DEEP-SEATED LANDSLIDE SUSCEPTIBILITY MAP OF NEW MEXICO

New Mexico Bureau of Geology and Mineral Resources Open-File Report OFR-594

Prepared by

Colin T. Cikoski, *Geologist*

Daniel J. Koning, *Senior Field Geologist*

New Mexico Bureau of Geology and Mineral Resources
New Mexico Tech, 801 Leroy Place
Socorro, NM 87801

Prepared for

**New Mexico Department of Homeland Security and Emergency Management**
13 Bataan Blvd.
Santa Fe, NM 87508

December, 2017

ACKNOWLEDGEMENTS

DISCLAIMER

Contents

Figures

Tables

Plates

Plate 1: Landslide susceptibility map of New Mexico

Appendices (Digital)

A) Input GIS files (Esri geodatabase)

B) Continuous data binning tables (Esri geodatabase)

C) Individual province model probability results (Esri geodatabase)

D) Statewide deep-seated landslide susceptibility (Esri geodatabase)

E) Scripts used for data processing (folder)

F) Indicator variable training area coverage assessment (folder)

G) Indicator variable independence evaluation (folder)

H) Spatial subsampling full results (folder)

I) Individual province models (folder)

1 Executive Summary

We used logistic regression methods to construct a map of deep-seated landslide susceptibility for the state of New Mexico. Deep-seated landslides involve slide planes deeper than unconsolidated surficial material, which is typically greater than 3 m in depth. As used here, "susceptibility" refers only to the propensity of a portion of the landscape to fail as a landslide, irrespective of driving forces such as heavy precipitation or earthquakes. We therefore only considered landscape indicator variables, and trained the logistic regression models using existing statewide maps of deep-seated landslide deposits.

Given the diversity in topographic and geologic settings across the state, we chose to divide the state into six distinct physiographic provinces and derive a separate model for each province. Within each province, we designated a specific region to be a training area, from which data would be extracted to derive the model, and a separate external validation area, from which data would be extracted to test the robustness of the model. An internal validation, using the same validation process but using data collected at random from the original training area, was also performed.

We considered numerous landscape variables as potential indicators for landslide susceptibility including elevation, slope angle, slope curvature, slope aspect, slope and surface roughness, geologic unit, precipitation, distance to a fault, distance to a major stream, and distance to a topographic escarpment. Each variable was binned or categorized in such a way as to compress the range of values in each input and to facilitate the capacity for the training areas to represent the remainder of each physiographic province. Each binned variable was tested for interdependence, for disproportionate influence, and for the consistency of the relationship between the variable and landslide occurrence at different locations within the training and validation areas. Test results led to revisions in indicator variable categorizations and to a final acceptable indicator variable list of (simplified) geologic unit, slope angle, slope curvature, slope aspect, and, locally, precipitation.

We used a statewide deep-seated landslide deposit map as the source of landslide presence/absence data. However, as this dataset proved to be the least precise in terms of spatial accuracy (lateral inaccuracies commonly 300 to 1,000 m), we relocated, using aerial imagery, much of the map elements in the training and validation areas to improve precision to within 200 m. We extracted data for the logistic regression model by grid-sampling the landslide deposit locations at a minimum 28 m spacing (the pixel resolution of the slope angle dataset) up to a maximum of 100,000 landslide-area training points, recording the values of the indicator variables occurring at each point,. We also collected an equal number of randomly-located non-landslide area points. These were exported to an R-based script (LAND-SE; Rossi and Reichenbach, 2016) to derive the logistic regression model and perform model diagnostics.

This method implicitly assumes that the landscape characteristics of the landslide deposits are also reflective of the characteristics of landslide-susceptible slopes. In certain provinces, particularly those characterized by plateaus and mesas, this assumption appears valid, and final models perform well according to standard measures of discriminatory capacity. For other provinces, particularly those characterized by high-relief mountainous terrain, this assumption may not be as valid. However, application of some slope stability-based restrictions on the parameterization of the slope angle indicator variable in these areas resulted in models that appear to accurately characterize landslide susceptibility.

Each model takes as inputs values of the indicator landscape parameters and outputs a model probability, a value between 0 and 1 that quantifies the potential for landslide susceptibility. Model accuracy was assessed by evaluating the receiver operating characteristic (ROC) curve, the distribution of model probabilities occurring in known landslide areas, and qualitative assessment of the spatial distribution of model probabilities in the training and validation areas. Models with validation results that indicate accurate portrayal of susceptibility were then extrapolated throughout the remainder of each province. Province-wide results were then merged together, with a gradation process applied to the probabilities occurring at province boundaries to produce a seamless map of landslide susceptibility probabilities.

Using the distribution of model probabilities occurring in known landslide areas as a guide, we determined threshold probabilities to distinguish four susceptibility classes: unlikely susceptible (probabilities between 0 and 0.285), potentially susceptible (0.285 to 0.485), moderately likely susceptible (0.485 to 0.685), and likely susceptible (0.685 to 1). We considered the precision of the various input data, and subsequently conservatively downsampled the susceptibility class raster to a 500 m pixel resolution. Based on the histogram of susceptibility classes occurring in known landslide areas across the state, the final map appears to accurately characterize susceptibility, as ~85% of known landslides lie in the likely susceptible class.

The final product is intended to be a coarse-scale, low-resolution guide for planning purposes. It is, however, not suitable for site-specific assessment of landslide susceptibility, and should not be used as a substitute for a focused, high-resolution geotechnical study.

2 Background and Theory

2.1 Purpose and Motivation

The purpose of this project was to produce a statewide map showing landscape susceptibility for deep-seated landsliding. The term susceptibility, as used here, means the propensity inherent in the landscape to experience deep-seated mass-wasting due to features that include slope steepness and rock type. Susceptibility does not consider driving forces such as precipitation events or earthquakes.

Types of deep-seated landslides are expounded on below, but are characterized by having their basal slide plane deeper than unconsolidated surficial material (such as colluvium), which is typically greater than 3 m (10 ft) in depth. In contrast, shallow landslides commonly involve only unconsolidated surficial sediment (such as soils or colluvium) that move downslope as earth slumps, earth-flows, debris avalanches, or debris slides -- which may possibly evolve to debris flows. Deep-seated landslides are ubiquitous in New Mexico's landscape, being especially common on the slopes of mesas in the northwest part of the state, northern Rio Grande valley, and in the northeastern part of the state. Limited field studies indicate that most landslides occurred in wetter climate regimes in the late Pleistocene (130,000 to 10,000 years ago) than those which dominated the Holocene (the past 10,000 years). Thus, it is possible that prolonged (multi-year) periods exhibiting above-average precipitation may simulate late Pleistocene conditions and increase the probability for localized deep-seated slides. Furthermore, construction activity at the foot of a landslide may weaken buttress-like resisting forces and induce deep-seated slope failure. For these reasons, having a deep-seated landslide susceptibility map is useful for construction and safety planning in New Mexico.

There were two main motivations for this project. One, in preparing for the 2018 update for the New Mexico state hazard mitigation plan for the Federal Emergency Management Agency (FEMA), it was realized that landslide hazard was poorly defined for New Mexico. A landslide susceptibility map would serve as a useful tool to identify regions in the state where this risk is non-negligible and should be taken into account by public and private planners. Two, a statewide, deep-seated landslide susceptibility map is a logical and useful derivation product of a pre-existing map that shows the spatial distribution of deep-seated landslides across the entire state (Cardinali et al., 1990). This map was constructed at a scale of 1:500,000 by Mauro Cardinali, Fausto Guzzetti (Research Institute for Hydrogeological Protection in Central Italy), and Earl E. Brabb (U.S. Geological Survey [USGS]). Having a single map of the state compiled by a limited number of authors who are experts in deep-seated landslides, presumably employing consistent criteria, is a major asset for constructing a statewide deep-seated landslide susceptibility map.

As discussed below, there are a number of methods that one could potentially use to produce a deep-seated landslide susceptibility map. Some methods are more appropriate for higher scales (more localized and detailed) and others more appropriate for lower scales (more regional and less detailed). Popular low-scale approaches include qualitative "best-expert" maps, weighted overlay maps and frequency ratios, logistic regression, ordinary least squares, and machine learning methods (cf., Guzzetti et al., 1999; Olsen et al., 2015). We chose to use logistic regression based on its objectiveness, the computational capability of the New Mexico Bureau of Geology and Mineral Resources (NMBGMR), and the availability of a landslide-specific logistic regression computer code (LAND-SE, Rossi and Reichenbach, 2016) that we could modify for our purposes.

2.2 Background

2.2.1 Style of landsliding

Varnes (1978) proposed a classification scheme for landslides that is still widely employed today, although some modifications have been proposed (Hungr et al., 2013). Four types of landslides are differentiated in New Mexico. Of these, rotational deep-seated landslide types are particularly common. These landslides have a concave-up slide plane and exhibit back-rotation of landslide-involved strata (Figure 2.1). A type of rotational landslide that is abundant in New Mexico is called a Toreva-block landslide. Toreva-block landslides are characterized by having resistant ridges of hard rock that are back-rotated in the larger rotational landslide (Figure 2.2). They are common along the slopes of mesas in the Colorado Plateau, in the Rio Grande rift, and in the Great Plains (northeast New Mexico). In these areas, hard caprocks (such as sandstones or lavas) overlie soft, failure-prone strata (e.g., rift basin fill or the Morrison Formation). Translational landslides are characterized by having a relatively planar slide plane, commonly coinciding with a dipping fracture plane or bedding plane (Figures 2.1 and 2.3). Rock avalanches occur when a large mass of rock peels off of a steep slope and disintegrates into boulder-dominated blocks (Figure 2.4). Lastly, earth-flows are locally differentiated; they also may form on the lower end of rotational slides (Figure 2.1).

2.2.2 Recent landslide incidents

A prime example of a recent deep-seated landslide is along the northwestern side of the Rio Grande gorge, at a location 6 km southwest of the town of Picuris (Figure 2.5). At this location, a combination of seeps near the head scarp and cutting of a preexisting landslide toe by the river has reactivated the preexisting landslide. Advancement of the landslide has constricted the Rio Grande, forming Sous rapid (Paul Bauer, pers. comm., 2016). Highway 68 is on the southeast side of the river and has not been damaged by this landslide.

Two other landslides were reported in New Mexico over the past 50 years, although whether they are deep-seated or shallow is ambiguous (NMDHSEM, 2013). One is a landslide that occurred on the Farmers Mutual Ditch in April of 2007 (Figure 2.6). This caused a complete obstruction of the main canal over a length of approximately 300 m. The total cost of repairing the canal was $263,408. Second, a landslide event in Taos, which occurred in June of 1977, caused $50,000 of property damage (NMDHSEM, 2013).

2.2.3 Previous landslides studies in New Mexico

Detailed studies of landslides have been surprisingly sparse in New Mexico and immediate environs. A quintessential study that first identified Toreva block landsliding was done in eastern Arizona (Reiche, 1937). Mega-landslides are present along the eastern escarpment of the Chuska Mountains in northwestern New Mexico. Originally, workers inferred that these were deposited by block glide processes, where individual blocks of capping sandstone separated along vertical joints and slid down on poorly cemented sandstone and shale without backward rotation (Watson and Wright, 1963). However, Cardinali et al. (1990) interpreted these as Toreva block type slides. Watson and Wright (1963) suggest a late Pleistocene age for the Chuska landslide complex. Reneau and Dethier (1996a,b) interpreted that most landslides in the White Rock gorge (Santa Fe County), comprised of rotational and slump type landslides (commonly Toreva block style), occurred in the late Pleistocene. In terms of age, practically all age-constrained landslides in New Mexico occurred in the Pleistocene.
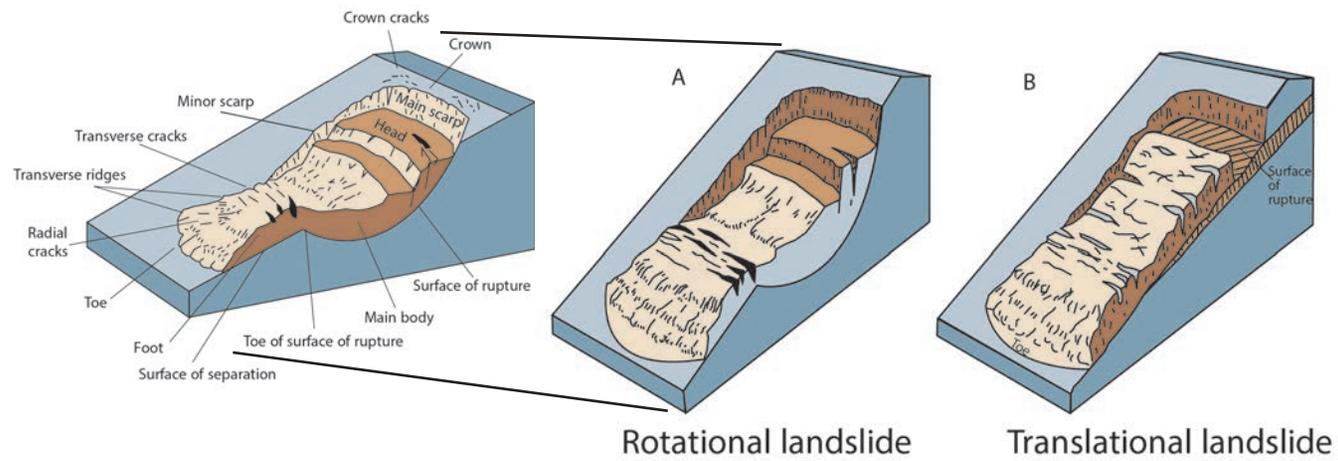
Figure 2.1. Drawings illustrating rotational versus translational landslides. Features within a landslide are noted on the leftmost drawing. From US Geological Survey (2004).
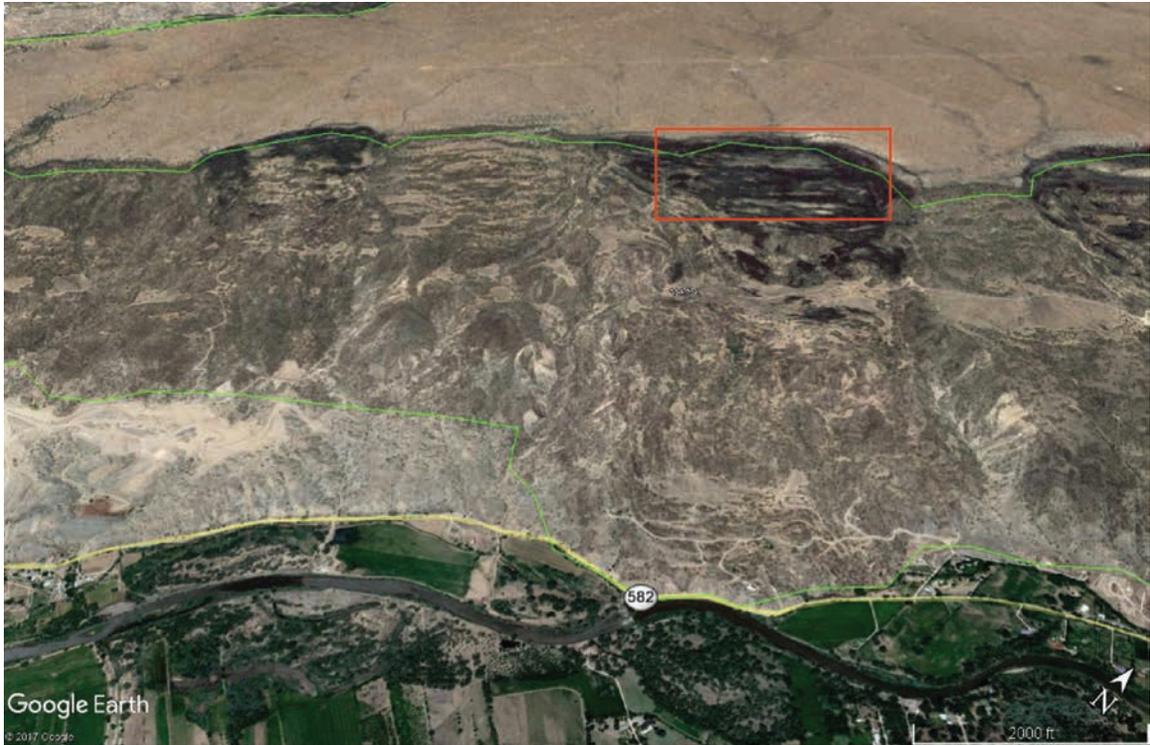
Figure 2.2. Top image: Examples of rotational landslide complexes north of Espanola, on the eastern slope of Black Mesa. This mesa is capped by basalt and underlain by poorly cemented Santa Fe Group basin fill consisting primarily of sand. Green line corresponds to our adjusted landslide boundary. Rio Grande is in the foreground. Red rectangle delineates the area of the lower image. Bottom image: Resistant ridges of basalt that have been displaced and back-rotated. Such resistant ridges are a hallmark of toreva-block landslides. Images courtesy of Google Earth, ©Google, Inc., 2017

6

Figure 2.3. A translational landslide near the crest of the San Andres Mountains, where strata dip steeply to the west.. Here, the bedding plain appears to be relatively planar (based on the lack of visible back-rotation) and probably coincides with a dipping bedding plane. Images courtesy of Google Earth, ©Google, Inc., 2017
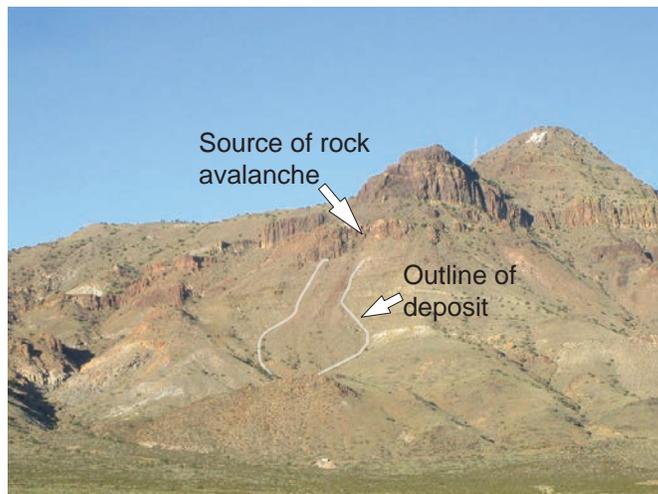
Figure 2.4. Photograph of a rock avalanche on the eastern side of Socorro Peak, immediately west of the town of Socorro. The deposit is demarcated by the thick, light gray line. The inferred source of the rock avalanche is also shown.

Figure 2.5. Landslide southwest of Picuris along the Rio Grande. Approximate extent of landslide deposit is shown in green; state highway 68 lies in the foreground in the lower-left corner of the image. Imagery courtesy of Google Earth, ©Google, Inc., 2017.

Figure 2.6. Example of a recent landslide that occurred along the Farmers Mutual Ditch near Farmington. The main canal was obstructed over about 300 m. This event lead to a State Disaster Declaration, and the total cost for repair was $263,408 (NMDHSEM, 2013).

2.2.4 Overview of regional approaches to assessing landslide susceptibility

There are many approaches available to construct a map showing regional susceptibility to deep-seated landsliding. These can be classified as direct or indirect (Guzzetti et al., 1999). Direct methods typically involve geomorphic mapping of landslides and associated hazards. Instability factors are ranked and weighted, either based on frequency ratios or *a priori* knowledge of the investigator (best-expert opinion). Indirect approaches attempt to predict future patterns of landslide instability from past and present distributions of landslide deposits (Guzzetti et al., 1999). This approach includes such methods as contouring of point densities of landslides, the weight of evidence method, the slope-angle threshold method, and multivariate approaches (e.g., ordinary least squares, discriminant function analysis, logistic regression, and machine learning). Deterministic methods commonly involve geotechnical process approaches, which determine safety factors using slope stability approaches. The high degree of simplification in the geotechnical process approach makes it suitable for fairly uniform ground conditions but not ideal for regional studies (Terlien et al., 1995; Wu and Sidle, 1995; Yilmaz, 2009; Olsen et al., 2015).

Within the direct approach, probably the most popular choice is using the weighted overlay method, commonly used in conjunction with landslide factor frequency ratios. In this method, a set of rasters are generated of particular landscape features, such as slope or rock type. Values are binned in each raster and weights are assigned to these subunits. The weighting is typically done by using frequency ratios (e.g., Lee and Sambath, 2006), although one could also do this by expert opinion. The rasters are overlain spatially in a geographic information system (GIS) and the weights (or frequency ratios) of the raster set over a particular grid is summed. Higher/lower values are then assigned corresponding higher/lower susceptibility values (Lee and Min, 2001; Lee and Sambath, 2006). Examples of the overlay method where weights were assigned using expert-opinion include: Anbalagan (1992), Pachauri and Pant (1992), and Sarkar et al. (1995). Other examples of probabilistic models under the direct approach include Jibson et al. (2000), Luzi et al. (2000), Parise and Jibson (2000), Donati and Turrini (2002), Lee et al. (2002a), Zhou et al. (2002), Lee and Choi (2003), and Lee et al. (2004).

Of the indirect approaches listed in Guzzetti et al., probably the simplest one is to contour the point density of landslides using GIS-processing techniques (Guzzetti et al., 1994, 1999; DeGraff and Canuti, 1998). The presence and density of past deep-seated landslides is a proxy for the susceptibility of a given area.

The weight of evidence method is a probabilistic approach based on a log linear form of Bayes' rule. This method uses the presence or absence of a landslide within an area to calculate a weight value for each landslide predictive factor (Bonham-Carter, 1988). Studies employing this method include Lee and Choi (2004), Lee et al. (2002b), Yilmaz (2009), Pradhan et al. (2010), Neuhauser et al. (2012), and Vakhshoori and Zare (2016).

In Utah, the Utah Geological Survey produced a statewide regional landslide susceptibility map using slope-angle thresholds (Giraud and Shaw, 2007). Landslide slope statistics were compiled as a function of generalized map units. For each generalized unit, the mean and standard deviation was calculated. Pre-existing landslides were categorized as highly susceptible

regardless of slope statistics. Otherwise, for a given generalized map unit, susceptibilities were assigned according to means and standard deviations of the data: 1) moderate (slopes above mean-minus-one-standard deviation); 2) low (slopes between mean-minus-two-standard-deviations and mean-minus-one-standard-deviations); and 3) very low (slopes lower than mean-minus-two-standard deviations).

Multivariate methods commonly involve multiple regression analysis, ordinary least squares analysis, discriminant function analysis, and logistic regression. Ordinary least squares analysis was used successfully by Gorsevski et al. (2000) and Olsen et al. (2015). Discriminant analysis was employed in Taiwan by Lee et al. (2008).

Another popular approach to determining regional susceptibility of landsliding involves using logistic regression modeling. The details of this technique are described below. Notable earlier works employing this technique include Carrara et al. (1991), Rowbotham and Dudycha (1998), Atkinson and Massari (1998), Dai et al. (2001), Dai and Lee (2002), Ohlmacher and Davis (2003), Lee (2004), and Lee and Sambath (2006).

Recently, machine learning has been employed to produce susceptibility maps (Ercamoglu and Gokceoglu, 2002; Pistocchi et al., 2002; Lee et al., 2003a,b, 2004b). The particular methods include artificial neural networks and fuzzy logic. These methods perform comparably or slightly better when compared to the commonly employed frequency ratio overlay model (Yilmaz, 2009; Vakhshoori and Zare, 2016).

2.2.5 Logistic regression modeling

Logistic regression seeks to quantify the relationship between a set of independent indicator variables and a categorical dependent variable, usually a binary dependent variable that can take values of only 0 or 1. Logistic regression determines a set of coefficients $B_0, B_1, B_2, …, B_n$ that, given a set of training data, best fit the following two equations:

$$P(Z) = (1 + e^{-Z})^{-1} \tag{1}$$
$$Z = B_0 + B_1X_1 + B_2X_2 + … + B_nX_n \tag{2}$$

Where:
$X_1, X_2, X_3, …, X_n$ are independent indicator variables.

P(Z) will vary between 0 and 1, as shown in Figure 2.7. The training data is a set of known binary responses and associated values of the selected indicator variables. The output of equation (1) is interpreted to be the probability that a given set of indicator variable values will be associated with a positive response (i.e., that the dependent variable will be 1). In the context of prediction, a threshold probability is typically assumed or otherwise selected (often taken to be 0.5) that categorizes the probability results of function P(Z) as a prediction of 0 (absence) or 1 (occurrence). Ideally, the function P(Z) has a sharp transition as shown in Figure 2.7 that well-defines the threshold.

In the context of landslide susceptibility mapping, the independent indicator variables are landscape or geologic parameters such as elevation, slope angle, slope aspect, underlying geologic unit, etc., that can each be defined at a single point on the landscape, while the binary

dependent variable is whether or not a given point on the landscape is susceptible to landsliding. The training data is compiled by collecting the values of landscape parameters at locations where landslides are known to have occurred, as well as collecting a set of values for landscape parameters where landslides have not occurred. The training data is then processed by regression analysis to determine the best values of coefficients $B_0, B_1, B_2, ..., B_n$ to fit the function P(Z) to the known occurrence or absence of landslides at each point sampled. The function P(Z) can then be evaluated at any point in the landscape where the values of the indicator variables are known to determine a probability that the given location is susceptible to landsliding. Typically, probability thresholds are defined to transform raw probability values into susceptibility classes. Over the past few decades, this general method has been used in several studies for the purpose of landslide susceptibility mapping (cf., Dai et al., 2003; Lee, 2005; Lei and Jing-feng, 2006).

The popularity of the logistic regression technique is the product of its several advantages over similar regression techniques. For one, logistic regression can directly incorporate categorical variables without mapping the categorical variable to a continuous quantity. For example, a variable such as geologic unit can be directly entered into the model without somehow converting the geologic unit category to a numerical value. This avoids the issue of determining an appropriate quantification of qualitative variables whose relationship to landslide susceptibility is not necessarily known *a priori*. The model is thus free to objectively determine the relationship. The logistic regression technique is also less sensitive to the distributions of both dependent and independent variables, as compared to other regression techniques.

We considered a variety of topographic, geographic, climatic, and geologic factors as potential indicators of landslide susceptibility. The input data collected is described in the following section.

$$P(x) = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}}$$

Model predicts occurrence of dependent variable

Model predicts absence of dependent variable
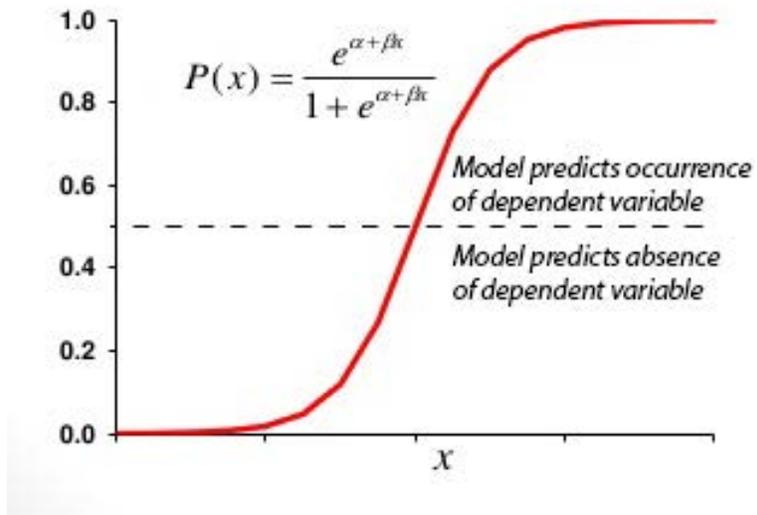
Figure 2.7: The logistic function with one indicator variable. The threshold value separating model prediction of occurrence versus absence (1 versus 0) is often assumed to be 0.5, but can be set to any value between 0 and 1.

3 Input Data

We collected several statewide datasets for potential inclusion in the logistic regression models. Data processing, as well as assessment for inclusion in the final models, is described in the following section.

3.1 Statewide map of deep-seated landslides and escarpments

The existence of a statewide map of deep-seated landslides, compiled by Cardinali et al. (1990), was one of the impetuses for this project. This map has a scale of 1:500,000 and categorizes the mapped landslides into rotational slides; translational slides (i.e., rock-block slides and combined rock slide or debris slide); Toreva block slides; "unclassified deep-seated landslides;" "unclassified complex landslide or slump-earth flows;" or "hummocky topography probably related to deep-seated landsliding." The specific category of landslide was not considered in our analyses. Cardinali et al. (1990) also produced maps of topographic escarpments, which we considered a potential indicator for predicting landslide susceptibility. Both Cardinali et al. (1990) maps were digitized by the Earth Data Analysis Center at the University of New Mexico and provided to us as GIS files for this project. Our assessment of the locational accuracy of these maps is described in the following section.

3.2 Geologic map

We considered that the mechanical properties of the underlying geology may be an indicator of landslide susceptibility. Geologic information was extracted from the 1:500,000-scale, statewide geologic map of New Mexico (NMBGMR, 2003). In addition, we considered that the distance to a mapped fault may also be an indicator variable; fault locations were extracted from the same statewide geologic map. Our experience with this dataset is that it is accurate to approximately 200 to 500 m.

3.3 Topographic data

We considered that elevation, slope angle, slope curvature, slope aspect, and "slope roughness" may be indicators of landslide susceptibility. We used a proprietary high-resolution (4.5 m pixel resolution) digital terrain model (Intermap, 2008) provided to the NMBGMR by the U.S. Army Corps of Engineers for research purposes, which we downsampled to 28 m pixel resolution, as a basis for all topography-related variables. Our experience with this dataset is that the limit of accuracy is below the 28 m pixel resolution. We processed the digital elevation model for slope angle, slope curvature, and slope aspect using Esri ArcGIS Spatial Analyst tools (Esri, Inc., 2016). Slope curvature was split into two inputs, positive slope curvature and negative slope curvature, each of which recorded only the magnitude of curvature as a positive value irrespective of the direction of convexity of the slope curvature. We also calculated "slope roughness" factors, in this case calculated as the standard deviation in elevation, slope angle, and slope curvature values occurring over any given 5 pixel by 5 pixel square area.

3.4 Precipitation data

We downloaded the 30-year average ("normal") annual precipitation model data at 800 m pixel resolution from the PRISM Climate Group at Oregon State University (PRISM Climate Group, 2016).

3.5 Hydrographic data

We considered that the distance to a major stream may be an indicator of landslide susceptibility. We used the National Hydrography Dataset available from the U.S. Geological Survey (USGS, 2013) for stream locations, keeping only the following major rivers: Canadian River, Cimarron River, Gila River, Jemez River, Mimbres River, Pecos River, Red River, Rio Chama, Rio Grande, Rio Puerco, San Juan River, and Zuni River. However, we also considered that the lateral extent of influence on landslide susceptibility may vary with valley width along each river. To capture this varying extent, we approximated the extent of each river's floodplain by extracting the extents of "young alluvium" and "young river alluvium" (map units A, Ay, AR, and ARy) from the surficial geologic map of Hawley et al. (2005) that occurred along each river. Where no alluvial deposit was found on the surficial geologic map, we gave the river segment an arbitrary 50 m width. Distance to stream was subsequently determined from the edges of the resulting floodplain polygons.

4   Methods

4.1   Modeling approach

4.1.1   Physiographic provinces

In a state as diverse as New Mexico, we anticipated that no single model could accurately portray the landslide susceptibility across the entire state. The rocky, rugged mountains of the basement-cored Sangre de Cristo Mountains or volcanic and volcaniclastic rock-dominated Gila Wilderness area are unequivocally different settings from the mesas and interbedded sedimentary rocks of the Colorado Plateau, and we anticipated that these characteristically different settings would each need their own regression model. Based on similarities in geomorphic and underlying geologic elements, the state is typically divided into four or more physiographic provinces (e.g., Pazzaglia and Hawley, 2004, and references therein), and we used the physiographic province concept as a guideline for dividing the state into six regions. Our regions were adapted from Hawley (2005), with the exception that we differentiated a northern Rio Grande rift province within his southern Rocky Mountains province. Our six physiographic provinces are: Colorado Plateau, Southern Rocky Mountains, northern Rio Grande rift ("North Rift"), Basin and Range, and Mogollon-Datil highlands (which Hawley called the "Transition Zone"); these provinces are shown in Figure 4.1. The Colorado Plateau is characterized by interbedded sedimentary strata of varying strengths, where stronger layers commonly form mesa tops and softer strata may underlie the mesa flanks. The Great Plains province also is underlain largely by sedimentary strata that create mesas or escarpments, particularly in its northern part. However, it is overall flatter than the Colorado Plateau and has a slightly denser grass cover. The Southern Rocky Mountains ranges from 6,000 to 13,000 ft elevation, has more abundant steep slopes but less common mesas than the aforementioned provinces, and is underlain by a wide variety of rock types (igneous, metamorphic, and sedimentary). The northern Rio Grande rift is characterized by mesas and low-gradient slopes underlain by either volcanic flows or weakly consolidated Cenozoic rift basin fill. Locally, the basin fill is highly dissected to produce locally high topographic relief. Vegetation ranges from semi-arid grassland to pinon-juniper to sagebrush. Located in the southern part of the state, the Basin and Range is mostly a low-relief, arid (sparsely vegetated Chihuahuan desert fauna) landscape sporting a diverse array of geologic formations and structures. Lastly, the Mogollon-Datil province was designated for the high-elevation landscape between the cities of Socorro and Silver City, including the Gila Wilderness area. It has a slightly less diverse array of rock types compared to the Basin and Range, with volcanic rocks being particularly common, and it is more hilly. The higher elevation of the Mogollon-Datil physiographic province allows it to receive more precipitation than the Basin and Range province.

4.1.2   Training areas and validation areas

It is desirable to have two distinct datasets of landslide occurrence or absence and associated landscape parameters in regression modeling, as one set can be used for training the model while the other provides an independent evaluation of the efficacy of the model. Multiple methods exist for deriving separate training and validation datasets; we chose to designate two non-overlapping areas within each province to serve as training and validation areas, from which the landslide occurrence and landscape parameter values would be extracted. Each area was chosen so as to be >2000 km^2 in areal extent, to contain a relatively high number of mapped

landslides, and to be representative of the physiographic province as a whole. Training and validation areas are shown in Figure 4.1.

4.1.3 Logistic regression modeling

We developed a data sampling, model generation, and model validation strategy that collects and processes the landscape parameters associated with mapped deep-seated landslide deposits to estimate landslide susceptibility across the state. The validity of using landslide deposits to characterize susceptibility is addressed in the Discussion section.

4.1.3.1 Data sampling

In order to characterize the landscape parameters associated with known landslide deposits, we first converted all landslide points to small circles with 150 m radius to generate small polygons of comparable size to the accuracy of the adjusted landslide deposit map (regarding adjustments: see "Input assessments and processing – Landslide map accuracy improvements"). We added these to the adjusted landslide deposit polygons, then densely point-sampled the combined landslide deposit polygons using a square grid geometry. The grid spacing was varied to provide even coverage of the landslide polygons in the training area with 100,000 landslide training points, with a minimum allowable spacing of 28 m (the resolution of the input topography-related rasters). As a consequence of the minimum allowable grid spacing, the Mogollon-Datil and Southern Rocky Mountain provinces, which have comparatively little total landslide area, collected fewer than 100,000 landslide training points, while the remaining four provinces collected the full number. We then collected an equal number of non-landslide area points. These points were randomly located throughout the remainder of the training area. Non-landslide points were not permitted to be located within a 300 m buffer around any landslide polygon, in order to accommodate the accuracy of the adjusted landslide map. A minimum point spacing for non-landslide points was set at 7 m; ideally, we would have used a 28 m minimum spacing (equal to the resolution of the topography-related rasters), but initial sampling experiments with a 28 m minimum spacing resulted in frequent crashing of the point-generating GIS tool, when the point density became too great to locate additional points. The data sampling procedure is illustrated in Figure 4.2.

We collected two validation datasets for each province: an internal validation set collected from the training area and an external validation set collected from the validation area. For the internal validation dataset, we collected twice as many sample points as were collected for the training dataset; for the external validation dataset, 400,000 points total were collected across the validation area. Points were located randomly with a minimum point spacing of 7 m (similar to the non-landslide training points). No restrictions on point locations were imposed.

At all training and validation points, the presence or absence of a landslide polygon (inclusive of the 150 m-radius circles about landslide points) was recorded as a 1 or 0, respectively. All specified indicator landscape parameter values were also recorded. These records were screened for null values, and any point with any null values was removed; this process removed less than a tenth of a percent of the points collected. Final training point totals ranged from 73,280 (36,734 landslide, 36,546 non-landslide) for the Mogollon-Datil area, to 158,678 (79,338 landslide, 79,340 non-landslide) in the Southern Rocky Mountains, to at least 199,884 (roughly evenly divided) in all other provinces. Training and validation data was then exported to shapefiles and text files as needed for the LAND-SE script, discussed below.
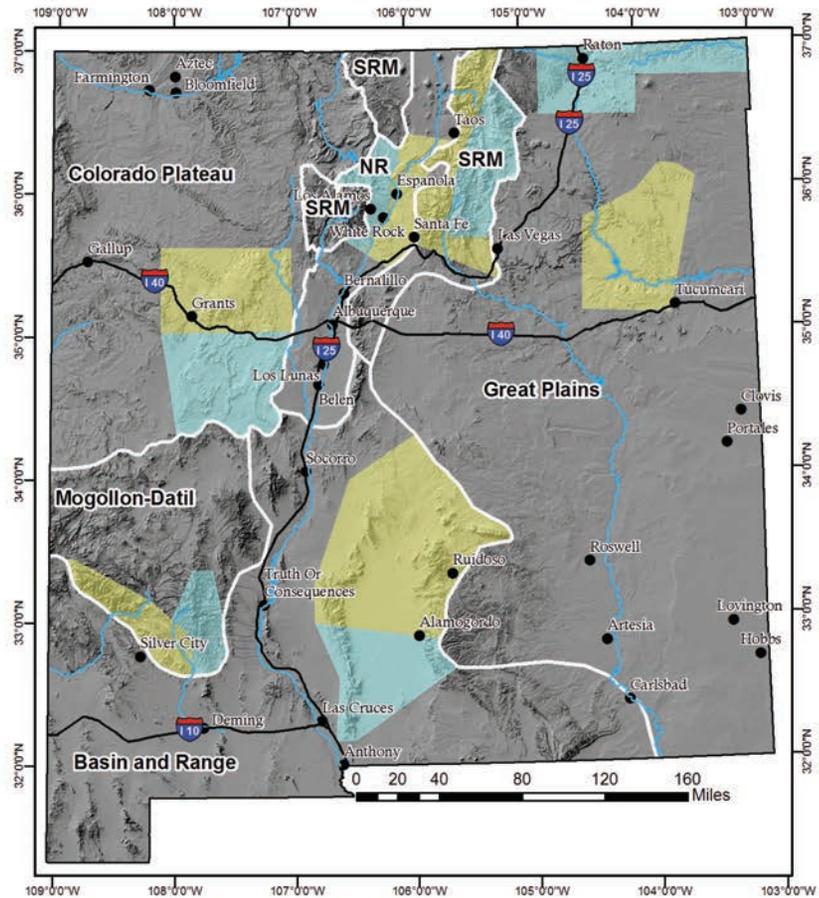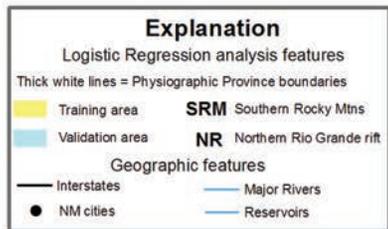
Figure 4.1: Map of physiographic provinces and associated training and validation areas. Southern Rocky Mountains province consists of three disconnected regions.

Figure 4.2: Illustration of the data sampling strategy. Landslide points are expanded to 150 m radius circles, and the combined extent of these circles and the mapped landslide polygons is grid-sampled for landslide data points. No-landslide data points are collected at random from the area outside a 300 m buffer zone around the landslide sampling area. Illustration extracted from a portion of the Colorado Plateau training area.

We used a modified version of the LAND-SE script (version 1, release 0, build 32) written by Rossi and Reichenbach (2016) to generate and preliminarily evaluate the logistic regression models. This script was written in the R programming language, an open source programming language with a particular niche in statistical computing (R Development Core Team, 2015). The LAND-SE ("LANDslide Susceptibility Evaluation") script was written to generate, evaluate, and map the results of multiple multivariate classification models specifically for landslide susceptibility. The script is designed to take two sets of inputs, one training dataset and another validation dataset, typically collected from the same area. The script outputs include tables of model coefficient estimates, coefficient estimate standard deviations, and coefficient Wald test statistics; tables of model evaluation results from evaluations performed at all training and validation points; and receiver operating characteristic (ROC) curves (described below) for both training and validation dataset evaluations. Our modifications to the script were minor, removing the multivariate classification models we were not using as well as removing some time-consuming processing and map generation steps that were not useful for our purposes.

4.1.3.3   Model evaluation methods

We found three model evaluation methods particularly useful: the ROC curve, the distribution of model probabilities determined for known landslide areas, and qualitative review of model probability maps.

The ROC curve graphs the false alarm rate against the hit rate for a validation dataset as a function of the threshold used to separate the prediction of landslide occurrence from absence (Figure 4.3). The false alarm rate is defined as the total number of non-landslide points for which the model predicted a landslide occurrence divided by the total number of non-landslide points, while the hit rate is the total number of landslide points for which the model predicted a landslide occurrence divided by the total number of landslide points. As the prediction threshold is raised from 0 to 1, the fraction of model probabilities that predict the occurrence of a landslide decreases from 100% to 0%, resulting in a decline in both the false alarm rate and hit rate. However, a model with good predictive ability will see a more rapid decline in the false alarm rate relative to the hit rate, resulting in a curve that bends toward the upper left corner (0% false alarm rate, 100% hit rate; Figure 4.3a). A model with poor predictive ability will lie close to the straight line from the lower left to the upper right (from 0% false alarm and hit rate to 100% false alarm and hit rate; Figure 4.3b); a model with this shape is little better than guessing. Since the area beneath the ROC curve increases as the model is bent toward the upper left corner, a commonly-used single-value measure of the discriminatory capacity of a model is simply the area beneath the ROC curve, which will be 0.5 for a poor model and will increase toward 1 as the model predictive ability increases.

A related measure of model efficacy is the distribution of model probabilities calculated for known landslide areas (Figure 4.4). An ideal model will result in high model probabilities in all known landslide areas (Figure 4.4a), while poorer models will show a greater percentage of low model probabilities in landslide areas (Figure 4.4b). These histograms were evaluated principally qualitatively, although the model probability distribution percentiles were calculated for each histogram. These histograms also proved useful in determining appropriate probability thresholds for defining susceptibility classes, as the distribution provides a measure of what fraction of known landslide areas would be captured in each class given various thresholds. This is discussed further in the "Final model results – Synthesis of province results" section.

ROC PLOT: VALIDATION LRM MODEL

ROC area = 0.86; Sample size = 399949; Bootstrap samples = 10

Model A   0.865 (0.841)

Black lines are the empirical ROC.
Red lines and symbols are the bi−normal ROC.
The area under the binormal curve is in parathesis.

Hit Rate

False Alarm Rate

(a)

ROC PLOT: VALIDATION LRM MODEL

ROC area = 0.61; Sample size = 399102; Bootstrap samples = 10

Model A   0.613 (0.598)

Black lines are the empirical ROC.
Red lines and symbols are the bi−normal ROC.
The area under the binormal curve is in parathesis.

Hit Rate

False Alarm Rate

(b)

Figure 4.3: Example ROC curves. In each case, the threshold separating predicted landslide occurrence from absence is varied from 0 to 1, and the hit rate and false alarm rate calculated at each threshold. Points corresponding to thresholds at every tenth (0.1, 0.2, etc.) are plotted and labelled. The bi-normal ROC curve is a smooth curve fitted to the empirical results. Both curves are from external validations. ROC plots generated by the LAND-SE script (Rossi and Reichenbach, 2016). (a) Example good model. (b) Example poor model.

Histogram of landslide−area probabilities

(a)

Histogram of landslide−area probabilities

(b)
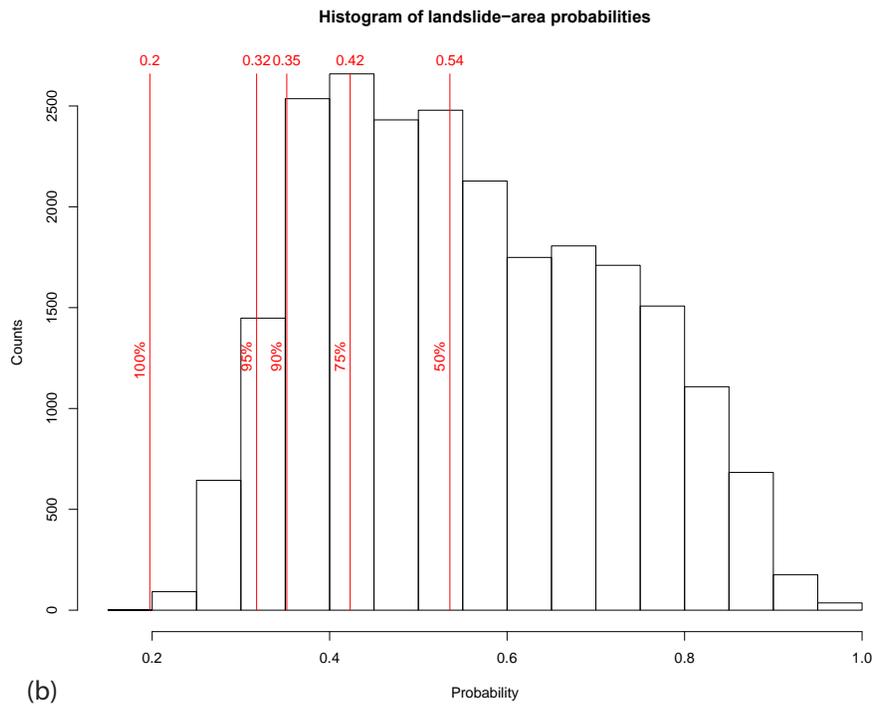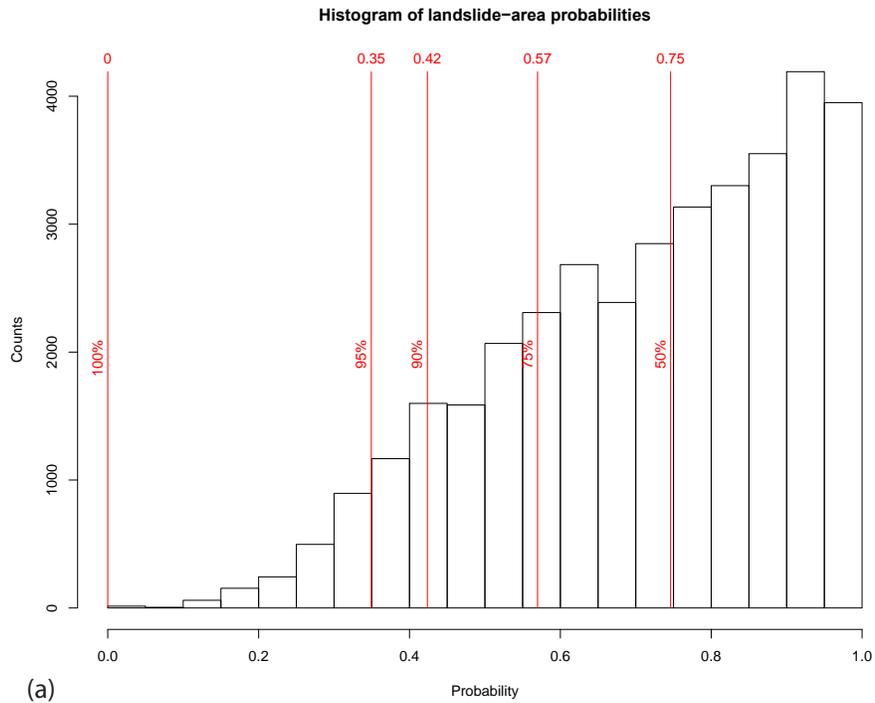
Figure 4.4: Example landslide-area model probability histograms. In each, red lines and labels locate distribution percentiles, where each percentage reports the percent of landslide-area sample points occurring above the threshold line. Both histograms are from external validations. Note the changing scales. (a) Example good distribution, showing a concentration of landslide area in the higher model probabilities, and more than 75% of the landslide areas have model probabilities greater than 0.5. (b) Example poor distribution, showing a concentration of landslide areas in the mid-range of model probabilities, and only a little more than 50% of the landslide areas have model probability greater than 0.5.

Qualitative review of model probability maps focused on map sensibility. For example, flat areas should not display high model probabilities, as a flat area should not be susceptible to landsliding (cf., Figure 4.5a). Maps of model probability were evaluated and landscape parameters associated with unusual probabilities were identified. This lead to some refinements in the treatment of geology and slope angle for some provinces.

4.2 Input preparation

4.2.1 Landslide map accuracy improvements

The Cardinali et al. (1990) landslide maps are a critical input to the logistic regression model. We evaluated the accuracy of the landslide maps by importing the GIS data into Google Earth (Google, Inc., 2015) and comparing the Cardinali et al. mapping to landslides interpreted from the Google Earth imagery. Based on this assessment, we determined lateral inaccuracies commonly in the range of 300 to 1,000 m, and locally >1 km (e.g., Figure 4.6). Given that this imprecision would be a limiting factor on the quality of the logistic regression models, we chose to manually improve the landslide mapping in all training and validation areas. We used Google Earth imagery and the digital elevation dataset to adjust landslide line and point data to achieve inaccuracies of <150 m (estimated 95% of adjusted linework locales) to <200 m (estimated 99% of adjusted linework locales). The scale used in adjustments ranged from 1:20,000 to 1:60,000. Adjoining polygons were lumped together if they were the same type of landslide. Locally, new polygons were drawn that were not on the original Cardinali et al. landslide map.

Improving the accuracy of point data proved more difficult than improving the accuracy of polygons, namely because the polygon shape allowed us to better constrain its location on the Google Earth imagery. With landslide points depicted in the original Cardinali et al. (1990) map, it was often difficult to know where the corresponding landslide was located in on the Google Earth imagery. For example, a landslide point from the Cardinali et al. map inaccurately located on flat ground could correlate to zero, one, or two landslides on the imagery, and sometimes it was difficult to know which one to choose. As much as possible, if two or more points were clustered within a radius of 3 km, then our correction-related movement of points within that cluster was kept consistent in terms of direction. In some locations, however, one was left with little choice but to have inconsistency within the correction-related movements. Occasionally, landslides depicted as points were >1 km^2. In these cases, new landslide polygons were mapped and the associated point deleted.

4.2.2 Indicator variable data binning and categorization

An advantage of the logistic regression modeling technique is its capacity to directly incorporate categorical variables as indicator inputs and its relative insensitivity to the distribution of indicator variable values. As a result, it would be possible to simply sample the actual values of all indicator variables, such as geologic unit, slope angle, and slope aspect, without any processing and use these values directly in the model generation. However, with over 150 separate geologic units on the statewide geologic map, and very unequal distributions of some continuous variables (e.g., low slopes are far more common than steep slopes), there is a risk of under-sampling the less common units and less frequent ends of continuous variable distributions. This becomes a particular problem when considering our methodology of defining certain regions of each province as training and validation areas, then subsequently applying the model from the training area to the remainder of the province; these training and validation

regions may not capture and adequately characterize the influence of less common elements of the province as a whole, resulting in an unreliable model.

4.2.2.1   Binning assessments

Binning strategies were assessed iteratively: initially by examining the area of each bin occurring in the training areas to ensure adequate representation; a second time through the spatial subsampling routine described later; and finally refined as a result of model evaluation. In the initial test, for each proposed bin, the percent of the province and the training area covered by that bin was determined, and the ratio between the training area proportional coverage and the province-wide proportional coverage was calculated (e.g., Table 4.1). This ratio defined whether the proposed bin was under-represented (<1) or over-represented (>1) in the training area relative to the province. Common under-representation of initial bins, particularly for initial geologic unit groupings, led to binning refinements. These ratios also supported the use of equal-frequency binning (described below) for continuous variables, as the ratios for the equal-frequency bins, particularly for the variables used in the final models, commonly averaged ~1 to 1.5 (Table 4.2).

The second binning assessment was performed with the spatial subsampling routine described below. Slope aspect and geologic unit grouping were particularly scrutinized during this procedure. For each variable, certain initial groupings were found to perform poorly in the later-described tests, resulting in grouping refinements.

Finally, geologic unit groupings were further refined as a part of final model evaluation. In some provinces, we found that certain geologic unit categories were consistently associated with under predicting landslide occurrence, indicating the need to refine the geologic unit grouping for that province.

Our final binning strategies are described below.

4.2.2.2   Continuous variables – Topography-related data, precipitation

For continuous variables (slope angle, curvature, elevation, precipitation, roughness, etc.), we used an equal-frequency binning strategy, with each bin nominally holding 1% of the total pixel counts, but with restrictions on the minimum bin width (must be at least 1% of the value's range) and minimum bin pixel count (must be >0.1% of the total pixel count). The minimum size restrictions were implemented to prevent disproportionate binning. For example, in the Colorado Plateau province, there are as many pixels in the slope angle range of 0 to 0.5° as in the range 19 to 77°, such that an unrestricted binning would result in numerous bins in the very low slope range and few bins in the steep range. Minimum bin width restrictions prevented the creation of disproportionately numerous low-slope bins.
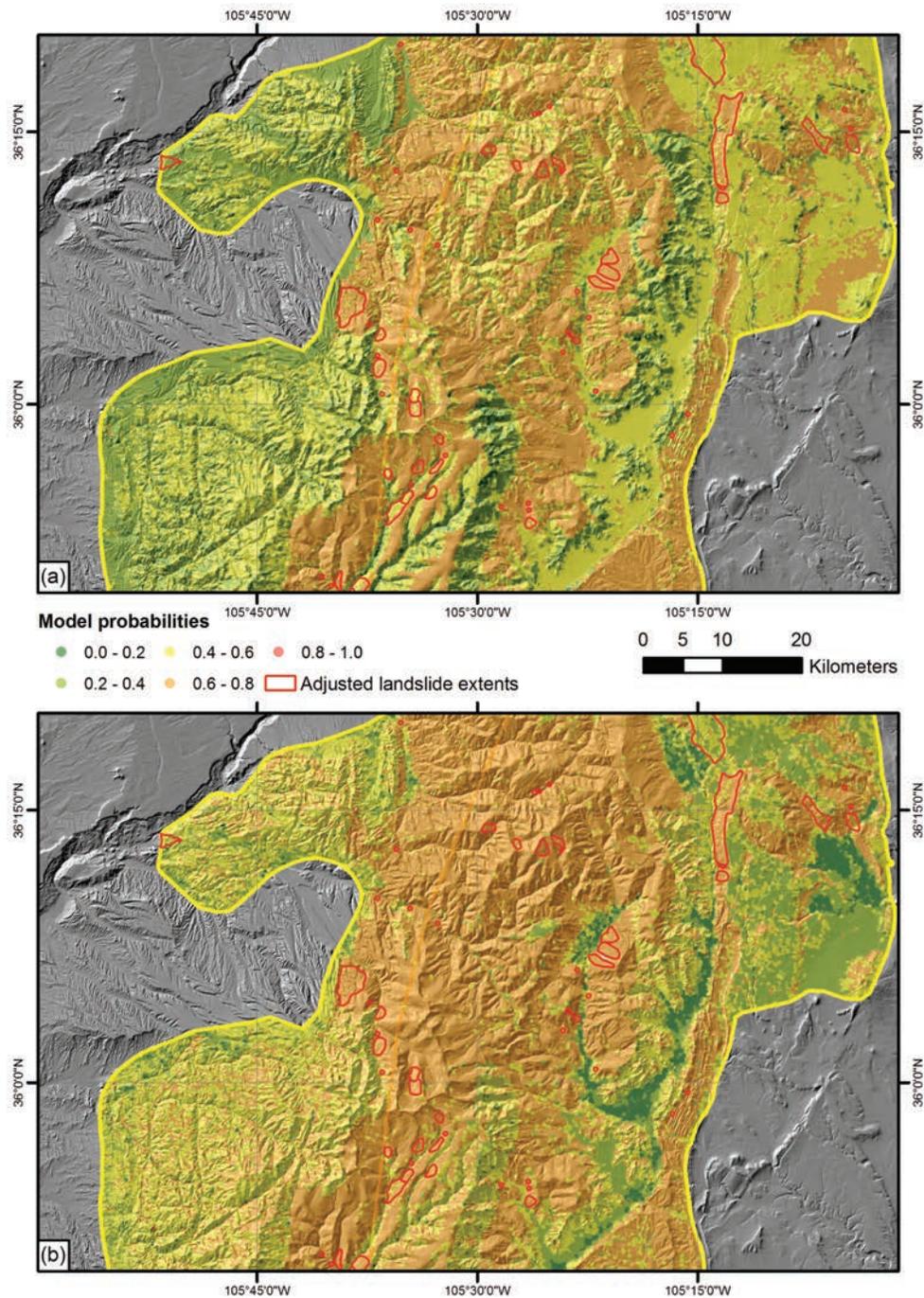
Figure 4.5: Example model probability maps used for qualitative review of model performance. Both maps are from the Southern Rocky Mountains area. (a) Example map showing poor sensibility. Note the relatively high model probabilities in low-relief areas along valley floors along the east (right) side of the map, and relatively low probabilities through the high-relief terrain in the south-central of the map. This results from models with negative coefficients for the slope parameter. (b) Example map showing the same area with a sensible parameterization. Note the low-relief valley floors show low model probabilities, while relatively high probabilities are determined for the high-relief areas in the south-central of the map. The lower map was produced using the alternative slope treatment applied to the Southern Rocky Mountains and Mogollon-Datil areas, described later.

26
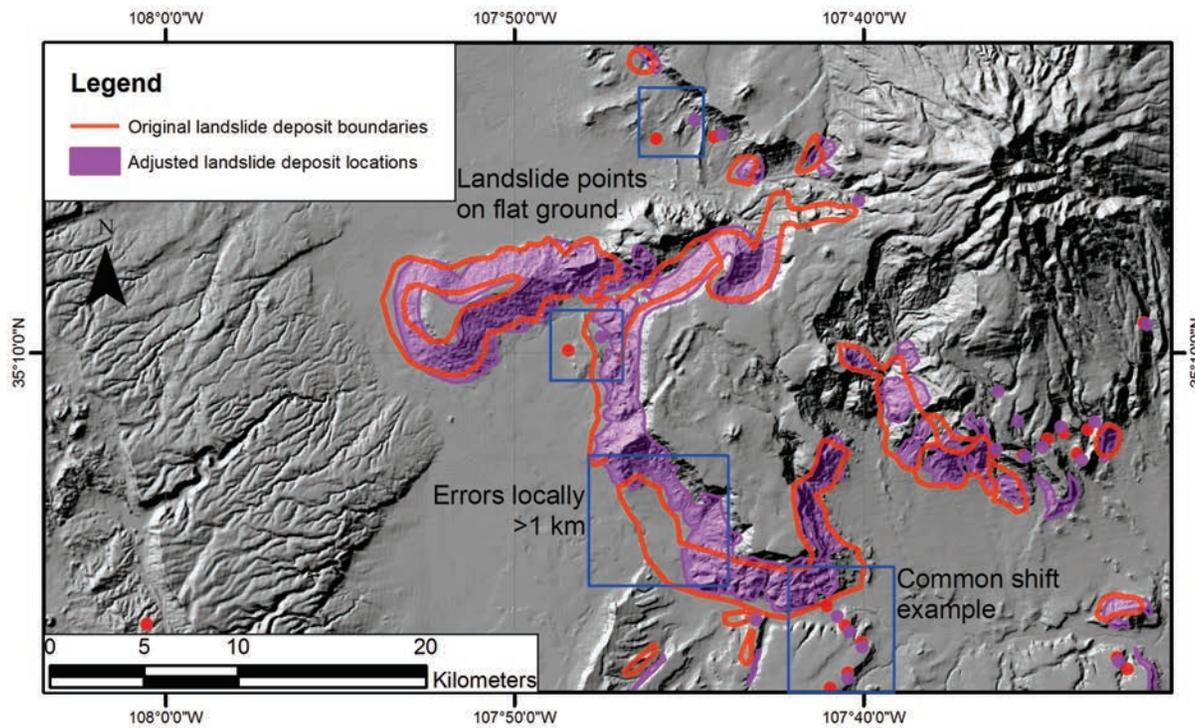
Figure 4.6: Example of landslide mapping adjustments, from the Colorado Plateau training area circa Mount Taylor. Note, in blue boxes: landslide points mapped on nearly flat ground (north-central of map); map errors locally >1 km (south-central); and landslide points given the same shift directions (south-central), as described in the text.

Table 4.1: Example bin coverage assessment - Preliminary geologic unit bins for the North Rift province

| Geo. Unit [1] | Province-wide | | Training Area | | |
	Area (km^2)	Fraction [2]	Area (km^2)	Fraction [2]	Coverage Ratio [3]
M1	123	0.006	1	0.000	0.056
S1	7,890	0.383	1,078	0.266	0.695
S2a	7,480	0.363	2,026	0.500	1.377
S2b	63	0.003	20	0.005	1.618
St1	4,510	0.219	911	0.225	1.027
St2	519	0.025	12	0.003	0.117
Totals	20,584		4,049		

Notes:

1: Geologic unit: see Table 4.4 for brief descriptions of units and final unit binning

2: Fraction of total area covered by unit

3: Ratio of training area fraction to province-wide fraction

Underrepresented units highlighted in orange.

Table 4.2: Summary of training area to province-wide bin coverage ratios

Variable	BR			CP			GP			MD			NR			SRM		
	Min	*Max*	*Mean*	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*	*Min*	*Max*	*Mean*
Neg. Curv.	0.99	1.23	1.16	0.98	1.31	1.21	0.94	2.79	2.28	0.91	2.43	1.69	0.93	1.57	1.47	0.95	1.62	1.42
Pos. Curv	0.98	1.30	1.18	0.98	1.24	1.17	0.94	3.79	2.64	0.89	2.55	1.82	0.93	1.72	1.54	0.93	1.49	1.37
Curv. Std. Dev.	0.89	1.23	1.15	0.92	1.26	1.18	0.78	3.67	2.61	0.28	3.03	1.89	0.50	1.70	1.51	0.54	1.60	1.33
Elev.	0.30	3.53	2.02	0.07	3.40	1.48	0.02	2.99	1.48	0.42	5.21	1.45	0.01	3.18	1.42	0.08	3.15	1.48
Elev. Std. Dev.	0.81	1.67	1.35	0.84	1.56	1.29	0.60	3.39	1.96	0.12	3.95	2.01	0.42	1.55	1.27	0.21	1.98	1.41
Geology[1]	0.76	1.74	1.32	0.60	3.12	1.66	0.24	2.96	1.31	0.55	1.24	0.90	0.69	1.36	1.00	0.70	1.18	0.93
Slope Aspect[2]	0.93	1.08	1.01	0.92	1.09	1.00	0.99	1.01	1.00	0.99	1.01	1.00	0.92	1.09	1.00	0.96	1.05	1.00
Precipitation	0.41	5.45	2.09	0.31	3.96	1.15	0.01	3.25	1.10	0.03	7.26	2.90	0.03	2.86	1.43	0.02	1.68	0.98
Slope	0.75	1.65	1.32	0.83	1.51	1.28	0.55	3.59	1.88	0.11	3.64	1.79	0.41	1.52	1.29	0.21	1.94	1.35
Slope Std. Dev.	0.80	1.30	1.18	0.86	1.53	1.26	0.70	3.96	2.40	0.09	2.35	1.55	0.45	1.48	1.35	0.24	1.41	1.21

Notes:

  1: Assesses province-specific geologic units, as shown in Table 4.4.

  2: Does not include flat aspect.

Abbreviations:

  BR: Basin and Range, CP: Colorado Plateau, GP: Great Plains, MD: Mogollon-Datil, NR: North Ri t, SRM: Southern Rocky Mountains

Curv.: Curvature, Elev: Elevation, Pos.: Positive, Neg.: Negative, Std. Dev.: Standard Deviation

Variables used in final models highlighted in orange

4.2.2.3   Categorical variables – Slope aspect, geology

Slope aspect was incorporated into final models categorized into north (azimuths from 247.5° through 0° to 67.5°), south (azimuths from 67.5° through 180° to 247.5°), and flat. These groupings resulted from assessments of initial binning trials. In the initial trials, slope aspect was grouped into 8 cardinal directions plus flat, and these bins were assessed through the spatial subsampling routine described below. These assessments determined that using 8 cardinal directions resulted in under-sampling of some directions and model coefficients that could be either positive or negative depending on the location of the subsample. However, within each province, coefficients for west through north to northeast were often similar, while coefficients for east through south to southwest were often similar (Table 4.3), leading us to use the aforementioned categorization scheme. This later categorization performed better in a subsequent spatial subsampling test.

Categorization of geologic units was defined and refined iteratively throughout the project. Initially, each geologic unit in the statewide geologic map was placed in one of 6 groups based on inferred mechanical properties (Table 4.4). However, binning tests, described above, determined that in each province there were groups that were under-represented in the training area (e.g., Figure 4.7), or performed poorly in the spatial subsampling tests. In addition, initial models for the Southern Rocky Mountains, Mogollon-Datil, and North Rift areas identified unit categories that were frequently associated with low model probabilities in known landslide areas, prompting us to further refine these groupings. These tests and assessment led to the province-specific unit categorizations shown in Table 4.4.

4.2.2.4   "Distance to" variables

We considered the potential indicator variables "distance to stream," "distance to fault," and "distance to escarpment." In deciding how to incorporate these variables into the logistic models, we considered: 1) the reasonable lateral extent of influence of each element from a mapped element location; 2) the locational uncertainty of each map element, in particular for the escarpments and geologic maps (described above); and 3) the large scale of the investigation. Given these considerations, we chose to incorporate each as a simple binary "close or not" variable. For escarpments and faults, the "close to" distance was chosen as 1 km, while for rivers the "close to" distance was 500 m from the edge of the floodplain. Each variable was sampled as either a 1 (if within the buffer distance of a feature) or a 0 (outside the buffer distance of all features).

4.3   Input assessment tests

After controlling for spatial accuracy and determining binning strategies for each potential indicator variable, we assessed the applicability of each input through a sequence of tests described below.

Table 4.3: Summary of aspect category coefficients from the spatial subsampling routine using 9 and 3 category binning

Cardinal	Azimuth	BR 9-cat. coef.	BR 3-cat. coef.	CP 9-cat. coef.	CP 3-cat. coef.	GP 9-cat. coef.	GP 3-cat. coef.	NR 9-cat. coef.	NR 3-cat. coef.	SRM 9-cat. coef.	SRM 3-cat. coef.
North	337.5° - 22.5°	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Northeast	22.5° - 67.5°	-0.194		-0.269		-0.273		0.022		-1.005	
East	67.5° - 112.5°	-0.878	-0.385	-0.386	-0.237	-0.267	-0.148	0.584	0.644	-0.682	-0.606
Southeast	112.5° - 157.5°	-0.938		-0.389		-0.465		0.917		-0.543	
South	157.5° - 202.5°	-0.379		-0.352		-0.523		0.848		-0.281	
Southwest	202.5° - 247.5°	-0.040		-0.137		-0.324		0.288		0.568	
West	247.5° - 292.5°	0.224	Ref.	0.068	Ref.	-0.229	Ref.	-0.243	Ref.	1.177	Ref.
Northwest	292.5° - 337.5°	0.278		0.227		-0.123		-0.124		1.106	

Notes:

  Cat.: category; coef.: coefficient; ref.: reference direction.

  Coefficient reported is the average of all non-outlier, well-constrained subsample coefficients.

  Mogollon-Datil area was combined with Basin and Range for the earlier 9-category trials, and hence is not reported here as a separate province.

Province abbreviations:

  BR: Basin and Range, CP: Colorado Plateau, GP: Great Plains, MD: Mogollon-Datil, NR: North Rift, SRM: Southern Rocky Mountains

Table 4.4: Summary of geologic unit groupings

General units				Province-specific units					
Unit	Description	Examples	Example units	BR	CP	GP	MD	NR	SRM
S1	Young sediment	Stream alluvium; eolian sands	mostly unnamed Quaternary deposits	S1	S1	S1	Su	S1	S1
S2a	Compacted, poorly cemented sediment	Rift basin fill	Santa Fe Grp, Gila Grp, Ogallala Fm.	S1	S2	S2a	Su	S2M	S2M
S2b	Weak sedimentary rocks	Shale-, mudstone-, and evaporite-rich units	Mancos Shale, Chinle Grp, Artesia Grp	S2b	S2	S2b	Su	S2M	S2M
M1	Interbedded weak and strong layers	Interbedded sandstones/limestones and shales/mudstones	Crevasse Canyon Fm, interbedded Dakota-Mancos and Gallup-Mancos, San Rafael Grp	M1	M1	M1	StM	S2M	S2M
St1	Strong, tabular rocks	Basalt flows, ignimbrites; thick sandstone/limestone units	Bandelier Tuff, San Andres Limestone	St	St1	St	StM	St	St
St2	Strong, massive rocks	Intrusive rocks; thick sections of volcanic rocks	Granitic plutons; caldera-fill pyroclastics and ring-fracture rhyolite domes	St	St2	St	StM	St	St

Notes:

BR: Basin and Range; CP: Colorado Plateau; GP: Great Plains; MD: Mogollon-Datil; NR: North Rift; SRM: Southern Rocky Mountains

Grp.: Group, Fm.: Formation
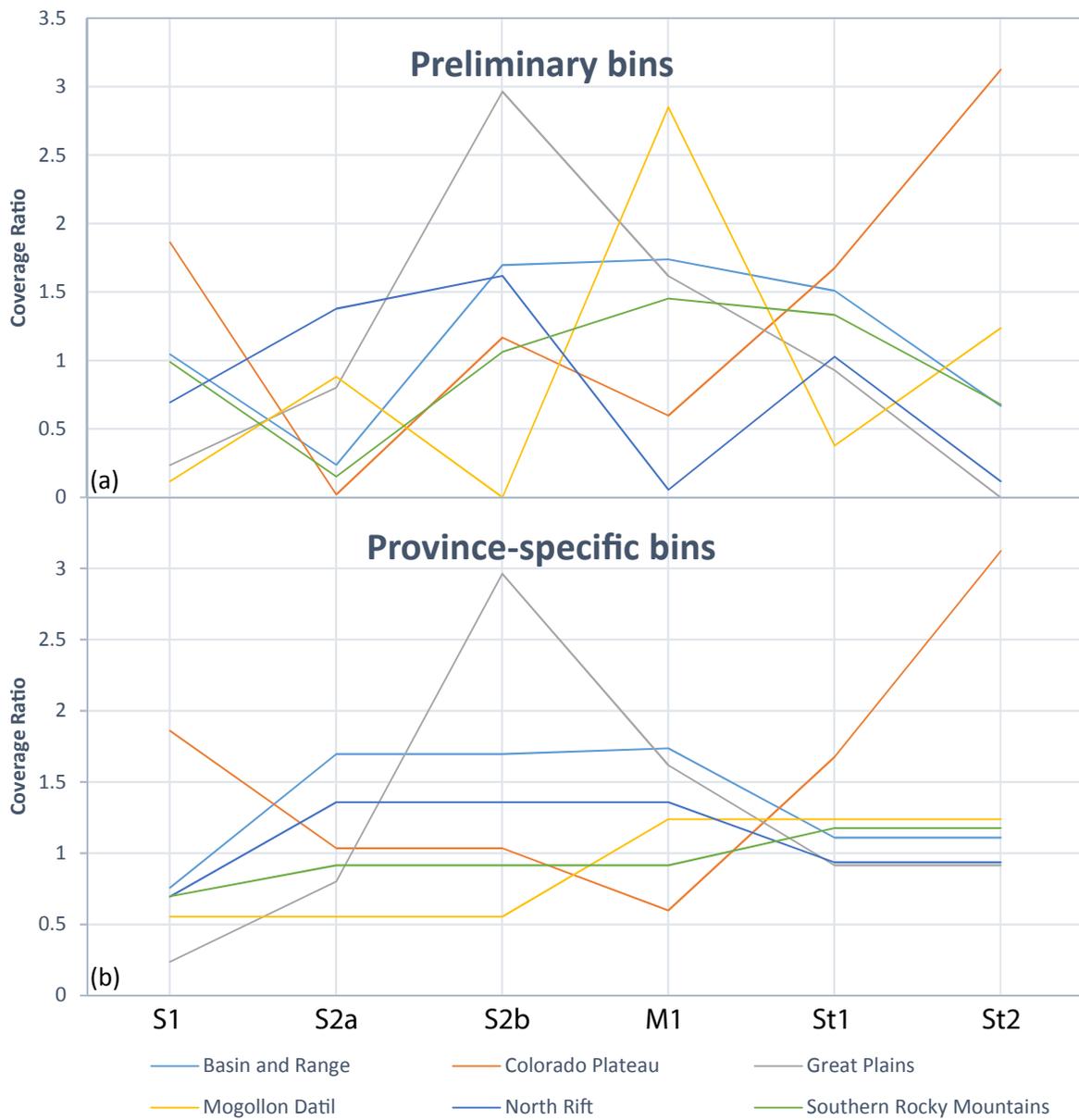
Figure 4.7: Comparison of the coverage ratios of preliminary geologic unit groups (a) and final province-specific groups (b). Coverage ratios are the ratio of the proportion of each province training area lying on a given geologic unit to the proportion of the province as a whole lying on that geologic unit. Grouping of preliminary bins into province-specific bins is given in Table 4.4.

4.3.1 Independence of indicator variables

The logistic regression technique assumes that the indicator parameters are independent of each other. Independence can be tested for by assessing the linear correlation between any two variables. We determined correlation coefficients between all combinations of continuous variables in each province (Table 4.5), which overall varied from 0 to 0.96. Values over 0.4 were considered too correlative and defined two groups of variables:

1) Slope group – Slope angle, slope standard deviation, elevation standard deviation, and curvature standard deviation; and
2) Elevation group – Elevation and precipitation

We decided that no more than one variable from each group should be allowed in subsequent models. From the slope group, we decided to subsequently only incorporate the slope angle, as this variable should be the most immediately related to landslide occurrence. From the elevation group, we chose to use only precipitation, as between the two choices the precipitation dataset provides the more independent complement to the slope angle.

Although not adequately quantified, we also considered that the distance to stream variable may not be independent of the elevation group variables, as major streams occur at the lowest elevations.

4.3.2 Disproportionately influential indicator variables

We considered indicator variables or elements of indicator variables to be disproportionately influential if high model probabilities only occurred in conjunction with these elements. Disproportionate influence results from particularly common colocation of mapped landslide deposits from the Cardinali et al. (1990) map and the influential element. Quantitatively, influence could manifest as a particularly large model coefficient as compared to similar indicator elements. Qualitatively, influence manifests in model evaluation map patterns as high model probabilities directly overlying only this element. Models including elements with strong influence are not necessarily "wrong," as they may very accurately model the distribution of landslides occurring within the training area, but such models will carry a risk: that landslide occurrence away from the influential element may be under predicted. This can be particularly problematic if the locational accuracy of the mapping of the influential element is questionable.

In initial models, we identified two elements that disproportionately affected model results: the geologic map unit Qls (landslide deposits) and the distance to escarpments variable. The influence of neither element is surprising; it should be expected that landslide deposits would be independently mapped by Cardinali et al. (1990) and New Mexico-area geologists in similar locations, while landslide deposits often lie along the base of and near to escarpments. Models incorporating either of these elements commonly only determined high model probabilities in areas overlying map unit Qls or in the vicinity of escarpments. We determined that this influence is problematic for two reasons: 1) the locational uncertainty of the mapping of these two elements (e.g., inaccuracies locally greater than 1,000 m in the Cardinali et al. maps); and 2) since the Qls map unit and, in many cases, the escarpments are the product of a previous landslide occurrence, such models would be useful for predicting where landslides have occurred in the past but may have little use in predicting where landslides may occur in the future. We thus eliminated both elements from subsequent models. We removed map unit Qls by using aerial imagery and expert knowledge to make "best guesses" as to the nature of the geologic map

unit lying beneath, deleting the Qls polygons and replacing them with polygons of grouped geologic map units, i.e. those listed in Table 4.4.

4.3.3 The ability of a training area to represent the province as a whole

In order to assess the ability of each province training area to accurately model the province as a whole, we considered 1) the range of landscape attributes occurring within the training area as compared to the province, and 2) the sensitivity of determined model coefficients to training area location. The first of these elements was addressed in the "Input data binning and categorization" section. The second of these was a particular concern, as our method is inherently extrapolatory and hence relies on the relationship between all indicator variables and landslide susceptibility being constant throughout the province. If the model coefficient for an indicator variable varies with training area location, it would imply that the effect of this indicator variable is not constant and the model coefficient should not be extrapolated throughout the province. We addressed the sensitivity of model coefficients to training area location through a spatial subsampling and subsample modeling routine.

4.3.3.1 Spatial subsample evaluation

Our spatial subsampling routine was designed to evaluate several elements of our methodology:
1) How sensitive are model coefficients to the training area location? Can we trust that coefficients created from a given training area location can be applied to another location?
2) Is our model binning strategy reliable?
3) What is a reasonable range of model coefficients for each indicator variable?
4) Which variables have significant effects on the models? That is, is the coefficient of the variable statistically significantly different from zero?

We decided that we wanted to exclude from our modeling efforts variables whose relationship to landslide occurrence varies with space, that is, whose model coefficient varies significantly with the location from which training data was acquired and the model was derived. We also sought to remove variables that had no significant effect on the model, which in this case we chose to mean variables whose coefficients were not significantly different from zero. To address these concerns, we created multiple models trained on spatial subsamples of the combined training and validation areas, which was the total area for which we had corrected landslide location data. We then compared the coefficients derived from each subsample model to each other.

We created subsample areas by randomly locating points in the combined training-validation area, then growing circles about each point until they were one tenth (±5%) of the total area of the training-validation area (e.g., Figure 4.8). Subsample areas were invalidated if they contained less than a specified area of landslide polygons. For most provinces, this minimum was 15.68 km^2, twice the nominal amount of area needed to grid-sample 10,000 points (one tenth the number collected in a full model) at 28 m spacing (the resolution of the topography-related raster datasets). For the Basin and Range and Mogollon-Datil areas, however, this restriction proved too limiting, and hence was relaxed to 7.84 km^2 and 3.92 km^2, respectively. For each province, 100 valid subsample areas were created. In each, landslide-area and non-landslide area points were located as described in the "Logistic regression – Data

sampling" section above, except that only one tenth as many points were collected, and the minimum point spacing restriction on randomly located non-landslide points was removed (due to frequent script crashes). Point-sampled data was exported to text files and processed in an R script using the same functions as used in the LAND-SE script to determine model coefficient estimates and coefficient estimate uncertainties. For each independent variable, the coefficient estimates and coefficient estimate uncertainties were collected and compared.

Our coefficient estimate assessment was based on the following hypotheses:
1) most coefficient estimates should be consistent (few outliers);
2) most coefficient estimates should be well constrained (low coefficient uncertainties);
3) most coefficient estimates should be statistically different from 0 (pass the Wald statistic test);
4) the coefficient distribution should not show systematic variability, such that the differences between coefficient estimates should be the result of random error, and hence the coefficient estimates should be normally distributed;
5) the mean of the coefficient distribution should be statistically different from 0; and
6) coefficient estimates should show a consistent relationship with landslide occurrence (be consistently positive or consistently negative, and relatively consistent in magnitude).

We tested these hypotheses using the following methods:
1) Counting the number of outliers in the coefficient distribution, with outliers defined by two criteria: the model coefficient itself must be outside the (1.5 X interquartile range of coefficients) outlier criteria and the model coefficient uncertainty must be outside the (1.5 X interquartile range of coefficient uncertainties) outlier criteria. Coefficient results were considered outliers only if they fell outside both ranges. Outliers were removed from the distribution prior to performing later tests.
2) Counting the number of estimates that had an uncertainty (standard deviation) greater than the total range of coefficient estimates in the distribution (after outliers had been removed). This was an arbitrary-yet-useful criteria for determining poorly-constrained coefficients, which was justified by examining the distribution of coefficient estimate standard deviations and recognizing that a large gap separated "well-constrained" and "poorly-constrained" coefficient estimates. Our criteria, although somewhat arbitrary, successfully and consistently identified coefficient estimates at the "poorly-constrained" end. Poorly-constrained coefficient estimates were removed from the distribution for later tests.
3) Counting the number of coefficient estimates for which the p-value of the Wald test statistic was less than 0.01. The Wald test statistic is calculated by R for each coefficient as a part of the model creation process, based on the coefficient estimate and the coefficient estimate uncertainty. The statistic, in this instance, tests if the coefficient estimate is significantly different from zero, and we set the criteria for that significance at the 0.01 level. Coefficients that failed this test were counted but were kept in the distribution for later tests.
4) We determined the Shapiro-Wilk test statistic as a test of the normality of the coefficient estimate distribution. We also fitted a normal distribution to each coefficient estimate distribution, and determined the Kolmogorov-Smirnov one-sample test statistic for goodness-of-fit. For each test, we flagged variables with statistic p-values <0.01 (i.e.,

those that were statistically different from the normal distribution at the 0.01 significance level).

5) We used the T-test on the coefficient distribution to assess whether the mean of the distribution was significantly different from zero, and flagged variables with T-test p-values >0.01 (i.e., those that were *not* statistically different from zero at the 0.01 significance level).

6) We evaluated graphs of the distribution of coefficient estimates to qualitatively assess the consistency of coefficient estimates (e.g., Figure 4.9).

Of these tests, the most crucial proved to be the last. For several variables, the coefficient estimates were well distributed to either side of zero, indicating that the relationship between the variable and landslide occurrence would switch polarity depending on the location on which the model was trained. This was a disconcerting observation, as it suggested that the variable could not be reliably trained in one location then applied to another. The coefficient thus could not be reliably extrapolated from the training area to the remainder of the province. Coefficient distributions with at least 25% of well-constrained, non-outlier values to both sides of zero were flagged as potentially unreliable. The precipitation and distance to fault variables, in particular, commonly spanned zero in this sense (Table 4.6). The T-test was useful as well, particularly in the sense that the test often identified coefficient distributions that spanned zero but also in identifying parameters with potentially insignificant influence on the model.

The first three tests we found useful in determining what variables were well-sampled and whose coefficients were well-constrained. This was particularly useful for evaluating the geology and slope aspect categorizations, as we discovered that many preliminary geologic unit groups and the preliminary 9-category slope aspect groups were undersampled and poorly-constrained. This led to refinements in these categorizations. These tests were also useful in determining the provinces for which the distance to major river variable was well-sampled and had a well-constrained coefficient. We flagged those entries where less than 50 (half of the total number of subsamples processed) model runs produced well-constrained, non-outlier results.

The tests for normality (fourth test above) proved least useful. Failure of the Shapiro-Wilk test for normality was fairly common, while failure of the Kolmogorov-Smirnov test for goodness-of-fit was fairly rare. The precipitation and distance to fault parameters most commonly failed the Shapiro-Wilk test across all provinces, as did some of the mechanically stronger geologic unit parameters. No reason for this pattern was established.

No single test was used unequivocally to discriminate between acceptable and unacceptable parameters. In the interest of having comparable models for each province, we sought to minimize the differences between the indicator variables used, and hence considered the test results for each parameter across provinces. We typically tried to use those parameters that showed good test results across the majority of provinces, while excluding parameters that showed poor results across the majority. Some exceptions were made, however, to improve individual model fits.

Table 4.5: Summary of correlation coefficients between all continuous variables across all provinces

Variable:	Slope angle			Pos. curvature			Neg. curvature			Elev. std. dev.			Slope std. dev.			Curv. std. dev.			Precipitation		
	Min.	Max.	Avg.	Min.	Max.	Avg.	Min.	Max.	Avg.	Min.	Max.	Avg.	Min.	Max.	Avg.	Min.	Max.	Avg.	Min.	Max.	Avg.
Elevation	0.00	0.29	0.07	0.00	0.02	0.01	0.00	0.01	0.01	0.00	0.30	0.08	0.00	0.20	0.04	0.00	0.13	0.04	0.47	0.81	0.66
Slope angle				0.04	0.07	0.06	0.02	0.08	0.04	0.88	0.96	0.93	0.22	0.61	0.43	0.26	0.64	0.45	0.01	0.23	0.07
Pos. slope curv.							0.01	0.32	0.13	0.05	0.08	0.07	0.08	0.13	0.10	0.10	0.14	0.13	0.00	0.02	0.00
Neg. slope curv.										0.03	0.08	0.06	0.08	0.13	0.11	0.09	0.20	0.15	0.00	0.01	0.00
Elev. std. dev.													0.24	0.76	0.53	0.26	0.74	0.50	0.02	0.24	0.08
Slope std. dev.																0.47	0.79	0.61	0.00	0.15	0.03
Curv. std. dev.																			0.00	0.10	0.03

Notes:

  Avg.: average; curv.: curvature; elev.: elevation; max.: maximum; min.: minimum; neg.: negative; pos.: positive; std. dev.: standard deviation

  Variables correlating with slope angle highlighted in green

  Variables correlating with elevation highlighted in orange
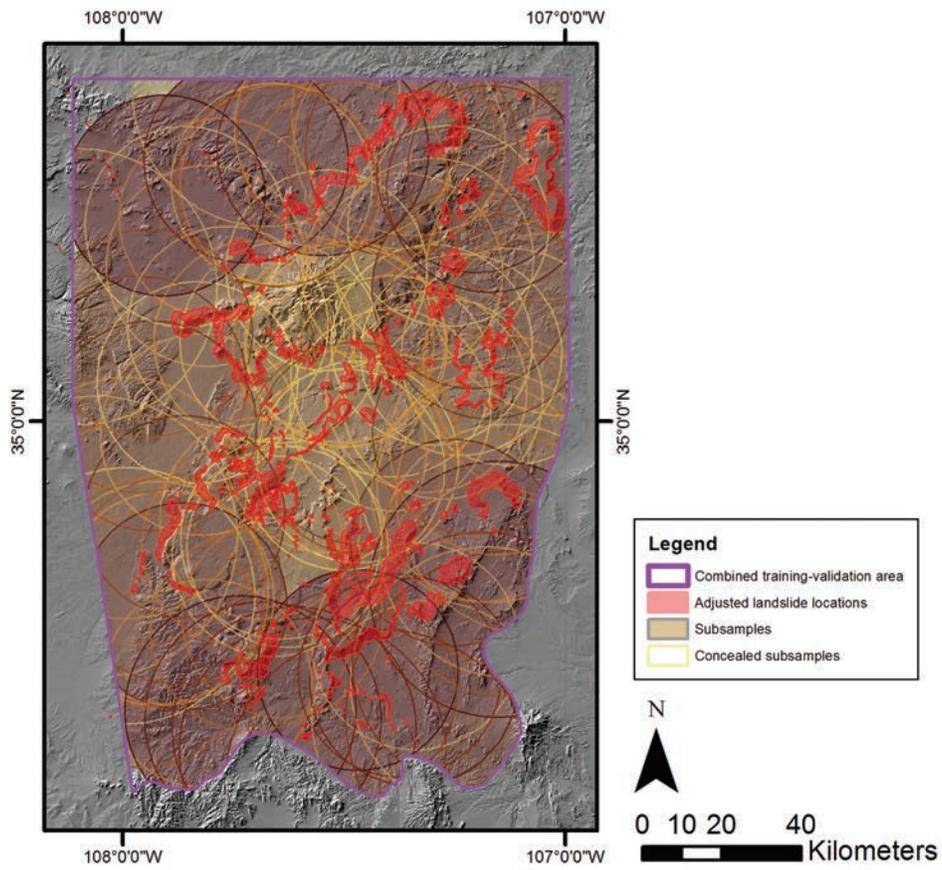
Figure 4.8: Example set of subsample areas, from the Colorado Plateau area. Each colored circle or circle outline is another subsample. Adjusted known landslide polygons are also shown.
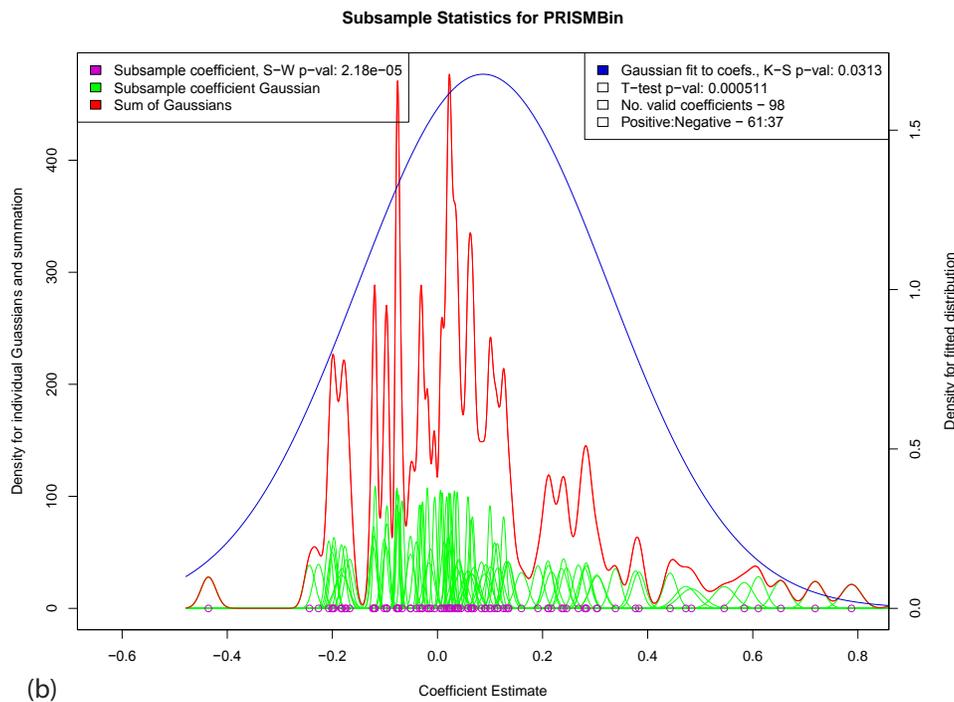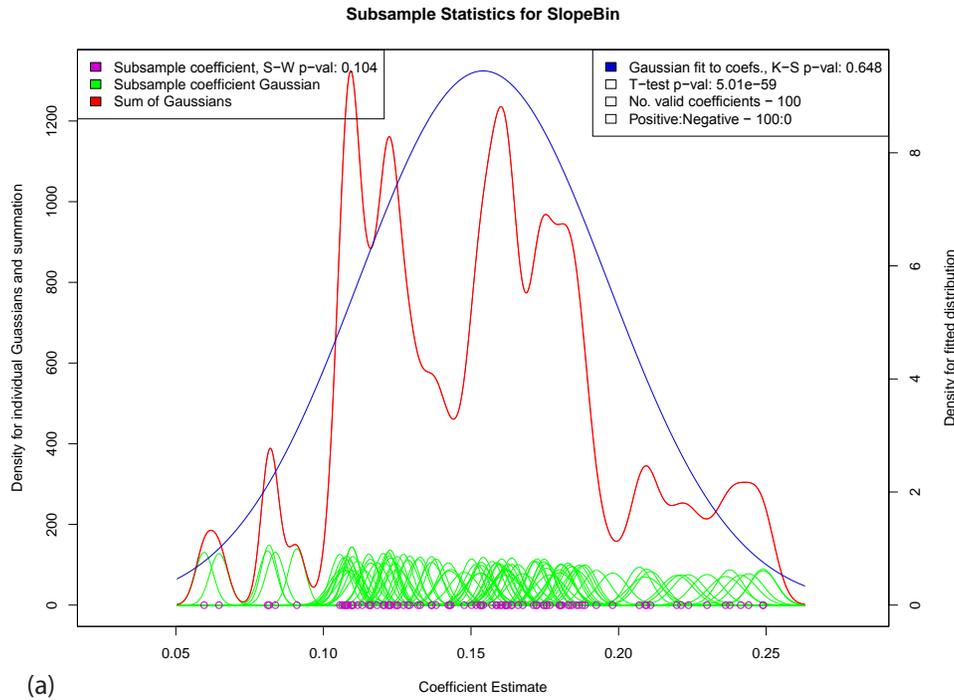
Figure 4.9: Example subsample parameter coefficient estimate variability evaluation plots. K-S: Kolmogorov-Smirnov goodness-of-fit test; S-W: Shapiro-Wilk test for normality. Individual coefficient estimates are shown as purple circles, while the uncertainty in each estimate is expressed as small green Gaussian distributions centered on the estimate. SlopeBin is the binned slope angle variable, PRISMBin is the binned precipitation variable. (a) Example evaluation plot for a variable with consistent relationship with landslide occurrence. (b) Example evaluation plot for a variable with inconsistent relationship with landslide occurrence.

Table 4.6: Summary of final subsampling routine results *Part 1: Continuous variables*

Precipitation

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR	100	0.000	0.001	0.125	0.002	0.920	No	Estimates are 40% positive, 60% negative
CP	98	0.000	0.087	0.238	0.031	0.001	No	Estimates are 62% positive, 38% negative
GP	100	0.000	0.041	0.462	0.128	0.379	No	Estimates are 63% positive, 37% negative
MD	99	0.000	-0.129	0.155	0.006	0.000	Yes	
NR	92	0.000	-0.002	0.305	0.023	0.949	No	Estimates are 42% positive, 58% negative
SRM	100	0.000	-0.050	0.083	0.000	0.000	No	Estimates are 38% positive, 62% negative

Slope angle

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR	100	0.000	0.094	0.067	0.038	0.000	Yes	
CP	100	0.104	0.154	0.042	0.648	0.000	Yes	
GP	96	0.035	0.149	0.048	0.520	0.000	Yes	
MD	100	0.000	-0.050	0.056	0.000	0.000	Yes	Consistently negative coefficients
NR	99	0.397	0.100	0.041	0.593	0.000	Yes	
SRM	100	0.027	-0.047	0.037	0.250	0.000	Yes	Consistently negative coefficients

Negative slope curvature

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR	100	0.000	0.162	0.116	0.020	0.000	Yes	
CP	100	0.248	0.201	0.069	0.841	0.000	Yes	
GP	100	0.002	0.174	0.057	0.295	0.000	Yes	
MD	100	0.012	-0.079	0.065	0.472	0.000	Yes	
NR	99	0.000	0.118	0.100	0.029	0.000	Yes	
SRM	97	0.188	-0.063	0.089	0.411	0.000	Yes	

Positive slope curvature

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR	100	0.000	0.097	0.128	0.034	0.000	No	Estimates are 68% positive, 32% negative
CP	100	0.989	0.135	0.064	0.994	0.000	Yes	
GP	99	0.002	0.216	0.106	0.177	0.000	Yes	
MD	100	0.393	-0.151	0.119	0.856	0.000	Yes	
NR	99	0.000	0.067	0.064	0.031	0.000	Yes	
SRM	97	0.040	-0.204	0.142	0.178	0.000	Yes	

Table 4.6: Summary of final subsampling routine results (continued) *Part 2: "Distance to" variables*

Proximal to a fault

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR	100	0.004	-0.058	0.975	0.123	0.559	No	Estimates are 46% positive, 54% negative
CP	87	0.008	-0.566	1.219	0.357	0.000	No	Estimates are 34% positive, 66% negative
GP	23	0.133	-0.890	1.245	0.547	0.003	No	Estimates are 30% positive, 70% negative
MD	92	0.000	0.779	0.741	0.001	0.000	Yes	
NR	78	0.000	1.516	2.473	0.010	0.000	Yes	
SRM	96	0.003	0.042	1.118	0.173	0.713	No	Estimates are 56% positive, 44% negative

Proximal to a river floodplain

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR	0	N/A	N/A	N/A	N/A	N/A		N/A
CP	8	0.391	-1.153	0.351	0.784	0.000	Yes	
GP	77	0.001	-0.239	1.203	0.040	0.087	No	Estimates are 38% positive, 62% negative
MD	25	0.002	0.774	0.779	0.109	0.000	Yes	
NR	89	0.025	3.000	1.219	0.193	0.000	Yes	
SRM	40	0.596	4.122	1.165	0.847	0.000	Yes	

Table 4.6: Summary of final subsampling routine results (continued)   *Part 3: Slope aspect*

Note that north aspect is used as a reference unit.

Flat aspect

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR	0	N/A	N/A	N/A	N/A	N/A	N/A	
CP	2	N/A	N/A	N/A	N/A	N/A	N/A	
GP	9	0.587	0.907	0.743	0.779	0.009	Yes	
MD	0	N/A	N/A	N/A	N/A	N/A	N/A	
NR	0	N/A	N/A	N/A	N/A	N/A	N/A	
SRM	9	0.275	-1.290	0.721	0.627	0.000	Yes	

South aspect

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR	100	0.004	-0.333	0.262	0.381	0.000	Yes	
CP	100	0.034	-0.218	0.498	0.266	0.000	No	Estimates are 26% positive, 74% negative
GP	100	0.162	-0.140	0.247	0.255	0.000	Yes	Estimates are 25% positive, 75% negative
MD	99	0.000	-0.644	0.797	0.010	0.000	Yes	
NR	98	0.206	0.878	0.413	0.652	0.000	Yes	
SRM	100	0.016	-0.591	0.511	0.567	0.000	Yes	

Table 4.6: Summary of final subsampling routine results (continued)   *Part 4: Geologic units*

Headings are general geologic unit, equivalent province-specific units are in parantheses. Note that unit S1 (or Su) is used as a reference

Unit S2a

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR (S1)								Combined with reference unit
CP (S2)	92	0.121	2.518	0.960	0.718	0.000	Yes	
GP	42	0.130	-2.724	1.871	0.537	0.000	Yes	
MD (Su)								Combined with reference unit
NR (S2M)	86	0.000	2.729	2.192	0.016	0.000	Yes	
SRM (S2M)	78	0.048	3.166	1.873	0.527	0.000	Yes	

Unit S2b

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR (S2)	48	0.000	1.839	2.018	0.009	0.000	Yes	
CP (S2)	92	0.121	2.518	0.960	0.718	0.000	Yes	
GP	91	0.000	1.236	1.994	0.035	0.000	Yes	
MD (Su)								Combined with reference unit
NR (S2M)	84	0.000	2.729	2.192	0.016	0.000	Yes	
SRM (S2M)	78	0.048	3.166	1.873	0.527	0.000	Yes	

Unit M1

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR	60	0.000	2.610	1.522	0.121	0.000	Yes	
CP	90	0.044	0.021	1.778	0.346	0.910	No	Estimates are 53% positive, 47% negative
GP	82	0.049	-0.410	2.361	0.680	0.122	No	Estimates are 49% positive, 51% negative
MD (StM)	43	0.001	2.717	1.577	0.146	0.000	Yes	
NR (S2M)	84	0.000	2.729	2.192	0.016	0.000	Yes	
SRM (S2M)	78	0.048	3.166	1.873	0.527	0.000	Yes	

Unit St1

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR (Stu)	60	0.112	-0.341	1.704	0.841	0.130	No	Estimates are 42% positive, 58% negative
CP	93	0.089	-0.400	1.320	0.878	0.005	No	Estimates are 35% positive, 65% negative
GP (Stu)	90	0.000	-0.374	2.425	0.050	0.149	No	Esimates are 52% positive, 48% negative
MD (StM)	43	0.001	2.717	1.577	0.146	0.000	Yes	
NR (St)	86	0.000	2.168	2.039	0.237	0.000	Yes	
SRM (St)	83	0.015	1.634	2.644	0.304	0.000	No	Estimates are 69% positive, 31% negative

Unit St2

	No. valid [1]	S.W. P-v [2]	Avg. [3]	Std. dev. [4]	K.S. p-v [5]	T-test p-v [6]	Cons. Pol. [7]	Comments
BR (Stu)	60	0.112	-0.341	1.704	0.841	0.130	No	Estimates are 42% positive, 58% negative
CP	34	0.019	-1.523	1.647	0.652	0.000	Yes	
GP (Stu)	90	0.000	-0.374	2.425	0.050	0.149	No	Esimates are 52% positive, 48% negative
MD (StM)	43	0.001	2.717	1.577	0.146	0.000	Yes	
NR (St)	86	0.000	2.168	2.039	0.237	0.000	Yes	
SRM (St)	83	0.015	1.634	2.644	0.304	0.000	No	Estimates are 69% positive, 31% negative

Notes:

1: Number of non-outlier, well-constrained subsamples

2: P-value of the Shapiro-Wilk test for normality

3: Average of the distribution of coefficient estimates

4: Standard deviation of the distribution of coefficient estimates

5: P-value of the Kolmogorov-Smirnov test for goodness-of-fit

6: P-value of the T-test for the mean of the distribution of coefficients being different from zero

7: Evaluation of distribution for constant polarity (positive or negative) relationship with landslide occurrence

Test failures are highlighted in orange

Province abbreviations:

BR: Basin and Range; CP: Colorado Plateau; GP: Great Plains; MD: Mogollon-Datil; NR: North Rift; SRM: Southern Rocky Mountains

4.3.4   Treatment of slope in Mogollon-Datil and Southern Rocky Mountains areas

Models created for the Mogollon-Datil and Southern Rocky Mountains provinces consistently determined negative coefficients for the slope variable (e.g., Table 4.6), suggesting that in these areas landslide susceptibility decreases with increasing slope angle, and resulting in susceptibility maps where the highest susceptibilities were determined for flat areas and the lowest susceptibilities for steep areas (e.g., Figure 4.5). This is physically unrealistic, and likely the result of our use of landslide deposit locations as training data for modeling landslide susceptibility (see "Discussion" section for further assessment). These two provinces are characterized by common steep, high-relief topography, but contain relatively few mapped deep-seated landslides. The landslides that are found commonly slid or rotated down to relatively shallow slope angles. As a consequence, as compared to other province training areas, the Mogollon-Datil and Southern Rocky Mountains training areas sample a greater proportion of steeper slope bins overall (Figure 4.10a) and a lower proportion of landslide areas in the steeper slope bins (Figure 4.10b). This apparently resulted in the logistic regression process determining a negative coefficient for the slope angle parameter.

Since this model coefficient result is physically unrealistic and qualitatively resulting in a poor representation of susceptibility, we chose to modify the treatment of slope angle in these two provinces. Since a landslide deposit is where a moving landslide came to rest, we reasoned that the slope angle of the deposit itself is a representation of a stable slope angle, and that stability should not increase with increasing slope angle above the dominant slope angle of the deposits. To incorporate this hypothesis in the logistic regression models, we first assessed the distributions of landslide deposit slope angles for each province training and validation area, and identified the slope bin associated with the peak of each distribution (Figure 4.11). We then used the slope angle parameter in subsequent logistic regression models as a categorical variable, with each slope bin as a separate category, and determined a coefficient for each slope bin category. Following the hypothesis that slope angles above the dominant slope angle of landslide deposits cannot be more stable than the dominant slope angle of the deposits, we applied the coefficient associated with the peak of the slope angle distribution as a minimum coefficient to all slope angles above that angle (e.g., Table 4.7). The end result is that model probability is prohibited from decreasing as a result of increasing slope, regardless of the coefficients determined for steeper slopes. Figure 4.5 displays the effect; the result is more physically realistic as model probabilities are higher for steeper slopes and lower for shallow slopes.

4.4   Summary of final input parameters

The results of the above data processing and indicator variable tests are summarized below. Note that a reference category needs to be defined for the categorical variables used in the regression model; the reference categories used are given below. Table 4.8 summarizes the indicator variables used.

4.4.1   Statewide map of deep-seated landslides

The linework and point data from the Cardinali et al. (1990) deep-seated landslide map were adjusted to reduce map inaccuracies to <200 m, and often <150 m, in the training and validation areas of all provinces. The adjusted lines and points were used to build the landslide deposit polygons that were subsequently used in the data sampling procedure to locate landslide points and non-landslide points for training and validating the regression models.

(a)

(b)
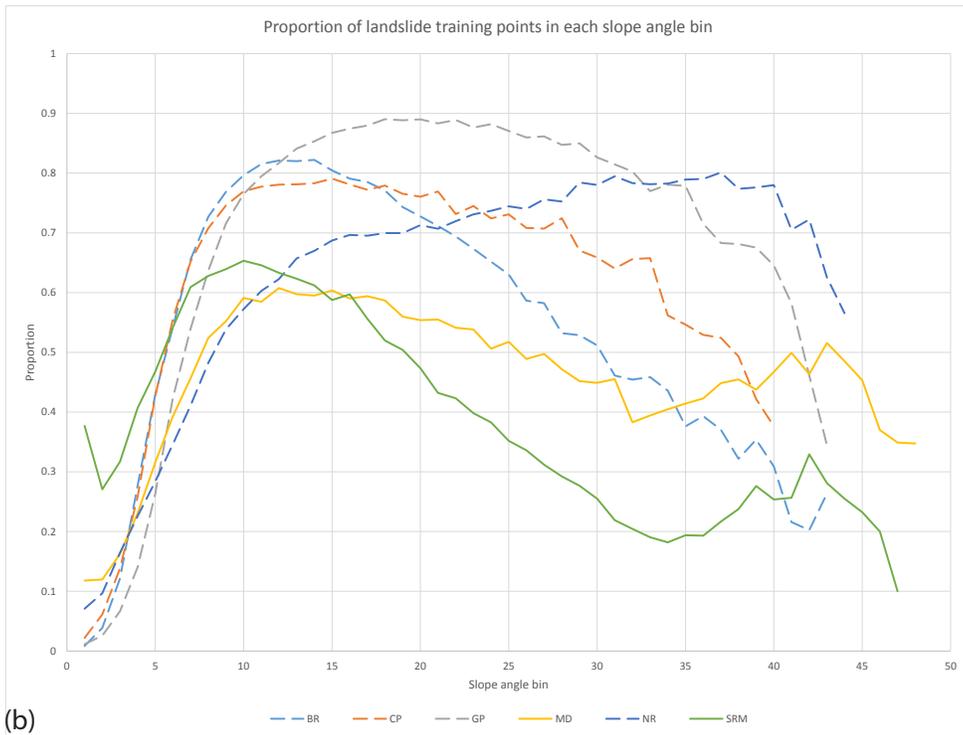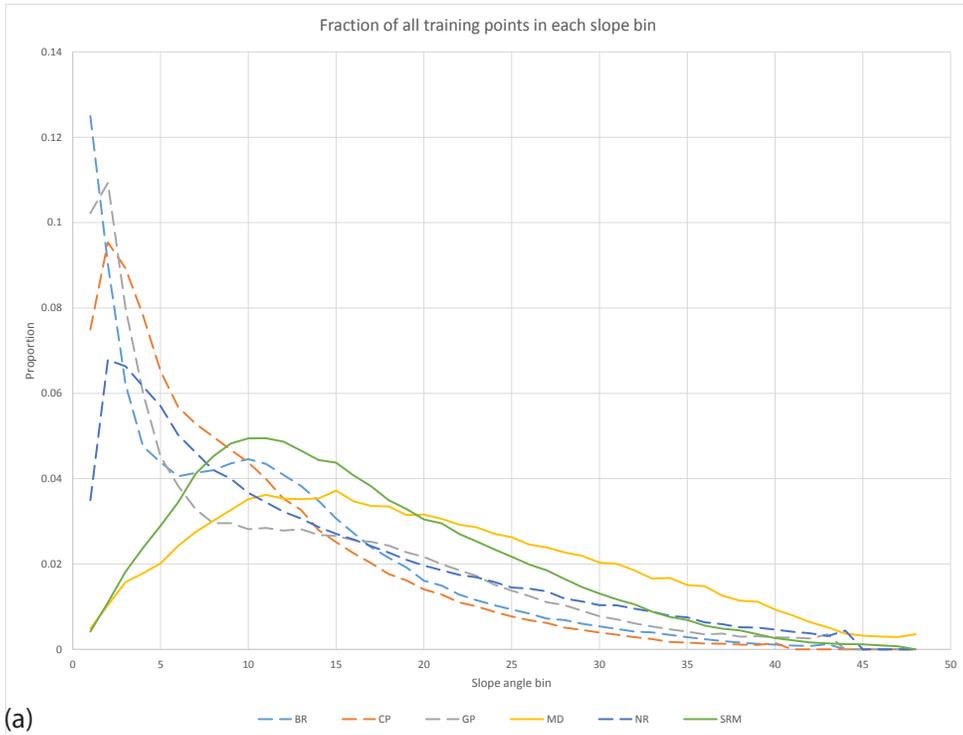
*Figure 4.10: Proportions of training data sample points in each slope angle bin by province. BR: Basin and Range; CP: Colorado Plateau; GP: Great Plains; MD: Mogollon-Datil; NR: North Rift; SRM: Southern Rocky Mountains. The Mogollon-Datil and Southern Rocky Mountains lines are shown solid, all others are shown dashed. (a) Proportion of all training points in each bin. (b) Proportion in each bin that were from landslide points.*
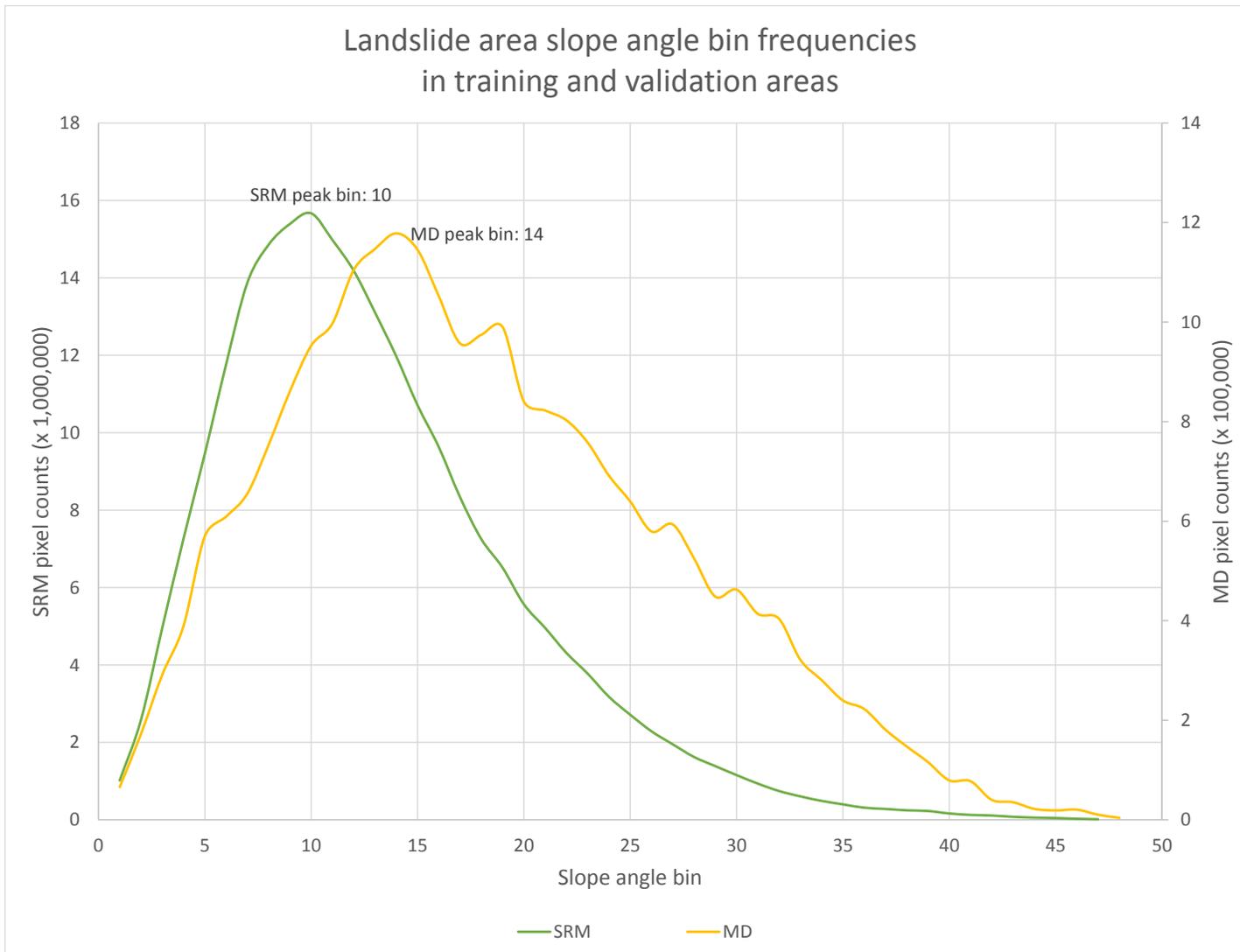
Figure 4.11: Distributions of slope angle bins in known landslide areas for the Mogollon-Datil (MD) and Southern Rocky Mountain (SRM) provinces. Data is counts of all landslide deposit pixels in both the training and validation areas of each province.

**Table 4.7: Final slope angle coefficients for the Mogollon-Datil and Southern Rocky Mountains models**

	Mogollon-Datil	Southern Rocky Mtns	
Slope angle bin	Coefficient	Coefficient	Comments
1	(Reference bin)	(Reference bin)	
2	-0.129	-0.297	
3	0.157	-0.048	
4	0.535	0.432	
5	0.991	0.699	
6	1.310	1.016	
7	1.560	1.275	
8	1.861	1.324	
9	1.959	1.325	
10	2.127	**1.379**	Peak bin for SRM
11	2.102	**1.379**	Uses peak bin coefficient (SRM)
12	2.204	**1.379**	Uses peak bin coefficient (SRM)
13	2.166	**1.379**	Uses peak bin coefficient (SRM)
14	**2.156**	**1.379**	Peak bin for MD
15	2.226	**1.379**	Uses peak bin coefficient (SRM)
16	2.161	**1.379**	Uses peak bin coefficient (SRM)
17	2.194	**1.379**	Uses peak bin coefficient (SRM)
18	2.186	**1.379**	Uses peak bin coefficient (SRM)
19	**2.156**	**1.379**	Uses peak bin coefficient (SRM, MD)
20 and greater	**2.156**	**1.379**	Uses peak bin coefficient (SRM, MD)

Notes:

  MD: Mogollon-Datil; SRM: Southern Rocky Mountains

Table 4.8: Summary of indicator variables considered

Variable	Type	Binning	Used?	Comments
Escarpments map	Proximity variable	1 if within 1 km of an escarpment, 0 otherwise	No	Disproportionately influential variable
Geology	Categorical	Province-specific	Yes	Binning iteratively optimized throughout
Fault map	Proximity variable	1 if within 1 km of a fault, 0 otherwise	No	Poor subsampling test performance
Elevation	Continuous	Restricted equal frequency	No	Colinear with precipitation variable
Slope angle	Continuous (BR, CP, GP, NR)	Restricted equal frequency	Yes	
	Modified categorical (MD, SRM)	Restricted equal frequency	Yes	
Slope curvature	Continuous	Restricted equal frequency	Yes	Divided into two separate inputs, negative and positive curvature
Elevation std. dev.[1]	Continuous	Restricted equal frequency	No	Colinear with slope angle variable
Slope angle std. dev.[1]	Continuous	Restricted equal frequency	No	Colinear with slope angle variable
Slope curv. std. dev.[1]	Continuous	Restricted equal frequency	No	Colinear with slope angle variable
Slope aspect	Categorical	North, south, flat	Yes	Excluded from final North Rift model
Precipitation	Continuous	Restricted equal frequency	No*	Poor subsampling test performance; *used in final Mogollon-Datil models
Stream map	Proximity variable	1 if within 500 m of a river floodplain, 0 otherwise	No	Poor subsampling test performance; potentially not independent of elevation

Notes:

 1: std. dev.: standard deviation within a 5 pixel by 5 pixel area

 Curv.: curvature

4.4.2 Map of escarpments

The distance-to-escarpment parameter was tested as a simple dichotomous parameter: a point recorded a value of 1 if within 1 km of an escarpment or a 0 if not. The parameter was determined to be disproportionately influential on model results, however, as high model probabilities generally only occurred within the region next to the mapped escarpments. The parameter was therefore not used in final models.

4.4.3 Geologic map

The geologic map was incorporated as a set of categorical variables, grouped according to Table 4.4. These units passed the above tests, and were used in final models. For consistency, the unit S1 (unconsolidated/uncompacted sediment deposits), which occurs in almost all provinces, was used as the reference category. In the Mogollon-Datil area, the lumped unit Su was used.

The distance-to-fault parameter was tested as a simple dichotomous parameter: a point recorded a value of 1 if within 1 km of a fault or a 0 if not. This parameter was assessed as a part of the spatial subsampling routine described above, which determined that for the Basin and Range, Colorado Plateau, Great Plains, and Southern Rocky Mountains areas the coefficient for the parameter varies between positive and negative with subsample location. In addition, the mean of the distribution of coefficients was not statistically different from 0 at the 0.01 significance level according to the T-test for the Basin and Range and Southern Rocky Mountains areas. Finally, less than 50% of subsamples provided well-constrained, non-outlier coefficient estimates in the Great Plains area. The parameter was therefore not used in subsequent models.

4.4.4 Topographic data

4.4.4.1 Elevation

Elevation was binned as a continuous variable, but not used due to collinearity with the precipitation data.

4.4.4.2 Slope angle

Slope angle was binned as a continuous variable and used in all provinces. For the Basin and Range, Colorado Plateau, Great Plains, and North Rift provinces, binned slope angle was incorporated into the logistic regression models directly as a continuous variable. For the Mogollon-Datil and Southern Rocky Mountains areas, binned slope was incorporated as a categorical variable, with the coefficient of the peak of the landslide area slope bin frequency distribution used as a minimum coefficient for all steeper slope bins, as described above. In each, the reference category was the shallowest slope bin.

4.4.4.3 Slope curvature

Curvature was divided into positive and negative curvatures, each was binned as a continuous variable, and each was used in all provinces.

4.4.4.4 Elevation standard deviation

The standard deviation of the elevation dataset was binned as a continuous variable, but not used due to collinearity with the slope angle data.

4.4.4.5   Slope angle standard deviation

The standard deviation of the slope angle dataset was binned as a continuous variable, but not used due to collinearity with the slope angle data.

4.4.4.6   Slope curvature standard deviation

The standard deviation of the slope curvature dataset was binned as a continuous variable, but not used due to collinearity with the slope angle data.

4.4.4.7   Slope aspect

Slope aspect was categorized as north, south, or flat, with the division between north and south placed at 67.5° and 247.5°, as described above. Slope aspect was evaluated using the spatial subsampling method described above, which found that the coefficient for the south aspect parameter varies somewhat between positive and negative depending on subsample location, but only for the Colorado Plateau area did the distribution of parameters fail our 25% guideline. Also, for all provinces the mean of the distribution of coefficients was statistically different from 0 at the 0.01 significance level according to the T-test, and for all provinces 99-100% of subsamples returned well-constrained, non-outlier coefficient estimates. We therefore chose to include the aspect parameter. Initially, we used aspect as an input for all provinces, but subsequent model results found that removing the aspect parameter from the North Rift model improved model fit significantly. For all provinces, the north aspect direction was used as the reference category.

4.4.5   Precipitation data

The precipitation dataset was binned as a continuous variable. The binned precipitation data was evaluated using the spatial subsampling method described above, which determined that in the Basin and Range, Colorado Plateau, Great Plains, North Rift, and Southern Rocky Mountains areas the coefficient for the precipitation parameter varies between positive and negative with subsample location. In addition, the mean of the distribution of coefficients was not statistically different from 0 at the 0.01 significance level according to the T-test for the Basin and Range, Great Plains, and North Rift areas. The precipitation parameter was therefore generally excluded. Incorporation of the precipitation parameter in the Mogollon-Datil area improved model performance, however, and hence the dataset was used in that province.

4.4.6   Hydrographic data

Distance-to-stream was incorporated as a simple dichotomous parameter: a point recorded a value of 1 if within 500 m of a major river floodplain or a 0 if not. This parameter was assessed as a part of the spatial subsampling routine described above, which determined that more than 50% of subsample models produced poorly-constrained or outlier coefficients in the Basin and Range, Colorado Plateau, Mogollon-Datil, and Southern Rocky Mountains areas. In addition, the distribution of parameter coefficients for the Great Plains area varied between positive and negative with subsample location, and the mean of the distribution was not statistically significantly different from 0 at the 0.01 significance level. We also were concerned that the distance-to-stream variable may not be independent of the elevation or precipitation variables. We therefore excluded the distance-to-stream parameter from subsequent models.

4.5   Model-driven refinements

Semi-quantitative and qualitative reviews of model results at various stages led to model refinements. Model evaluation methods are described above. Refinements incorporated as a result of the input parameter tests are described above. Some additional refinements post-dated the indicator variable tests. For example, some geologic units in the Mogollon-Datil, North Rift, and Southern Rocky Mountains areas were further grouped as a result of poor sampling and unit-specific poor model results. Also, for some provinces we experimented with training a model on the original validation area and validating on the original training area. This proved to provide better model performance in the Great Plains, Mogollon-Datil, and Southern Rocky Mountains areas. Finally, post-test models found better fits for the North Rift model when aspect was excluded, and better fits for the Mogollon-Datil model when precipitation was included. These changes were incorporated after indicator variable testing was completed.

5 Final Model Results

5.1 Individual model results

Individual logistic regression models were constructed for each province, using the indicator variables described above. Due to variable binning and model parameterization strategies, the coefficients of individual models (Table 5.1) are not strictly comparable. Some trends can be teased out, however, that appear to reflect the nature of the individual provinces. Some commonalities amongst the provinces:

1) Coefficients for the stronger geologic units (St1, St2, Stu, St) are consistently less than the coefficients for soft or interbedded units. Where split out, coefficients for interbedded strong and weak layer units (M1) are typically less than those for soft units (S2a, S2b, S2), except for in the Basin and Range. Where split out, soft sedimentary rocks (e.g., shales: unit S2b) have stronger positive relationship with landslide occurrence than sediment and very weakly cemented sedimentary rocks (units S1, S2a). These trends suggest that the most landslide-prone terrain is underlain by soft sedimentary rocks, while the least landslide-prone terrain is that underlain by strong geologic units, as would be expected.

2) Northerly slope aspects are somewhat more prone to landslide occurrence than southerly slope aspects, as evidenced by the weak negative coefficients for the south slope aspect direction across the five provinces that included aspect as a parameter.

3) Higher magnitude curvature, both positive and negative, is generally positively correlated to landslide occurrence, but not in the Mogollon-Datil and Southern Rocky Mountains areas. This may be a product of the topographic setting, as described in more detail in the "Treatment of slope in Mogollon-Datil and Southern Rocky Mountains" section above. In these two regions, common rugged topography may result in many non-landslide sample points recording high curvatures, resulting in negative correlations between landslide occurrence and magnitude of curvature. In the remaining provinces, landslide-covered slopes are some of the more rugged topography, leading to positive coefficients.

5.1.1 Basin and Range

As compared to other provinces, the Basin and Range model coefficients exhibit a relatively weak relationship between slope angle and landslide susceptibility and a relatively strong relationship between susceptibility and geologic unit (Table 5.1). The implication is that even relatively shallow slopes may be susceptible to landslide generation, provided the slope is underlain by a soft geologic unit or interbedded soft and weak rocks. This perhaps reflects the structural setting of the Basin and Range, in that many of the mountain ranges in this province are underlain by tilted fault blocks and dipping strata. Dipping sedimentary beds are perhaps particularly susceptible to landslide generation, as the inclined bedding planes can become slip surfaces for the landslide to fail along. A steep slope is then not necessary for slope failure, provided the underlying strata contain numerous inclined bedding planes. The strong relationship between geologic unit and landslide susceptibility appears to provide an excellent predictor for landslide occurrence, as the Basin and Range model has the highest area under the ROC curve of any province for both internal and external validations (Table 5.2, Figure 5.1).

Table 5.1: Summary of coefficients for final models

Variable	BR	CP	GP	MD	NR	SRM
	Coefficients					
(Model intercept)	-2.714	-2.252	-2.843	-1.775	-2.200	-1.102
Precipitation	Not used			-0.037	Not used	
Slope angle	0.070	0.142	0.138	***	0.091	***
Min***	N/A			-0.129	N/A	-0.297
Max***				2.226		1.379
Neg. Curv.	0.122	0.161	0.197	-0.041	0.089	-0.064
Pos. Curv.	0.087	0.098	0.292	-0.085	0.042	-0.172
<u>*Aspect*</u>						
Flat	-8.820	-5.192	-11.614	-8.704	Not used	1.707
North	(Ref.)	(Ref.)	(Ref.)	(Ref.)		(Ref.)
South	-0.400	-0.335	-0.078	-0.313		-0.182
<u>*Geology*</u>	*(Colors indicate province-specific groupings)*					
<u>*S1*</u>	(Ref.)	(Ref.)	(Ref.)	(Ref.)	(Ref.)	(Ref.)
S2a	(Ref.)	1.747	-2.763	(Ref.)	1.268	1.217
S2b	2.689	1.747	1.123	(Ref.)	1.268	1.217
M1	3.121	0.858	0.413	1.325	1.268	1.217
St1	0.190	-0.522	0.265	1.325	-0.390	0.444
St2	0.190	-3.391	0.265	1.325	-0.390	0.444
Province-specific groups:	S1			Su		
		S2			S2M	S2M
	Stu		Stu	StM	St	St

Notes:

  BR: Basin and Range; CP: Colorado Plateau; GP: Great Plains; MD: Mogollon-Datil; NR: North Rift; SRM: Southern Rocky Mountains

  ***: MD and SRM areas use a modified categorical binning for slope. See Table 4.7 for full coefficients

  Ref.: reference category

Table 5.2: Summary of final model evaluation measures

	Evaluators		BR	CP	GP	MD	NR	SRM	Avgs
Internal validation	**ROC AUC**[1]	*Empirical*	0.87	0.83	0.87	0.68	0.79	0.65	0.78
		Binormal	0.86	0.81	0.86	0.68	0.79	0.66	0.78
	Landslide area percentiles[2]	*100%*	0.06	0.01	0.01	0.13	0.09	0.10	0.07
		95%	0.17	0.26	0.31	0.33	0.25	0.40	0.29
		90%	0.38	0.38	0.39	0.41	0.37	0.46	0.40
		75%	0.69	0.54	0.56	0.52	0.48	0.56	0.56
		50%	0.77	0.70	0.77	0.62	0.64	0.68	0.70
External validation	**ROC AUC**[1]	*Empirical*	0.91	0.87	0.83	0.66	0.76	0.64	0.78
		Binormal	0.91	0.84	0.82	0.64	0.78	0.63	0.77
	Landslide area percentiles[2]	*100%*	0.09	0.02	0.00	0.11	0.10	0.14	0.08
		95%	0.20	0.30	0.35	0.39	0.39	0.39	0.34
		90%	0.27	0.44	0.42	0.52	0.44	0.49	0.43
		75%	0.71	0.58	0.57	0.61	0.53	0.60	0.60
		50%	0.84	0.71	0.75	0.67	0.65	0.74	0.73

Notes:

1: ROC AUC: Receiver operator characteristic curve, area under the curve:

Empirical: integrated based on actual results; binormal: best-fit smoothed curve

2: Percentage refers to proportion of landslide area validation points with model probability above

the given threshold probability

BR: Basin and Range; CP: Colorado Plateau; GP: Great Plains; MD: Mogollon-Datil; NR: North Rift;

*SRM: Southern Rocky Mountains; Avgs: average results across provinces*
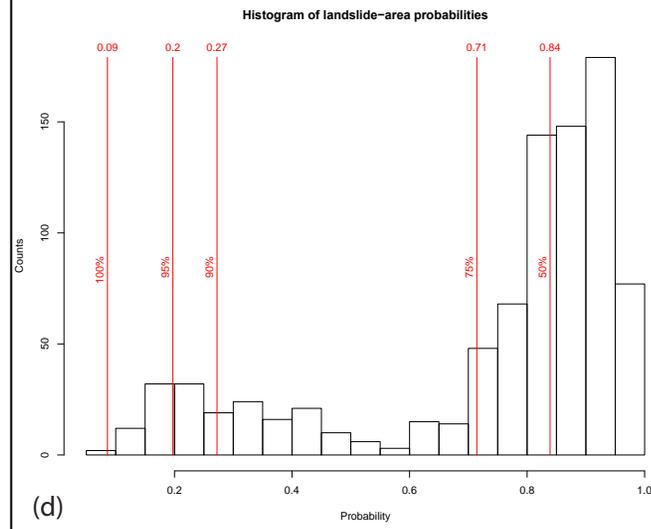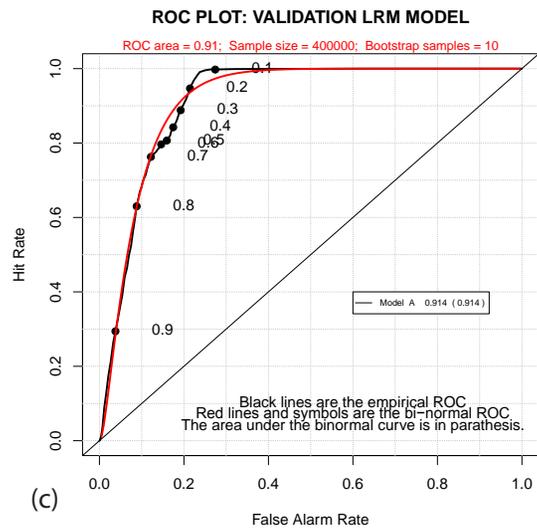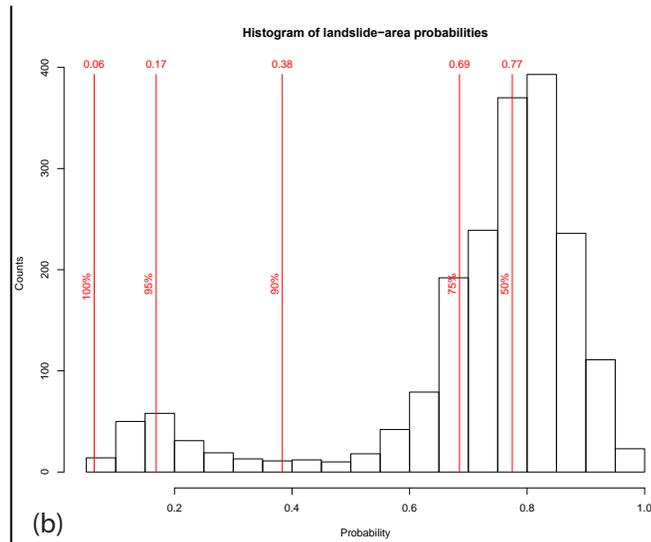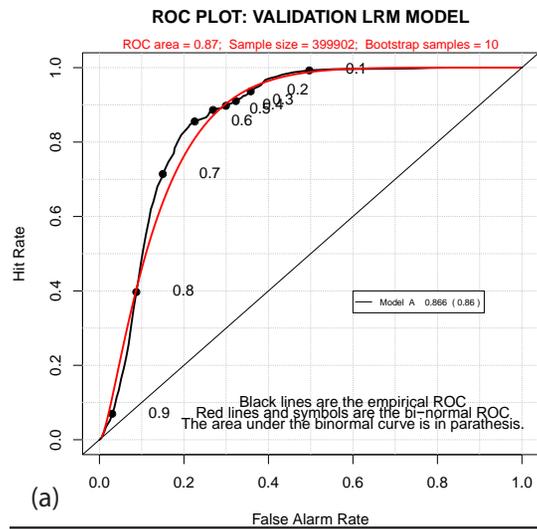
Figure 5.1: Model evaluation plots for the final Basin and Range model. ROC plots generated by the LAND-SE script (Rossi and Reichenbach, 2016). (a) Internal validation ROC plot. (b) Internal validation histogram of known landslide-area model probabilities. (c) External validation ROC plot. (d) External validation histogram of known landslide-area model probabilities.

5.1.2 Colorado Plateau

As compared to the other provinces, the Colorado Plateau model coefficients demonstrate a relatively strong relationship between slope angle and landslide susceptibility, and a relatively strong suppression of susceptibility by strong geologic units (St1, St2; Table 5.1). Steep slopes underlain by soft or interbedded-soft-and-hard rocks would show the highest susceptibilities given these coefficients. This is not surprising; the Colorado Plateau area is characterized by mesas and plateaus underlain by interbedded shales/mudstones and sandstones, and one would expect the soft shales and mudstones beneath mesa flanks to be particularly susceptible to landsliding. The strong geologic and topographic controls apparently result in an efficacious predictor of landslide occurrence, as the Colorado Plateau model has an above average area under the ROC curve for both internal and external validations (Table 5.2, Figure 5.2).

5.1.3 Great Plains

As compared to other provinces, the Great Plains model coefficients demonstrate a relatively strong relationship between slope angle and landslide susceptibility, and relatively weak relationship between geologic unit and susceptibility (Table 5.1). The range in coefficient values for soft sedimentary rocks, interbedded soft and hard rocks, and hard rocks (units S2b, M1, and Stu, respectively) is not as wide as for other provinces, suggesting that rock type is less a factor here. However, it should be noted that the coefficient for the consolidated sediment unit, unit S2a, is particularly strongly negative, indicating a suppression of landslide occurrence over this unit. The coefficient for slope angle, meanwhile, is one of the largest in magnitude. The province consists of broad plains and plateaus, and landslides typically occur only along the steep-sided flanks of plateaus, leading to this model parameterization. The relationship is apparently an efficacious predictor of landslide occurrence, as the Great Plains model has an above average area under the ROC curve for both internal and external validations (Table 5.2, Figure 5.3).

5.1.4 Mogollon-Datil area

The Mogollon-Datil area model is difficult to compare to the other provinces due to the unique parameterization, using a modified slope angle input and additionally incorporating precipitation data. The area also has the most simplified geologic binning. One difference that is apparent is the large, positive coefficient for the strongest geologic unit, a feature not seen in any other province. This observation is not a surprise, however, as the strong geologic unit underlies and upholds the majority of the high-relief terrain where landslides occur. The coefficients for the slope parameter show a greater range than those in the Southern Rocky Mountains province, suggesting a stronger relationship between slope angle and landslide occurrence in this province. This model appears to have below average discriminatory capacity versus other models, as demonstrated by the low area under the ROC curve (Table 5.2, Figure 5.4), but one cause for the low area may simply be the general lack of landslide deposits in this province, resulting in many false positives.

5.1.5 North Rift

As compared to other provinces, the North Rift model coefficients demonstrate a relatively weak relationship between slope angle and landslide occurrence, and relatively strong relationship between geologic unit and landslide occurrence (Table 5.1). The province has a relatively less diverse geologic categorization scheme, which was the result of the relative

homogeneity of the rock types found in the province. The North Rift area is underlain mainly by strong tabular volcanic rocks (basalts, tuffs) capping consolidated sediments or weakly cemented sedimentary rocks (combined into unit S2M), both of which are inset upon by young sediment (unit S1). Not surprisingly, the capping rocks tend to suppress model probabilities, while unit S2M, which often underlies steep hillsides flanking stream valleys where landslides occur, tends to increase model probabilities. The weak relationship with slope is surprising, but perhaps reflects the weak mechanical strength of the sediments within which landslides are occurring; particularly weak strata may not need as steep slopes to fail as landslides. In terms of discriminatory capacity, this model is one of the weaker models, but the area under the ROC curve for both internal and external validations is about average for the six models compared (Table 5.2, Figure 5.5).

5.1.6   Southern Rocky Mountains

The Southern Rocky Mountains area is difficult to compare to other provinces due to the alternative treatment of the slope angle parameter. The geologic unit coefficients suggest a relatively weak relationship between geologic unit and landslide occurrence, as the range in coefficients is relatively narrow. The range in slope coefficient values is also narrower than that for the Mogollon-Datil area. The implication is that landslide susceptibility through the Southern Rocky Mountains area is relatively constant, which would be a sensible result given the common high-relief terrain in the province. The consequence of a relatively invariant susceptibility model, however, is a lack of discriminatory capacity, as is reflected by the poor area under the ROC curve (Table 5.2, Figure 5.6). Another possible cause for the poor ROC curve may be the relative paucity of mapped landslide deposits, resulting in many false positives.

5.1.7   Tableland vs mountainous provinces

Several aspects of the final model results naturally group the provinces into "tableland" (Colorado Plateau, Great Plains, North Rift, and Basin and Range) and "mountainous" (Mogollon-Datil and Southern Rocky Mountains) provinces:

1) The coefficient of a continuous slope parameter – positive for tableland, negative for mountainous
2) The coefficient of the curvature parameters – positive for tableland, negative for mountainous
3) Quantitative measures of model performance – consistently lower areas under the ROC curve for mountainous versus tableland

Our hypothesis for these differences was touched on above in the "Treatment of slope in Mogollon-Datil and Southern Rocky Mountains," namely that the more mountainous provinces have common rugged terrain with relatively low slope angle landslide deposits (Figure 4.10), resulting in non-landslide point samples often recording steeper slopes than the landslide deposit point samples. In tableland provinces, non-landslide areas are often low-slope benches, valley floors, or plateau tops, while landslide deposits blanket the intervening steep slopes. The result is positive correlations between slope angle and curvature and landslide occurrence. Because of this strong correlation between slope angle and landslide occurrence in tableland provinces, landslide occurrence can be well-predicted and model fits are generally good. In contrast, the abundance of steep terrain without landslides seen in the mountainous provinces results in difficulty determining a clear predictor of landslide occurrence and poorer model fits.
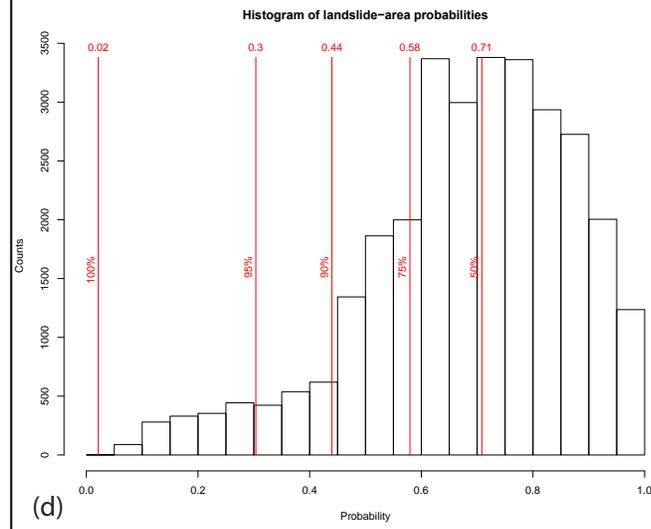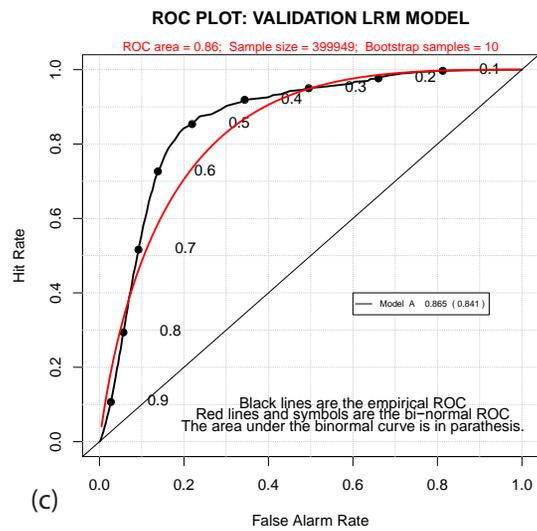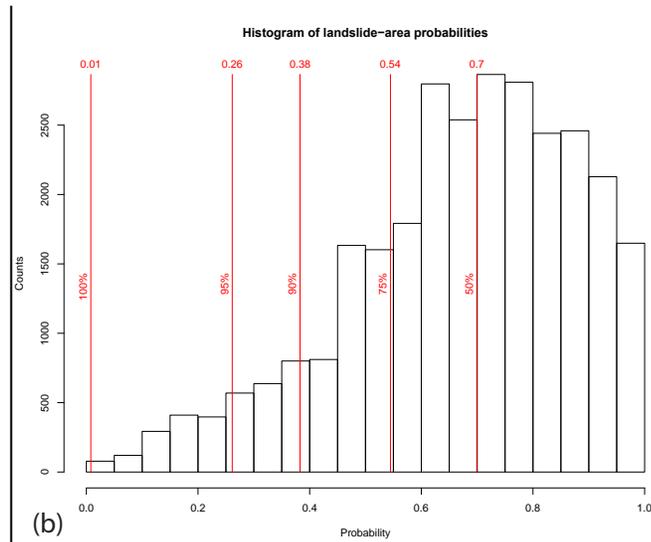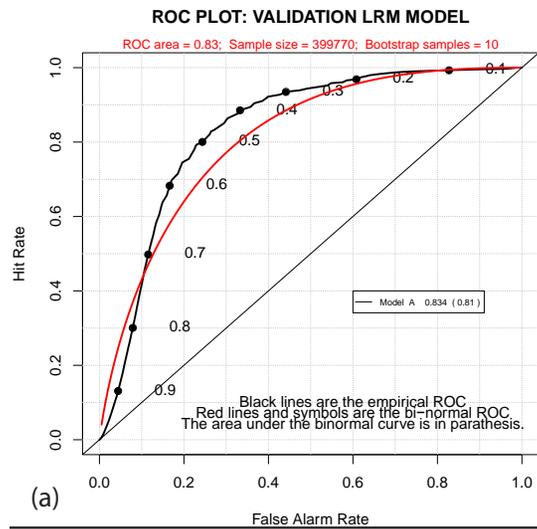
Figure 5.2: Model evaluation plots for the final Colorado Plateau model. ROC plots generated by the LAND-SE script (Rossi and Reichenbach, 2016). (a) Internal validation ROC plot. (b) Internal validation histogram of known landslide-area model probabilities. (c) External validation ROC plot. (d) External validation histogram of known landslide-area model probabilities.
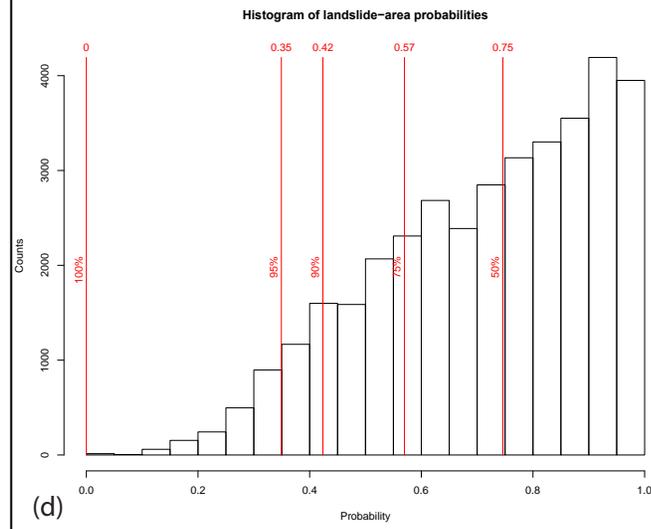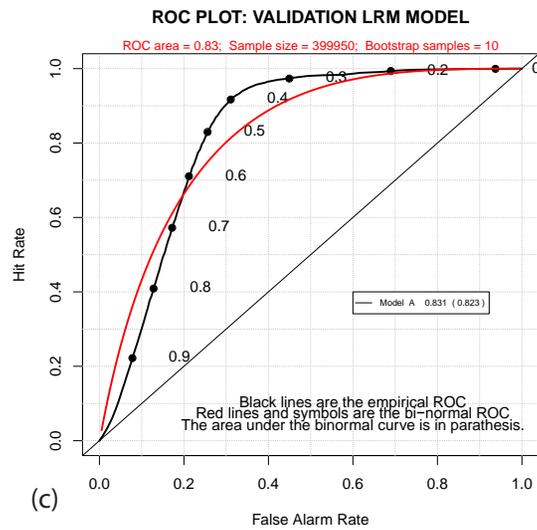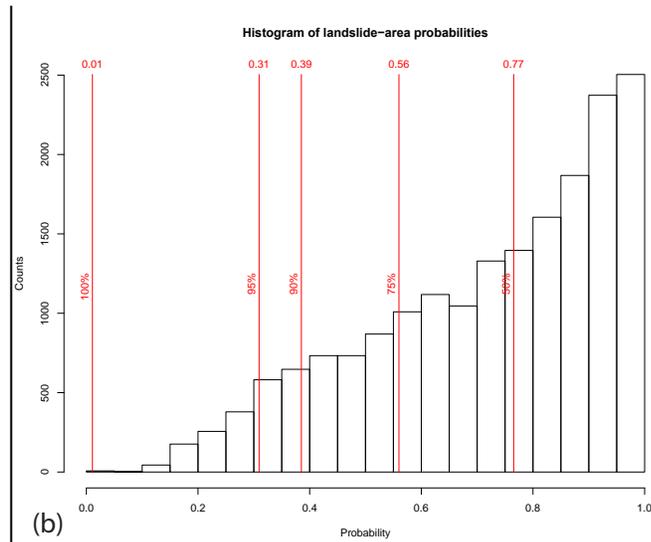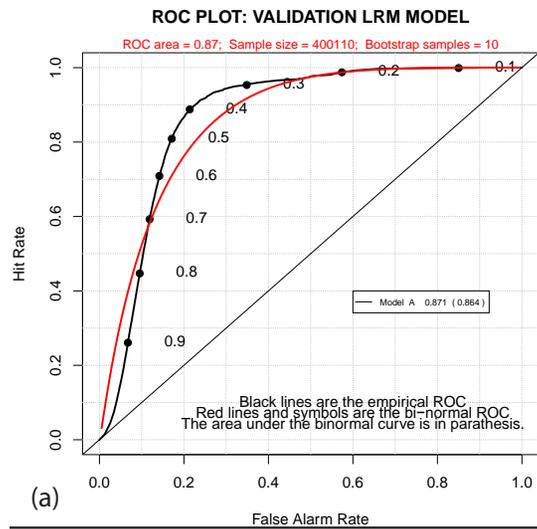
Figure 5.3: Model evaluation plots for the final Great Plains model. ROC plots generated by the LAND-SE script (Rossi and Reichenbach, 2016). (a) Internal validation ROC plot. (b) Internal validation histogram of known landslide-area model probabilities. (c) External validation ROC plot. (d) External validation histogram of known landslide-area model probabilities.
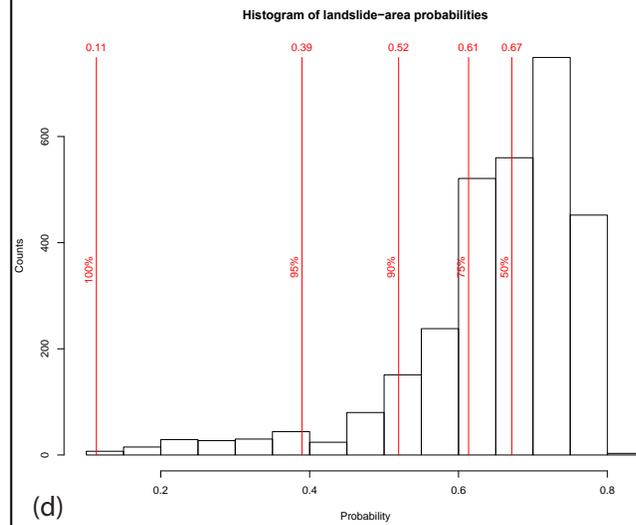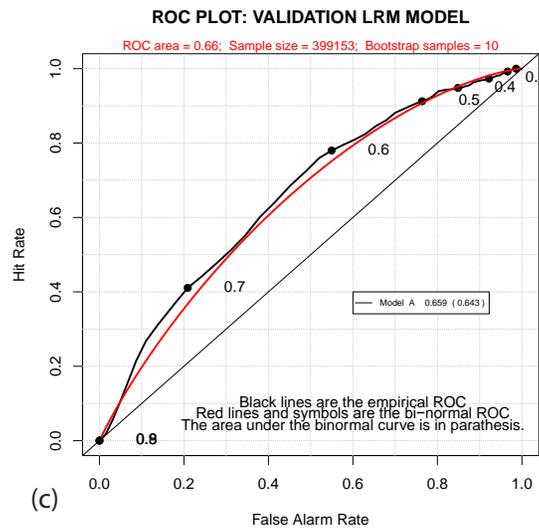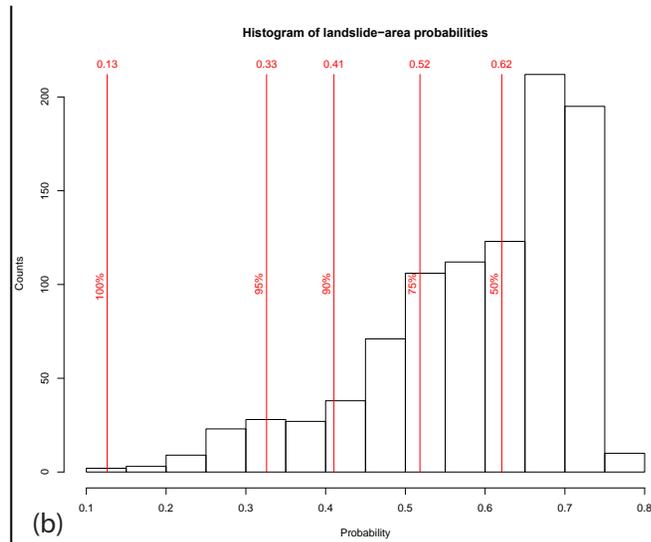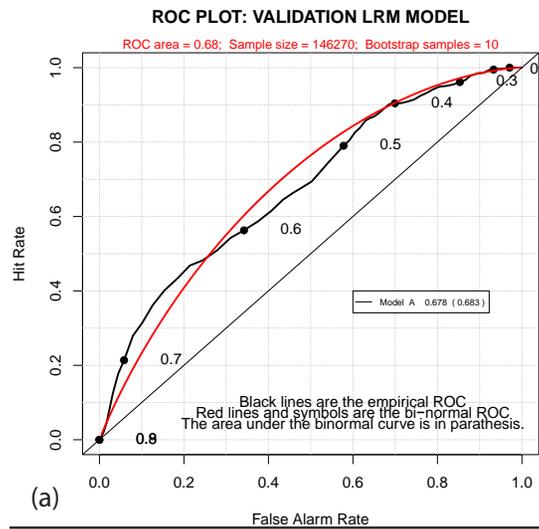
Figure 5.4: Model evaluation plots for the final Mogollon-Datil model. ROC plots generated by the LAND-SE script (Rossi and Reichenbach, 2016). (a) Internal validation ROC plot. (b) Internal validation histogram of known landslide-area model probabilities. (c) External validation ROC plot. (d) External validation histogram of known landslide-area model probabilities.
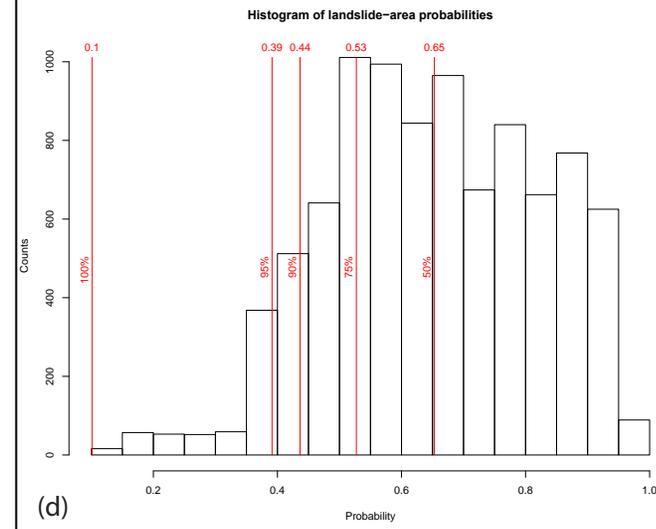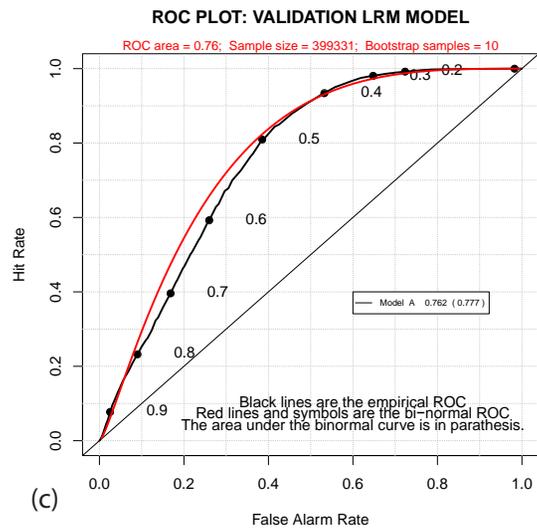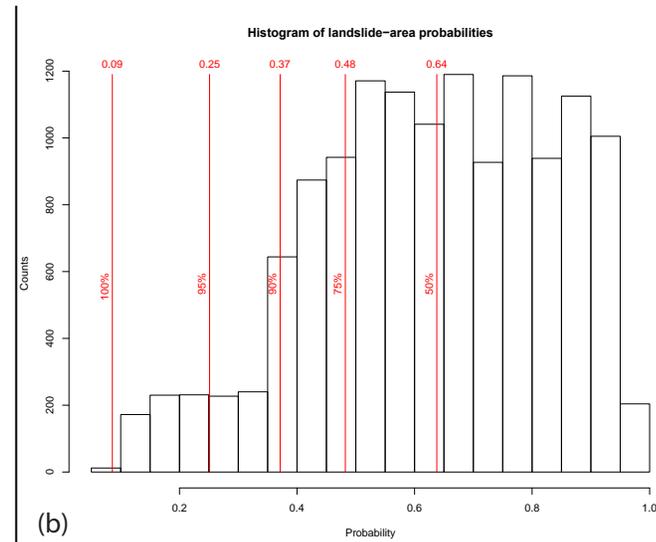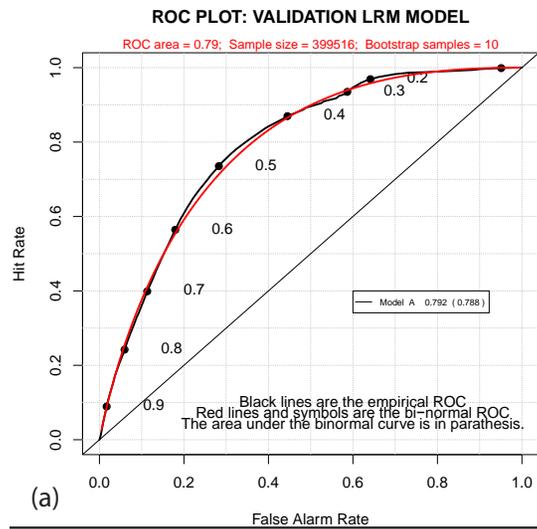
Figure 5.5: Model evaluation plots for the final North Rift model. ROC plots generated by the LAND-SE script (Rossi and Reichenbach, 2016). (a) Internal validation ROC plot. (b) Internal validation histogram of known landslide-area model probabilities. (c) External validation ROC plot. (d) External validation histogram of known landslide-area model probabilities.
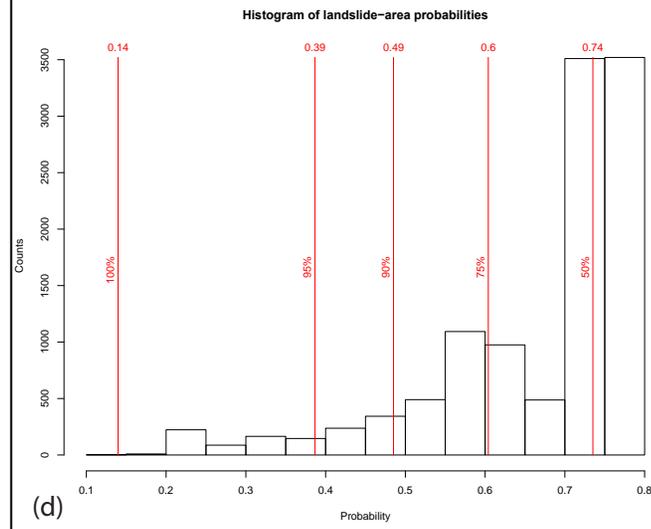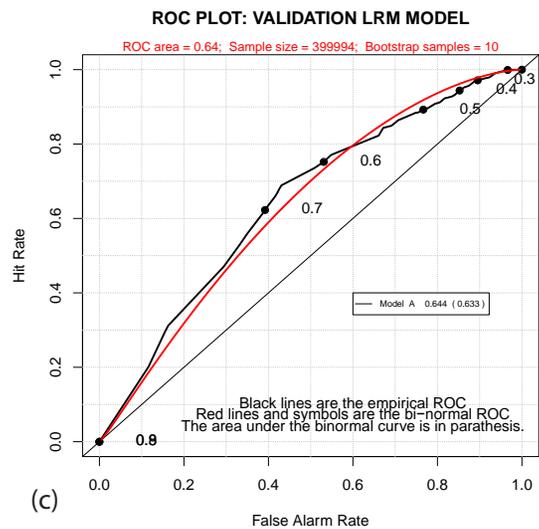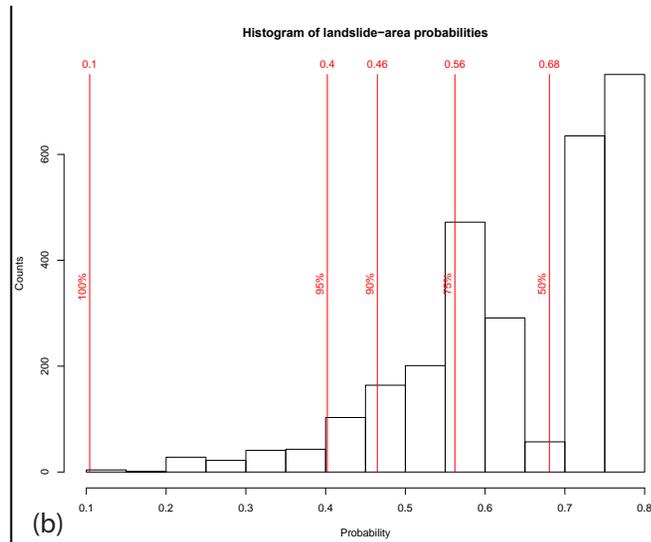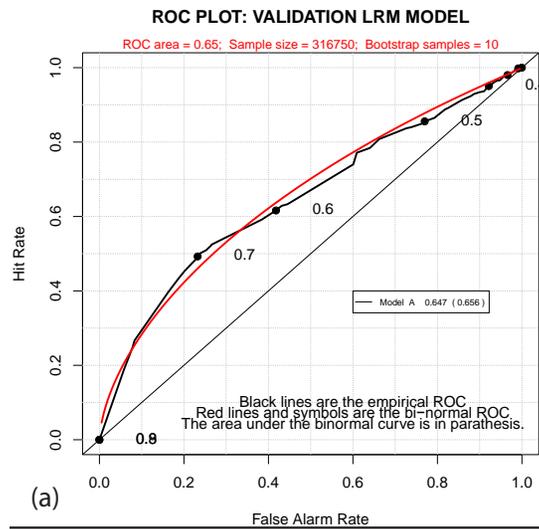
Figure 5.6: Model evaluation plots for the final Southern Rocky Mountains model. ROC plots generated by the LAND-SE script (Rossi and Reichenbach, 2016). (a) Internal validation ROC plot. (b) Internal validation histogram of known landslide-area model probabilities. (c) External validation ROC plot. (d) External validation histogram of known landslide-area model probabilities.

5.2 Synthesis of province results

Individual province-specific logistic regression model coefficients were applied to the remainder of each province using 28 m pixel resolution rasters to produce province-wide model probability maps. We then processed the province-wide maps to produce gradational boundaries, classify into susceptibility classes, and finally downsampled to a target 500 m pixel resolution raster, as described below.

5.2.1 Province boundary mismatches

A benefit of constructing separate models for separate physiographic provinces is better fitting models to each individual province; a consequence is disparity between model results at the boundaries between provinces. In particular, the weightings of different geologic units can change rapidly at the boundary, resulting in sharp discontinuities in model probabilities in a statewide map. Since susceptibility should not change sharply at arbitrary boundaries such as physiographic province boundaries, we chose to apply a smoothing process to provide gradational boundaries between model results at these boundaries.

Our method was to apply the model coefficients for each province to the landscape characteristics of all surrounding provinces up to 10 km from the boundary, then use a distance-weighted, row-standardized sum of overlapping boundary models to create 20 km-wide gradational zones centered on the boundary. The distance-weighting algorithm gave equal weight to all adjoining provinces at the boundary itself, then linearly varied weights such that at 10 km from the boundary into any single province the sum of weighted probabilities equaled the model probability of that single province. Weights were row-standardized, meaning the sum of weights applied at any single point would equal 1. In the instance where 3 or more 10 km boundary zones overlapped, all overlapping model results were weighted and summed for the purpose of creating a gradational boundary. An example of the effect is shown in Figure 5.7.

5.2.2 Developing susceptibility classes

Although model probabilities themselves are useful in estimating landslide susceptibility, a more categorical measure of potential for susceptibility is desired. In many cases, with logistic regression models it is assumed *a priori* that model probabilities above 0.5 imply the model predicts the occurrence of the model dependent variable, while probabilities below 0.5 imply absence; in the context of landslides, this would mean that 0.5 separates susceptible terrain from non-susceptible terrain. We consider this idealistic and not necessarily reflective of the precision of the model, and in fact our analysis of model probabilities determined for known landslide areas as a part of our model validation process determined that as much as 25% of known landslide areas had model probabilities less than 0.5 (e.g., Figure 5.5b). We therefore sought to determine less arbitrary thresholds for defining susceptibility classes.

We chose to place susceptibility class thresholds based on the distribution of model probabilities occurring in known landslide areas. Our goal was to determine thresholds that accurately classified the known landslide areas, with the assumption that in doing so we would accurately classify landslide potential in the remainder of the map. This approach is conservative, in the sense that it maximizes accurately characterizing "true positives" with only implicit regard to preventing "false positives."

We collected the model probabilities of points that fell on known landslide areas from all validation results, both internal and external, for all six final models (Figure 5.8). Using R, we then fitted several curves to the distribution of landslide area model probabilities, fitting the

curves specifically to (1 – model probability) as this produces a right-tailed distribution that can be fitted by a larger range of functions. Of the three fitted curves, that belonging to a Weibull function best fit the actual distribution. This fitted curve has shape and scale parameters of 1.613 and 0.352, respectively, and an estimated mean and standard deviation of 0.315 and 0.200 (in terms of 1 – model probability). As the mean and various standard deviations from the mean provide meaningful characteristics of the distribution itself, we chose to place thresholds at the mean (0.685), at 1 standard deviation from the mean (0.485), and at 2 standard deviations from the mean (0.285; Figure 5.8; Table 5.3). The statewide model probability map was subsequently classified into four susceptibility classes based upon these selected thresholds. An example of the effect of classification is shown in Figure 5.9.

We define our susceptibility classes as follows. For areas with model probabilities between 0 and 0.285, the landscape and geologic setting is generally dissimilar to known landslide-affected areas, and these areas are not likely to encompass locations that are conducive to deep-seated landsliding. However, certain destabilizing activities such as excavations or certain driving events such as earthquakes could still cause these areas to either experience or be impacted by a deep-seated landslide. These areas are classified as 'unlikely susceptible.'

For areas with model probabilities between 0.285 and 0.485, the landscape and geologic setting is weakly comparable to known landslide-affected areas, and at the scale of the final product these areas may encompass some locations that are conducive to deep-seated landsliding. These areas are classified as 'potentially susceptible.'

For areas with model probabilities between 0.485 and 0.685, the landscape and geologic setting is moderately comparable to known landslide-affected areas, and at the scale of the final product these areas are moderately likely to encompass locations that are conducive to deep-seated landsliding. These areas do not necessarily encompass locations that are susceptible, however. These areas are classified as 'moderately likely susceptible.'

For areas with model probabilities between 0.685 and 1, the landscape and geologic setting is comparable to known landslide-affected areas, and at the scale of the final product these areas likely encompass locations that are conducive to deep-seated landsliding. These areas are not necessarily susceptible in their entirety, however. These areas are classified as 'likely susceptible.'

5.2.3   Low relief susceptibility classification

Following Olsen et al. (2015), we suppressed model classifications in low relief areas with no known nearby landslide activity. Susceptibility classification was forced to the lowest susceptibility category if the following conditions were met:
1) The pixel is not within 300 m of a mapped landslide deposit;
2) The elevation range is no more than 5 m in a 300 m by 300 m window, centered on the pixel;
3) The slope angle range is no more than 2° in a 300 m by 300 m window, centered on the pixel; and
4) The slope angle at the pixel is less than 2°.

This low relief filter changed some pixels classified as "potentially susceptible" to "unlikely susceptible."

5.2.4   Downsampling

Province-wide model probabilities were originally calculated at the resolution of the elevation dataset, that is, at a 28 m pixel resolution. Although this resolution provides excellent maps for model evaluation, it does not accurately portray the overall precision of the final map, as several datasets are at coarser scales. This resolution is also impractical for a statewide dataset. Given the level of precision of the geologic map, as well as the intended final product scale, we chose 500 m as the final pixel resolution. We downsampled to 500 m conservatively. First, we determined the maximum susceptibility class occurring in any given 250 m-radius circle, producing another 28 m-resolution raster with expanded regions of higher susceptibility, then we resampled this new raster to 500 m using the nearest neighbor method. The first process was used to ensure that small areas of high susceptibility class, such as narrow canyons, were not lost to the resampling process. We subsequently applied a 3 pixel by 3 pixel majority filter to the resampled raster to remove isolated single pixels of anomalous class. Figure 5.9 shows an example of the effect of downsampling on model results.

5.3   Statewide results

Plate 1 shows the final results of the classified and downsampled model results. Figure 5.10 shows the histogram of susceptibility class pixels occurring in known landslide areas. A consequence of the conservative downsampling procedure was to expand the area of highest susceptibility, resulting in ~85% of known landslide area pixels lying in the "likely susceptible" classification (Table 5.4). This indicates that the downsampling procedure was successful in preserving small areas of high susceptibility classifications, and also indicates that the final susceptibility map accurately categorizes known landslide areas as susceptible to landslide risk.

The final map classifies 26% of the state as "likely susceptible" to deep-seated landslides (Table 5.5). These areas are principally high-relief areas or the flanks of plateaus and stream valleys, as would be expected. Qualitative review of the final map shows the conservative downsampling procedure has expanded the area of this classification onto adjacent low relief areas, but we suggest this is an accurate reflection of the potential for landslide hazards as a deep-seated landslide could travel downslope or propagate upslope into adjacent low-relief terrain. About 13% of the state classified as "moderately likely susceptible," and these appear to be high-relief areas with less susceptible underlying geologic units, or lower relief areas underlain by susceptible units. Another 14% classified as "potentially susceptible," and these appear to reflect even lower relief areas underlain by susceptible geologies, and locally moderately high relief areas underlain by less susceptible geologic units. Qualitative review of the map shows that these later two categories locally extend into low relief zones that topographically do not appear at risk for landslide generation. However, these areas may be underlain by susceptible geologic units that could generate deep-seated landslide failure planes if subject to excavation. The final map classifies the remaining 48% as "unlikely susceptible." These are dominantly valley floors.
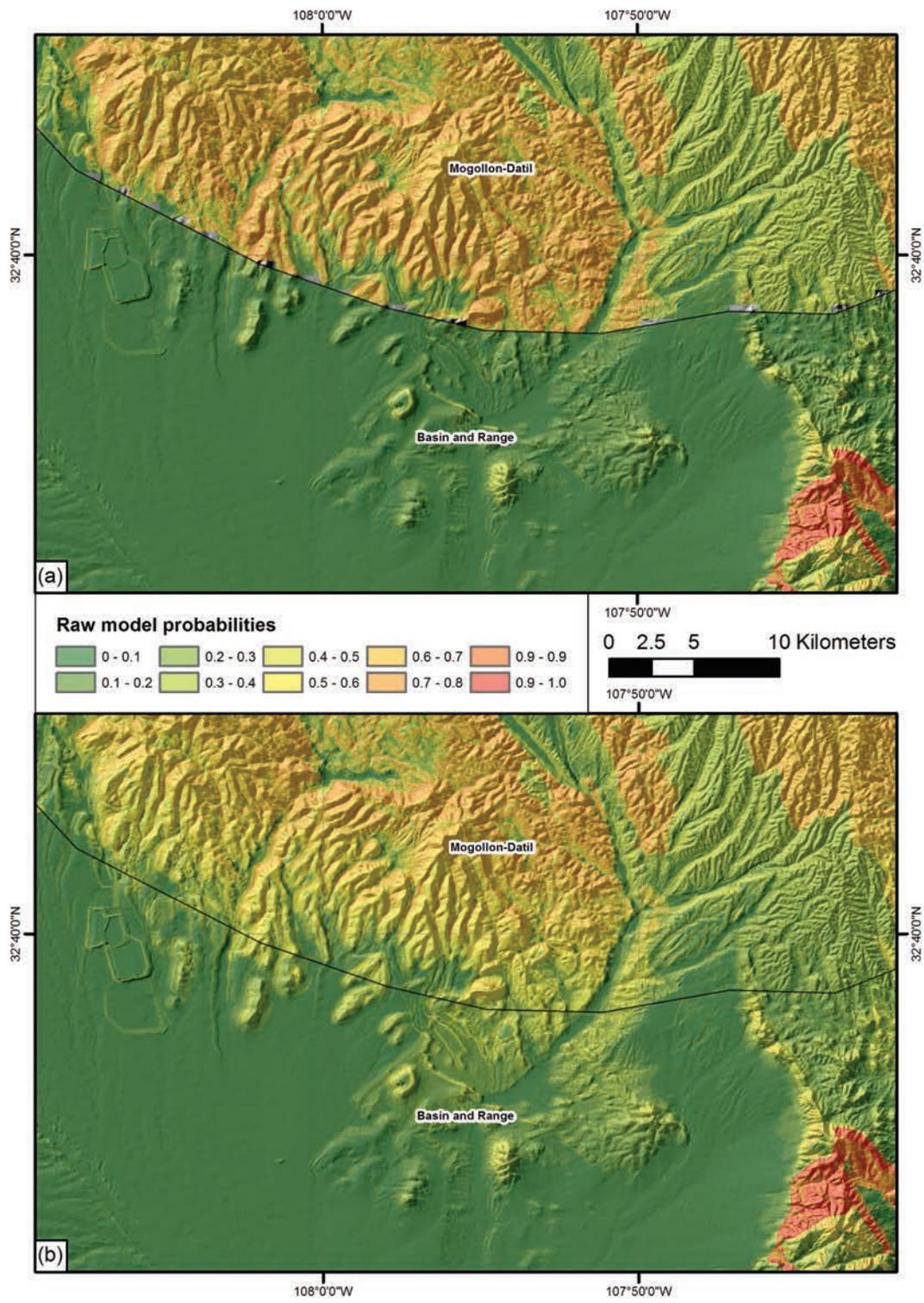
Figure 5.7: Boundary gradation example. (a) Mogollon-Datil/Basin and Range boundary without gradation applied. (b) Same boundary, with gradation applied.

68
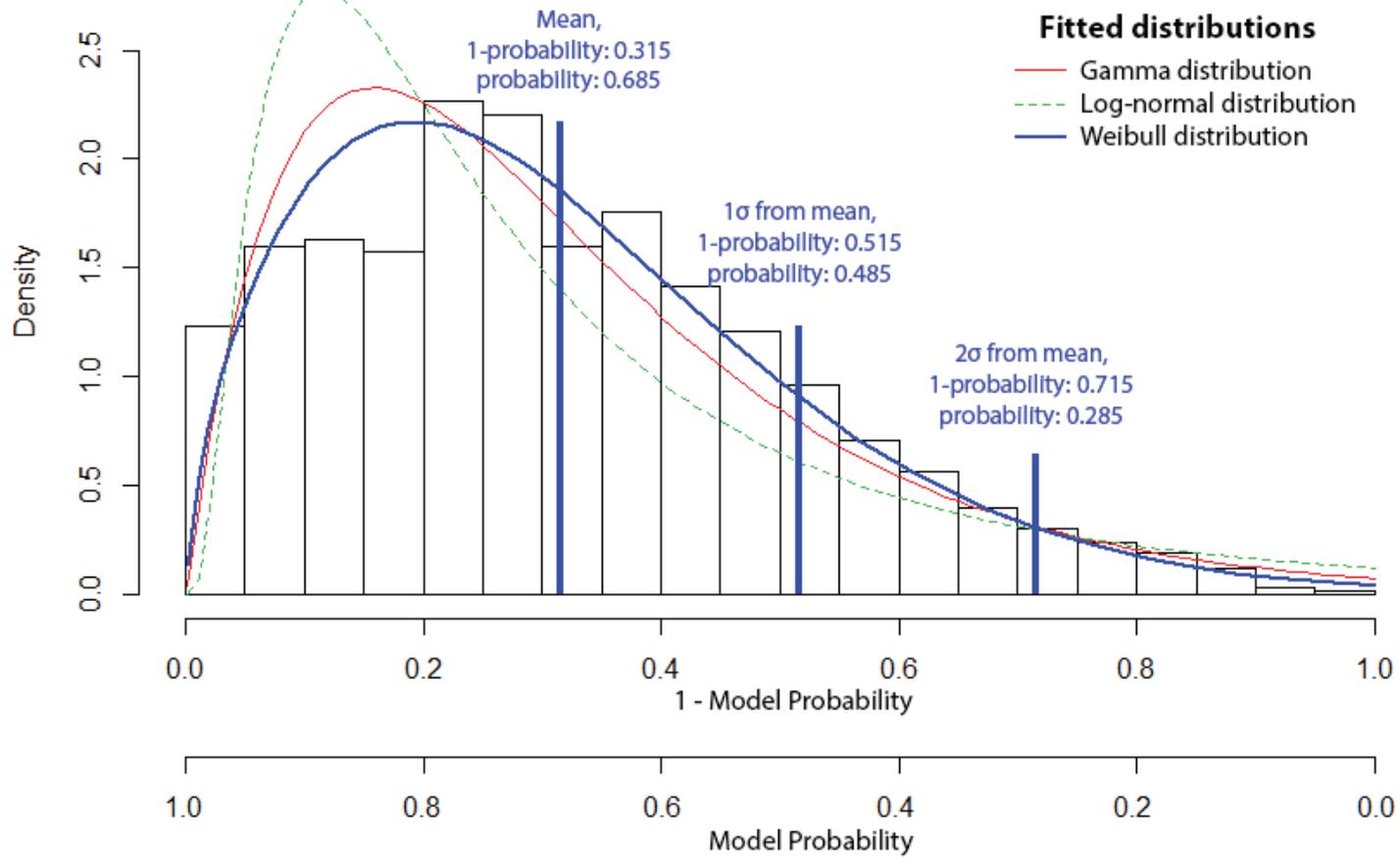
Figure 5.8: Histogram of model probabilities from known landslide areas, as plotted for fitting various distribution curves. Mean and standard deviations from the mean are those for the Weibull distribution.

Table 5.3: Summary of Weibull distribution-based susceptibility classification

Weibull parameters	
[in terms of (1 - probability)]	
Parameter	**Value**
Shape	1.613
Scale	0.352
Est. mean	0.315
Est. std. dev.	0.200

Notes:

Est.: Estimated

Std. dev.: standard deviation

Weibull-based classification			
Thresholds [in terms of probability]		*Classification*	*Classification description*
Max	1.000		
Mean	-0.685-	Likely susceptible	Landscape setting is comparable to known landslide-affected areas; 'likely susceptible' areas likely include locations that are susceptible to deep-seated landsliding.
Mean - 1σ	-0.485-	Moderately likely susceptible	Landscape setting is moderately comparable to known landslide-affected areas; 'moderately likely susceptible' areas are moderately likely to include locations that are susceptible to deep-seated landsliding.
Mean - 2σ	-0.285-	Potentially susceptible	Landscape setting is weakly comparable to known landslide-affected areas; 'potentially susceptible' areas may include locations that are susceptible to deep-seated landsliding.
Min	0.000	Unlikely susceptible	Landscape setting is generally dissimilar to known landslide-affected areas; 'unlikely susceptible' areas are unlikely to include locations that are susceptible to deep-seated landsliding.

Figure 5.9: Susceptibility classification and downsampling example. Example area is the Zuni Mountains-Mount Taylor area of the Colorado Plateau. (a) Map of final merged model probabilities. (b) Map of susceptibility classifications at 28 m pixel resolution. (c) Map of susceptibility classifications following the conservative downsampling to 500 m pixel resolution.
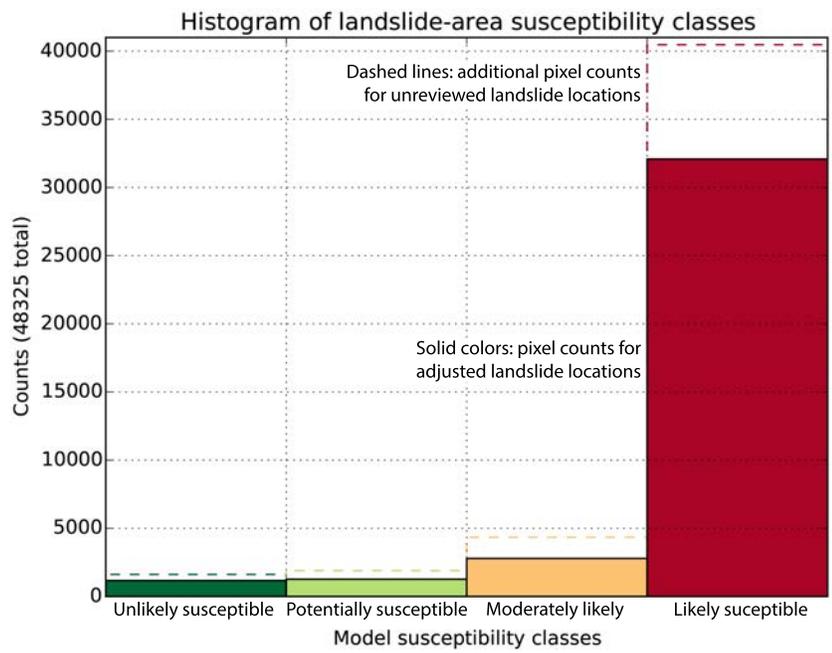
Figure 5.10: Histogram of pixel counts of landslide susceptibility classes from final map in known landslide areas.

Table 5.4: Summary of known landslide area susceptibility classifications in final susceptibility map

Susceptibility class	Landslide areas with reviewed and adjusted locations			All landslide areas, including unreviewed locations		
	Pixel count	Areal coverage (sq. km)	Perc. of total	Pixel count	Areal coverage (sq. km)	Perc. of total
Unlikely susceptible	1167	291.75	3%	1619	404.75	3%
Potentially susceptible	1267	316.75	3%	1896	474.00	4%
Mod. likely susceptible	2791	697.75	7%	4340	1085.00	9%
Likely susceptible	32086	8021.50	86%	40470	10117.50	84%
Totals	37311	9327.75	100%	48325	12081.25	100%

Notes:

  Mod.: Moderately; perc.: percent; sq. km: square kilometers

  Area determined by multiplying pixel count by nominal pixel dimensions

Table 5.5: Summary of final map susceptibility class coverage

Susceptibility class	Pixel count	Areal coverage (sq. km)	Perc. of total
Unlikely susceptible	884498	221124.50	48%
Potentially susceptible	258656	64664.00	14%
Mod. likely susceptible	239954	59988.50	13%
Likely susceptible	476641	119160.25	26%
Totals	1859749	464937.25	100%

Notes:

Mod.: Moderately; perc.: percent; sq. km: square kilometers

Area determined by multiplying pixel count by nominal pixel dimensions

6 Discussion

6.1 Methodology

6.1.1 Predicting landslide susceptibility by assessing landslide deposits

Our method relies on using the landscape characteristics of landslide deposits to create models for predicting landslide susceptibility. This is non-ideal, as the characteristics of a landslide deposit are the characteristics of material moved and deposits by a landslide failure and not necessarily the characteristics of a slope that is about to fail. Thus the efficacy of our method hinges on the similarity between the landslide deposits and the slopes that generated the landslide. This may be a contributing factor to the issues described above pertaining to tableland versus mountainous provinces. In the tableland provinces, particularly the Colorado Plateau, Great Plains, and North Rift areas, the landslide deposits are often found blanketing the very slopes that failed to produce the deposits. Many of the landslides occur along the flanks of mesas and plateaus where strong, broad, tabular cap rocks overlie soft mesa-flank rocks, and the landslide deposits are broad masses of cap rock material sliding down the mesa flank. The slide deposits have thus not traveled far from their origin and the characteristics of the deposit reflect the characteristics of the susceptible slope that underlies them. In contrast, many of the landslide deposits in the more mountainous Southern Rocky Mountains and Mogollon-Datil areas appear to have slid to notably lower slope angle positions, and the deposit characteristics may not be as representative of the characteristics of susceptible slopes. Thus, the efficacy of our method appears to be a function of the geomorphic setting to which it is applied, and is possibly more accurate for tableland-like settings than for high-relief mountainous settings.

In addition, the method requires us to adapt the concepts of "true negatives" and "false positives," as our validation datasets are also landslide deposits, despite our actual interest in landslide susceptibility. The model validation would register a false positive error in any area where the model predicted a landslide but found no landslide deposit, even though the area may still be susceptible to future landslide occurrence, and hence the prediction would actually be a success. Likewise, the model validation would register a true negative success in any area where the model did not predict a landslide and found no landslide deposit, even though the area may actually be susceptible to future landslide occurrence, and hence the model result would actually be a false negative error. This potential issue means that strictly quantitative measures of model performance need to be checked against qualitative assessments. Indeed, the "poor model fits" of the mountainous provinces, as measured by the area under the ROC curves, may actually be a reflection of the validation process incorrectly classifying model results as false positives where high model probabilities are calculated for landslide-susceptible slopes but no landslide deposit is found.

Our use of landslide deposit characteristics for training landslide susceptibility models is a major reason why we did not rely solely on quantitative measures of model performance in selecting model input parameters or for model optimization. In addition, the potential for false positives to actually be accurate model predictions was a major factor in our decision to use histograms of model probabilities in known landslide areas for defining susceptibility classes, rather than using the ROC plot.

6.1.2 Accuracy of model extrapolation

Our method relies on using a model trained on data from a select portion of a province to determine the susceptibility throughout the province. The accuracy of this extrapolatory

75

approach hinges on determining a set of indicator variables that have a consistent relationship with landslide occurrence throughout the province that can be well-characterized in the training area. Our main test for this was the spatial subsampling routine described above. By determining which variables have consistent relationships and are well characterized in subsamples of the training and validation area, we believe we were successful in identifying the most reliable input variables to use for extrapolation across the province.

6.2   Model results

At the regional scale of 1:750,000, there is a relative consistency of moderate susceptibility along steep slopes of mesas in the northwestern, north-central, and northeastern part of the state. In northwestern New Mexico, areas of relatively consistent moderate susceptibility include Mesa Chivato near the Grants area, steep slopes near the I-40 corridor west of Grants, and steep slopes flanking mesas south of Grants. To the north, in the Farmington area and south-southwest of Farmington, moderate susceptibility is mapped for steep slopes developed in weak rocks in the eastern San Juan basin. It is here that a slope failure occurred along approximately 300 m of the Farmers Mutual Ditch in San Juan County (Figure 2.6). Also of potential concern should be steep slopes near the Rio Chama valley, Rio Jemez, and the northern flanks of the Jemez Mountains. Near the Rio Grande valley, there is consistent moderate susceptibility mapped for steep slopes along the gorges of the Rio Grande, including White Rock Canyon and its tributaries, as well as the Rio Grande gorge near Taos. Planners wary of deep-seated slope failure should also take note of the steep slopes in the Española valley, where a small landslide occurred in the 1970s near the town of Rio Chiquito (located 17 km [10-11 mi] east of Española). The Sangre de Cristo Mountains, especially where these mountains are underlain by Paleozoic strata (east and south of Taos), tend to have relatively consistent values of moderate susceptibility. The Sangre de Cristo Mountains west of Raton, which are underlain by low-strength sedimentary rocks of the Raton Basin, also have large areas of moderate susceptibility. In northeastern New Mexico, moderate susceptibility areas are prevalent in steep slopes along the Dry Cimmaron Valley and its tributaries; the northern, western, and southwestern sides of Johnson Mesa; the Canadian Gorge; and along the Canadian escarpment that runs east-northeast from the town of Anton Chico towards Clayton. The east-trending bluffs located ~30 km south of Tucamcari, known as The Caprock and which extend eastwards to the Texas boundary, also have moderate susceptibilities. The flanks of mesas between Santa Rosa and Tucumcari exhibit moderate susceptibilities, including Luciano Mesa and Mesa Rica.

In southern New Mexico, deep-seated landslide susceptibility is overall lower because of large regions with very low slopes. Basin and Range mountains, which tend to have steep slopes, are generally associated with moderate susceptibilities, especially where sedimentary strata are present. For example, the western San Andres Mountains, the mountains northeast of Carrizozo, and the Sacramento Mountains have moderate susceptibilities. The Mogollon-Datil plateau has landslides mostly on its southern flanks. Extending the resulting logistic regression model throughout the province resulted in low to moderate susceptibilities.

6.3   Use of map and associated limitations

The resulting deep-seated landslide susceptibility map (Plate 1), combined with mapped locations of preexisting landslides, should be a useful planning tool for regional endeavors related to land use, public safety, transportation and utility corridors, and construction projects. Examples of study sizes suitable for using this map (at a scale 1:750,000 and with a raster pixel

resolution of 500 m) include the entire state, counties, the larger Indian reservations, or large municipalities (e.g., Albuquerque, Santa Fe, Las Cruces). The final map indicates areas where deep-seated landslide potential is not negligible and where reasonable probabilities may exist for deep-seated landsliding. Given the broad-brush methodology and low-scale of these maps, they are not a substitute for a site-specific geologic or geotechnical study.

Limitations mainly involve issues of scale and time-based risk assessment. The map should not be utilized for projects involving relatively small areas (<10 km^2), except to alert planners where a site-specific study may be warranted. Furthermore, this map does not convey information regarding frequency of occurrence. Interpretations regarding frequency would require detailed mapping of multiple-aged landslides over an appreciable area in addition to age control for landsliding events. This susceptibility map does not contain such information. Nor does this map contain information pertaining to societal costs (including human injury or death) that might arise in a future landslide event. Lastly, it should be reiterated that this map only interprets relative susceptibility for deep-seated landslides, rather than shallow landslides involving only surficial material or debris flows.

6.3.1   Land use

Our landslide susceptibility map could play a role in regional land use studies. For future residential or commercial development, regional zoning maps (e.g., county level) could stipulate that detailed site studies be conducted in areas of likely susceptibility. In National Forests, these maps could guide where to potentially allow such commercial operations as ski resorts or intensive logging.

6.3.2   Public safety

The lack of temporal data precludes reasonable estimations of recurrence intervals of landslide events. This inhibits these maps being used directly for risk assessment. These maps could be used to compare relative landslide hazard in different parts of the state, however. For example, New Mexico has been subdivided into Preparedness Areas for the purposes of emergency planning (NMDHSEM, 2013). The susceptibilities shown on Plate 1 could be used to compare the relative differences in landslide hazards between Preparedness Areas. Consequently, this susceptibility map would have notable value in updating the Hazard Identification/Risk Assessment Section and Vulnerabilities Section in the New Mexico State Hazard Mitigation Plan.

6.3.3   Transportation and utility corridors

Our deep-seated landslide susceptibility map could be an asset in the planning of long-distance transportation or utility corridors across New Mexico. Specifically, this map could be used to identify large regions that are likely susceptible to deep-seated landsliding, which could be avoided in various planning scenarios. If a higher susceptibility area must be crossed by the corridor, then a site-specific study employing an engineering geologist would be warranted.

6.3.4   Construction Projects

Most construction projects involve areas less than <10 km^2, and so this susceptibility map would not be useful except for alerting planners where a site-specific study should be conducted. However, this map would be useful in long-distance construction projects, as explained in the preceding sub-section. In addition to site-specific studies, extra vigilance may be warranted in

areas of moderately likely to likely susceptibility. Features such as ground fissures, leaning trees or fence posts, or bulges of the ground surface could signal landslide activation or reactivation.

7 References

Anbalagan, D., 1992, Landslide hazard evaluation and zonation mapping in mountainous terrain: Engineering Geology, v. 32, p. 269-277.

Atkinson, P.M., and Massari, R., 1998, Generalized linear modeling of susceptibility to landsliding in the central Apennines, Italy: Computer Geoscience, v. 24, p. 373-385.

Bonham-Carter, G., Agterberg, F., and Wright, D., 1988, Integration of geological datasets for gold exploration in Nova Scotia: Photogrammetric Engineering and Remote Sensing, v. 54, p. 1585-1592., et al., 1988

Cardinali, M., Guzzetti, F., and Brabb, E.E., 1990, Preliminary maps showing landslide deposits and related features in New Mexico: U.S. Geological Survey, Open-file Report 90-293, scale 1:500,000.

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., and Reichenbach, P., 1991, GIS techniques and statistical models in evaluating landslide hazard: Earth Surface Processes and Landforms, v. 16, p. 427-445.

Dai, F.C., Lee, C.F., Li, J., and Xu, Z.W., 2001, Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong: Environmental Geology, v. 40, p. 381-391.

Dai, F.C., and Lee, C.F., 2002, Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong: Geomorphology, v. 42, p. 213-228.

Dai, F.C., Lee, C.F., Tham, L.G., Ng, K.C., and Shum, W.L., 2003, Logistic regression modelling of storm-induced shallow landsliding in time and space on natural terrain of Lantau Island, Hong Kong: Bulletin of Engineering Geology and the Environment, v. 63, p. 315-327.

DeGraff, J.V., and Canuti, P., 1998, Using isopleth mapping to evaluate landslide activity in relation to agricultural practices: Bulletin of the International Association of Engineering Geology-Bulletin de l'Association Internationale de Géologie de l'Ingénieur, v. 38, no. 1, p. 61-71.

Donati, L., and Turrini, M.C., 2002, An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology: application to an area of the Apennines (Valnerina; Perugia, Italy): Engineering Geology, v. 63, p. 277-289.

Ercanoglu, M., and Gokceoglu, C., 2002, Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkey) by fuzz approach: Environmental Geology, v. 41, p. 720-730.

Esri, Inc., 2016, ArcGIS 10.4.1 for Desktop (Version 10.4.1.5686) [Software].

Giraud, R.E., and Shaw, L.M., 2007, Landslide susceptibility map of Utah: Utah Geological Survey Map 228DM: scale 1:500,000.

Google, Inc., 2015, Google Earth Pro (Version 7.1.4.1529) [Software]. Available from http://www.google.com/earth/.

Gorsevski, P.V., Gessler, P., and Foltz, R.B., 2000, Spatial prediction of landslide hazard using logistic regression and GIS: 4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4).

Guzzetti, F., Cardinali, M., and Reichenbach, P., 1994, The AVI Project: A bibliographical and archival inventory of landslides and floods in Italy: Environmental Management, v. 18, issue 4, p. 623-633.

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

Hawley, J.W., 2005, Five million years of landscape evolution in New Mexico: an overview based on two centuries of geomorphic conceptual model development, *in* Lucas, S.G., Morgan, G.S., and Zeigler, K.E., *eds.*, New Mexico's Ice Ages: New Mexico Museum of Natural History and Science Bulletin No. 28, p. 9-93.

Hawley, J.W., McCraw, D.J., Love, D.W., and Connell, S.D., 2005, The map of surficial geologic materials of New Mexico: New Mexico Bureau of Geology and Mineral Resources, Open-file Report 462, scale 1:500,000.

Hungr, O., Leroueil, S., and Picarelli, L., 2014, The Varnes classification of landslide types, an update: Landslides, v. 11, p. 167-194. DOI 10.1007/s10346-013-0436-y.

Intermap (Intermap Technologies, Inc.), 2008, Digital Terrain Models, Core Product Version 4.2, Edit Rule Version 2.2: Englewood, CO.

Jibson, W.R., Edwin, L.H., and John, A.M., 2000, A method for producing digital probabilistic seismic landslide hazard maps: Engineering Geology, v. 58, p. 271-289.

Lee, S., 2004, Application of likelihood ratio and logistic regression model for landslide susceptibility mapping using GIS: Environmental Management, v. 34, p. 223-232.

Lee, S., 2005, Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data: International Journal of Remote Sensing, v. 26, p. 1477-1491.

Lee, S., and Min, K., 2001, Statistical analysis of landslide susceptibility at Yongin, Korea: Environmental Geology, v. 40, p. 1095-1113. DOI 10.1007/s002540100310.

Lee, S., and Choi, U., 2003, Development of GIS-based geological hazard information system and its application for landslide analysis in Korea: Journal of Geosciences, v. 7, p. 243-252.

Lee, S., and Choi, U., 2004, Application of a weight-of-evidence model to landslide susceptibility analysis: International Journal of Informational Science, v. 18, p. 789-814.

Lee, S., and Sambath, T., 2006, Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models: Environmental Geology, v. 50, p. 847-855.

Lee, S., Chwae, U., and Min, K., 2002a, Landslide susceptibility mapping by correlation between topography and geologic structure: the Janghung area, Korea: Geomorphology, v. 46, p. 49-162.

Lee, S., Choi, J., and Min, K., 2002b, Landslide susceptibility analysis and verification using the Bayesian probability model: Environmental geology, v. 43, issue 1-2, p. 120-131.

Lee, S., Ryu, J.H., and Min, K.D., and Won, J.S., 2003a, Landslide susceptibility analysis using GIS and artificial neural network: Earth and Surficial Processes and Landforms, v. 27, p. 1361-1376.

Lee, S., Ryu, J.H., Lee, M.J., and Won, J.S., 2003b, Landslide susceptibility analysis using artificial neural network at Boun, Korea: Environmental Geology, v. 33, p. 820-833

Lee, S., Ryu, J.H., Won, J.S., and Park, H.J., 2004, Determination and application of the weights for landslide susceptibility mapping using an artificial neural network: Engineering Geology, v. 71, p. 289-302.

Lee, S., Lee, C., Huang, C., Lee, J., Pan, K., Lin, M., and Dong, J., 2008, Statistical approach to earthquake induced landslide susceptibility: Engineering Geology, v. 100, p. 43-58.

Lei, Z., and Jing-feng, H., 2006, GIS-based logistic regression method for landslide susceptibility mapping in regional scale: Journal of Zhejiang University SCIENCE A, v. 7, p. 2007-2017.

Luzi, L., Pergalani, F., and Terlien, M.T.J., 2000, Slope vulnerability to earthquakes at subregional scale, using probabilistic techniques and geographic information systems: Engineering Geology, v. 58, p. 313-336.

Neuhäuser, B., Damm, B, and Terhorst, B., 2012, GIS-based assessment of landslide-susceptibility on the base of the weights-of-evidence model: Landslides, v. 9, p. 511-528.

NMDHSEM (New Mexico Department of Homeland Security and Emergency Management), 2013, New Mexico State Hazard Mitigation Plan: unpublished report by the New Mexico Department of Homeland Security and Emergency Management, 620 p. << http://www.nmdhsem.org/Mitigation.aspx, last accessed Nov. 3, 2017>>

NMBGMR (New Mexico Bureau of Geology and Mineral Resources), 2003, Geologic Map of New Mexico, scale 1:500,000.

Olsen, M.J., Ashford, S.A., Mahlingam, R., Sharifi-Mood, M., O'Banion, M., and Gillins, D.T., 2015, Impacts of potential seismic landslides on lifeline corridors: Oregon Department of Transportation, Report SPR-740: 176 pp.

Ohlmacher, G.C., and Davis, J.C., 2003, Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kanas, USA: Engineering Geology, v. 69, p. 331-343.

Pachauri, A.K., and Pant, M., 1992, Landslide hazard mapping based on geological attributes: Engineering Geology, v. 32, p. 81-100.

Parise, M., and Jibson, W.R., 2000, A seismic landslide susceptibility rating of geologic units based on analysis of characteristics of landslides triggered by the 17 January, 1994 Northridge, California earthquake: Engineering Geology, v. 58, p. 251-270.

Pazzaglia, F.J., and Hawley, J.W., 2004, Neogene (rift flank) and Quaternary geology and geomorphology: *in* Mack, G.H., and Giles, K.A., *eds*., The Geology of New Mexico, A Geologic History: New Mexico Geological Society, Special Publication 11: p. 407-437.

Pistocchi, A., Luzi, L., and Napolitano, P., 2002, The use of predictive modeling techniques for optimal exploitation of s partial databases: a case study in landslide hazard mapping with expert system-like methods: Environmental geology, v. 41, p. 765-775.

Pradhan, B., Sezer, E.A., Gokceoglu, C., and Buchroithner, M.F., 2010, Landslide susceptibility mapping by neuro-fuzzy approach in a landslide-prone area (Cameron Highlands, Malaysia): IEEE Transactions on Geoscience and Remote Sensing, v. 48, p. 4164-4177.

PRISM Climate Group, 2016, Oregon State University, http://prism.oregonstate.edu, accessed Fall 2016.

R Development Core Team, 2015, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.

Reiche, P., 1937, The Toreva-block, a distinctive landslide type: Journal of Geology, v. 45, issue 5, p. 538-548.

Reneau, S.L., and Dethier, D.P., 1996a, Pliocene and Quaternary history of the Rio Grande, White Rock Canyon and vicinity, New Mexico, *in* Goff, F., Kues, B.S., Rogers, M.A., McFadden, L.D., and Gardner, J.N., eds., The Jemez Mountains Region: New Mexico Geological Society, 47th Annual Field Conference, Guidebook, p. 317-324.

Reneau, S.L., and Dethier, D.P., 1996b, Late Pleistocene landslide-dammed lakes along the Rio Grande, White Rock Canyon, New Mexico: Geological Society of America Bulletin, v. 108, issue 11, p. 1492-1507.

Rossi, M., and Reichenbach, P., 2016, LAND-SE: a software for statistically based landslide susceptibility zonation, version 1.0: Geoscience Model Development 9, p. 3533-3543.

Rowbotham, D.N., and Dudycha, D., 1998, GIS modelling of slope stability in Phewa Tal watershed, Nepal: Geomorphology, v. 26, p. 151-170.

Sarkar, S., Kanungo, D.P., and Mehorotra, G.S., 1995, Landslide hazard zonation: a case study in Garhwal Himalaya, India: Mountain Research and Development, v. 15, p. 301-309.

Terlien, M.T.J., Van Asch, T.W.J., and Van Western, C.J., 1995, Deterministic modelling in GIS-based landslide hazard assessment, *in* Carrara, A., and Guzzetti, F., *eds*., Geographical Information Systems in Assessing Natural Hazards: Kluwer Academic Publishers, The Netherlands, p. 57-77.

U.S. Geological Survey, 2013, National Hydrography Geodatabase: The National Map, viewer available on the World Wide Web (https://viewer.nationalmap.gov/viewer/nhd.html?p=nhd).

U.S. Geological Survey, 2004, Landslide types and processes: Fact Sheet 2004-3072, 4 p.

Vakhshoori, V., and Zare, M., 2016, Landslide susceptibility mapping by comparing weight of evidence, fuzzy logic, and frequency ratio methods: Geomatics, Natural Hazards and Risk, v. 7, issue 5, p. 1731-1752. DOI: 10.1080/19475705.2016.1144655.

Varnes, D.J., 1978, Slope movement types and processes, *in* Schuster, R.L., and Krizek, R.J., *eds.,* Landslides, analysis and control, Special Report 176: Transportation Research Board, National Academy of Sciences, Washington, D.C., p. 11-33.

Watson, R.A., and Wright, H.E., Jr., 1963, Landslides on the east flank of the Chuska Mountains, northwestern New Mexico: American Journal of Science, v. 261, issue 6, p. 525-548

Wu, W., and Sidle, R.C., 1995, A distributed slope stability model for steep forested basins: Water Resources Research, v. 31, p. 2097-2110.

Yilmaz, I., 2009, Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat-Turkey): Computers and Geosciences, v. 35, no. 6, p. 1125-1138.

Zhou, C.H., Lee, C.F., Li, J., Xu, Z.W., 2002, On the spatial relationship between landslides and causative factors on Lantau Island, Hong Kong: Geomorphology, v. 43, p. 197-207.

8 Descriptions of Digital Appendices

A) Input GIS files (Esri geodatabase)

Geodatabase containing adjusted landslide location feature classes, modified geologic map unit polygons (with Qls and water removed), geologic map unit look-up table, province extents, training and validation area extents, major rivers with floodplains feature class, and the extent of "low-relief" ground as defined by the criteria given in the text (after Olsen et al., 2015).

B) Continuous data binning tables (Esri geodatabase)

Geodatabase of reclassification tables for converting continuous raster datasets to the bins used in this project.

C) Individual province model probability results (Esri geodatabase)

Geodatabase of province-specific model probability map rasters, as well as rasters of model probabilities in the gradational zones between provinces.

D) Statewide deep-seated landslide susceptibility (Esri geodatabase)

Geodatabase with statewide deep-seated landslide susceptibility map as a raster.

E) Scripts used for data processing (folder)

Scripts used during the course of this project. Includes a copy of the LAND-SE script of Rossi and Reichenbach (2016) and copies of various project-specific Python- and R-based scripts.

F) Indicator variable training area coverage assessment (folder)

Excel tables of indicator variable coverage in training areas relative to province-wide extents.

G) Indicator variable independence evaluation (folder)

Excel tables and graphs of results from the evaluation of indicator variable independence.

H) Spatial subsampling full results (folder)

Data tables and pdf plots of the results of the final spatial subsampling tests.

I) Individual province models (folder)

Data tables and pdf plots of the final province-specific regression models and their evaluations.