we had about 500 DEM cells containing mapped initiation points and 15 million cells with no initiation points. This is an extremely unbalanced sample. The plots above were made using all 15 million of these DEM cells, but such an unbalanced sample could lead to rounding errors and long processing times for classification algorithms. Hence, studies using classification models for landslide susceptibility use a sample of points selected randomly from the non-initiation DEM cells. The calculated probabilities then

depend on what sample balance was used. We are interested in the spatial distribution of probabilities, not the magnitudes, so this is not a problem as long as the sampled points adequately represent conditions across the entire study area.

2.1.1.3. Model Performance

We need measures of model performance for several tasks.

- To compare different algorithms (e.g., logistic regression versus random forest).

- To compare models built with different sets of predictors.

- To determine how well the model can reproduce what was observed; i.e., how well can it reproduce the spatial distribution of mapped landslide initiation points.

- To estimate how well it will predict the spatial distribution of landslide densities when extrapolated to new areas.

To accomplish these tasks, there are five things measures of model performance need to quantify:

3. How well the sample of nonlandslide points characterizes the joint distributions of predictor values across the entire study area.

4. How well the chosen model algorithm characterizes the distribution of landslide densities within the data space defined by the predictors.

5. How well the choice of predictors resolves spatial variations in landslide density.

6. How sensitive model results are to the predictor values, and

7. Geomorphic plausibility.

Typical measures of model performance include receiver operating characteristic ([ROC](#)) curves and the consequent area under the ROC curve (AUC), the precision recall curve (PRC, Yordanov and Brovelli, 2020) and the Brier score, among others. All of these are based on some measure of classification success. Given that for our entire sample, we might expect one out of every 30,000 DEM cells to contain a landslide initiation point and there is complete overlap of the distributions of initiation and non-initiation sites within the predictor data space (because we truncated that space to include only values within the range of mapped landslide initiation points), measures based on classification success may be difficult to interpret. Here is an interpretable alternative that can be used for the measures listed above.

The "success-rate" curve was introduced by Chung and Fabbri (2003). To construct a success-rate curve, we rank DEM cells by the modeled probability that they contain a landslide initiation point. We then plot the proportion of mapped landslides versus the proportion of area, ranked by modeled probability. With the calculated probability, we can also plot the proportion of landslides predicted by the model. Classification algorithms preserve the marginal probability of the response (predicted) variable, here landslide density. Integrating density (modeled probability) over area, the same here as summing over DEM cells, gives the number of

landslides in the training data. At least it should, if the model is working properly. If the nonlandslide area has been subsampled, which is almost always the case, then the predicted probabilities will be higher than the actual values. This is not a problem, because we plot the *proportion* of modeled landslides, not the number. Within any increment of modeled probability, integrating the modeled probability values over the area included within those values will give the number, translated to proportion by dividing by the total, of landslides observed within that increment. The curve of proportion of landslides versus proportion of area made by summing ranked probability values over all DEM cells should match exactly the curve for the observed landslides. These curves can be used to examine the first three measures listed above.

The curves below show results for a logistic regression model using only gradient as a predictor. Each point corresponds to one landslide in the inventory. The cumulative plots in the left panel show how study-site area and mapped landslides are distributed across the range of modeled probability. The success-rate curve in the right panel is built using the proportion of area and proportion of landslide values for each landslide point. If gradient provided no information about the spatial distribution of landslide points, or if the distribution were uniform, the points in the success-rate curve would fall along the diagonal. The degree to which the points curve toward the lower-right corner indicates the degree to which the model resolves spatial variation in landslide density. Smaller area-under-the-curve indicates better resolution.
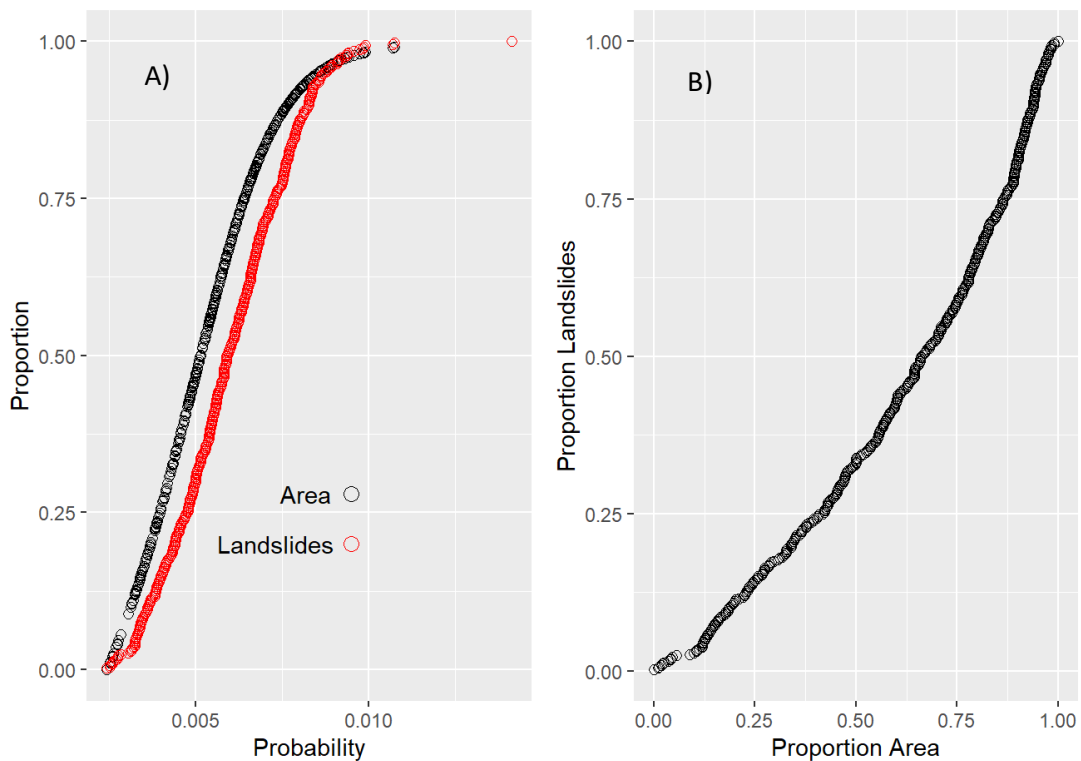

Figure 8: A) Cumulative distributions of DEM area and inventoried landslides ordered by modeled probability of initiation. B) The resulting success-rate curve.

The plot below shows success-rate curves for four logistic regression models:

8. Gradient only

9. Gradient + tangential curvature

10. Gradient + tangential curvature + log(contributing area)

11. Gradient + gradient^2 + curvature + curvature2 + log(CA) + log(CA)2 + log(total contributing area) + log(total contributing area)2 + gradient*curvature + gradient*log(contributing area) + curvature*log(contributing area)

The fourth model includes both total contributing area and quadratic terms. Each additional predictor improves the model's ability to resolve spatial variability in landslide density, but the amount of increased resolution decreases with each additional predictor. Indeed, although model 4 includes eight more terms than model 3, the success-rate curve indicates little improvement in resolution of spatial variation in landslide density.
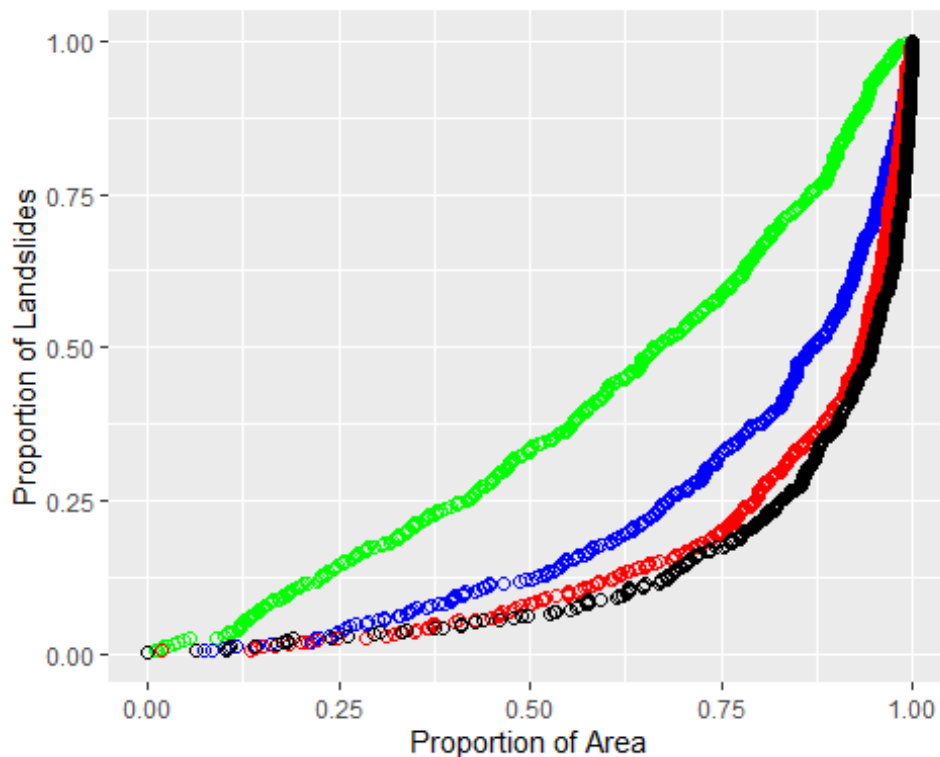
Figure 9: Success-rate curves for four logistic regression models with different sets of predictors.

At any point along the success-rate curve, the tangent to the curve gives landslide density normalized by the mean density:

$$\frac{d(ProportionLandslides)}{d(ProportionArea)} = \frac{(\Delta Landslides/TotalLandslides)}{(\Delta Area/TotalArea)} = \frac{\rho_{Obs}(x)}{\rho bar}$$

Here ρ_{Obs} indicates the "observed" density indicated by the tangent to the empirical success-rate curve. For a well-performing model, this value will equal the modeled probability. The degree to which the observed density and the modeled probability match provides a measure of how well the set of predictors, the sampled values, and the model algorithm represents actual spatial variability in landslide density. The four plots below compare the modeled probability and the "empirical" landslide density found from the tangent to the success-rate curve. The tangent was estimated by fitting a quadratic across 11 points. Each plot shows results for a different number of sampled non-landslide points, i.e., different sample balances. The sample balance affects the magnitude of the modeled probability, so the modeled values were normalized by the number of landslides in the inventory divided by the cumulative sum of all modeled DEM-cell probabilities.
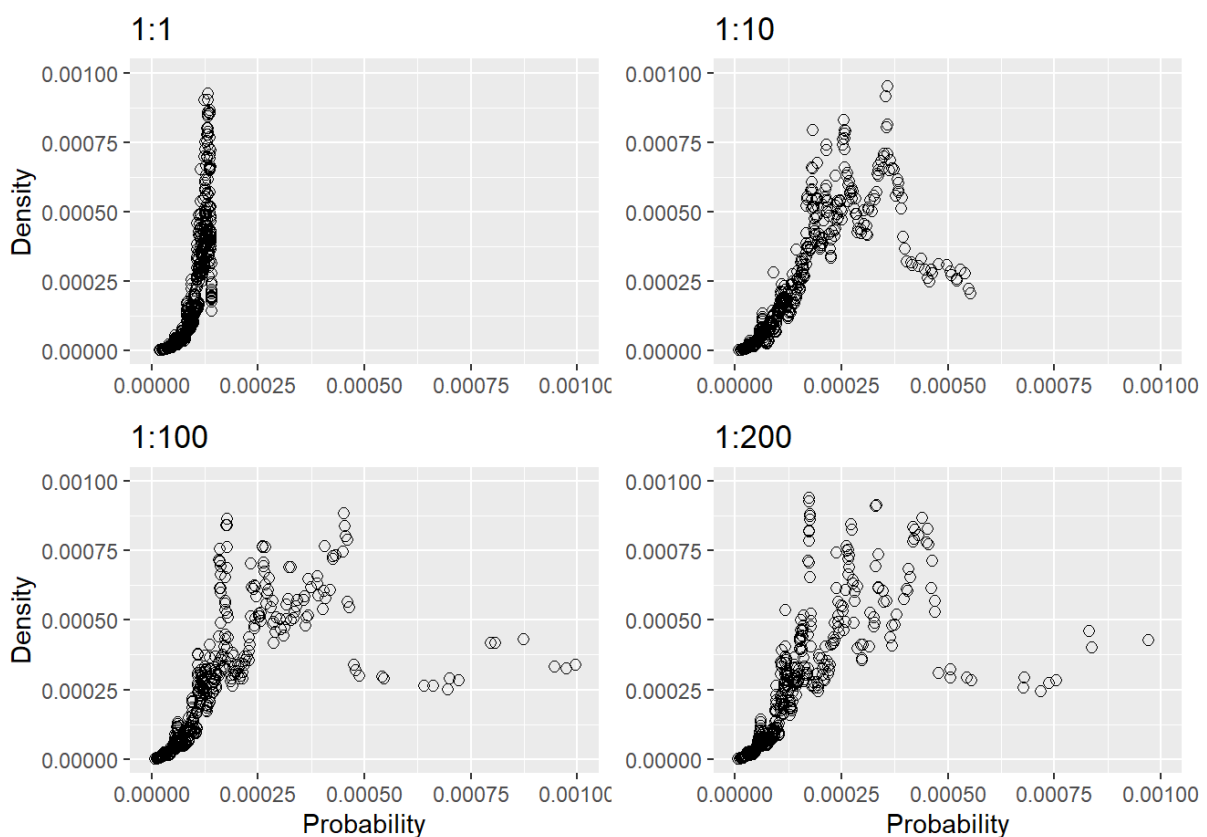

Figure 10: Landslide density versus modeled probability, linear model terms.

For a well performing model, the points should plot along a diagonal line with the empirical points equal to the modeled probability. We expect scatter about the diagonal; the inventory provides only a partial sample of the total population of potential landslide locations. For a balanced sample (equal number of landslide and nonlandslide elements in the training sample), the points follow the diagonal only at the very beginning and then the empirical values are much larger than the modeled. The success-rate curve in Figure 9 shows that the area associated with low landslide densities is large compared to the area associated with high

landslide densities. The balanced sample has insufficient points in the high-density zones to adequately characterize the distribution of predictor values in those zones. The model was producing probabilities near 1.0 (normalized in the plot above). This is implausible: for even the least stable zones, most sites will not have landslides in the inventory, in part because many will not have sufficient accumulated soil since the last time they failed. A modeled value near 1.0 indicates a lack of nonlandslide sample points in these zones. With larger samples, the points approach, but never quite reach, the diagonal and, at the largest modeled probabilities, the empirical values fall off to lower values. This reflects the inability of the linear terms in this model to match the nonlinear shape of the response term (landslide density) in the data space defined by the predictors, as illustrated Figure 4 and Figure 5. The next set of plots compares the modeled probability and empirical density for model 4 above.
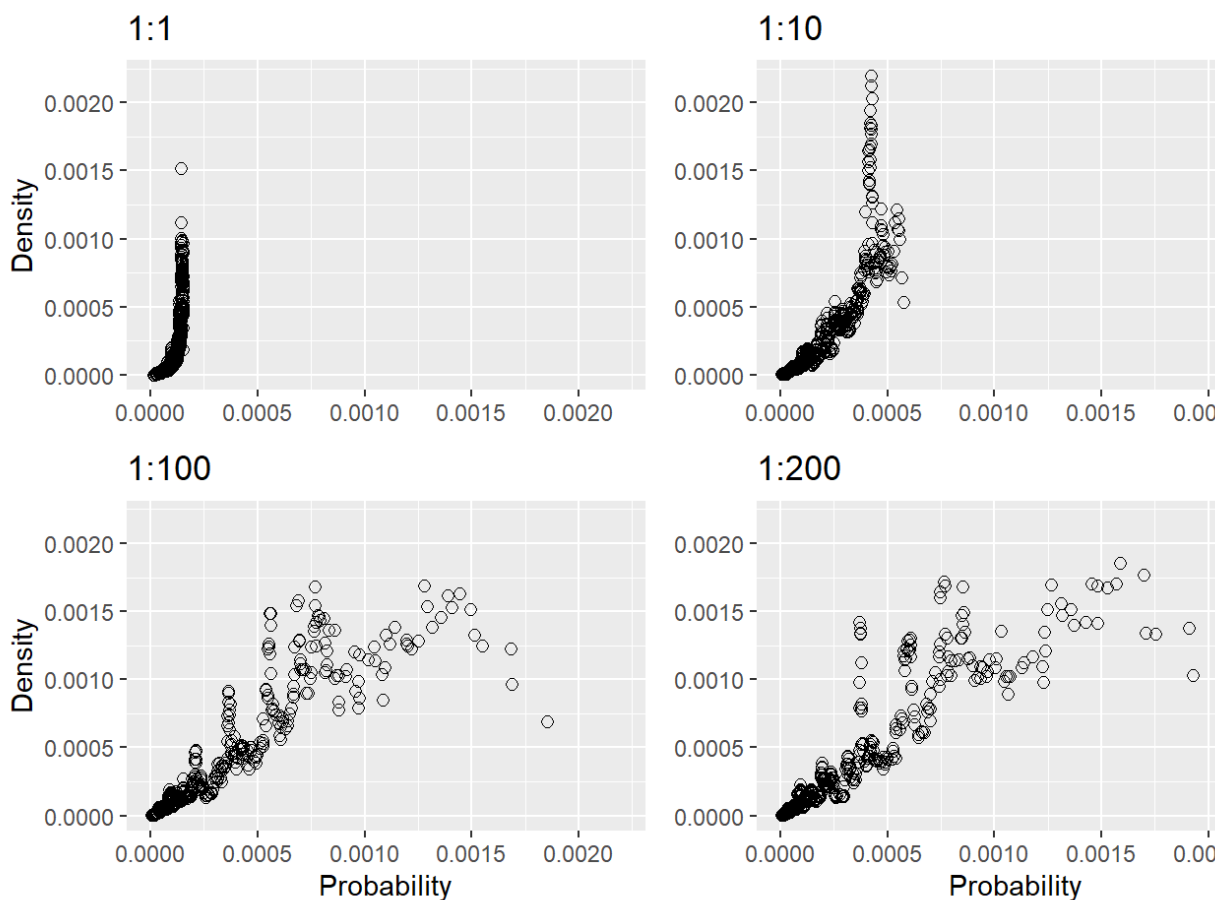

Figure 11: Landslide density versus modeled probability, quadratic model terms.

These show a similar pattern with increasing sample size. With small samples, the high-density zones are under sampled. With increasing sample size, the points approach the diagonal across the entire range of modeled probabilities indicating both that the sample size is becoming sufficient to characterize conditions in high-density zones and that the terms in the model can more closely match the shape of the target surface in the multidimensional data space.

This variation in model performance with sample size is seen in comparing success-rate curves built from modeled data (by integrating modeled probability over area) and the observed data from the landslide inventory. Here is an example using the quadratic model described above. The modeled success-rate curves approach closer to the observed curve as the number of nonlandslide points in the sample increases.
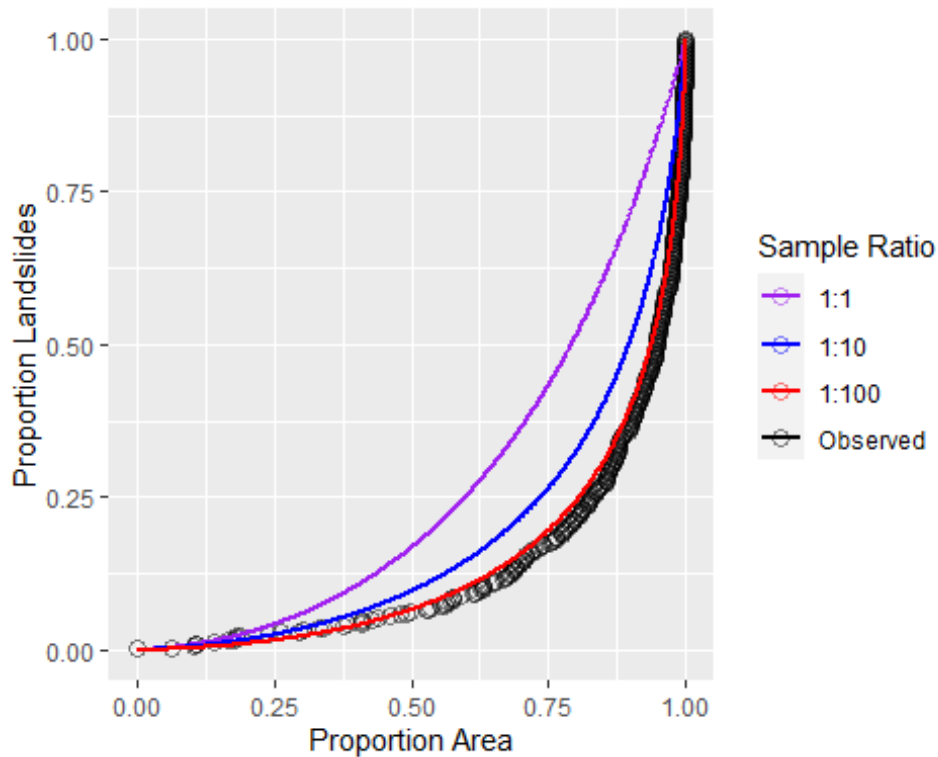

Figure 12: Success-rate curves for different ratios of landslide to non-landslide points in the training data compared to the curve for the inventoried landslides

The observed success-rate curves in Figure 9 suggest that models 3 and 4 perform about the same, yet Figure 10 and Figure 11 suggest that model 4 better reflects landslide locations in the high-density zones. The high-density zones occupy a relatively small proportion of the study area, so improved resolution there has relatively little influence on the success-rate curve. It is useful to use multiple measures of model performance, since different measures reflect different aspects of how well a model works. As shown with the linear-term model (#3), a numeric measure, such as area under the success-rate curve, may indicate good performance overall yet miss some small but important aspect of model performance. In this case, modeled probabilities for large values of gradient, curvature, and contributing area that were too high.

We seek models that can best resolve spatial variability in landslide density. The success-rate curve provides one measure of that ability, but it does not provide a direct prediction by which to gauge model performance when applied to test (rather than training) data. A testable prediction is provided by comparing the proportion of modeled and observed landslides over some specified range of modeled probability. This comparison also provides a measure of

model performance. To illustrate, Figure 13 below shows modeled initiation probability for a small basin in the Siuslaw Watershed in Oregon. The model was trained using the DOGAMI Special Paper 53 inventory with a 1:200 ratio of nonlandslide to landslide points.
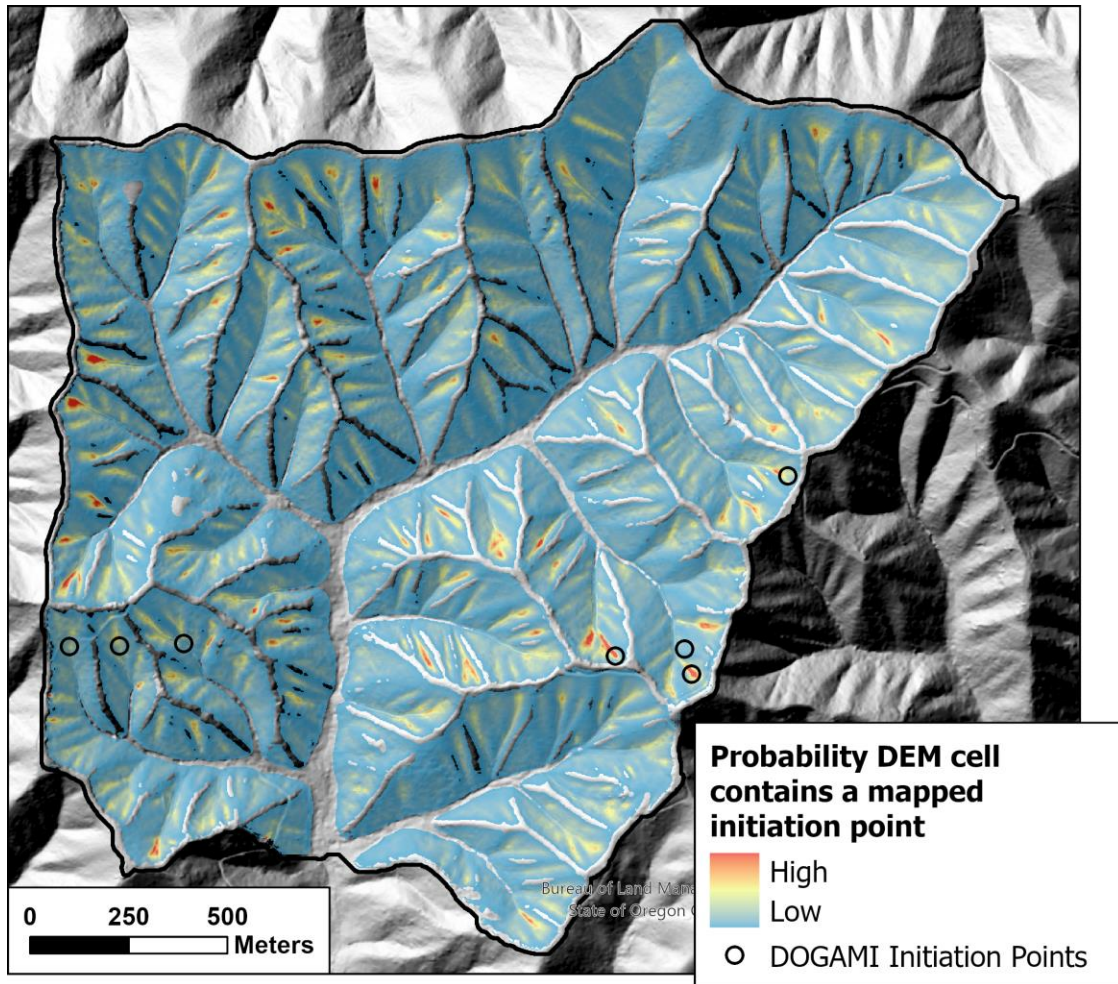

Figure 13: Probability of initiation for a small basin in the Siuslaw Watershed.

As discussed above, the integral of modeled probability over space, or summed over DEM cells, gives the modeled number of landslides within those cells. If the sum is over the entire study area, this sum gives the total number of observed landslides. Ordering by modeled probability, summing, and dividing by the total gives a cumulative frequency distribution for the number of landslides ordered by initiation probability, as shown in panel A of Figure 14 below for the linear model with a 1:10 sample ratio.
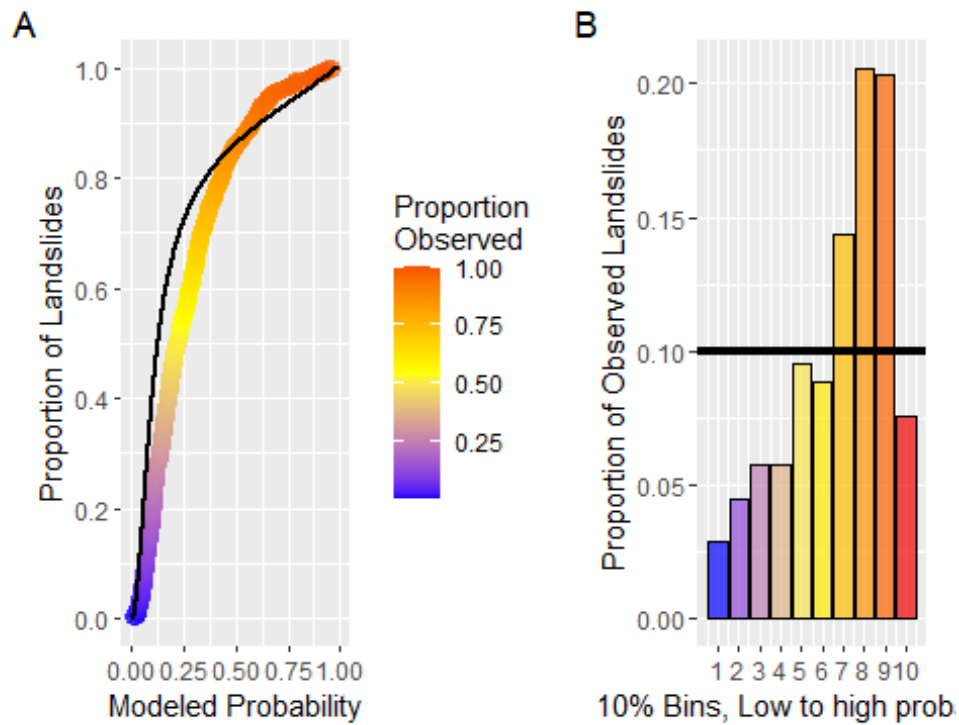
Figure 14: A) Cumulative distribution of modeled landslides. B) Proportion of observed landslides in each 10% bin.

For each DEM cell, there is a one-to-one correspondence between modeled probability and the proportion of modeled and observed landslides. This is translated to a map showing those areas encompassing a given proportion of all landslides as predicted by the model, as shown in Figure 15 below. Over the study area, the area contained within each colored zone is predicted to contain 10% of all observed landslides. (Note that Figure 15 shows only a small portion of the entire study area.)
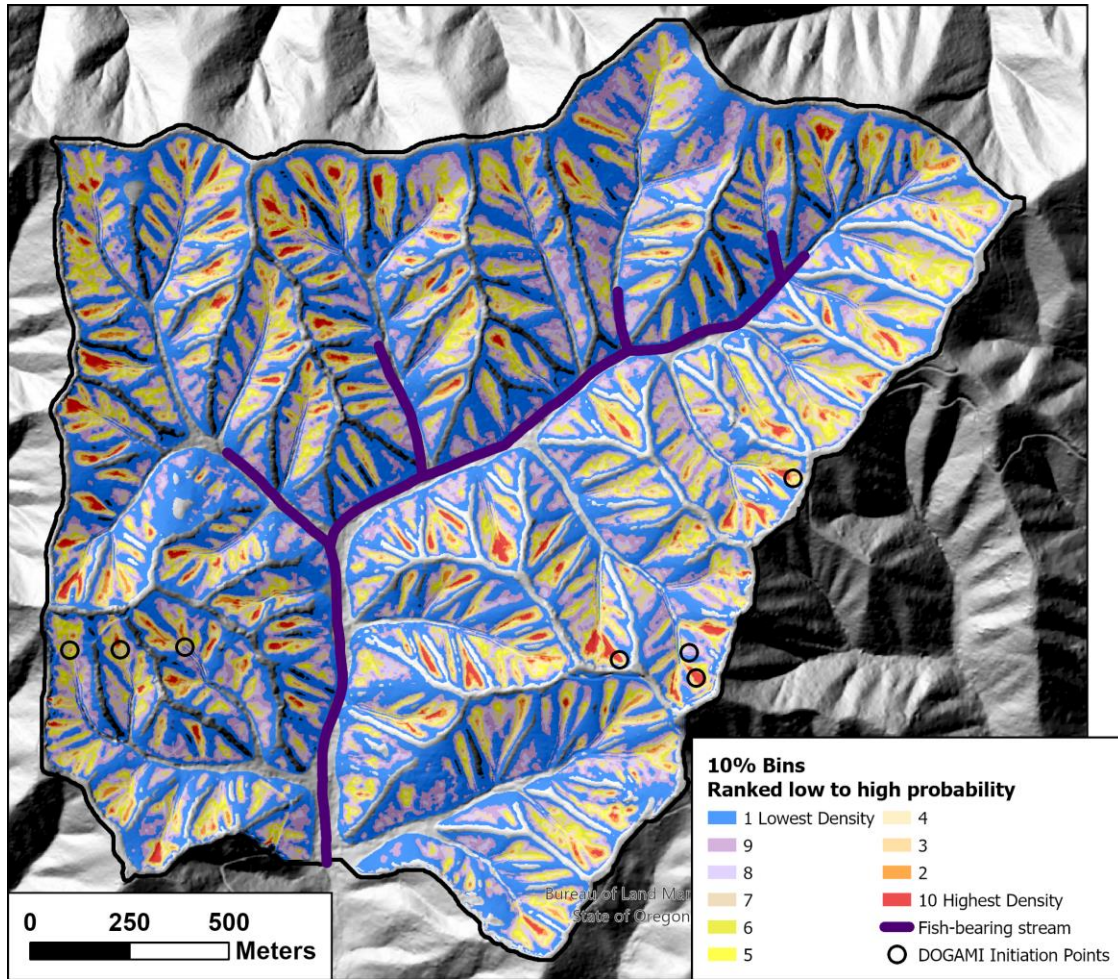
Figure 15: Modeled proportion of all landslide initiation points

Over any increment of modeled probability, a good model will predict the same proportion of landslides as observed. Panel B in Figure 14 shows the proportion of observed landslides laying within each 10% increment of modeled landslides over the entire study area used for the DOGAMI inventory. These bins correspond to the 10 rankings in Figure 15. Where the bars fall below 0.1, the model has over-predicted the proportion of landslides; where the bars fall above 0.1, the model has under-predicted the proportion of observed landslides. The sum of the absolute value of these differences provides a measure of model performance based on its ability to match spatial patterns in observed landslide density.

Figure 14 was based on the linear model with a 1:10 ratio of landslide to nonlandslide points in the training data. Figure 16 below shows the misfit, the modeled minus the observed proportion for each modeled 10% bin for the linear and quadratic models with both 1:10 and 1:100 sample ratios.
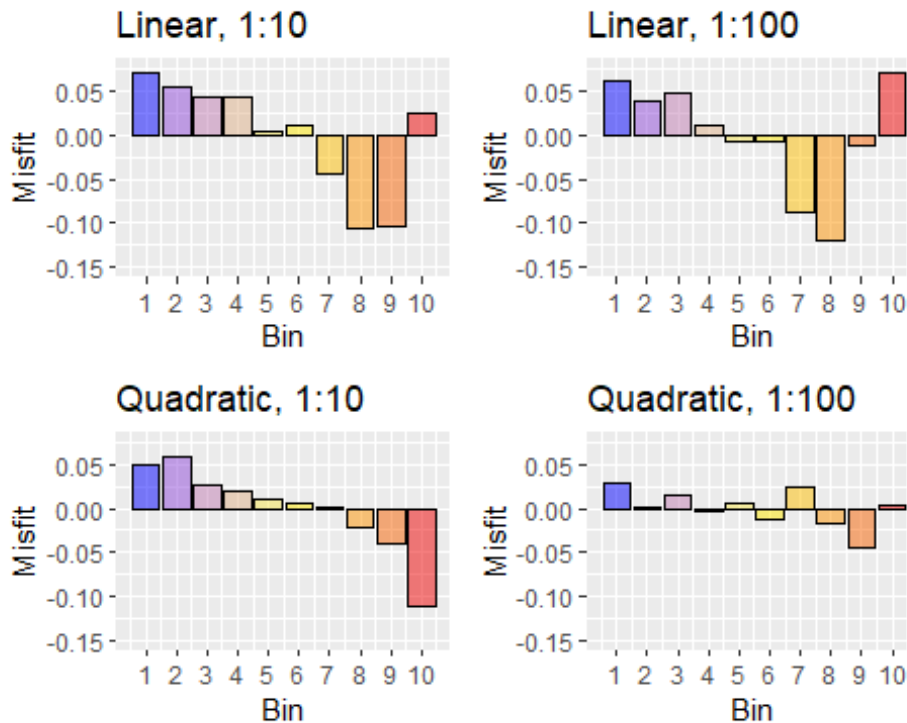
Figure 16: Misfit of modeled landslide proportions within each 10% bin.

The linear model exhibits systematic differences at both low and high sample ratios. The quadratic model exhibits systematic differences at a low sample ratio, but more random variations at a high sample ratio. These variations may reflect unavoidable noise in the landslide data itself.

The difference sums are listed in Table 1 below:

Table 1: Summed absolute differences

	Linear	Quadratic
1:10	0.506	0.347
1:100	0.465	0.156

Based on ability of each model to replicate observed spatial variations in landslide density. Higher sample ratios perform better and the quadratic model performs better than the linear model.

This gives us two measures of model performance: the success-rate curve, which compares the ability of different models to resolve spatial variability in landslide density, and the proportion differences illustrated above, which compare ability of different models to replicate observed spatial variability in landslide density.

We can compare these density-based measures to a measure based on classification success. The ROC curve plots the True Positive Rate versus the False Positive Rate for a range of threshold values spanning the range of modeled probability. The area under the ROC curve (AUC) is used as a single-valued measure of model performance; a higher AUC value indicates a better-performing model. For a given threshold in modeled probability, the True Positive Rate is the number of correctly classified points (True Positives) divided by the total number of landslide points. The False Positive Rate is the number of nonlandslide points incorrectly classified as landslides divided by the total number of nonlandslide points. To build an ROC curve, the TPR and FPR are calculated over the full range of modeled probabilities and the TPR plotted as a function of the FPR. At a very low threshold probability, all points are classified as landslides and both TPR and FPR equal one. At a very high threshold probability, no points are classified as landslides and TPR and FPR are both equal to zero. A model that is good at correctly classifying points in the training sample will have high TPR values and low FPR values, so the plotted curve will approach the upper left corner of the plot. The area under the ROC curve, referred to as AUC, is then a single-valued measure of model performance: higher AUC values indicate a better-performing model.

For a landslide susceptibility model, there are two factors that hinder use of ROC and AUC: 1) most terrain locations where landslides could occur do not have landslides in the inventory but will have a high modeled probability, so the FPR is large no matter how good the model, and 2) ROC should be calculated for the entire DEM, not just the training data, so the FPR becomes a function also of the proportion of DEM area in low and high modeled probability values. Here are ROC curves for the models discussed above:
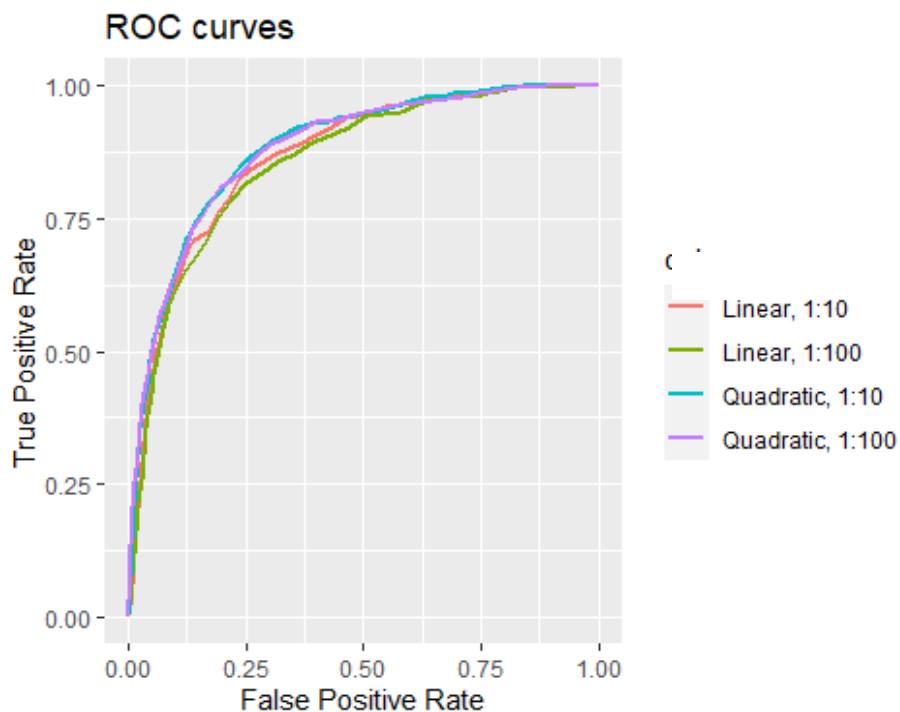

Figure 17: ROC curves

AUC for each model is listed below.

Table 2: AUC values

	Linear	Quadratic
1:10	0.867	0.883
1:100	0.857	0.882

The AUC values suggest that the quadratic model is slightly better than the linear model, but also indicates that the 1:10 ratio sample is better than the 1:100-ratio sample for the linear model and that the two sample ratios have almost no effect on model results for the quadratic model. These results are misleading; Table 1 clearly shows that, based on modeled ability to replicate observed spatial patterns in landslide density, the quadratic model is better than the linear model and higher sample ratios are better than lower sample ratios. Other measures of model performance have also been used for evaluating models of susceptibility to landslide initiation, such as the precision-recall curve (Yordanov and Brovelli 2020) and the Brier Score (Woodard et al. 2023). Precision is defined as the number of true positives divided by the sum of true positives and false positives. If the model is accurate, this gives the proportion of potential landslide sites where landslides were observed, which is a function of the sequence of storms over the time period sampled by the landslide inventory. Even a good model will have a precision that is very small because the number of potential landslide sites is far larger than the number of observed landslides. The Brier Score is defined as

$$B = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2$$

where p_i is the modeled probability for the ith DEM cell, o_i is one if the DEM contains a landslide initiation point and zero otherwise, and the sum is over all N DEM cells. Smaller Brier Scores indicate better performing models. However, the vast majority of DEM cells have no initiation point, so the Brier Score essentially gives the average squared probability over the DEM, which is not a useful measure of model success. Such measures can be used, but need to be evaluated similarly to how ROC and AUC were examined above.

We have looked so far at model performance in terms of the adequacy of the sample and selected predictors at resolving spatial variability in landslide density using logistic regression. A thorough analysis would apply additional algorithms and use the same measures to compare performance across model types. These could include general additive models, random forest, vector machine, and neural networks, among others. Each model type uses different approaches for characterizing the response (landslide density) surface within the data space defined by the chosen predictors. Some algorithms will work better than others. We have described above some methods by which models can be evaluated, but the actual choice of what predictors to try, what models to use, and how to test them will depend on what is found during the process of data collection and analysis.

Ultimately, however, a model must also be judged by the geomorphic plausibility of the results, which requires making maps and comparing model results to the expectations of people with experience in the area. Discrepancies between model predictions and expectations provide opportunities to re-examine the input data in light of those discrepancies. If the data do not support model results, then the model needs to be modified or the landslide inventory and expert opinions evaluated for errors.

2.1.1.4. Confidence

So far we have looked at ways to evaluate goodness of fit: how well a model can reproduce the data it was trained with. We want a model that can accurately reproduce observed spatial distributions of landslide density with which to calculate the proportion of landslides expected within different landform types. We also want to know how much confidence to place in the calculated proportions. This is true both for the study area and time period on which the model was trained and also when a model is extrapolated to predict landslide proportions into the future or for locations outside those covered by the landslide inventory. It is not possible to calculate this confidence directly without additional data, but it is possible to estimate it by looking at how sensitive model predictions are to the input data. This sensitivity can be assessed using cross validation, in which a model is trained using a subset of the data and then tested against the remaining data not in that subset. Subsets may be obtained using a random sample from the inventory, or by dividing the study area into different zones, or by subsetting landslides by date of occurrence. This process can be repeated multiple time using different subsets of the inventory with each iteration. The set of model predictions can then be used to calculate confidence intervals for model predictions, in this case, the modeled probability (landslide density) rasters. Each iteration also provides a measure of model performance, such as the difference sums shown in Table 1 above, when a model calibrated to the training subset of the inventory is applied to the remaining test data. The distribution of these performance measures provides an estimate of how poorly the model might perform when applied to new data.

The examples above looked solely at logistic regression. However, each type of model might work well for some cases and not so well for others. It is useful, therefore, to apply more than one model type in an analysis of confidence. These ideas are explored by Fabbri and Patera (2021).

2.1.2. Temporal Probability

The predictors in the discussion above involved relatively immutable topographic elements of the terrain; the predictors themselves do not change over the time scales of interest here (except for the influence of forest roads). This is not true for other potentially primary controls on landslide location: vegetation cover, antecedent moisture, and storm characteristics. We briefly described issues involved in associating forest cover to landslide density: specifically, that the spatial distribution of forest-cover characteristics (e.g., stand age) needs to be known for the date of each landslide.There are a variety of gridded precipitation data sets available from which to estimate rainfall, as discussed in the study design. These are derived from a variety of data sources and offer a range of spatial and temporal resolutions and accuracy. The

storm characteristics that can be measured are constrained by the resolution and accuracy of the precipitation data.

We can, however, see what guidance is offered by other studies. In examining landslides associated with an intense storm in Washington, Turner et al. (2010) found that landslide density varied nonlinearly with 24-hour rainfall intensity measured relative to a 24-hour, 100-year-recurrence-interval storm. They interpolated rainfall amounts between a local network of precipitation gauges. For a large typhoon in Japan, Marc et al. (2019) found that landslide density varied with event rainfall measured relative to the 10-year-return-interval event rainfall. They used data from a weather radar system in Japan to estimate rainfall amounts. Rossi et al. (2017) describe use of gridded precipitation data from the NASA Tropical Rainfall Measuring Mission (TRMM) to identify duration-intensity thresholds for landslide initiation in Italy. The TRMM provides rainfall amounts over a 0.25-degree grid at 3-hour increments. They looked solely at thresholds for landslide initiation, not at landslide density. Stanley et al. (2020) used National Climate Assessment Land Data Assimilation System (NCA-LDAS) data to build a "landslide hazard indicator" for the Pacific Northwest. The NCA-LDAS provides a variety of weather indicators, including rainfall, over a 0.125-degree grid with daily time step. Their landslide hazard indicator is intended to estimate the probability that a landslide would occur within a 0.125-degree grid cell on any day of the year. This is approximate, because their landslide inventory was a serendipitous collection of events from several inventories, not a census of landslides. Kirschbaum and Stanley (2018) use data from the Global Precipitation Measurement system (0.1-degree grid, 30-minute interval) and TRMM to estimate 7-day running totals of rainfall. These are correlated with landslide susceptibility maps to provide real-time estimates of landslide potential world wide.

How well these different precipitation data sets resolve spatial variations in storm intensity remains to be determined. There is a large literature on the strengths and weaknesses of different precipitation data sets (e.g., Li et al., 2020; Molter, Collins and Risser, 2021; Ombadi et al., 2021; Rajulapati et al., 2020; Thornton et al., 2021); we listed those we thought most relevant in the study design, but others may also prove useful. The maps of 24-hour intensity, normalized to the 24-hour, 100-year recurrence interval storm, presented by Turner et al. (2010) for the 2007 storm in southwest Washington provide a baseline to compare against.

3 Runout

As described in the study design, this study is focused on shallow landslides, those most likely to be influenced by forest practices. These landslides typically involve failure of soil overlying a more competent substrate. Numerous studies have examined controls on runout extent of shallow landslides, particularly those that evolve into debris flows. Some specific to the Pacific Northwest include Benda and Cundy (1990); Hofmeister et al. (2002); Fannin and Rollerson (1993); Robison et al. (1999); May (2002); Lancaster et al. (2003); Miller and Burnett (2008); Guthrie et al. (2010a); Coe et al. (2011), and Reid, Coe and Brien (2016). These and other studies consistently point to several factors that influence runout lengths:

- channel gradient and confinement,

- abrupt changes in flow direction at channel junctions,

- the volume of mobilized material, and

- the size and number of trees encountered and the amount of large wood incorporated into the mobilized material.

Empirical models to predict runout length seek to calibrate observed runout lengths to measurements or estimates of these factors. Benda and Cundy (1990) used channel slope and tributary junction angles. Fannin and Wise (2001) use channel gradient and confinement, estimated volume, and changes in flow direction. Miller and Burnett (2008) used channel gradient and confinement, estimated volume, tributary junction angles, and stand-age brackets (as indicators of tree size and wood availability). Guthrie et al. (2010) used gradient, changes in flow direction, and presence/absence of mature timber. These are all one-dimensional models; they look solely at runout length, not inundation area.

Another class of empirical models seeks to relate simple measures of deposit geometry (length, cross-section and planimetric area) to deposit volume and other controlling factors (Griswold and Iverson, 2008; McDougall, 2017; Schilling, 1998). Such models have been applied for predicting debris-flow runout extent and inundation area (Berti and Simoni, 2014; Hofmeister and Miller, 2003; Reid, Coe and Brien, 2016).

Another approach is to iteratively extend a debris-flow downslope from an initiation site, DEM cell by DEM cell, applying mass balance with empirical rules for scour and deposition. These include that by Lancaster, Hayes and Grant (2003), which used a triangular mesh; "random-walk" models (Mergili, Schwarz and Kociu, 2019); "agent-based" models (Guthrie and Befus, 2021); the "ProDF" model (Gorr et al., 2022); and the "MassWastingRunout" model (Keck et al.).

MassWastingRunout (MWR) is a cellular-automata landslide runout model designed for probabilistic sediment transport, topographic change and landslide hazard assessment applications. MWR is coded in Python and implemented as a component for the Landlab earth surface modeling toolkit (Hobley et al., 2017). MWR includes a Markov Chain Monte Carlo calibration utility that determines the best-fit parameter values for a site as well as empirical Probability Density Functions (PDF) of the parameter values. It includes a utility called MWR Probability that takes the PDF output from the calibration utility to model runout probability.

The MWR model includes only the key controls on landslide runout, such as the influence of slope on erosion and deposition rates of the landslide runout material. Other processes, such as momentum, the impacts of woody debris or grain-to-grain forces, are not explicitly represented in the model, but through calibration, the effects of those excluded processes can be implicitly represented.

The MWR calibration utility automatically samples MWR model parameter space and tunes the model to match observed patterns of landslide runout extent, deposition and erosion. Observed patterns in landslide runout extent, deposition and erosion can be determined from lidar DEM differencing using lidar DEMs recorded before and after the observed landslide.

Once MWR is calibrated to an observed-runout dataset, it may work as a predictive tool for assessing the runout extent of neighboring potential landslides if regional landslide processes are relatively uniform. If a region consists of a wide diversity of landslide types, a more rigorous calibration may be required that might involve calibrating the model to each landslide type. An illustration of the modeled runout of a large debris avalanche in the Cascade Mountains is shown below:
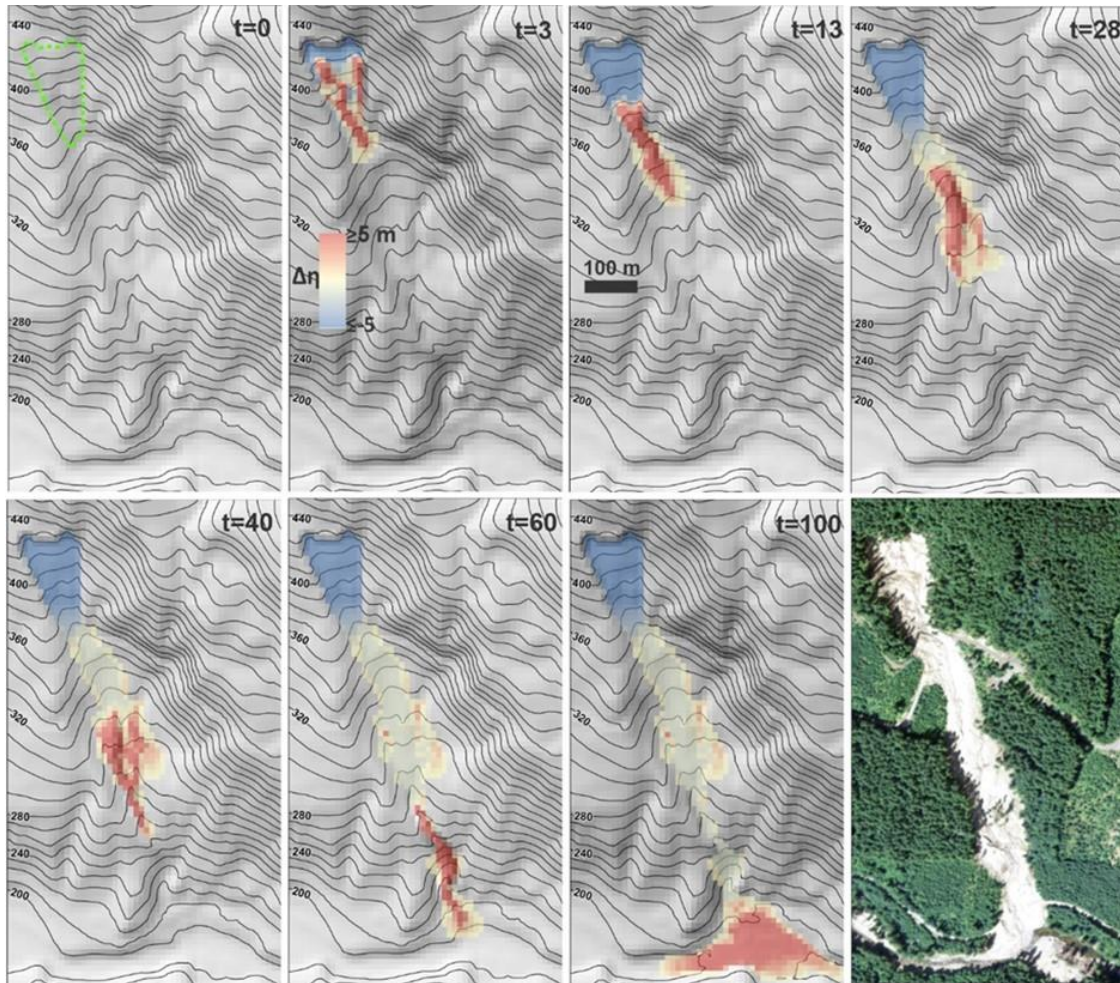

Figure 18: Illustration of modeled runout and topographic evolution of a hillslope following the initiation of a large debris avalanche in the Cascade Mountains of Washington State. Red indicates a gain and blue indicates a decrease in topographic elevation. Topography lines reflect the underlying terrain, which is updated after each model iteration. The last image is an aerial photo of the observed runout.

As described in the study design, we seek a method to estimate probability that a landslide will runout to any point downslope of the initiation site. The model used must be calibrated and applied using available data and must be computationally efficient, because it will be run for every potential initiation point in every area where the model is applied. We need only runout extent, for which a more computationally efficient one-dimensional model will suffice. We

included reference to survival analysis because it is such a one-dimensional model that can be readily calibrated with available open source tools. It is described and illustrated below. However, as with initiation, each type of model might work well for some cases and not so well for others. It is useful, therefore, to apply more than one model type for an analysis of runout as well.

3.1 Survival Analysis

Empirical estimates of probable lifespan for individuals in a population are based on the distribution of lifespans measured for some sample from that population. Effects of factors that might influence lifespan are then evaluated by their effects on the shape of that distribution. Here, "lifespan" refers to the increment of time until some event of interest. In medical applications, this might involve the time that patients remain cancer free following different types of treatment. In engineering applications, this might involve how long until a machine part fails. In social applications, it might involve the length of time that a couple remains married as a function of income. See the introductory article by Emmert-Streib and Dehmer (2019) for other examples. In a geomorphic context, Furbish and coworkers characterize downslope movement of soil particles and rock clasts in terms probable travel lengths, similar to our characterization of probable debris-flow travel length (Furbish and Roering, 2013; Furbish et al., 2021; Williams and Furbish, 2021). A key aspect of survival analysis is the ability to incorporate "censured" data; that is, to use measured time spans for individuals in a sample for which the expected event does not occur prior to the end of the observation period.

Here we look at the "lifespan" of a debris flow, not in terms of time, but in terms of distance. We want to know the expected runout length as a function of environmental factors encountered along the runout path. We are interested in the factors listed in the introduction above: channel gradient and confinement, changes in flow direction, the volume of material mobilized, and the quantity of large woody debris incorporated.

3.1.1. Survival Curves

A survival curve gives the proportion of individuals in a population that survive beyond a given time. It varies from zero to one on the y axis and 0 to the length of the period of observation on the x axis. A survival curve for debris flows indicates the proportion of events in a population of debris flows that runout beyond a given distance. An empirical estimate of the survival curve is obtained from the cumulative distribution of measured debris-flow-track lengths. For a given sample of debris-flow runout lengths, the survival curve is estimated as the proportion of tracks in the sample longer than a given length:

$$S(x) \approx \frac{\# \; tracks \; longer \; than \; x}{N} \qquad (4)$$

where N is the total number of tracks (see Emmert-Streib and Dehmer 2019 for a concise description). The shape of this distribution is determined by the probability that any individual debris flow will stop in the next interval of travel along its travel path. This probability is called the hazard rate and is defined as

$$h(x) = \lim_{\Delta x \to 0} \frac{P(x \leq X < x + \Delta x | X \geq x)}{\Delta x}, \quad (5)$$

or estimated empirically as $h(x) \approx n/(x + \Delta x)$, where x is distance from the initiating landslide and n is the number of tracks with lengths between $x + \Delta x$. The cumulative hazard function H "describes the accumulated risk up to time t" (Emmert-Streib and Dehmer 2019), or to distance x, and is defined as the integral of the hazard rate:

$$H(x) = \int_0^x h(\tau) d\tau. \quad (6)$$

The survival curve is then determined from the cumulative hazard function as

$$S(x) = \exp(-H(x)). \quad (7)$$

If the hazard rate is constant with distance, then the frequency distribution of track lengths will follow an exponential distribution. Remarkably, observed distributions of debris-flow travel lengths are fairly well approximated with an exponential distribution (Miller and Burnett 2008). Other parametric distributions can also be used fit empirical survival curves (Emmert-Streib and Dehmer 2019). For example, if the hazard rate increases or decreases with time (distance), the survival curve is described with a Weibull distribution. If the hazard rate follows a U-shaped curve, decreasing at first and then increasing, the survival curve can be described with a log-normal distribution. We expect that the hazard rate will vary with conditions along the travel path; that is, that the probability that a debris flow will continue through any increment of length along a potential travel path will vary with channel gradient and confinement, changes in flow direction, the volume of material mobilized, the amount of large woody debris, and other factors we have not yet considered. Thus, we expect the hazard rate to vary uniquely for every potential debris-flow track. We can use the shape of the empirical survival curve to infer how the hazard rate changes in response to these conditions. Measures of these conditions, e.g., of the gradient, are referred to as covariates (the independent variables) and the value of these covariates varies along the debris-flow travel path. To estimate the effect of these distance-varying covariates on the hazard rate, we use a relative risk model, typically referred to as a Cox model Emmert-Streib and Dehmer (2019), with time-varying (distance-varying in this case) covariates. The hazard rate is then defined as

$$h(x, Z(x)) = h_0(x) \exp\left(\sum_{i=1}^{p} \beta_i Z(x)_i\right) \quad (8)$$

where $Z(x)$ is a vector of distance-varying covariates (e.g., gradient), β is a vector of coefficients, one for each covariate, p is the number of covariates, and $h_0(x)$ is a baseline hazard rate (i.e., the hazard rate when all the covariates Z are zero). We fit the empirical survival curve, Equation 4, with a parametric distribution to define $h_0(x)$. Once values for the coefficients β are estimated, the change in cumulative hazard function at any point x along a potential debris-flow track is estimated as

$$\Delta H\big(x|Z(x)\big) = 1 - \big(1 - \Delta H_0(x)\big)^{\exp(x\beta(x))} \qquad (9)$$

where ΔH_0 is the baseline cumulative hazard function defined in <u>Equation 6</u> using a parametric-distribution fit to the empirical survival curve (Kalbfleisch and Prentice 2002; Ruhe 2018). The survival curve is then determined as

$$S\big(x|Z(x)\big) = \exp\Big(-\sum_{i=1}^{n} \Delta H\big(x_i|Z(x_i)\big), \qquad (10)$$

where the x_i are the locations where the covariates $Z(x_i)$ have been measured.

For the examples below, the "flexSurv" R package was used to fit a parametric distribution to the debris-flow-track lengths measured from the DOGAMI Special Paper 53 inventory. A variety of digital data products were then used to obtain covariate values at intervals along each inventoried track. Descriptions of these covariates and how they were measured are provided in a following section. Tabulated values for all inventoried tracks were then used with the "survival" R package to obtain coefficient estimates for each covariate.

3.1.2. Censored Data

An important capability of survival-analysis methods is the ability to use information from samples for which the event of interest is not observed. In this case, the event of interest is the terminal end point of a debris-flow deposit. This might occur when the distal end of a deposit has been removed by stream erosion or where the end of the deposit is not visible when mapped from aerial photographs. These cases still provide useful information because we know the debris flow traveled at least as far as the furthest point observable. Samples for which the event of interest - the terminal end of the debris-flow deposit - are not observed are referred to has being "censored". These censored samples are included with the uncensored samples, those for which the complete debris-flow track lengths are known, in estimating the survival curve. An empirical estimate of the survival curve based solely on the observed censored and uncensored flow-path lengths is obtained with the Kaplan-Meier estimate (Kalbfleisch and Prentice 2002, pg 16):

$$\hat{S}(x) = \prod_{j|x_j \le x} \frac{n_j - d_j}{n_j}. \qquad (11)$$

Here, n_j indicates the number of tracks in the sample with censored or uncensored lengths greater than x_j and d_j is the number of track terminal endpoints observed at x_j. Because changes in the $\hat{S}$ curve value occur only at lengths (x values) corresponding to observed complete (uncensored) debris-flow tracks, the curve consists of a series of steps. The corresponding cumulative hazard function is obtained with the Nelson-Aalen estimate:

$$\hat{H}_0(x) = \sum_{x_i \le x} \frac{d_i}{n_i}, \qquad (12)$$

Landslide susceptibility, revised draft study design, September 7, 2023

The zero subscript indicates that this estimate can be used as the baseline cumulative hazard function for the relative-risk model (Equation 8), as it does not include the influence of any covariates.

3.1.3. Covariates

The list of observed influences on debris-flow runout length listed at the beginning of Section 3 provide a starting point: channel gradient and confinement, abrupt changes in flow direction at channel junctions, the volume of mobilized material, and the size and number of trees encountered and the amount of large wood incorporated into the mobilized material. For the PFA modeling, I found no influence with changes in flow direction at tributary junctions. This may be an idiosyncrasy of the DOGAMI inventory as other studies have found a relationship (e.g., Benda and Cundy 1990; R. J. Fannin and Wise 2001; Miller and Burnett 2008; R. H. Guthrie et al. 2010). The volume of mobilized material could not be calculated directly, so an empirical proxy was used instead. Following Miller and Burnett (2008), the ratio of volume deposited to volume scoured was estimated at increments along all mapped debris-flow tracks. These volumes were calculated by integrating the modeled probability of scour and deposition along each track. Probabilities were based on multinomial logistic regression of field mapped zones of scour, transitional flow, and deposition from the ODF 1996 Storm Study (Robison et al. 1999) with gradient and curvature calculated over a 15-m radius from 2-meter lidar-derived DEMs. Estimated stand age at the time of each debris flow was based on data from the LEMMA project and this was used as an indicator of large wood availability.

Values for each of these covariates (gradient, tangential curvature, volume ratio, stand age) were measured at 2-meter increments along each debris-flow track, coincident with the DEM grid. These provided "time-varying" inputs to the "coxph" function from the R "survival" package, which uses maximum likelihood to solve for the coefficients in Equation 9, with which coefficient values for each covariate were obtained.

4 Combining initiation and runout probabilities

Once coefficients $\beta(x)$ have been calibrated for a relative-risk survival model, a survival curve can be calculated along any potential flow path to provide the probability that a debris flow will travel to any downslope point along that path. For any debris-flow initiation site, the probability of runout to any location along the downslope flow path traced on the DEM is determined using Equation 9 and Equation 10. Any potential flow path may have multiple initiation sites that feed into it. Following Miller and Burnett (2008), the probability P_{DF} that a debris-flow from any upslope initiation site will reach a point x along that path is

$$P_{DF}(x) = 1 - \prod_{i=1}^{n}(1 - S_i(x)P_{Ii}) \qquad (13)$$

where $S_i(x)$ is the survival-curve value at point x (from Equation 10) and P_{Ii} the probability of initiation for the ith initiation site, with the product over all n upslope initiation sites. The spatial distribution modeled for P_I was shown in Figure 13 for a small drainage in the Siuslaw

basin. To implement calculation of Equation 13 over a DEM, surface-flow paths are traced from every DEM grid point with a nonzero initiation probability and Equation 13 calculated for every point along the flow path until the survival-curve value goes to zero.

The probability that a hillslope point is traversed by a debris flow that originates upslope and continues flowing downslope to deposit material into a stream channel is determined in a similar fashion (see Burnett and Miller 2007). The travel path from each DEM point with a nonzero probability of initiation is traced downslope until it intersects a channel. Label that intersection location as x_c. The probability calculated for that debris flow at that point of intersection is $S_i(x_c)P_{Ii}$, where $S_i(x_c)$ is the survival-curve value at x_c and P_{Ii} is the initiation probability for the ith DEM cell (where the flow path originated). $S_i(x_c)$ gives the probability that a debris flow from the ith DEM cell will travel to a channel. This value can be mapped back to the initiating DEM cell to create a map of delivery probabilities, an example of which is shown in Figure 19 below calculated for delivery to fish-bearing channels for the PFA. Note in Figure 19 the extent of the debris-flow tracks from the DOGAMI inventory (red lines): all originated in areas with a low modeled probability of delivery and none of them extend to the fish-bearing channel.
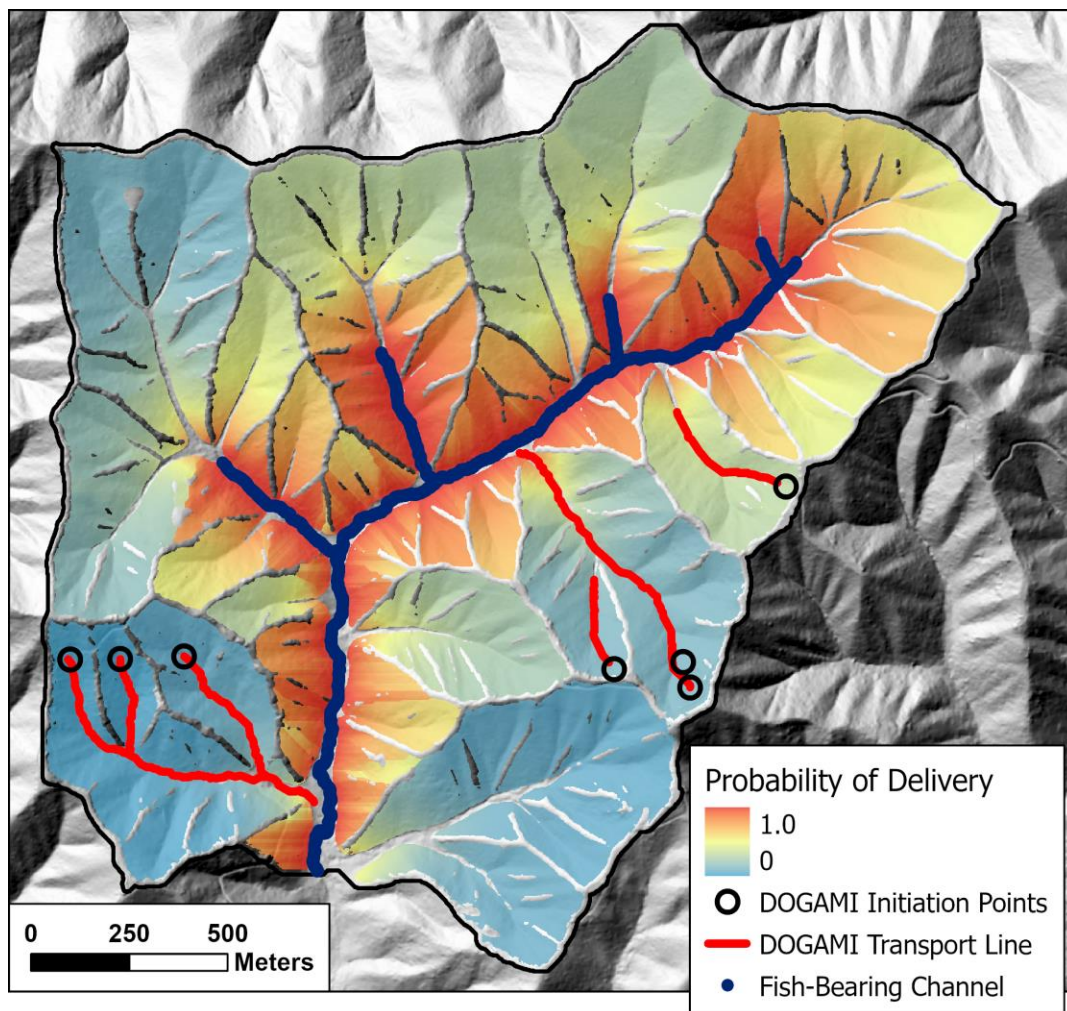
Figure 19: Modeled probability of delivery to a fish-bearing stream.

The quantity $S_i(x_c)P_{Ii}$ gives the probability that a debris flow from the ith DEM cell delivers material to a stream channel. This value too can be mapped back to the DEM cell where each flow path originates to create a map showing the modeled probability that a debris flow will be initiated and travel to a fish-bearing stream, illustrated in Figure 20 below.
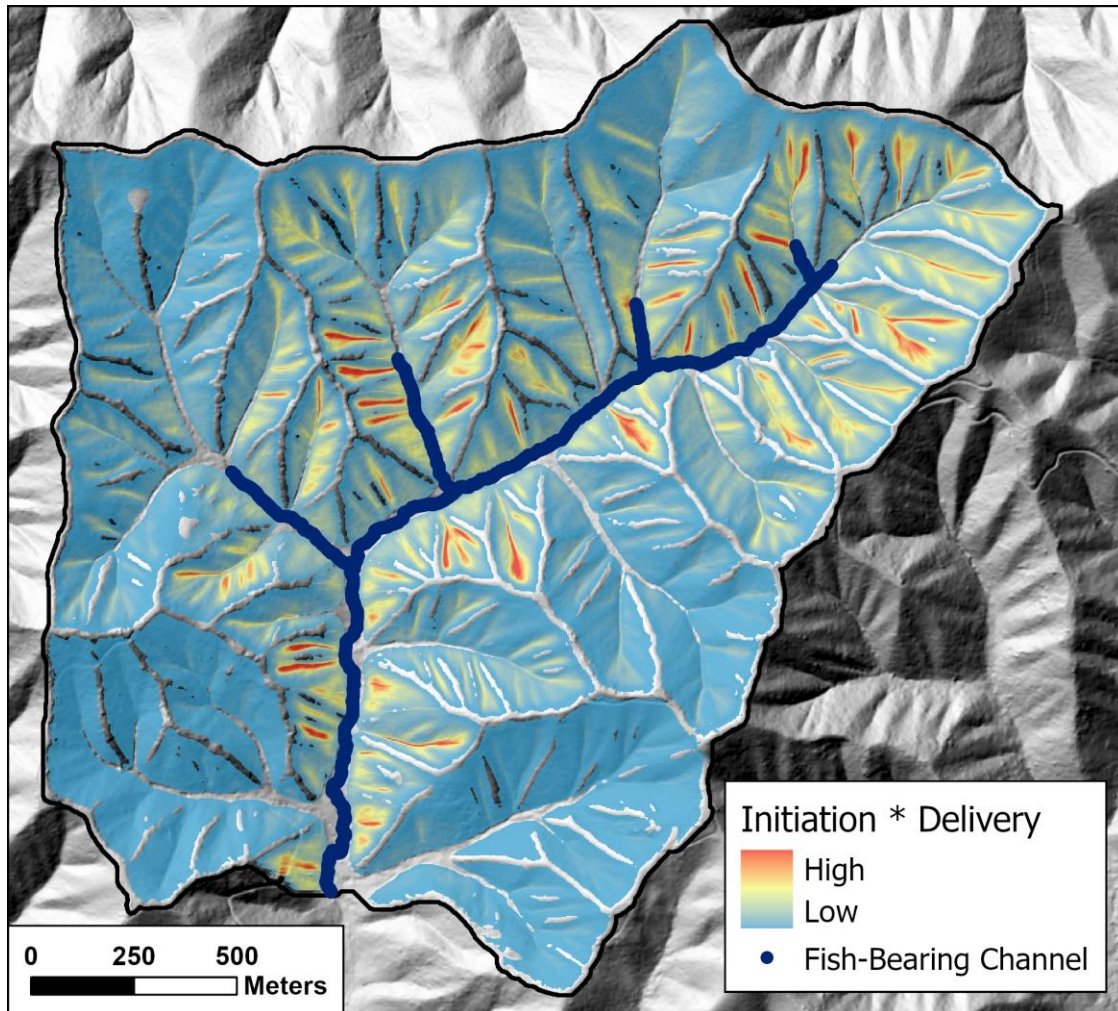

Figure 20: Modeled probability of initiation and delivery.

These values indicate the probability that a DEM cell contains a mapped initiation point for a landslide that ran out to a fish-bearing channel in the landslide inventory used to calibrate the initiation and runout models. Hence, the calculated probability indicates the landslide density and integration of that probability over any specified portion of the study area gives the number of delivering landslides initiated within that area. Dividing that number by the total number of landslides in the inventory gives the proportion of delivering landslides found in that specified area. This is illustrated in Figure 21 below.
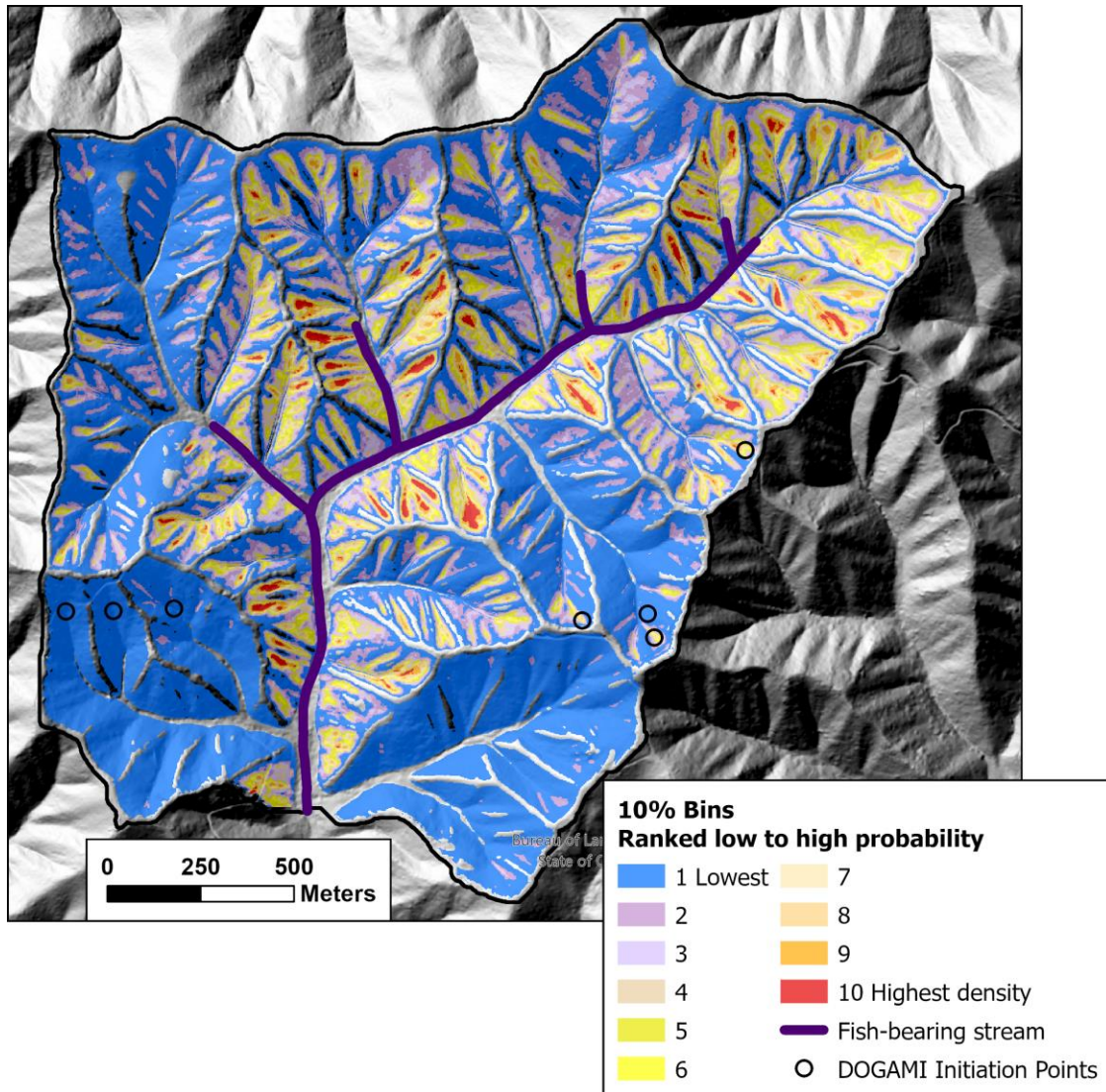
Figure 21: Proportion of delivering landslide initiation sites

Figure 13 and Figure 15 show the modeled source areas for all landslides; Figure 20 and Figure 21 show modeled source areas for those landslides that can deliver material directly to fish-bearing streams. These examples show how linking modeled probabilities of initiation and delivery can identify and rank terrain locations in the context of downslope resources potentially impacted by landslides. These locations can be ranked both by the density and proportion of landslides potentially affecting downslope resources. In this example from the PFA modeling, delivery was calculated for fish-bearing streams. To evaluate RILs in Washington, delivery would be calculated to any channel.

Here we have parsed the landscape into zones based on ranked probability of initiation and delivery. The landscape can also be parsed into landform polygons and modeled probability integrated over each polygon to estimate the proportion of landslides originating from each landform type as was shown in Table 3 of the study design. Those modeled proportions can be

compared to the observed proportions from the landslide inventory the same as was done above in Figure 16. These methods thus provide a means for evaluating linked models of initiation and delivery, for comparing landform types in terms of landslide density and the proportion of landslide-delivered material originating from each, and for testing the accuracy of these comparisons against inventory data. Rasters created by these linked models, such as shown in Figure 20 and Figure 21, can be used along with other terrain attributes as input to image segmentation programs to generate new landform polygons. This provides a means of evaluating potential changes to RIL criteria.

9 References

Aguilar, F., Agüera, F., Aguilar, M. A., and Carvajal, F., 2005, Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy: Photogrammetric Engineering & Remote Sensing, v. 71, no. 7, p. 805-816.

Amato, G., Eisank, C., Castro-Camilo, D., and Lombardo, L., 2019, Accounting for covariate distributions in slope-unit-based landslide susceptibility models. A case study in the alpine environment: Engineering Geology, v. 260.

Anderson, S. W., 2019, Uncertainty in quantitative analyses of topographic change: error propagation and the role of thresholding: Earth Surface Processes and Landforms, v. 44, no. 5, p. 1015-1033.

Bailey, G., Li, Y., McKinney, N., Yoder, D., Wright, W., and Washington-Allen, R., 2022, Las2DoD: Change Detection Based on Digital Elevation Models Derived from Dense Point Clouds with Spatially Varied Uncertainty: Remote Sensing, v. 14, no. 7.

Benda, L., Veldhuisen, C., Miller, D., and Miller, L. R., 1998, Slope Instability and Forest Land Managers: A Primer and Field Guide, v. 1, p. 98.

Benda, L. E., and Cundy, T. W., 1990, Predicting deposition of debris flows in mountain channels: Canadian Geotechnical Journal, v. 27, p. 409-417.

Bernard, T. G., Lague, D., and Steer, P., 2021, Beyond 2D landslide inventories and their rollover: synoptic 3D inventories and volume from repeat lidar data: Earth Surface Dynamics, v. 9, no. 4, p. 1013-1044.

Berti, M., and Simoni, A., 2014, DFLOWZ: A free program to evaluate the area potentially inundated by a debris flow: Computers & Geosciences, v. 67, p. 14-23.

Booth, A. M., Sifford, C., Vascik, B., Siebert, C., and Buma, B., 2020, Large wood inhibits debris flow runout in forested southeast Alaska: Earth Surface Processes and Landforms.

Borga, M., Fontana, G. D., and Cazorzi, F., 2002, Analysis of topographic and climatic control on rainfall-triggered shallow landsliding using a quasi-dynamic wetness index: Journal of Hydrology, v. 268, p. 56-71.

Brardinoni, F., Hassan, M. A., and Slaymaker, O. H., 2002, Complex mass wasting response of drainage basins to forest management in coastal British Columbia: Geomorphology, v. 49, p. 109-124.

Brardinoni, F., Slaymaker, O., and Hassan, M. A., 2003, Landslide inventory in a rugged forested watershed: a comparison between air-photo and field survey data: Geomorphology, v. 54, p. 179-196.

Buban, M. S., Lee, T. R., and Baker, C. B., 2020, A Comparison of the U.S. Climate Reference Network Precipitation Data to the Parameter-Elevation Regressions on Independent Slopes Model (PRISM): Journal of Hydrometeorology, v. 21, no. 10, p. 2391-2400.

Bull, J. M., Miller, H., Gravley, D. M., Costello, D., Hikuroa, D. C. H., and Dix, J. K., 2010, Assessing debris flows using LIDAR differencing: 18 May 2005 Matata event, New Zealand: Geomorphology, v. 124, no. 1-2, p. 75-84.

Burnett, K. M., and Miller, D. J., 2007, Streamside policies for headwater channels: an example considering debris flows in the Oregon Coastal Province: Forest Science, v. 53, no. 2, p. 239-253.

Burns, W. J., Coe, J. A., Kaya, B. S., and Ma, L., 2010, Analysis of elevation changes detected from multi-temporal lidar surveys in forested landslide terrain in western Oregon: Environmental & Engineering Geoscience, v. 16, no. 4, p. 315-341.

Burns, W. J., Franczyk, J. J., and Calhoun, N. C., 2022, Protocol for channelized debris flow susceptibility mapping: Oregon Department of Geology and Mineral Industries.

Cai, S., Yu, S., Hui, Z., and Tang, Z., 2023, ICSF: An Improved Cloth Simulation Filtering Algorithm for Airborne LiDAR Data Based on Morphological Operations: Forests, v. 14, no. 8.

Cannon, S. H., 1993, An empirical model for the volume-change behavior of debris flows, *in* Shen, H. W., Su, S. T., and Wen, F., eds., Proceedings, Hydraulic Engineering '93, Volume 2: New York, American Society of Civil Engineers, p. 1768-1777.

Catani, F., Segoni, S., and Falorni, G., 2010, An empirical geomorphology-based approach to the spatial prediction of soil thickness at catchment scale: Water Resources Research, v. 46, no. W05508.

Cățeanu, M., and Ciubotaru, A., 2021, The Effect of LiDAR Sampling Density on DTM Accuracy for Areas with Heavy Forest Cover: Forests, v. 12, no. 3.

Cavalli, k. M., Goldin, B., Comiti, F., Brardinoni, F., and Marchi, L., 2017, Assessment of erosion and deposition in steep mountain basins by differencing sequential digital terrain models: Geomorphology, v. 291, p. 4-16.

Cekic, S., Aichele, S., Brandmaier, A. M., Köhncke, Y., and Ghisletta, P., 2021, A tutorial for joint modeling of longitudinal and time-to-event data in R: Quantitative and Computational Methods in Behavioral Sciences, v. 1.

Chen, C., Wang, M., Chang, B., and Li, Y., 2020, Multi-Level Interpolation-Based Filter for Airborne LiDAR Point Clouds in Forested Areas: IEEE Access, v. 8, p. 41000-41012.

Cheng, L., Chen, S., Liu, X., Xu, H., Wu, Y., Li, M., and Chen, Y., 2018, Registration of Laser Scanning Point Clouds: A Review: Sensors (Basel), v. 18, no. 5.

Chung, C. J., and Fabbri, A. G., 2003, Validation of spatial prediction models for landslide hazard mapping: Natural Hazards, v. 30, p. 451-472.

Coe, J. A., Bessette-Kirton, E. K., Brien, D. L., and Reid, M. E., 2021, Debris-flow growth in Puerto Rico during Hurricane Maria: preliminary results from analyses of pre- and post-event lidar data: Proceedings fo the 13th International Symposium on Landslides.

Coe, J. A., Reid, M. E., Brien, D. L., and Michael, J. A., 2011, Assessment of topographic and drainage network controls on debris-flow travel distance along the west coast of the United States: Italian Journal of Engineering Geology and Environment.

Conoscenti, C., Rotigliano, E., Cama, M., Caraballo-Arias, N. A., Lombardo, L., and Agnesi, V., 2016, Exploring the effect of absences selection on landslide susceptibility models: a case study of Sicily, Italy: Geomorphology, v. 261, p. 222-235.

Croke, J., Todd, P., Thompson, C., Watson, F., Denham, R., and Khanal, G., 2013, The use of multi temporal LiDAR to assess basin-scale erosion and deposition following the catastrophic January 2011 Lockyer flood, SE Queensland, Australia: Geomorphology, v. 184, p. 111-126.

Cucchiaro, S., Maset, E., Cavalli, M., Crema, S., Marchi, L., Beinat, A., and Cazorzi, F., 2020, How does co-registration affect geomorphic change estimates in multi-temporal surveys?: GIScience & Remote Sensing, v. 57, no. 5, p. 611-632.

D'Odorico, P., and Fagherazzi, S., 2003, A probabilistic model of rainfall-triggered shallow landslides in hollows: A long-term analysis: Water Resources Research, v. 39, no. 9, p. 1262.

Dao, D. V., Jaafari, A., Bayat, M., Mafi-Gholami, D., Qi, C., Moayedi, H., Phong, T. V., Ly, H.-B., Le, T.-T., Trinh, P. T., Luu, C., Quoc, N. K., Thanh, B. N., and Pham, B. T., 2020, A spatially explicit deep learning neural network model for the prediction of landslide susceptibility: Catena, v. 188.

DeLong, S. B., Hammer, M. N., Engle, Z. T., Richard, E. M., Breckenridge, A. J., Gran, K. B., Jennings, C. E., and Jalobeanu, A., 2022, Regional-Scale Landscape Response to an Extreme Precipitation Event From Repeat Lidar and Object-Based Image Analysis: Earth and Space Science, v. 9, no. 12.

Dhakal, A. S., and Sullivan, K., 2014, Shallow groundwater response to rainfall on a forested headwater catchment in northern coastal California: implications of topography, rainfall, and throughfall intensities on peak pressure head generation: Hydrological Processes, v. 28, p. 446-463.

DiBiase, R. A., and Lamb, M. P., 2019, Dry sediment loading of headwater channels fuels post-wildfire debris flows in bedrock landscapes: Geology, v. 48, no. 2, p. 189-193.

Dietrich, W. E., Bellugi, D., and de Asua, R. R., 2001, Validation of the shallow landslide model, SHALSTAB, for forest management, *in* Wigmosta, M. S., and Burges, S. J., eds., Land Use and Watersheds: Washington, D.C., American Geophysical Union, p. 195-227.

Dietrich, W. E., Dunne, D., Humphrey, N. F., and Reid, L. M., 1982, Construction of sediment budgets for drainage basins.

Dietrich, W. E., Reiss, R., Hsu, M.-L., and Montgomery, D. R., 1995, A process-based model for colluvial soil depth and shallow landsliding using digital elevation data: Hydrological Processes, v. 9, p. 383-400.

Dieu, J., 2015, North fork Calawah Watershed Analysis, Module K 2015 Mass Wasting Prescription Reanalysis Level 2: Rayonier, Pacific Resource Unit.

Dornik, A., Dragut, L., Oguchi, T., Hayakawa, Y., and Micu, M., 2022, Influence of sampling design on landslide susceptibility modeling in lithologically heterogeneous areas: Sci Rep, v. 12, no. 1, p. 2106.

Dunne, T., 1991, Stochastic aspects of the relations between climate, hydrology and landform evolution: Transactions of the Japanese Geomorpphological Union, v. 12, p. 1-24.

Eitel, J. U. H., Höfle, B., Vierling, L. A., Abellán, A., Asner, G. P., Deems, J. S., Glennie, C. L., Joerg, P. C., LeWinter, A. L., Magney, T. S., Mandlburger, G., Morton, D. C., Müller, J., and Vierling, K. T., 2016, Beyond 3-D: The new spectrum of lidar applications for earth and ecological sciences: Remote Sensing of Environment, v. 186, p. 372-392.

Emmert-Streib, F., and Dehmer, M., 2019, Introduction to Survival Analysis in Practice: Machine Learning and Knowledge Extraction, v. 1, no. 3, p. 1013-1038.

Erdbrügger, J., van Meerveld, I., Bishop, K., and Seibert, J., 2021, Effect of DEM-smoothing and-aggregation on topographically-based flow directions and catchment boundaries: Journal of Hydrology, v. 602, p. 126717.

Fabbri, A. G., and Patera, A., 2021, Spatial Uncertainty of Target Patterns Generated by Different Prediction Models of Landslide Susceptibility: Applied Sciences, v. 11, no. 8.

Fannin, R. J., and Rollerson, T. P., 1993, Debris flows: some physical characteristics and behavior: Canadian Geotechnical Journal, v. 30, no. 1, p. 71-81.

Fannin, R. J., and Wise, M. P., 2001, An empirical-statistical model for debris flow travel distance: Canadian Geotechnical Journal, v. 38, no. 5, p. 982-994.

Fernández, T., Pérez-García, J. L., Gómez-López, J. M., Cardenal, J., Moya, F., and Delgado, J., 2021, Multitemporal Landslide Inventory and Activity Analysis by Means of Aerial Photogrammetry and LiDAR Techniques in an Area of Southern Spain: Remote Sensing, v. 13, no. 11.

Fleischer, F., Haas, F., Piermattei, L., Pfeiffer, M., Heckmann, T., Altmann, M., Rom, J., Stark, M., Wimmer, M. H., Pfeifer, N., and Becht, M., 2021, Multi-decadal (1953–2017) rock glacier kinematics analysed by high-resolution topographic data in the upper Kaunertal, Austria: The Cryosphere, v. 15, no. 12, p. 5345-5369.

Furbish, D. J., and Roering, J. J., 2013, Sediment disentrainment and the concept of local versus nonlocal transport on hillslopes: Journal of Geophysical Research: Earth Surface, v. 118, no. 2, p. 937-952.

Furbish, D. J., Roering, J. J., Doane, T. H., Roth, D. L., Williams, S. G. W., and Abbott, A. M., 2021, Rarefied particle motions on hillslopes – Part 1: Theory: Earth Surface Dynamics, v. 9, no. 3, p. 539-576.

Gaidzik, K., and Ramirez-Herrera, M. T., 2021, The importance of input data on landslide susceptibility mapping: Sci Rep, v. 11, no. 1, p. 19334.

Ghuffar, S., Székely, B., Roncat, A., and Pfeifer, N., 2013, Landslide Displacement Monitoring Using 3D Range Flow on Airborne and Terrestrial LiDAR Data: Remote Sensing, v. 5, no. 6, p. 2720-2745.

Glennie, C. L., Hinojosa-Corona, A., Nissen, E., Kusari, A., Oskin, M. E., Arrowsmith, J. R., and Borsa, A., 2014, Optimization of legacy lidar data sets for measuring near-field earthquake displacements: Geophysical Research Letters, v. 41, no. 10, p. 3494-3501.

Goetz, J. N., Guthrie, R. H., and Brenning, A., 2015, Forest harvesting is associated with increased landslide activity during an extreme rainstorm on Vancouver Island, Canada: Natural Hazards and Earth System Science, v. 15, no. 6, p. 1311-1330.

Goodwell, A. E., Zhu, Z., Dutta, D., Greenberg, J. A., Kumar, P., Garcia, M. H., Rhoads, B. L., Holmes, R. R., Parker, G., Berretta, D. P., and Jacobson, R. B., 2014, Assessment of floodplain vulnerability during extreme Mississippi River flood 2011: Environ Sci Technol, v. 48, no. 5, p. 2619-2625.

Gorr, A. N., McGuire, L. A., Youberg, A. M., and Rengers, F. K., 2022, A progressive flow-routing model for rapid assessment of debris-flow inundation: Landslides, v. 19, no. 9, p. 2055-2073.

Griswold, J. P., and Iverson, R. M., 2008, Mobility statistics and automated hazard mapping for debris flows and rock avalanches: U.S. Geological Survey.

Guilinger, J. J., Foufoula-Georgiou, E., Gray, A. B., Randerson, J. T., Smyth, P., Barth, N. C., and Goulden, M. L., 2023, Predicting Postfire Sediment Yields of Small Steep Catchments Using Airborne Lidar Differencing: Geophysical Research Letters, v. 50, no. 16.

Guo, Q., Li, W., Yu, H., and Alvarez, O., 2010, Effects of topographic variability and lidar sampling density on several DEM interpolation methods: Photogrammetric Engineering & Remote Sensing, v. 76, no. 6, p. 701-712.

Gupta, S., Hengl, T., Lehmann, P., Bonetti, S., and Or, D., 2021, SoilKsatDB: global database of soil saturated hydraulic conductivity measurements for geoscience applications: Earth System Science Data, v. 13, no. 4, p. 1593-1612.

Gupta, S. K., 2020, Data imbalance in landslide susceptibility zonation: under-sampling for class-imbalance learning: The International Archives of Photogrammetry. Remote Sensing and Spatial Information Sciences, v. 42, p. 51-57.

Guthrie, R., and Befus, A., 2021, DebrisFlow Predictor: an agent-based runout program for shallow landslides: Natural Hazards and Earth System Sciences, v. 21, no. 3, p. 1029-1049.

Guthrie, R. H., Hockin, A., Colquhoun, L., Nagy, T., Evans, S. G., and Ayles, C., 2010a, An examination of controls on debris flow mobility: Evidence from coastal British Columbia: Geomorphology, v. 114, no. 4, p. 601-613.

Guthrie, R. H., Mitchell, S. J., Lanquaye-Opoku, N., and Evans, S. G., 2010b, Extreme weather and landslide initiation in coastal British Columbia: Quarterly Journal of Engineering Geology and Hydrogeology, v. 43, no. 4, p. 417-428.

Guzzetti, F., Malamud, B. D., Turcotte, D. L., and Reichenbach, P., 2002, Power-law correlations of landslide areas in central Italy: Earth and Planetary Science Letters, v. 195, p. 169-183.

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., and Chang, K.-T., 2012, Landslide inventory maps: New tools for an old problem: Earth-Science Reviews, v. 112, p. 42-66.

Habib, M., Alzubi, Y., Malkawi, A., and Awwad, M., 2020, Impact of interpolation techniques on the accuracy of large-scale digital elevation model: Open Geosciences, v. 12, no. 1, p. 190-202.

Heckmann, T., and Vericat, D., 2018, Computing spatially distributed sediment delivery ratios: inferring functional sediment connectivity from repeate high-resolution digital elevation models: Earth Surface Process and Landforms, v. 43, no. 7, p. 1547-1554.

Hobley, D. E. J., Adams, J. M., Nudurupati, S. S., Hutton, E. W. H., Gasparini, N. M., Istanbulluoglu, E., and Tucker, G. E., 2017, Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics: Earth Surface Dynamics, v. 5, no. 1, p. 21-46.

Hofmeister, R. J., and Miller, D. J., 2003, GIS-based modeling of debris-flow initiation, transport and deposition zones for regional hazard assessments in western, Oregon, USA, *in* Reickenmann, and Chen, eds., Debris-Flow Hazards Mitigation: Mechanics, Prediction, and Assessment: Rotterdam, Millpress, p. 1141-1149.

Hofmeister, R. J., Miller, D. J., Mills, K. A., Hinkle, J. C., and Beier, A. E., 2002, Hazard map of potential rapidly moving landslides in western Oregon, Interpretive Map Series - 22: Oregon Department of Geology and Mineral Industries.

Horton, P., Jaboyedoff, M., Rudaz, B., and Zimmermann, M., 2013, Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale: Natural Hazards and Earth System Sciences, v. 13, p. 869-885.

Hu, Q., Li, Z., Wang, L., Huang, Y., Wang, Y., and Li, L., 2019, Rainfall Spatial Estimations: A Review from Spatial Interpolation to Multi-Source Data Merging: Water, v. 11, no. 3.

Hui, Z., Jin, S., Xia, Y., Nie, Y., Xie, X., and Li, N., 2021, A mean shift segmentation morphological filter for airborne LiDAR DTM extraction under forest canopy: Optics & Laser Technology, v. 136.

Hungr, O., Corominas, J., and Eherhardt, E., 2005, Estimating landslide motion mechanism, travel distance and velocity, *in* Hungr, O., Fell, R., Couture, R., and Eberhardt, E., eds., Landslide Risk Management: London, Taylor & Francis Group.

Hungr, O., Leroueil, S., and Picarelli, L., 2014, The Varnes classification of landslide types, an update: Landslides, v. 11, p. 167-194.

Hunter, G., and Fell, R., 2003, Travel distance angle for "rapid" landslides in constructed and natural soil slopes: Canadian Geotechnical Journal, v. 40, p. 1123-1141.

Hussin, H. Y., Zumpano, V., Reichenbach, P., Sterlacchini, S., Micu, M., van Westen, C., and Bălteanu, D., 2016, Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model: Geomorphology, v. 253, p. 508-523.

Hyyppä, H., Yu, X., Hyyppä, J., Kaartinen, H., Kaasalainen, S., Honkavaara, E., and Rönnholm, P., 2005, Factors affecting the quality of DTM generation in forested areas, Laser Scanning 2005, Volume 3: Enschede, the Netherlands, ISPRS, p. 85-90.

Iida, T., 1999, A stochastic hydro-geomorphological model for shallow landsliding due to rainstorm: Catena, v. 34, p. 293-313.

Imaizumi, F., Sidle, R. C., and Kamei, R., 2008, Effects of forest harvesting on the occurrence of landslides and debris flows in steep terrain of central Japan: Earth Surface Processes and Landforms, v. 33, no. 6, p. 827-840.

Iverson, R. M., 2000, Landslide triggering by rain infiltration: Water Resources Research, v. 36, no. 7, p. 1897-1910.

Jasiewicz, J., and Stepinski, T. F., 2013, Geomorphons - a pattern recognition approach to classification and mapping of landforms: Geomorphology, v. 182, p. 147-156.

Jones, J. N., Boulton, S. J., Bennett, G. L., Stokes, M., and Whitworth, M. R. Z., 2021, Temporal variations in landslide distributions following extreme events: implications for landslide susceptibility modelling: Journal of Geophysical Research Earth Surface.

Ju, L.-Y., Zhang, L.-M., and Xiao, T., 2023, Power laws for accurate determination of landslide volume based on high-resolution LiDAR data: Engineering Geology, v. 312.

Kamp, N., Krenn, P., Avian, M., and Sass, O., 2023, Comparability of multi-temporal DTMs derived from different LiDAR platforms: Error sources and uncertainties in the application of geomorphic impact studies: Earth Surface Processes and Landforms, v. 48, no. 6, p. 1152-1175.

Keaton, J. R., Wartman, J., Anderson, S. A., Benoît, J., deLaChapelle, J., Gilbert, R., and Montgomery, D. R., 2014, The 22 March 2014 Oso Landslide, Snohomish County, Washington: National Science Foundation.

Keck, J., Istanbulluoglu, E., Campforts, B., Tucker, G., and Horner-Devine, A.

Keim, R. F., and Skaugset, A. E., 2003, Modelling effects of forest canopies on slope stability: Hydrological Processes, v. 17, p. 1457-1467.

Kharroubi, A., Poux, F., Ballouch, Z., Hajji, R., and Billen, R., 2022, Three Dimensional Change Detection Using Point Clouds: A Review: Geomatics, v. 2, no. 4, p. 457-485.

Kidd, C., and Levizzani, V., 2022, Satellite rainfall estimation, *in* Morbidelli, R., ed., Rainfall.

Kirschbaum, D., and Stanley, T., 2018, Satellite-based assessment of rainfall-triggered landslide hazard for situational awareness: Earth's Future.

Kleinbaum, D. G., and Klein, M., 2012, Survival Analysis. A Self-Learning Text, Springer.

Kong, J., Rios, T., Kowalczyk, W., Menzel, S., and Bäck, T., 2020, On the performance of oversampling techniques for class imbalance problems, *in* Lauw, H., Wong, R., Ntooulas, A., Lim, E., Ng, S., and Pan, S., eds., Advances in Knowledge Discovery and Data Mining. PAKDD 2020. Lecture Notes in Computer Science, Volume 12085, Springer.

Kusari, A., Glennie, C. L., Brooks, B. A., and Ericksen, T. L., 2019, Precise Registration of Laser Mapping Data by Planar Feature Extraction for Deformation Monitoring: IEEE Transactions on Geoscience and Remote Sensing, v. 57, no. 6, p. 3404-3422.

Lague, D., Brodu, N., and Leroux, J., 2013, Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z): ISPRS Journal of Photogrammetry and Remote Sensing, v. 82, p. 10-26.

Lallias-Tacon, S., Liébault, F., and Piégay, H., 2014, Step by step error assessment in braided river sediment budget using airborne LiDAR data: Geomorphology, v. 214, p. 307-323.

Lancaster, S. T., Hayes, S. K., and Grant, G. E., 2003, Effects of wood on debris flow runout in small mountain watersheds: Water Resources Research, v. 39, no. 6, p. doi:10.1029/2001WR001227.

Lane, S. N., Westaway, R. M., and Murray Hicks, D., 2003, Estimation of erosion and deposition volumes in a large, gravel-bed, braided river using synoptic remote sensing: Earth surface processes and landforms: the journal of the British Geomorphological Research Group, v. 28, no. 3, p. 249-271.

Larsen, I. J., Montgomery, D. R., and Korup, O., 2010, Landslide erosion controlled by hillslope material: Nature Geoscience, v. 3, no. 4, p. 247-251.

Lee, C.-T., 2017, Landslide trends under extreme climate events: Terrestrial, Atmospheric and Oceanic Sciences, v. 28, no. 1, p. 33-42.

Legros, F., 2002, The mobility of long-runout landslides: Engineering Geology, v. 63, p. 301-331.

Li, Z., Chen, M., Gao, S., Hong, Z., Tang, G., Wen, Y., Gourley, J. J., and Hong, Y., 2020, Cross-Examination of Similarity, Difference and Deficiency of Gauge, Radar and Satellite Precipitation Measuring Uncertainties for Extreme Events Using Conventional Metrics and Multiplicative Triple Collocation: Remote Sensing, v. 12, no. 8.

Lima, P., Steger, S., Glade, T., and Murillo-García, F. G., 2022, Literature review and bibliometric analysis on data-driven assessment of landslide susceptibility: Journal of Mountain Science, v. 19, no. 6, p. 1670-1698.

Liu, J., Chen, X., Lin, H., Liu, H., and Song, H., 2013, A simple geomorphic-based analytical model for predicting the spatial distribution of soil thickness in headwater hillslopes and catchments: Water Resources Research, v. 49, no. 11, p. 7733-7746.

Mahalingam, R., Olsen, M. J., and O'Banion, M. S., 2016, Evaluation of landslide susceptibility mapping techniques using lidar-derived conditioning factors (Oregon case study): Geomatics, Natural Hazards and Risk.

Malamud, B. D., Turcotte, D. L., Guzzetti, F., and Reichenbach, P., 2004, Landslide inventories and their statistical properties: Earth Surface Processes and Landforms, v. 29, no. 6, p. 687-711.

Marc, O., Gosset, M., Saito, H., Uchida, T., and Malet, J. P., 2019, Spatial Patterns of Storm-Induced Landslides and Their Relation to Rainfall Anomaly Maps: Geophysical Research Letters, v. 46, no. 20, p. 11167-11177.

Marden, M., Basher, L., Phillips, C., and Black, R., 2015, Should detailed terrain stability or erosion susceptibility mapping be mandatory in erodible steep lands?: New Zealand Journal of Forestry, v. 59, no. 4, p. 32-42.

Martin, Y., Rood, K., Schwab, J. W., and Church, M., 2002, Sediment transfer by shallow landsliding in the Queen Charlotte Islands, British Columbia: Canadian Journal of Earth Science, v. 39.

Massey, C. I., Townsend, D., Jones, K., Lukovic, B., Rhoades, D., Morgenstern, R., Rosser, B., Ries, W., Howarth, J., Hamling, I., Petley, D., Clark, M., Wartman, J., Litchfield, N., and Olsen, M., 2020, Volume Characteristics of Landslides Triggered by the M

W

7.8 2016 Kaikōura Earthquake, New Zealand, Derived From Digital Surface Difference Modeling: Journal of Geophysical Research: Earth Surface, v. 125, no. 7.

Matasci, G., Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Hobart, G. W., and Zald, H. S. J., 2018, Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots: Remote Sensing of Environment, v. 209, p. 90-106.

Maxwell, A. E., and Shobe, C. M., 2022, Land-surface parameters for spatial predictive mapping and modeling: Earth-Science Reviews, v. 226.

May, C. L., 2002, Debris flows through different forest age classes in the central Oregon Coast Range: Journal of the American Water Resources Association, v. 38, no. 4, p. 1-17.

May, C. L., and Gresswell, R. E., 2003, Processes and rates of sediment and wood accumulation in headwater streams of the Oregon Coast Range, USA: Earth Surface Processes and Landforms, v. 28, no. 4, p. 409-424.

McDougall, S., 2017, Landslide runout analysis - current practice and challenges: Canadian Geotechnical Journal, v. 54, p. 605-620.

Menardi, G., and Torelli, N., 2012, Training and assessing classification rules with imbalanced data: Data Mining and Knowledge Discovery, v. 28, no. 1, p. 92-122.

Meng, X., Currit, N., and Zhao, K., 2010, Ground Filtering Algorithms for Airborne LiDAR Data: A Review of Critical Issues: Remote Sensing, v. 2, no. 3, p. 833-860.

Mergili, M., Schwarz, L., and Kociu, A., 2019, Combining release and runout in statistical landslide susceptibility modeling: Landslides, v. 16, no. 11, p. 2151-2165.

Miller, D. J., 1995, Coupling GIS with physical models to assess deep-seated landslide hazards: Environmental & Engineering Geoscience, v. 1, no. 3, p. 263-276.

Miller, D. J., and Burnett, K. M., 2007, Effects of forest cover, topography, and sampling extent on the measured density of shallow, translational landslides: Water Resources Research, v. 43, p. 23.

-, 2008, A probabilistic model of debris-flow delivery to stream channels, demonstrated for the Coast Range of Oregon, USA: Geomorphology, v. 94, p. 184-205.

Minár, J., Evans, I. S., and Jenčo, M., 2020, A comprehensive system of definitions of land surface (topographic) curvatures, with implications for their application in geoscience modelling and prediction: Earth-Science Reviews, v. 211.

Mirus, B. B., Smith, J. B., and Baum, R. L., 2017, Hydrologic impacts of landslide disturbances: implications for remobilization and hazard persistence.: Water Resources Research, v. 53, p. 8250-8265.

Miskovic, T., and Powell, J., 2009, Mass wasting prescription-scale effectiveness monitoring project (post-mortem) quality assurance/quality control (QA/QC) report: WA Department of Natural Resources.

Molter, E. M., Collins, W. D., and Risser, M. D., 2021, Quantitative Precipitation Estimation of Extremes in CONUS With Radar Data: Geophysical Research Letters, v. 48, no. 16.

Montealegre, A. L., Lamelas, M. T., and de la Riva, J., 2015, A Comparison of Open-Source LiDAR Filtering Algorithms in a Mediterranean Forest Environment: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, v. 8, no. 8, p. 4072-4085.

Montgomery, D. R., and Dietrich, W. E., 1994, A physically based model for the topographic control on shallow landsliding: Water Resources Research, v. 30, no. 4, p. 1153-1171.

Montgomery, D. R., Schmidt, K. M., Greenberg, H. M., and Dietrich, W. E., 2000, Forest clearing and regional landsliding: Geology, v. 28, no. 4, p. 311-314.

Morell, K. D., Alessio, P., Dunne, T., and Keller, E., 2021, Sediment Recruitment and Redistribution in Mountain Channel Networks by Post-Wildfire Debris Flows: Geophysical Research Letters, v. 48, no. 24.

Morino, C., Conway, S. J., Balme, M. R., Hillier, J., Jordan, C., Saemundsson, Þ., and Argles, T., 2019, Debris-flow release processes investigated through the analysis of multi-temporal LiDAR datasets in north-western Iceland: Earth Surface Processes and Landforms, v. 44, no. 1, p. 144-159.

Müller, L., Rutzinger, M., Mayr, A., and Kollert, A., 2022, Airborne laser scanning change detection for quantifying geomorphological processes in high mountain regions: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXIV ISPRS Congress, p. 391-398.

Murphy, B. O., Sarikhan, I., and Slaughter, S. L., 2013, Southern Willapa Hills retrospective study: Washington State Department of Natural Resources, Forest Practices Division.

Nissen, E., Krishnan, A. K., Arrowsmith, J. R., and Saripalli, S., 2012, Three-dimensional surface displacements and rotations from differencing pre- and post-earthquake LiDAR point clouds: Geophysical Research Letters, v. 39, no. 16, p. n/a-n/a.

Ohmann, J. L., and Gregory, M. J., 2002, Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, USA: Canadian Journal of Forest Resources, v. 32, p. 725-741.

Okyay, U., Telling, J., Glennie, C. L., and Dietrich, W. E., 2019, Airborne lidar change detection: An overview of Earth sciences applications: Earth-Science Reviews, v. 198.

Ombadi, M., Nguyen, P., Sorooshian, S., and Hsu, K.-l., 2021, How much information on precipitation is contained in satellite infrared imagery?: Atmospheric Research, v. 256, p. 105578.

Oommen, T., Baise, L. G., and Vogel, R. M., 2010, Sampling Bias and Class Imbalance in Maximum-likelihood Logistic Regression: Mathematical Geosciences, v. 43, no. 1, p. 99-120.

Orem, C. A., and Pelletier, J. D., 2015, Quantifying the time scale of elevated geomorphic response following wildfires using multi-temporal LiDAR data: An example from the Las Conchas fire, Jemez Mountains, New Mexico: Geomorphology, v. 232, p. 224-238.

Orlandini, S., Moretti, G., Corticelli, M. A., Santangelo, P. E., Capra, A., Rivola, R., and Albertson, J. D., 2012, Evaluation of flow direction methods against field observations of overland flow dispersion: Water Resources Research, v. 48, no. W10523.

Otto, J., and Smith, M. J., 2013, Geomorphological mapping, in Cook, S. J., Clarke, L., E., and Nield, J. M., eds., Geomorpholocial Techniques: London, British Society for Geomorphology.

Pack, R. T., Tarboton, D. G., and Goodwin, C. N., 1998, The SINMAP approach to Terrain Stability Mapping, 8th Congress of the International Association of Engineering Geology: Vancouver, British Columbia, Canada.

Passalacqua, P., Belmont, P., Staley, D. M., Simley, J. D., Arrowsmith, J. R., Bode, C. A., Crosby, C., DeLong, S. B., Glenn, N. F., Kelly, S. A., Lague, D., Sangireddy, H., Schaffrath, K., Tarbon, D. G., Waklewicz, T., and Wheaton, J. M., 2015, Analyzing high resolution topography for advancing the understanding of mass and entergy transfer through landscapes: a review: Earth-Science Reviews, v. 148, p. 174-193.

Patton, N. R., Lohse, K. A., Godsey, S. E., Crosby, B. T., and Seyfried, M. S., 2018, Predicting soil thickness on soil mantled hillslopes: Nat Commun, v. 9, no. 1, p. 3329.

Paulin, G. L., Mickelson, K. A., Contreras, T. A., Gallin, W., Jacobacci, K. E., and Bursik, M., 2022, Assessing landslide volume using two generic models: application to landslides in Whatcom County, Washington, USA: Landslides, v. 19, no. 4, p. 901-912.

Pelletier, J. D., and Orem, C. A., 2014, How do sediment yields from post-wildfire debris-laden flows depend on terrain slope, soil burn severity class, and drainage basin area? Insights from airborne-LiDAR change detection: Earth Surface Processes and Landforms, v. 39, no. 13, p. 1822-1832.

Petras, V., Petrasova, A., McCarter, J. B., Mitasova, H., and Meentemeyer, R. K., 2023, Point Density Variations in Airborne Lidar Point Clouds: Sensors (Basel), v. 23, no. 3.

Pierce, D. W., Su, L., Cayan, D. R., Risser, M. D., Livneh, B., and Lettenmaier, D. P., 2021, An extreme-preserving long-term gridded daily precipitation data set for the conterminous United States: Journal of Hydrometeorology.

Piovan, S. E., Hodgson, M. E., Mozzi, P., Porter, D. E., and Hall, B., 2023, LiDAR-change-based mapping of sediment movement from an extreme rainfall event: GIScience & Remote Sensing, v. 60, no. 1.

Prat, O. P., and Nelson, B. R., 2015, Evaluation of precipitation estimates over CONUS derived from satellite, radar, and rain gauge data sets at daily to annual scales (2002–2012): Hydrology and Earth System Sciences, v. 19, no. 4, p. 2037-2056.

Prokešová, R., Kardoš, M., Tábořík, P., Medveďová, A., Stacke, V., and Chudý, F., 2014, Kinematic behaviour of a large earthflow defined by surface displacement monitoring, DEM differencing, and ERT imaging: Geomorphology, v. 224, p. 86-101.

Pyles, M. R., and Froehlich, H. A., 1987, Discussion of "Rates of landsliding as impacted by timber managment activities in northwestern California", by M. Wolfe and J. Williams: Bulletin of the Association of Engineering Geologists, v. 24, no. 3, p. 425-431.

Qin, R., Tian, J., and Reinartz, P., 2016, 3D change detection - approaches and applications: ISPRS Journal of Photogrammetry and Remote Sensing, v. 122, p. 41-56.

Rajulapati, C. R., Papalexiou, S. M., Clark, M. P., Razavi, S., Tang, G., and Pomeroy, J. W., 2020, Assessment of Extremes in Global Precipitation Products: How Reliable Are They?: Journal of Hydrometeorology, v. 21, no. 12, p. 2855-2873.

Rechberger, C., Fey, C., and Zangerl, C., 2021, Structural characterisation, internal deformation, and kinematics of an active deep-seated rock slide in a valley glacier retreat area: Engineering Geology, v. 286.

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., and Guzzetti, F., 2018, A review of statistically-based landslide susceptibility models: Earth-Science Reviews, v. 180, p. 60-91.

Reid, L. M., 1998, Calculation of average landslide frequency using climatic records: Water Resources Research, v. 34, no. 4, p. 869-877.

Reid, L. M., and Dunne, T., 1996, Rapid Evaluation of Sediment Budgets, Reiskirchen, Germany, Catena Verlag, 164 p.:

Reid, L. M., and Page, M. J., 2002, Magnitude and frequency of landsliding in a large New Zealand catchment: Geomorphology, v. 49, p. 71-88.

Reid, M. E., Coe, J. A., and Brien, D. L., 2016, Forecasting inundation from debris flows that grow volumetrically during travel, with application to the Oregon Coast Range, USA: Geomorphology, v. 273, p. 396-411.

Reneau, S. L., Dietrich, W. E., Rubin, M., Donahue, D. J., and Jull, A. J. T., 1989, Analysis of hillslope erosion rates using dated colluvial deposits: The Journal of Geology, v. 97, no. 1, p. 45-63.

Rengers, F. K., McGuire, L. A., Kean, J. W., Staley, D. M., Dobre, M., Robichaud, P. R., and Swetnam, T., 2021, Movement of Sediment Through a Burned Landscape: Sediment Volume Observations and Model Comparisons in the San Gabriel Mountains, California, USA: Journal of Geophysical Research: Earth Surface, v. 126, no. 7.

Rickenmann, D., 2005, Runout predition methods, *in* Jokob, M., and Hungr, O., eds., Debris-flow Hazards and Related Phenomena: Heidelbert, Springer.

Rizopoulos, D., 2012, Joint Models for Longitudinal and Time-to-Event Data with Applications in R, CRC Press, Biostatistics Series.

-, 2016, The R Package JMbayes for Fitting Joint Models for Longitudinal and Time-to-Event Data Using MCMC: Journal of Statistical Software, v. 72, no. 7.

Robison, G. E., Mills, K. A., Paul, J., Dent, L., and Skaugset, A., 1999, Storm impacts and landslides of 1996: final report: Oregon Department of Forestry, 4.

Rossi, M., Luciani, S., Valigi, D., Kirschbaum, D., Brunetti, M. T., Peruccacci, S., and Guzzetti, F., 2017, Statistical approaches for the definition of landslide rainfall thresholds and their uncertainty using rain gauge and satellite data: Geomorphology, v. 285, p. 16-27.

Sailer, R., Bollmann, E., Hoinkes, S., Rieg, L., Sproß, M., and Stötter, J., 2012, Quantification of geomorphodynamics in glaciated and recently deglaciated terrain based on airborne laser scanning data: Geografiska Annaler: Series A, Physical Geography, v. 94, no. 1, p. 17-32.

Schaffrath, K. R., Belmont, P., and Wheaton, J. M., 2015, Landscape-scale geomorphic change detection: Quantifying spatially variable uncertainty and circumventing legacy data issues: Geomorphology, v. 250, p. 334-348.

Scheidl, C., Rickenmann, D., and Chiari, M., 2008, The use of airborne LiDAR data for the analysis of debris flow events in Switzerland: Natural Hazards and Earth System Science, v. 8, p. 1113-1127.

Scheip, C., and Wegmann, K., 2022, Insights on the growth and mobility of debris flows from repeat high-resolution lidar: Landslides, v. 19, no. 6, p. 1297-1319.

Scheip, C. M., and Wegmann, K. W., 2021, HazMapper: a global open-source natural hazard mapping application in Google Earth Engine: Natural Hazards and Earth System Sciences, v. 21, no. 5, p. 1495-1511.

Schilling, S. P., 1998, LAHARZ: GIS programs for automated mapping of lahar-inundation hazard zones: U.S. Geological Survey, Open-File Report 98-638.

Schmidt, K. M., Roering, J. J., Stock, J. D., Dietrich, W. E., Montgomery, D. R., and Schaub, T., 2001, The variability of root cohesion as an influence on shallow landslide susceptibility in the Oregon Coast Range: Canadian Geotechnical Journal, v. 38, p. 995-1024.

Schwab, J. W., and Geertsema, M., 2010, Terrain stability mapping on British Columbia forest lands: an historical perspective: Natural Hazards, v. 53, no. 1, p. 63-75.

Scorpio, V., Cavalli, M., Steger, S., Crema, S., Marra, F., Zaramella, M., Borga, M., Marchi, L., and Comiti, F., 2022, Storm characteristics dictate sediment dynamics and geomorphic changes in mountain channels: A case study in the Italian Alps: Geomorphology, v. 403.

Scott, C., Phan, M., Nandigam, V., Crosby, C., and Arrowsmith, J. R., 2021, Measuring change at Earth's surface: On-demand vertical and three-dimensional topographic differencing implemented in OpenTopography: Geosphere, v. 17, no. 4, p. 1318-1332.

Scott, C. P., Beckley, M., Phan, M., Zawacki, E., Crosby, C., Nandigam, V., and Arrowsmith, R., 2022, Statewide USGS 3DEP Lidar Topographic Differencing Applied to Indiana, USA: Remote Sensing, v. 14, no. 4.

Scott, C. P., DeLong, S. B., and Arrowsmith, J. R., 2020, Distribution of Aseismic Deformation Along the Central San Andreas and Calaveras Faults From Differencing Repeat Airborne Lidar: Geophysical Research Letters, v. 47, no. 22.

Shao, X., Ma, S., Xu, C., and Zhou, Q., 2020, Effects of sampling intensity and non-slide/slide sample ratio on the occurrence probability of coseismic landslides: Geomorphology, v. 363.

Sidle, R. C., 1991, A conceptual model of changes in root cohesion in response to vegetation management: Journal of Environmental Quality, v. 20, p. 43-52.

Silva, C. A., Klauberg, C., Hentz, Â. M. K., Corte, A. P. D., Ribeiro, U., and Liesenberg, V., 2018, Comparing the Performance of Ground Filtering Algorithms for Terrain Modeling in a Forest Environment Using Airborne LiDAR Data: Floresta e Ambiente, v. 25, no. 2.

Sîrbu, F., Drăguţ, L., Oguchi, T., Hayakawa, Y., and Micu, M., 2019, Scaling land-surface variables for landslide detection: Progress in Earth and Planetary Science, v. 6, no. 1.

Skempton, A. W., and deLory, F. A., 1957, Stability of Natural Slopes in London Clay, 4th International Conference on Soil Mechanics and Foundation Engineering, Earth Dams, Slopes and Open Excavations: London.

Slaughter, S. L., 2015, Landslide inventory in Washington State: the past, present, and future, AEG Professional Forum: Time to Face the Landslide Hazard Dilemma: Briding Science, Policy, Public Safety, and Potential Loss: Seattle, WA, p. 51-54.

Slaughter, S. L., Burns, W. J., Mickelson, K. A., Jacobacci, C. E., Biel, A., and Contreras, T. A., 2017, Protocol for landslide inventory mapping from lidar data in Washington state: Washington Geological Survey.

Sokol, Z., Szturc, J., Orellana-Alvear, J., Popová, J., Jurczyk, A., and Célleri, R., 2021, The Role of Weather Radar in Rainfall Estimation and Its Application in Meteorological and Hydrological Modelling—A Review: Remote Sensing, v. 13, no. 3.

Stanley, T. A., Kirschbaum, D. B., Sobieszczyk, S., Jasinski, M. F., Borak, J. S., and Slaughter, S. L., 2020, Building a landslide hazard indicator with machine learning and land surface models: Environmental Modelling & Software, v. 129.

Stark, C. P., and Hovius, H., 2001, The characterization of landslide size distributions: Geophysical Research Letters, v. 28, no. 6, p. 1091-1094.

Steger, S., Brenning, A., Bell, R., and Glade, T., 2016, The propagation of inventory-based positional errors into statistical landslide susceptibility models: Natural Hazards and Earth System Sciences, v. 16, no. 12, p. 2729-2745.

Steger, S., Brenning, A., Bell, R., and Glade, T., 2017, The influence of systematically incomplete shallow landslide inventories on statistical susceptibility models and suggestions for improvements: Landslides, v. 14, no. 5, p. 1767-1781.

Stewart, G., Dieu, J., Phillips, J., O'Connor, M., and Velduisen, C., 2013, The Mass Wasting Effectiveness Monitoring Project: An examination of the lanslide response to the December 2007 storm in Southwestern Washington: Cooperative Monitoring, Evaluation and Research committee of the Washington State Forest Practices Board.

Strauch, R., Istanbulluoglu, E., Nudurupati, S. S., Bandaragoda, C., Gasparini, N. M., and Tucker, G. E., 2018, A hydroclimatological approach to predicting regional landslide probability using Landlab: Earth Surface Dynamics, v. 6, no. 1, p. 49-75.

Su, Y., Guo, Q., Fry, D. L., Collins, B. M., Kelly, M., Flanagan, J. P., and Battles, J. J., 2015, A Vegetation Mapping Strategy for Conifer Forests by Combining Airborne LiDAR Data and Aerial Imagery: Canadian Journal of Remote Sensing, v. 42, no. 1, p. 1-15.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K. L., 2018, A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons: Reviews of Geophysics, v. 56, no. 1, p. 79-107.

Swanson, F. J., and Dyrness, C. T., 1975, Impact of clearcutting and road construction on soil erosion by landslides in the western Cascade Range, Oregon: Geology, v. 3, no. 7, p. 393-396.

Szypuła, B., 2017, Geomorphometric comparison of DEMs built by different interpolation methods: Landform Analysis, v. 32, p. 45-58.

Tarboton, D. G., 1997, A new method for the determination of flow directions and upslope areas in grid digital elevation models: Water Resources Research, v. 33, no. 2, p. 309-319.

Temme, A., Guzzetti, F., Samia, J., and Mirus, B. B., 2020, The future of landslides' past—a framework for assessing consecutive landsliding systems: Landslides, v. 17, no. 7, p. 1519-1528.

Theler, D., Reynard, E., Lambiel, C., and Bardou, E., 2010, The contribution of geomorphological mapping to sediment transfer evaluation in small alpine catchments: Geomorphology, v. 124, no. 3-4, p. 113-123.

Therneau, T., 2022, A package for survival analysis in R, Volume https://cran.r-project.org/web/packages/survival/vignettes/survival.pdf.

Thomas, M. A., Lindsay, D. N., Cavagnaro, D. B., Kean, J. W., McCoy, S. W., and Graber, A. P., 2023, The Rainfall Intensity-Duration Control of Debris Flows After Wildfire: Geophysical Research Letters, v. 50, no. 10.

Thornton, P. E., Shrestha, R., Thornton, M., Kao, S. C., Wei, Y., and Wilson, B. E., 2021, Gridded daily weather data for North America with comprehensive uncertainty quantification: Sci Data, v. 8, no. 1, p. 190.

Timmermans, B., Wehner, M., Cooley, D., O'Brien, T., and Krishnan, H., 2019, An evaluation of the consistency of extremes in gridded precipitation data sets: Climate Dynamics, v. 52, no. 11, p. 6651-6670.

Titti, G., van Westen, C., Borgatti, L., Pasuto, A., and Lombardo, L., 2021, When Enough Is Really Enough? On the Minimum Number of Landslides to Build Reliable Susceptibility Models: Geosciences, v. 11, no. 11.

Turner, T. R., Duke, S. D., Fransen, B. R., Reiter, M. L., Kroll, A. J., Ward, J. W., Bach, J. L., Justice, T. E., and Bilby, R. E., 2010, Landslide densities associated with rainfall, stand age, and topography on forested landscapes, southwestern Washington, USA: Forest Ecology and Management, v. 259, no. 12, p. 2233-2247.

UPSAG, 2006, Landslide Hazard Zonation Project Protocol: Olympia, Washington Department of Natural Resources.

van den Goorbergh, R., van Smeden, M., Timmerman, D., and Van Calster, B., 2022, The harm of class imbalance corrections for risk prediction models: illustration and simulation using logistic regression: J Am Med Inform Assoc, v. 29, no. 9, p. 1525-1534.

Viedma, O., 2022, Applying a Robust Empirical Method for Comparing Repeated LiDAR Data with Different Point Density: Forests, v. 13, no. 3.

Wallis, J. R., Schaefer, M. G., Barker, B. L., and Taylor, G. H., 2007, REgional precipitation-frequency analysis and spatial mapping for 24-hour and 2-hour durations for Washington State: Hydrology and Earth System Science, v. 11, no. 1, p. 415-442.

Washington Forest Practices Board, 2011, Board Manual: Standard Methodology for Conducting Watershed Analysis.

Werner, A. T., Schnorbus, M. A., Shrestha, R. R., Cannon, A. J., Zwiers, F. W., Dayon, G., and Anslow, F., 2019, A long-term, temporally consistent, gridded daily meteorological dataset for northwestern North America: Nature, v. 6, p. 180299.

Wheaton, J. M., Brasington, J., Darby, S. E., and Sear, D. A., 2010, Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets: Earth Surface Processes and Landforms, v. 35, p. 136-156.

Wieczorek, G. F., 1984, Preparing a detailed landslide-inventory map for hazard evaluation and reduction: Bulletin of the Association of Engineering Geologists, v. 21, p. 337-342.

Wieczorek, G. F., 1987, Effect of rainfall intensity and duration on debris flows in central Santa Cruz Mountains, California, *in* Costa, J. E., and Wieczorek, G. F., eds., Debris Flows / Avalanches: Process, Recognition, and Mitigation, Volume 7: Boulder, CO, The Geological Society of America, p. 93-104.

Williams, S. G. W., and Furbish, D. J., 2021, Particle energy partitioning and transverse diffusion during rarefied travel on an experimental hillslope: Earth Surface Dynamics, v. 9, no. 4, p. 701-721.

Wilson, J. P., Aggett, G., Yongxin, D., and Lam, C. S., 2008, Water in the landscape: a review of contemporary flow routing algorithms, *in* Zhou, Q., Lees, B., and Tang, G.-a., eds., Advances in digital terrain analysis: Berlin, Springer, p. 213-236.

Winiwarter, L., Anders, K., and Höfle, B., 2021, M3C2-EP: Pushing the limits of 3D topographic point cloud change detection by error propagation: ISPRS Journal of Photogrammetry and Remote Sensing, v. 178, p. 240-258.

Wood, S. N., 2017, Generalized Additive Models; An Introduction with R, CRC Press, Statistical Science.

Wu, W., and Sidle, R. C., 1995, A distributed slope stability model for steep forested basins: Water Resources Research, v. 31, no. 8, p. 2097-2110.

Yordanov, V., and Brovelli, M. A., 2020, Comparing model performance metrics for landslide susceptibility mapping: ISPRS Congress-Technical Commission III, v. 43, no. 3.

Zhang, X., Glennie, C., and Kusari, A., 2015, Change Detection From Differential Airborne LiDAR Using a Weighted Anisotropic Iterative Closest Point Algorithm: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, v. 8, no. 7, p. 3338-3346.