Evaluation of landslide susceptibility mapping techniques using lidar-derived conditioning factors (Oregon case study)

Rubini Mahalingam, Michael J. Olsen and Matt S. O’Banion

School of Civil and Construction Engineering, Oregon State University, Corvallis, OR, USA

ABSTRACT
Landslides are a significant geohazard, which frequently result in significant human, infrastructure, and economic losses. Landslide susceptibility mapping using GIS and remote sensing can help communities prepare for these damaging events. Current mapping efforts utilize a wide variety of techniques and consider multiple factors. Unfortunately, each study is relatively independent of others in the applied technique and factors considered, resulting in inconsistencies. Further, input data quality often varies in terms of source, data collection, and generation, leading to uncertainty. This paper investigates if lidar-derived data-sets (slope, slope roughness, terrain roughness, stream power index, and compound topographic index) can be used for predictive mapping without other landslide conditioning factors. This paper also assesses the differences in landslide susceptibility mapping using several, widely used statistical techniques. Landslide susceptibility maps were produced from the aforementioned lidar-derived data-sets for a small study area in Oregon using six representative statistical techniques. Most notably, results show that only a few factors were necessary to produce satisfactory maps with high predictive capability (area under the curve > 0.7). The sole use of lidar digital elevation models and their derivatives can be used for landslide mapping using most statistical techniques without requiring additional detailed data-sets that are often difficult to obtain or of lower quality.

1. Introduction
Landslides pose a great threat to lives and also lead to socio-economic losses across the globe. Worldwide, significant funds are spent mitigating the effects of landslides (Burns et al. 2013). With an increase in population and urban land use, landslide-triggered losses are expected to increase. In order to minimize economic and human losses, potential landslide areas need to be identified and communicated. Landslide susceptibility mapping (LSM) is considered as a critical component in disaster management and mitigation (Paudel et al. 2003; Pearce 2003; Chen, Liu, et al. 2006; Coppola 2006; Mansourian et al. 2006; Sassa et al. 2006; Burns & Madin 2009; Smith 2013). LSM techniques have evolved greatly over the last few decades and numerous techniques have been proposed and applied by many researchers. Each technique has its own unique advantages and disadvantages.

Landslides are triggered by a variety of factors, which makes them complex to analyse and predict. For example, topography, geomorphology, geology, and hydrological conditions are some of the most important factors to consider for LSM. However, the data-sets representing specific conditioning factors (e.g., slope, geology, aspect, terrain roughness, vegetation index, distance to rivers,
distance to roads, etc.) included in the analysis for LSM are often not consistent between each study. Data source and quality can also vary widely.

With the emergence of light detection and ranging (lidar) technology, high-resolution digital elevation models (DEMs) have become increasingly important in landslide studies (Gutierrez et al. 2001; Hsiao et al. 2004; Chen, Chang, et al. 2006; Schulz 2007; Kasai et al. 2009). Jaboyedoff, Choffet, et al. (2012) have documented the importance of lidar DEMs in landslide research and suggest that the trend of using lidar data for landslide studies will increase substantially over the next 5–10 years given wider availability of the data. Programs such as the 3D Elevation Plan (Sugarbaker et al. 2014) are in place to collect seamless lidar data across the entire United States given its tremendous value. The potential information available from lidar when combined with advanced statistical techniques can strengthen the accuracy of LSM by improving data precision, resolution, accuracy, and completeness, especially in forested areas (Gutierrez et al. 2001; Haneberg et al. 2009; Lato et al. 2009; Ventura et al. 2011).

While these previous studies have shown the effective use of lidar data in landslide delineation and inventory (Guzzetti et al. 2012; Jaboyedoff, Oppikofer, et al. 2012), these studies also indicate that much work is still needed to evaluate the use of lidar in automated, predictive mapping. Given the rapid state of lidar technology evolution, many features of the data have not yet been exploited such as the ability to quantify topographic attributes at a fine scale and link those with hydrological parameters. Further, while many landslide mapping techniques have been applied on individual lidar data-sets with success (Schulz 2007; Burns 2010; Derron & Jaboyedoff 2010; Duplantis 2011), few studies have examined the same study area using multiple techniques to evaluate the consistency of results between techniques. Further, it is important to exploit the complete benefits of lidar data by performing LSM analysis on solely lidar-derived conditioning factors. Independent of our study, Jebur et al. (2014, 2015) recently compared LSM using solely lidar-derived data-sets to those produced considering additional factors from other data sources. In their study, they used three different techniques for LSM. Their inventory contained 31 landslide samples (21 training and 10 testing samples) within a small catchment area (31 km²). However, these landslide samples in the inventory were not mapped or derived from lidar data; rather they were derived from aerial photographs, satellite imagery, and field surveys. Their study concluded that solely using lidar-derived factors ‘may possibly be sufficient’ to produce LSM with acceptable accuracy and other factors did not significantly improve the maps. Jebur et al. (2015) performed similar work using more advanced methods.

Our study builds on this prior work to test the performance of LSM utilizing solely lidar-derived data-sets with a much expanded scope. First, our study considers a significantly larger area (134 km²) with a detailed landslide inventory (581 landslide deposits and 672 scarps) generated directly from lidar data. The improved inventory, which contains polygon delineations of both scarps and deposits, enables the generation of robust training and test samples, which are independent of each other. Next, our study also considers several additional techniques (including advanced machine learning) beyond those considered in the Jebur et al. (2014) study. Finally, we analyse these techniques and the suitability of lidar-derived condition factors in a location with a much higher landslide susceptibility than the area considered in the Jebur et al. (2014, 2015) study.

To this end, the primary objectives of this paper are to:

1. Investigate the applicability and performance of terrain factors from lidar (slope, slope roughness, terrain roughness, stream power index (SPI), and compound topographic index (CTI)) in LSM in a landslide prone location. Are these factors sufficient alone to produce a satisfactory landslide susceptibility map without requiring additional data-sets representing more complex conditioning factors? While Jebur et al. (2014, 2015) provide preliminary results to indicate their suitability, additional studies are needed to confirm these results in a wide variety of terrain types as well as using a variety of techniques.
2. Evaluate the differences in LSM using six widely used and representative statistical techniques in a regional-level analysis. Can more advanced machine learning techniques provide higher predictive capability?
While additional, high-quality data-sets representing other features can surely improve the results of any LSM effort, they were not included in this study so that the effectiveness of lidar data alone could be evaluated. It is hypothesized that the high-resolution topographic information fingerprinted into the lidar data encodes critical geologic and hydrologic properties that are useful for determining susceptibility. For this reason, this study did not directly incorporate the detailed surficial geologic mapping that was available for the study area. Additionally, a proper treatment of geology is not straightforward in many mapping techniques; sometimes it is incorporated as a numerical parameter and in other cases as a category. This, in turn, renders it difficult to compare the techniques themselves when it is included. In an evaluation of various factors for western Oregon, Sharifi-Mood et al. (2013) show that many of the topographic-based conditioning factors varied significantly by lithological unit and recommended that analyses be separated by geologic or lithological units.

1.1. Statistical techniques

Statistical techniques employ mathematical functions under rigorous conditions to predict areas that would be susceptible based on landslide inventories. By using the spatial information of landslide presence and non-presence (i.e., absence) for calibration, the statistical techniques can be effectively applied for regional LSM (Pourghasemi et al. 2013; Pradhan 2013; Pradhan, Abokharama, et al. 2014).

Both deterministic and probabilistic methods are widely used for producing landslide susceptibility maps (Carrara 1983; Carrara et al. 1995; Terlien et al. 1995; Westen & Terlien 1996; Terlien 1998; Chung & Fabbri 1999; Jibson et al. 2000; Reische & Capolongo 2002; Lee et al. 2004; Choi et al. 2010; Frattini et al. 2010; Rampini et al. 2013). Probabilistic methods have gained popularity in the past three decades amongst landslide researchers (Xie et al. 2004; Guzzetti et al. 2005; Lee & Pradhan 2006; Lee & Pradhan 2007; Neuhäuser & Terhorst 2007; Manzo et al. 2013; Pradhan 2013).

Landslide researchers have also used quantitative methods in bivariate, multivariate, and nonlinear approaches. Most published literature regarding LSM using bivariate methods use weights of evidence (WoE) and frequency ratio (FR) (Constantin et al. 2011; Yalcin et al. 2011; Conforti et al. 2012; Tehrany et al. 2013; Althuwaynee et al. 2014a). Logistic regression (LR) and discriminant analysis (DA) are widely used when multivariate methods are preferred for LSM (Pradhan 2010a; Akgun 2012; Chen et al. 2013; Althuwaynee et al. 2014a; Umar et al. 2014). Artificial neural network (ANN) and support vector machine (SVM) techniques have gained popularity amongst researchers who use nonlinear methods for LSM (Choi et al. 2010; Pradhan & Lee 2010; Yilmaz 2010; Marjanović et al. 2011; Zare et al. 2012; Li & Kong 2013; Yao et al. 2013; San 2014).

Despite the wide array of techniques available, it should be noted that few studies utilize more than one technique for a given study area; hence, it is difficult to compare the effectiveness of each individual technique. More detailed information on each of the specific statistical techniques selected for this study is presented in the methodology section.

1.2. Conditioning factors

Based on prior literature, landslide conditioning factors are often grouped into two categories, namely intrinsic and extrinsic factors. Intrinsic factors are slope, aspect, hydrological characteristics, and geologic conditions, while extrinsic factors include precipitation, earthquakes, and human activities (Guzzetti 2000; Lee et al. 2004; Guzzetti et al. 2005; Raia et al. 2011; Santangelo et al. 2012; Ardizzone et al. 2013). Numerous studies have recorded the importance of slope, elevation, aspect, flow rates, flow direction, normalized difference vegetative index (NDVI), SPI, topographic wetness index, distance from streams, and distance from roads (Sidle 1984; Van Westen et al. 1997; Hewitt 1998; Noguchi et al. 1999; Althuwaynee et al. 2014b; Nampak et al. 2014; Pradhan, Hagemann, et al. 2014).
Extensive work is involved in the preparation of data layers to represent these conditioning factors before using them for landslide mapping. Collection, pre-processing, post-processing, and field validations are common protocols adopted in generation of the data-sets used in landslide mapping. Field collection of data is also dependent on the accessibility to the area and prevailing weather conditions. Often, for many types of data, field observations are transferred to a digital format through digitization or manual data entry. The manual nature of this stage can lead to errors, which can adversely affect the quality of the resulting maps (Ardizzone et al. 1999; Ardizzone et al. 2002; Booth et al. 2009).

Because of the difficulty in acquiring reliable data for many factors, information that can be extracted from remote sensing data is of high value. While such information may not be able to yield as accurate individual point measurements, remote sensing data provides broader spatial coverage. In particular, the high-resolution capabilities provided by lidar help infer information about these broad ranges of conditioning factors within a single data-set. For example, hydrological properties such as CTI and SPI can be derived from a DEM and provide information about the presence and activity levels of water shaping the topographic surface.

2. Study area and inventory

Landslides are one of the most common and most devastating geohazards in Oregon (North 1964; Swanson and Dyrness 1975; Palmer 1977; Bottom et al. 1985; Miles & Swanson 1986; Sessions et al. 1987; Frissell & Nawa 1992; Komar & Shih 1993; Robinson et al. 1999; Hoffmeister et al. 2002; Wang et al. 2002; Burns & Madin 2009). In Oregon, landslides contribute over $10 million of economic losses every year, and Oregon is one of the seven pilot states which derive funding from the U.S. Geological Survey for loss estimation. With the increase of population and urban land use, landslide-triggered losses are expected to increase (Wang et al. 2002). The Department of Geology and Mineral Industries (DOGAMI) has mapped approximately 37,226 landslides within the state of Oregon (Burns, Mickelson, et al. 2012). Recent technological advances such as lidar, enabling more detailed topographic information have aided DOGAMI to map more landslides in the last 5 years than those mapped over the last 60 years (Burns, Duplantis, et al. 2012).

Figure 1 shows the study area, the Gales Creek Quadrangle, which is located in north-west Oregon (123°07’30” W to 123°15’00” W and 45°30’00” N to 45°37’30” N). The 7.5’ USGS quad covers approximately 144 km²; however, it was cropped to a rectangular area of 134 km² since a small
portion of the quad did not contain lidar data. Approximately one-third of the study area contains high topographic relief. The mean annual precipitation (over the last 30 years) ranged between 110 and 170 cm/year (PRISM Climate Group 2011) from east to west across the study area. Based on the Oregon Geologic Data Compilation (OGDC V5, Ma 2009), the generalized surficial geology of the study area consists of 41% general sediment deposits (predominately Missoula Flood deposits), 37% volcanics (predominately Tillamook volcanics), 14% marine sediments, and 8% intrusive rocks (predominately Columbia River Basalt). Landslides occur frequently in most of the generalized geologic units with the exception of the Missoula flood sediment deposits, which exhibits minimal topographic relief and shallow slopes.

An overview of the lidar elevation data used in this study can be found in Watershed Sciences (2009). The DEM for this quadrangle was part of a much larger data collection effort spanning 6420 km². Because statistics were not computed by quadrangle, the results from the overall collection are described herein. The native spatial resolution of the bare earth lidar DEM for the collection was 0.91 m. The point cloud data-set was estimated to have an absolute (open-terrain) vertical accuracy of 4 cm (RMS) based on a closest-point comparison with nearly 17,000 RTK GPS coordinates. The total point density was typically around 7 points per m², with approximately 0.7 ground points per m². Data were acquired from a flight height of approximately 900 m above ground with a scan angle of ±14° from nadir and 50% flight-line side-lap (100% overlap).

The original lidar DEM was resampled to 10 m using a bilinear interpolation technique. The data were resampled so that derivative products from the DEM such as slope are more representative of overall trends and topographic features of interest rather than smaller, localized features. For example, at high resolution small drainage ditches can be falsely identified as landslide prone because of the steep slope, but are generally small enough to not be of concern for landslide studies. Based on the findings of Mahalingam and Olsen (2015), a 10-m spatial resolution was selected for further analysis in this study. While higher resolution provides additional detail, it often results in many false positives for LSM across a larger scale. However, for inventory mapping, the higher resolution data are critical.

DOGAMI has performed detailed landslide mapping for this study area using lidar data (Burns, Mickelson, et al. 2012). The landslide inventory used in this study was obtained from the Statewide Landslide Information Database of Oregon (SLIDO) v3.0 (Burns et al. 2008, 2012). Debris flow deposits were removed from the data-set because of their differing failure mechanism. Small deposits measuring less than 0.0007 km² were filtered from the database since 70% of these small deposits were subsequent failures mapped on the top of larger landslide deposits. After the processing and filtering, the SLIDO database for this area contains 499 complex movement type deposits, 45 earth flow deposits, 3 rotational earth slide deposit, and 34 rotational rock slide deposits. Of these deposits, 532 were classified as deep-seated landslides and 47 were classified as shallow landslides. Table 1 documents the details of the landslide inventory used in the analysis. Landslide scarps and deposits cover approximately 20% of the study area.

| Parameters                        | Full study area | Landslide deposits | Landslide scarps |
|-----------------------------------|-----------------|--------------------|------------------|
| Number of features                | 581             | 672                |
| Total area                        | 134 km²         | 21 km² (16.0%)     | 4.6 km² (3.4%)   |
| Average slope                     | 10°             | 14°                | 23°              |
| Std. dev. slope                   | 9.2°            | 7.1°               | 7.5°             |
| Slope range (Min—Max)             | 0°—61°          | 0.025°—51°         | 1°—59°          |
| Individual feature area Statistics|                 |                    |                  |
| Average area                      | —               | 0.035 km²          | 0.006 km²        |
| Std. dev. area                    | —               | 0.085 km²          | 0.013 km²        |
| Area range (Min-Max)              | —               | 0.00072—1.1 km²    | 4.9E5—0.16 km²   |

Table 1. Details of the landslide inventory and topography for the Gale’s Creek Quadrangle after filtering. The slope values are based on the resampled DEM. Debris flow fans were not included.
3. Methodology

The key stages and steps in the workflow to evaluate the landslide hazard mapping methodology consists of four primary stages: (1) prepare, (2) analyse, (3) produce, and (4) validate. ArcGIS 10.1 software was used for data processing throughout all stages.

3.1. Prepare

3.1.1. Landslide conditioning factors

For this study, we evaluate the ability to derive representative intrinsic and extrinsic factors from the resampled (10 m) lidar bare earth DEM in order to simplify the data needed for landslide assessment. Table 2 describes each of these factors, how they are calculated, their influence on landslide susceptibility, their value ranges in the study area, and references that have utilized this parameter. Figure 2 shows the DEM-derived data layers representing these landslide conditioning factors for the study area. Of all of these factors, slope plays the most significant role in landslide assessment.

3.1.2. Training and test samples

In order to produce landslide susceptibility maps using bivariate, multivariate, and soft computing techniques, landslide training and testing samples are needed (Pradhan 2010b; Yalcin et al. 2011; Zare et al. 2012). Training samples are used to develop the model and testing samples are used for validation.

These samples are chosen based on the completeness of the inventory, size of the study area, and the type of approach used. While there are currently no standard protocols for the selection of training and testing samples, it is important that the training and testing samples be independent of each other (Guzzetti et al. 2012; Yilmaz et al. 2012; Yilmaz et al. 2013). Basheer and Hajmeer (2000) have documented different splitting ratios used for various statistical methods. According to their summary, Swingler (1996) and Gopal et al. (1999) used an 80:20 ratio for training and testing of their models. More recently, Guzzetti et al. (2012) used a 70:30 ratio for delineating landslides from high-resolution images and Bui et al. (2012) used the same ratio for predicting landslides using soft computing techniques. Random sampling techniques are commonly employed in all of the above-mentioned studies. In order to make the models computationally robust, a 70:30 ratio was chosen.

The landslide inventory used in this study was created rigorously from high-quality lidar data and is a comprehensive inventory of where landslides have occurred. Training and testing samples were gathered from landslide deposit polygons, and scarp flank polygons, respectively (figure 3(a) and 3(b), respectively). The use of points sampled from the landslide deposits for training is expected to result in a conservative landslide susceptibility prediction model. In general, the landslide deposit polygons represent material which has reached equilibrium, forming a post-landslide terrain in which the slope angle is relaxed in comparison to the pre-landslide terrain. Unlike the landslide deposits, the scarp flanks represent over-steepened portions of the post-landslide terrain. Given the anticipated conservative nature of the resulting landslide susceptibility prediction models, we expect the scarp flank samples to serve as a suitable candidate for independent validation.

Initially, one sample point per pixel occupied by each landslide deposit and scarp was created. However, the large data-set proved too cumbersome for processing using many of the techniques. Hence, randomly sampled points were taken from the deposits and scarp flanks with an approach to balance the representation of all individual landslide features as well as reflect the internal variability within larger landslide deposits. Therefore, landslide deposit polygons smaller than the mean deposit area (0.044 km²) were represented with two sample points and polygons larger than the mean area were represented with 18 points per polygon. In total, 1720 training points were
Table 2. Parameters derived from lidar bare earth model used in this study and their significance in landslide mapping (modified from Gorsevski et al. (2000)).

| Conditioning factor          | Calculation process                                                                 | Influence                                                                                   | Value range in study area | References                                                                 |
|------------------------------|-------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|----------------------------|---------------------------------------------------------------------------|
| Slope (S)                    | The maximum rate of change between each pixel or cell in the elevation data-set and its neighbours (ESRI 2011). | Governs stresses and strains in the soil, geomorphologic processes (magnitude and direction of sediment transport) and hydrology (subsurface velocity, run-off flow rate and direction) | 0°—80°                    | Noguchi et al. (1999), Khazai and Sitar (2000), Mowen et al. (2003), Jaboyedoff et al. (2012), Sharifi-Mood et al. (2013) |
| Slope roughness              | The standard deviation of the slope values within neighbouring cells                 | Erosion and deposition rate, activity rates, and age of deposits. Changes in gradient result in increased rainfall accumulation and infiltration | 0°—18°                    | Fernández et al. (2003), Baeza et al. (2010)                               |
| Elevation                    | Directly obtained from the bare-earth DEM                                            | Captures climate and vegetation growth characteristics                                        | 14—166 m                  | Carrara et al. (1995), Chen et al. (2006), Chen et al. (2013), Pourghasemi et al. (2013), Pradhan (2013), Yu et al. (2006) |
| Terrain roughness            | The standard deviation of the elevation values within neighbouring cells             | Convergence and divergence of water flow, discontinuities of topography                      | 0—4 m                     | Lee et al. (2008)                                                         |
| Stream power index (SPI)     | Calculated by determining the flow accumulation at each cell: SPI = As × tan S, where As = upslope contributing area of flow accumulation | Potential for erosion from water flow, sediment yield and concentration time. As can be readily calculated from ArcGIS hydrological tools. | −11—14                    | Santacana et al. (2003)                                                   |
| Compound topographic index (CTI) | Calculated by determining the flow accumulation at each cell: CTI = ln (As / tan S) | A steady-state wetness index representing flow accumulations down the slope, runoff velocity, hydrologic conditions. The highest values occur in areas of lower slope (<10°) | 2—10                      | Moore et al. (1994), Gessler et al. (2000)                                |
generated from landslide deposits for the training data-set. The testing samples were similarly generated from the scarp flanks. However, because the range of areas is relatively small compared to the landslide deposits, each scarp flank was represented with three point features, resulting in a total of 837 testing samples for validation.

Figure 2. Thematic maps used in this study (Slope, slope roughness, elevation, terrain roughness, SPI, and CTI). These landslide conditioning factors are derived from lidar 10 m bare earth models.
Once the testing and training protocol was developed for samples representing the presence of landslides, an equal number of testing and training points were randomly generated outside of the landslide scarps and deposits to represent locations with the absence of landslides. These point features were then combined into a sample database. Values for all conditioning factors including slope, elevation, slope roughness, terrain roughness, CTI, and SPI were then normalized between 0 and 1, so the coefficients could be used to evaluate the relative importance of each factor in the derived models. For each sample, values for each conditioning factor were extracted from the raster layers using a GIS extract multi-valued routine.

3.2. Analyse

The following techniques were applied individually to generate a series of susceptibility maps.

3.2.1. Frequency ratio (FR)

FR provides the probabilities of presence and absence of an event for individual conditioning factors by generating weights based on the ratio of area which experienced landslides in the past to the total study area (Lee & Pradhan 2007; Pradhan 2010a; Umar et al. 2014). This technique helps in understanding the spatial relationship between landslides and individual conditioning factors because the values in the classes of each conditioning factor indicate the strength of correlation with the event. In order to compute the FR weights, the area ratio for landslide occurrence points was identified for the class of each conditioning factor considered in the current study. The landslide inventory deposit polygons were rasterized and overlaid with the conditioning factors. The area ratio for the class of each factor to the total area was obtained. The FR weights can be obtained by dividing the landslide occurrence ratio in a particular class to the area ratio in that class. Table 3 presents an example of calculated FR weights for slope. A final susceptibility map can be produced by summing up each factor’s weights using equation (1). FR weights greater than one indicate higher correlation of that class in triggering landslides:

\[
\text{LSM} = FR_{w1} + FR_{w2} + FR_{w3} + \cdots + FR_{wn}. 
\]
3.2.2. Weights of evidence method (WofE)

WofE employs a log-linear form of Bayesian probability to estimate the spatial association of landslide conditioning factors and landslide locations. This is a popular method among bivariate analyses because of its flexibility in using thematic factors; in addition, it is often used in expert-based mapping techniques (Neuhäuser & Terhorst 2007; Vijith et al. 2013; Tehrany et al. 2014). WofE functions on the principle that landslide occurrence can be calculated by understanding the event’s prior probability and the occurrences of the conditioning factors. In this current study, WofE was performed by overlaying a binary map with pixels indicating the presence or absence of landslides with each of the conditioning factor classes. Positive and negative weights evaluating the ratios of the presence and absence of landslides for classes for each factor are computed as follows (Akgun 2012):

\[ W^+ = \ln \frac{P(B/D)}{P(B/D)}, \]  
\[ W^- = \ln \frac{P(B/D)}{P(B/D)}, \]

where \( P \) denotes the probability of occurrence of an event (e.g., landslide), \( B \) and \( \overline{B} \) are the presence and absence of the conditioning factors, and \( D \) and \( \overline{D} \) are the presence and absence of a landslide, respectively. The difference between these two weights is called contrast \( C \), which represents the overall weight, showing the association of landslide presence within each class for each conditioning factor. The landslide susceptibility (LS) value is then calculated for each pixel \( j \) as follows:

\[ \text{LS}_j = \sum \limits_{i}^{n} C_{i,j}, \]

where \( C \) represents the contrast value for the class of each conditioning factor \( i \) for pixel \( j \), and \( n \) defines the number of conditioning factors used.

An example of the weights calculated for the slope conditioning factor is shown in Table 4. The sign of the \( C \) value for each class indicates whether there is positive or negative correlation between
the conditioning factor and landslide presence. The weights from all the conditioning factors were normalized and summed in the raster calculator to generate the final LSM.

### 3.2.3. Logistic regression (LR)

LR is a multivariate regression between landslides (dependent variable) and the conditioning factors (independent variables) within a logit function. In LR, a dependent variable is dichotomous and independent variables can be categorical or continuous (Lee 2004). The regression is based on maximizing likelihood. In this current study, the dependent variable is dichotomous, indicating either the presence or absence of a landslide. However, the independent variables are continuous in nature. LR can be expressed with the logit function as follows:

\[
P = \frac{1}{1 + e^{-z}},
\]

where \( P \) is the probability of landslide occurrence which varies between 0 and 1. \( z \) is defined with the linear equation (6), which is obtained through an LR, using a maximum likelihood estimation:

\[
z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n,
\]

where \( \beta_0 \) represents the coefficient of the model, and \( \beta_1, \beta_2, \ldots, \beta_n \) represent the regression coefficients. \( X_1, X_2, \ldots, X_n \) represent the independent variables used in the study (e.g., slope, elevation, slope roughness, terrain roughness, CTI, and SPI). If the values are normalized, the regression parameter indicates the relative contribution of each independent variable towards the event happening (e.g., landslide occurrence).

The LR was performed using Stat Tools\textsuperscript{®}. The resulting coefficients were then multiplied by conditioning factor values for each pixel using the GIS raster calculator to produce the final susceptibility map as follows:

\[
z = -1.340 + 0.1373\times S + 0.0006\times E + 0.0082\times SR + 0.0021 \\
\times SPI + 0.6155\times CTI + 0.4340\times TR,
\]

### Table 4. Example of weights of evidence values for slope map. Higher \( C \) values represent the slope sections which are highly vulnerable to landslides.

| Class | Number of pixels | \( W^+ \) | \( W^- \) | \( C \) |
|-------|-----------------|----------|----------|-------|
| 0–3   | 908             | -3.66    | 0.19     | -3.86 |
| 3–6   | 5836            | -1.16    | 0.07     | -1.23 |
| 6–9   | 5125            | -0.74    | 0.03     | -0.77 |
| 9–12  | 5470            | -0.39    | 0.01     | -0.41 |
| 12–15 | 6024            | -0.19    | 0.01     | -0.19 |
| 15–18 | 7098            | -0.06    | 0.00     | -0.06 |
| 18–21 | 8682            | 0.09     | 0.00     | 0.09  |
| 21–24 | 10,221          | 0.22     | -0.01    | 0.23  |
| 24–27 | 11,692          | 0.35     | -0.02    | 0.37  |
| 27–30 | 13,071          | 0.49     | -0.03    | 0.51  |
| 30–33 | 13,466          | 0.56     | -0.03    | 0.59  |
| 33–36 | 13,615          | 0.60     | -0.03    | 0.63  |
| 36–39 | 13,543          | 0.65     | -0.03    | 0.68  |
| 39–42 | 13,680          | 0.71     | -0.04    | 0.75  |
| 42–45 | 13,317          | 0.71     | -0.04    | 0.75  |
| 45–48 | 12,667          | 0.67     | -0.03    | 0.71  |
| 48–51 | 11,639          | 0.60     | -0.03    | 0.62  |
| 51–54 | 9943            | 0.46     | -0.02    | 0.48  |
| 54–57 | 8055            | 0.30     | -0.01    | 0.31  |
| 57–60 | 5941            | 0.08     | 0.00     | 0.08  |
where $S$ is slope in degrees, $E$ is elevation in metres, $SR$ is slope roughness in degrees, $SPI$ is the stream power Index, $CTI$ is the compound topographic index, and $TR$ is terrain roughness in metres.

### 3.2.4. Discriminant analysis (DA)

DA is a statistical method which explores the combination of independent variables that contribute to the presence and absence of landslide events (Reger 1979; Carrara 1983; Gorsevski et al. 2000; Dong et al. 2009). Stepwise DA enables the addition or removal of the independent variables based on the outcome of analysis. In the application for landslide mapping, two groups are established such as stable and unstable terrains. It is assumed that two groups are distinct in nature and mapping units can be categorized into one of these groups. In the context of LSM, DA can help determine the group membership of the mapping unit. This is achieved by identifying the linear combination of landslide conditioning factors which maximizes the difference between stable and unstable zones. Each conditioning factor plays a role in segmenting the mapping unit into its respective zones. The relative contribution of these factors in the identification process can be studied from equation (8). Application of this function for the unidentified units allows for identification of areas susceptible to landslides (Van Den Eeckhaut et al. 2009). DA optimizes a linear combination of the independent variables (e.g., conditioning factors) through multilinear regression based on a least-squares error minimization approach. DA follows a similar process to LR; however, in LR, a logit function is used and a maximum-likelihood approach is utilized for the regression. Dichotomous points were selected representing the presence and absence of landslide events and normalized values for each conditioning factor at those locations were extracted. The following shows the linear form of the resulting model (Baeza & Corominas 2001):

$$z = a + C_1X_1 + C_2X_2 + \cdots + C_nX_n;$$

where $z$ represents the DA score, $C_1-C_n$ represent coefficients/weights for each conditioning factor, and $a$ is an intercept coefficient.

The DA regression was performed using Stat Tools®. Several models were generated for various combinations of each conditioning factor. Using testing samples for validation, a success rate curve is plotted to identify each model’s ability to correctly identify the given input samples. The model with the highest success rate was chosen and the coefficients from equation (9) were used to produce the final susceptibility map:

$$z = -0.0701 + 1.504 \times S + 0.859 \times TR + 0.2714 \times CTI + 0.055 \times SR + 0.014 \times SPI + 0.003 \times E,$$

where $S$ is slope, $TR$ is terrain roughness, $CTI$ is compound topographic index, $SR$ is slope roughness, $SPI$ is stream power Index, and $E$ is elevation.

### 3.2.5. Artificial Neural Network (ANN)

ANN is a ‘computational mechanism able to acquire, represent and compute a mapping from one multivariate space of information to another, given a set of data representing that mapping’ (Garrett, 1994). A back propagation algorithm is widely used and has proven successful in LSM. The model generates weight by generalizing and predicting the outcomes from a set of inputs that were not previously seen (Choi et al. 2010). ANN assigns weights to the layers by understanding the errors between the actual and target output values. A multilayered network used in landslide mapping consists of one input layer, hidden layers, and an output layer. Back propagation algorithms train a network until minimum error is reached between the desired and the targeted output values. When the network is trained, a feed-forward algorithm is used to classify the entire data-set. The prior stage is
training and the latter is a classification stage. ANN has numerous interconnected nodes; each node represents a processing element which responds to the weights it receives from other input nodes. As such, each node receives a weighted value (Basheer & Hajmeer 2000; Kiymik et al. 2004; Ermini et al. 2005).

In this study, the ANN technique has three layers, namely the input layer which represents the presence and absence of landslides, hidden layers which represent conditioning factor values, and an output layer. The back propagation algorithm employed in this technique enabled training by determining the weights of each factor. The trained model was then used for classification of the test data, which was omitted during training of the model. Weights are primarily determined by modifying the number of hidden nodes and the learning rate between the input and hidden layers, and between the hidden and output layers. RapidMiner Studio® was used to run neural network modeling by importing the data from GIS for the landslide presence and absence samples. All of the previously discussed conditioning factors were used independently in the model for analysis. When the RMSE goal (0.01) was not met, the model was set to terminate at 6000 epochs. Initially, a learning rate of 0.01 and random weights between 0.1 and 0.9 were used. An iterative process was implemented to refine those values and ensure an appropriate level of goodness of fit of the model. The final weights were then computed from the trained network with the highest accuracy and normalized to generate the final susceptibility map.

3.2.6. Support vector machine (SVM)
SVM is a popular machine learning technique which employs a supervised binary classifier based on statistical learning theory. Marjanović et al. (2011) and San (2014) documented that SVM implicitly maps the original input areas into a high-dimensional feature space, using the training data-set. The optimal hyperplane in the feature space is determined by maximizing the boundaries (Abebe et al. 2010; Pradhan 2013). Support vectors are defined as the training points that are closest to the optimal hyperplane. Classification of the new data is achieved after obtaining the decision surface, which is based on the hyperplane.

In a landslide application, consider a set of training samples with instance label pairs \(x_i, y_i\) with \(x_i \in \mathbb{R}^n\), \(y_i \in \{1, -1\}\), and \(i = 1, ..., m\), where \(x\) is a vector in the input section that encompasses other independent variables. The presence and absence of landslide points are denoted by \(\{1, -1\}\) in the input space. The two core concepts of SVM are

- to distinguish the landslide and non-landslide points by developing an optimum separating hyper plane; and
- the use of kernel functions (table 5) to convert originally nonlinear data to a linearly separable data-set (Yao et al. 2013).

Yao et al. (2008) documented that two-class SVM methods produce higher accuracy landslide susceptibility maps when compared to results using more than two classes. The two-class SVM method used in this study utilizes a hyperplane to segment linearly separable data. Mathematically, the orientation of the hyperplane in the feature space is represented by \(w\), a coefficient vector. The offset of the hyperplane from the origin is denoted by \(b\). Lagrangian multipliers are used to determine an optimal hyperplane using equation (10) and is subject to the constraints mentioned in

**Table 5. Variable functions used in the support vector machine algorithm (modified from Pradhan, 2000).**

| Kernel                     | Equation | Variables |
|----------------------------|----------|-----------|
| Linear                     | \(K(x_i, x_j) = x_i^T \cdot x_j\) | \(y, r, \) and \(d\) denote parameters of kernel functions. |
| Polynomial                 | \(K(x_i, x_j) = (y \cdot x_i^T \cdot x_j + r)^d, y > 0\) | |
| Radial basis function (RBF)| \(K(x_i, x_j) = e^{-\gamma \cdot \|x_i - x_j\|^2}, y > 0\) | |
| Sigmoid                    | \(K(x_i, x_j) = \tanh(y \cdot x_i^T \cdot x_j + r)\) | |
equation (11):

\[
L = \frac{1}{2} \|w\|^2 - \sum_{i=1}^{n} \lambda_i (y_i (w \times x_i + b) - 1)
\]

\[
y_i (w \times x_i + b) \geq 1.
\]

Positive slack of the variables are used to form non-separable cases and is denoted by \(\xi_1\) (Cortes and Vapnik 1995) in the following equation:

\[
y_i (w \times x_i + b) \geq 1 - \xi_1.
\]

Hence, for a non-separable case, equation (10) can be modified as follows:

\[
L = \frac{1}{2} \|w\|^2 - \frac{1}{\nu \xi_1} \sum_{i=1}^{n} \xi_1,
\]

where \(\nu(0, 1)\) accounts for misclassification (Cortes & Vapnik 1995; Schölkopf 2000; LaConte et al. 2005).

Xu et al. (2012) reported that selection of the kernel function is very important when implementing the SVM method. Four commonly used kernel functions are linear, polynomial, radial, and sigmoid (table 5). Detailed explanation of these kernel functions can be found in many published research works (Yılmaz 2010; Marjanović et al. 2011; Ballabio & Sterlacchini 2012; Xu et al. 2012; Li & Kong 2013; Pradhan 2013; Yao et al. 2013; San 2014; Tehrany et al. 2014).

A complexity constant, \(C\), is used to set the tolerance of misclassification and is defined for all types of kernels. Higher \(C\) values allow softer boundaries (over-generalization) and lower values of \(C\) create harder boundaries (over-fitting) (Pourghasemi et al. 2013).

Rapid Miner Studio® was used to perform the SVM operations and a cross-validation function was used to find the optimal kernel parameters. To begin, ranges were defined for each kernel before running the model. Based on the aforementioned research, each kernel has different impacts on the accuracy of the landslide susceptibility map. Hence, the current study applied the four above kernels to evaluate which kernel was most suitable for the lidar-derived conditioning factors. In addition to those four methods, a novel ensemble method was created by combining normalized FR values from the bivariate analysis with a radial basis function (RBF) kernel. The RBF used in this method utilized the following parameters: \(\gamma = 0.3, C = 10, \) convergence epsilon = 0.001, and max iterations = 100,000.

The results of each of the five scenarios were subsequently analysed to determine the optimal kernel resulting in the highest prediction rate. In this study, the combined FR and RBF method was chosen for the development of the landslide susceptibility map.

3.3. Produce

Each technique resulted in a raster data-set stored in a geo-database. The landslide susceptibility maps from the six techniques were converted to a common scale of 0–1 and reclassified to very low, low, medium, high, and very high zones based on a natural break classification method to normalize all maps to a similar scale. Figure 4 shows the landslide susceptibility maps produced from six statistical techniques. The test samples from the scarp flanks were overlaid on the output maps to visually understand the distribution of landslide samples within the hazard zones.
3.4. Validate

In order to evaluate and compare the analysis and mapping techniques, several validation techniques were implemented including (1) receiver operating characteristic (ROC), (2) area under the curve (AUC), and (3) a histogram analysis.

Figure 4. Output landslide susceptibility maps for each technique.
3.4.1. Receiver operating characteristic (ROC)
An ROC describes the true positive and false positive proportions along with the quality of the deterministic and probabilistic methods (Swets, 1988). An ROC curve is shown with sensitivity and specificity values. Sensitivity is the proportion of landslide points correctly classified as susceptible areas. Specificity is the proportion of the absence of landslide points correctly classified as very low susceptibility. An ROC can be constructed using training or test samples. Training samples are used to identify the model’s goodness of fit and test samples are used for determining prediction skill (Van Den Eeckhaut et al. (2009). In this current study, ROC was calculated for validation samples to determine the prediction rate. For a model to classify the presence and absence point samples correctly, the proportion of true positives should be higher than the proportion of false positives.

\[
\text{Specificity} = 1 - \frac{\text{TN}}{\text{TN} + \text{FP}} \\
\text{Sensitivity} = \frac{\text{TN}}{\text{TP} + \text{FN}}
\]

3.4.2. Area under the curve (AUC)
The AUC is a quantitative measure of the predictive performance of the model. AUC ranges between 0 and 1, with higher values indicating better performance. Swets (1988) considers classification models with AUC values greater than 0.90 as highly accurate and those with AUC values less than or equal to 0.5 as poor classifiers. AUC values above 0.7 are generally considered satisfactory in landslide research (Pourghasemi et al. 2012; Pradhan 2013; Pradhan et al. 2014; Althuwaynee et al. 2014a).

3.4.3. Histogram analysis
A histogram analysis was done using the test samples generated from the scarp flanks. The equivalent numbers of absence samples were generated from undisturbed morphological study sections. Values from the output susceptibility maps were extracted on the corresponding spatial location of presence and absence samples. The values were then reclassified using the same bins as used for the output map classification.

4. Results and discussion
The reclassified landslide susceptibility maps from the six statistical classification techniques are presented in figure 4. All maps, with exception to the one created using ANN, appear very similar.

Figure 5 shows the ROC curves and AUC values for each technique. All curves are well above the 1:1 line and AUC values are well above 0.5, indicating that all the six methods could be successfully used with these conditioning factors alone and achieve satisfactory performance. The SVM (AUC = 0.89) and ANN (AUC = 0.72) techniques showed the best and worst performances, respectively. All other techniques shared similar ROC curves and AUC values.

For additional validation, the histogram analysis outputs (figure 6) plot the number of landslide presence (red) and absence (green) samples that fall within each hazard bin (e.g., very high). Ideally, all landslide presence points should plot within the high and very high bins while absence sample points should plot within the low and very low bins. These plots help compare the individual techniques and evaluate their ability to distinguish between landslide prone and stable locales. FR, WofE, LR, DA, and SVM show a clear separation of landslide samples placed in high and very high bins. Except for ANN, all the models classified more than 75% of the test samples as high and very high zones. Further, the ANN output more or less equally distributes landslide presence samples throughout all bins, leading to ambiguous results. A similar trend is noticed for absence samples. Visual analysis of figure 4 shows the ANN map is washed out and does not provide sufficient contrast to distinguish between susceptible and non-susceptible areas. Visual analysis of ANN results also indicate
unstable zones derived from scarp test samples classified as stable zones. Therefore, it is noted that the ANN technique needs more ancillary data-sets for improved training and resulting LSM.

In addition to binning and comparing the test sample points, the original polygons can be used for validation. The total area of the scarp polygons in the study area is 4.6 km² (table 1). FR, WofE, LR, DA, and SVM identified more than 3.0 km² (65%) of the scarps as high hazard zones where as the ANN technique identified only 1.7 km² (37%) of scarps as high hazard.

All methods showed that slope was the dominant conditioning factor in landslide susceptibility for the study area. For example, results of using the FR technique indicate a contribution of each slope class in triggering landslides. It shows that landslide occurrences are typically between 21° and 45°. Although less than 4% of the total area has a slope greater than 45°, there is a positive correlation for failure of those steep slopes. C values (table 4) from WofE show similar results with a strong positive correlation between landslides and slope classes ranging from 20° to 51° (table 4). In the LR and DA analyses, the coefficients for slope were significantly larger in the regression analysis using the normalized conditioning factors.

After slope, the individual FR values of each factor indicated that elevation and SPI had the next most significant contribution. Other factors showed correlation with landslide occurrence; however, they were not as strong as those observed with slope. From the FR analysis, elevation values from 350 to 450 m show higher contribution to landslide occurrence. However, values above 450 m begin to show a negative correlation with landslides. SPI demonstrates a positive correlation from 0 to 12, which starts to decrease gradually in value after that range. Slope roughness demonstrates a positive correlation from 0° to 8° but shows a negative trend for greater than 8°. The entire range of CTI and terrain roughness values for the study area results in a negative correlation with landslide occurrence.

Results were similar with other techniques; however, the relative contribution of factors other than slope varied in order of importance. For example, the positively signed coefficients for LR indicated that slope, slope roughness, SPI, and elevation are correlated with landslide occurrence. For DA, ANN, and SVM, weights indicated slope and slope roughness as the strongest indicators when defining landslide prone areas. LR and DA also indicated that all six techniques were found to be statistically significant using both techniques.

Figure 5. ROCs and AUC values for the landslide susceptibility maps produced by each technique.
Each technique has its own advantages in terms of its ability to incorporate input parameters and information provided in interim stages or the final output that can be used for analysis. Bivariate (FR and WofE) analyses help provide a clear understanding of the details of classes in each of the conditioning factors, which were well reflected in the output maps. The multivariate (LR and DA) output maps had the ability to highlight susceptible regions near the streams even in regionally flat regions. The two soft computing techniques, ANN and SVM, had polarized results. On the one hand, the ANN technique produced a smooth output with limited details to delineate the hazard zones. On the other hand, in the SVM map, the details of other factors such as SPI and CTI can be seen in the output map, indicating that it captures these morphological features well in the output maps. When compared to the other models, the slope pattern and elevation dominate the susceptibility maps and show minimal contribution from the other factors. The SVM strength is attributed to the combination of using a RBF with the FR output to generate the output map.

5. Conclusion

The primary aim of this paper is to analyse powerful statistical techniques using only lidar-derived topographic variables for predictive mapping of areas susceptible to landslides. Most notably, this research shows the advantage of selecting fewer critical conditioning factors for LSM when using a lidar DEM similar to the results of Jebur et al. (2014, 2015). A single-lidar DEM and its derivative products (e.g., slope) proved sufficient to be used effectively and reliably in LSM using bivariate, multivariate, or soft computing techniques. Additional detailed data-sets (e.g., hydrology) that may be difficult to obtain or poor in quality were not required. The ability of a lidar DEM to preserve morphological features enables other conditioning factors to be accounted for in the analysis indirectly, as evidenced by the satisfactory AUC values and ROC curves.
Few studies are available in geomorphological literature to evaluate the reliability of various statistical analysis techniques. In this study, with the exception of ANN, all techniques appeared to produce statistically similar and reliable results, despite some minor differences. Based on a north-west Oregon study area, the SVM technique performed relatively well for LSM compared to FR, WofE, LR, DA, and ANN. The overall performance of ANN in LSM was low. The poor performance of the ANN method is attributed to the need for ancillary data to train the data-sets and generate weights.

It should be noted that all statistical techniques are sensitive to quality of the landslide inventory data used for calibration. The current study used a lidar-derived inventory, which may not always be available. Hence, in situations where such an inventory and high-resolution lidar are not available, data layers representing other conditioning factors may be necessary. In particular, geologic mapping can help distinguish areas susceptible to landslides. Geologic information pertaining to material type, age, orientation, and tectonic setting of in situ soil and rock can provide insight into the potential for slope instability.

The information derived from this study can assist researchers, planners, scientists, engineers, and government organizations to use a minimum selection of conditioning factors for LSM with the applied statistical techniques. It is envisioned that these techniques can be evolved in the future for disaster response to rapidly map susceptibility for specific events such as earthquakes, storms, or heavy rainfall provided that the lidar data and landslide inventory are available. While the current study indicates that for the study area, the use of a few conditioning factors from lidar can enable fast and reliable generation of susceptibility maps, it is recommended to use sufficient validation techniques in other locations before relying solely on this data-set for LSM for disaster mitigation purposes. It is also recommended to apply several techniques when mapping an area to compare and contrast the overall results since each technique has its own strengths and weaknesses. Combination of the results from each technique could prove to provide a more realistic assessment and minimize the artefacts of each technique.

Acknowledgments
This paper is based on work funded by the Oregon Department of Transportation (SPR-740). The lidar data used in this study was provided by the Oregon Lidar Consortium. The Oregon Department of Geology and Mineral Industries provided the SLIDO database. We thank them for their feedback and assistance with this project. We also thank the anonymous reviewers for their helpful feedback.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This paper is based on work funded by the Oregon Department of Transportation (SPR-740).

Notes on contributors

Rubini has a background in natural hazards and GIS. She is currently working on her PhD in geomatics engineering at Oregon State University. Her research interests include geohazards, GIS and remote sensing.

Michael Olsen is an associate professor of geomatics in the School of Civil and Construction Engineering at Oregon State University. He is currently serving as the editor-in-chief for the ASCE Journal of Surveying Engineering. He has BS and MS degrees in civil engineering from the University of Utah and a PhD from the University of California, San Diego. He has also worked as an engineer in training for West Valley City. His current areas of research include terrestrial laser scanning, remote sensing, GIS, geohazard engineering, hazard mapping, and 3D visualization.
Matt O’Banion has a background in geoscience and geotechnical engineering. He is currently working on his PhD in geomatics engineering at Oregon State University. Current Research interests include geohazards, lidar analysis, and immersive visualization.

References

Abebe B, Dramis F, Fubelli G, Umer M, Asrat A. 2010. Landslides in the Ethiopian highlands and the Rift margins. J Afr Earth Sci. 56:131–138.

Akgun A. 2012. A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at Izmir, Turkey. Landslides. 9:93–106.

Althuwaynee OF, Pradhan B, Park H-J, Lee JH. 2014a. A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping. Catena. 114:21–36.

Althuwaynee OF, Pradhan B, Park H-J, Lee JH. 2014b. A novel ensemble decision tree-based CHi-squared Automatic Interaction Detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping. Landslides. 11:1063–1078.

Ardizzone F, Cardinali M, Carrara A, Guzzetti F, Reichenbach P. 1999. Impact of mapping errors on the reliability of landslide hazard maps. Nat Hazards Earth Syst Sci. 2:3–14.

Ardizzone F, Cardinali M, Carrara A, Guzzetti F, Reichenbach P. 2002. Uncertainty and errors in landslide mapping and landslide hazard assessment. Nat Hazards Earth Syst Sci. 2:3–14.

Ardizzone F, Fiorucci F, Santangelo M, Cardinalli M, Mondini AC, Rossi M, Reichenbach P, Guzzetti F. 2013. Very-high resolution stereoscopic satellite images for landslide mapping. Landslide science and practice. Berlin: Springer; p. 95–101.

Baeza C, Corominas J. 2001. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. Earth Surf Process Landforms. 26:1251–1263.

Baeza C, Lantada N, Moya J. 2010. Influence of sample and terrain unit on landslide susceptibility assessment at La Pobla de Lillet, Eastern Pyrenees, Spain. Environ Earth Sci. 60:155–167.

Ballabio C, Sterlacchini S. 2012. Support vector machines for landslide susceptibility mapping: the Staffora River Basin case study, Italy. Math Geosci. 44:47–70.

Basheer IA, Hajmeer M. 2000. Artificial neural networks: fundamentals, computing, design, and application. J Microbiol Methods. 43:3–31.

Booth AM, Roering JJ, Perron JT. 2009. Automated landslide mapping using spectral analysis and high-resolution topographic data: Puget Sound lowlands, Washington, and Portland Hills, Oregon. Geomorphology. 109:132–147.

Bottom, DL, Howell, PJ, Rogers, JD. 1985. The effects of stream alterations on salmon and trout habitat in Oregon. Portland: Oregon Department of Fish and Wildlife.

Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB. 2012. Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): a comparative assessment of the efficacy of evidential belief functions and fuzzy logic models. Catena. 96:28–40.

Burns WJ. 2010. Using lidar to locate and attribute landslides in Oregon. Paper presented at: 2010 GSA Denver Annual Meeting.

Burns WJ, Duplantis S, Mickelson KA, Spritzer JM, Wells RE. 2012. Landslide inventory maps of the Gales Creek quadrangle, Washington County. 3rd ed. Portland: Oregon Department of Geology and Mineral Industries.

Burns WJ, Madin, IP. 2009. Protocol for inventory mapping of landslide deposits from light detection and ranging (lidar) imagery. (DOGAMI Special Paper 42, 36 p).

Burns WJ, Madin IP, Ma L. 2008. Statewide Landslide Information Database for Oregon (SLIDO), Release 1. Paper presented at: 2008 Joint Meeting of the Geological Society of America, Soil Science Society of America, American Society of Agronomy, Crop Science Society of America, Gulf Coast Association of Geological Societies with the Gulf Coast Section of SEPM.

Burns WJ, Mickelson KA, Saint-Pierre EC. 2012. Statewide Landslide Information Database for Oregon Release 2 (SLIDO-3), scale 1: 750,000. Department of Geology and Mineral Industries.

Burns WM, Mickelson KA, Saint-Pierre CE. 2013. Landslide inventory maps in.

Carrara A. 1983. Multivariate models for landslide hazard evaluation. J Int Assoc Math Geol. 15:403–426.

Carrara A, Cardinali M, Guzzetti F, Reichenbach P. 1995. GIS technology in mapping landslide hazard. Geographical information systems in assessing natural hazards. Springer Netherlands; p. 135–175. doi:10.1007/978-94-015-8404-3_8

Chen LC, Liu YC, Chan KC. 2006. Integrated community-based disaster management program in Taiwan: a case study of Shang-An village. Nat Hazards. 37:209–223.

Chen RF, Chang KJ, Angelier J, Chan YC, Deffontaines B, Lee CT, Lin ML. 2006. Topographical changes revealed by high-resolution airborne LiDAR data: The 1999 Tsaoling landslide induced by the Chi—Chi earthquake. Eng Geol. 88:160–172.
Chen SC, Chang CC, Chan HC, Huang LM, Lin LL. 2013. Modeling typhoon event-induced landslides using GIS-based logistic regression: a case study of Alishan Forestry Railway, Taiwan. Math Probl Eng. Available from: http://dx.doi.org/10.1155/2013/728304

Choi J, Oh HJ, Won JS, Lee S. 2010. Validation of an artificial neural network model for landslide susceptibility mapping. Environ Earth Sci. 60:473–483.

Chung C-JF, Fabbri AG. 1999. Probabilistic prediction models for landslide hazard mapping. Photogrammet Eng Remote Sens. 65:1389–1399.

Conforti M, Robustelli G, Muto F, Critelli S. 2012. Application and validation of bivariate GIS-based landslide susceptibility assessment for the Vitravo river catchment (Calabria, south Italy). Nat Hazard. 61:127–141.

Constantin M, Bednarik M, Jurchescu MC, Vlaicu M. 2011. Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). Environ Earth Sci. 63:397–406.

Coppola, D.P. 2006. Introduction to international disaster management. Oxford: Butterworth-Heinemann.

Cortes C, Vapnik V. 1995. Support-vector networks. Mach Learn. 20:273–297.

Derron MH, Jaboyedoff M. 2010. Preface: LIDAR and DEM techniques for landslides monitoring and characterization. Nat Hazards Earth Syst. 10:1877–1879.

Dong JJ, Tung YH, Chen CC, Liao JJ, Pan YW. 2009. Discriminant analysis of the geomorphic characteristics and stability of landslide dams. Geomorphology. 110:162–171.

Duplantis S. 2011. Landslide inventory mapping and dating using LiDAR-based imagery and statistical comparison techniques in Milo McIver State Park, Clackamas County, Oregon [dissertation]. Portland (OR): Portland State University.

Ermini L, Catani F, Casagli N. 2005. Artificial neural networks applied to landslide susceptibility assessment. Geomorphology. 66:327–343.

Fernández T, Irigaray C, El Hamdouni R, Chacón J. 2003. Methodology for landslide susceptibility mapping by means of a GIS. Application to the Contraviesa area (Granada, Spain). Nat Hazards. 30:297–308.

Frattini P, Crosta G, Carrara A. 2010. Techniques for evaluating the performance of landslide susceptibility models. Eng Geol. 111:62–72.

Frissell CA, Nawa RK. 1992. Incidence and causes of physical failure of artificial habitat structures in streams of western Oregon and Washington. N Am J Fish Manag. 12:182–197.

Garrett JH, Jr. 1994. Where and why artificial neural networks are applicable in Civil Engineering. J Comp Civil Eng. 8:129–130.

Gessler PE, Chadwick OA, Chamran F, Althouse L, Holmes K. 2000. Modeling soil–landscape and ecosystem properties using terrain attributes. Soil Sci Soc Amer J. 64:2046–2056.

Gopal S, Woodcock CE, Strahler AH. 1999. Fuzzy neural network classification of global land cover from a 1 AVHRR data set. Remote Sens Environ. 67:230–243.

Gorsevski PV, Gessler P, Foltz RB. 2000. Spatial prediction of landslide hazard using discriminant analysis and GIS. GIS in the Rockies 2000: Applications for the 21st Century; Denver (CO).

Gutierrez R, Gibeaut J, Smyth R, Hepner T, Andrews J, Weed C, Gutelius W, Mastin M. 2001. Precise airborne lidar surveying for coastal research and geo-hazards applications. Int Arch Photogramm Remote Sens Spat Inform Sci. 34:185–194.

Guzzetti F. 2000. Landslide fatalities and the evaluation of landslide risk in Italy. Eng Geol. 58:89–107.

Guzzetti F, Mondini AC, Cardinale M, Fiorucci F, Santangelo M, Chang K-T. 2012. Landslide inventory maps: new tools for an old problem. Earth Sci Rev. 112:42–66.

Guzzetti F, Reichenbach P, Cardinale M, Galli M, Ardizzone F. 2005. Probabilistic landslide hazard assessment at the basin scale. Geomorphology. 72:272–299.

Haneberg WC, Cole WF, Kasali G. 2009. High-resolution lidar-based landslide hazard mapping and modeling, UCSF Parnassus Campus, San Francisco, USA. Bull Eng Geol Environ. 68:263–276.

Hewitt K. 1998. Catastrophic landslides and their effects on the Upper Indus streams, Karakoram Himalaya, northern Pakistan. Geomorphology. 26:47–80.

Hoffmeister R J, Miller DJ, Mills KA, Hinkle JC, Beier AE. 2002. GIS overview map of potential rapidly moving landslide hazards in western Oregon. Oregon Department of Geology and Mineral Industries Interpretive Map Series IMS-22.

Hsiao KH, Liu JK, Yu MF, Tseng YH. 2004. Change detection of landslide terrains using ground-based LiDAR data. XXth ISPRS Congress, Istanbul. (Commission VII, WG, Vol. 7, No. 5, p. 5).

Jaboyedoff M, Choffet M, Derron MH,orton P, Aoye A, Longchamp C, Mazotti B, Michoud C, Pedrazzini A. 2012. Preliminary slope mass movement susceptibility mapping using DEM and LiDAR DEM. Terrigenous mass movements. Berlin: Springer. p. 109–170.

Jaboyedoff M, Oppikofer T, Abellán A, Derron M-H, Aoye A, Metzger R, Pedrazzini A. 2012. Use of LIDAR in landslide investigations: a review. Nat Hazards. 61:5–28.

Jebur MN, Pradhan B, Tehrany MS. 2014. Optimization of landslide conditioning factors using very high-resolution airborne laser (LiDAR) scanning data at catchment scale. Remote Sens Environ. 152:150–165.
Jebur MN, Pradhan B, Tehrany MS. 2015. Manifestation of LiDAR derived parameters in spatial prediction of landslides using a novel ensemble evidential belief functions and support vector machine models in GIS. IEEE J Sel Top Appl Earth Obs Remote Sens. 8:674–689.

Jibson RW, Harp EL, Michael JA. 2000. A method for producing digital probabilistic seismic landslide hazard maps. Eng Geol. 58:271–289.

Kasai M, Ikeda M, Asahina T, Fujisawa K. 2009. LiDAR-derived DEM evaluation of deep-seated landslides in a steep and rocky region of Japan. Geomorphology. 113:57–69.

Khazai B, Sitar N. 2000. Landsliding in native ground: a GIS-based approach to regional seismic slope stability assessment. report.

Kiymik MK, Akin M, Subasi A. 2004. Automatic recognition of alertness level by using wavelet transform and artificial neural network. J Neurosci Method. 139:231–240.

Komar PD, Shih S-M. 1993. Cliff erosion along the Oregon coast: a tectonic-sea level imprint plus local controls by beach processes. J Coast Res. 9:747–765.

LaConite S, Strother S, Cherkassky V, Anderson J, Hu X. 2005. Support vector machines for temporal classification of block design fMRI data. NeuroImage. 26:317–329.

Lato M, Hutchinson J, Diederichs M, Ball D, Harrap R. 2009. Engineering monitoring of rockfall hazards along transportation corridors: using mobile terrestrial LiDAR. Nat Hazards Earth Syst Sci. 9:935–946.

Lee S. 2004. Application of likelihood ratio and logistic regression models to landslide susceptibility mapping using GIS. Environ Manag. 34:223–232.

Lee S, Choi J, Min K. 2004. Probabilistic landslide hazard mapping using GIS and remote sensing data at Boun, Korea. Int J Remote Sens. 25:2037–2052.

Lee CT, Huang CC, Lee JF, Pan KL, Lin ML, Dong JJ. 2008. Statistical approach to earthquake-induced landslide susceptibility. Eng Geol. 100:43–58.

Lee S, Pradhan B. 2006. Probabilistic landslide hazards and risk mapping on Penang Island, Malaysia. J Earth Syst Sci. 115:661–672.

Lee S, Pradhan B. 2007. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. Landslides. 4:33–41.

Li XZ, Kong JM. 2013. Application of GA-SVM method with parameter optimization for landslide development prediction. Nat Hazard Earth Syst Sci Discuss. 1:5295–5322.

Ma L. 2009. OGDC-5: statewide coverage of the Oregon geologic data compilation. Paper presented at: 2009 Portland GSA Annual Meeting; Portland, OR

Mahalingam R, Olsen MJ. 2015. Evaluation of the influence of source and spatial resolution of DEMs on derivative products used in landslide mapping. Geomat Nat Hazard Risk. 1–21. Available from: http://www.tandfonline.com/doi/10.1080/19475705.2015.1115431

Mansourian A, Rajabifard A, Valadan Zoej MJ, Williamson I. 2006. Using SDI and web-based system to facilitate disaster management. Comput Geosci. 32:303–315.

Manzo G, Tofani V, Segoni S, Battistini A, Catani F. 2013. Landslide science and practice: application of GIS techniques for landslide susceptibility assessment at regional scale. Berlin: Springer. p. 459–465.

Marjanović M, Kovačević M, Bajat B, Voženšek V. 2011. Landslide susceptibility assessment using SVM machine learning algorithm. Eng Geol. 123:225–234.

Miles DWR, Swanson FJ. 1986. Vegetation composition on recent landslides in the Cascade Mountains of western Oregon. Can J Forest Res. 16:739–744.

Moore JG, Normark WR, Holcomb RT. 1994. Giant hawaiian landslides. Ann Rev Earth and Planetary Sci. 22:119–144.

Mowen X, Esaki T, Guoyun Z. 2003. GIS method for slope-unit-based 3D landslide hazard evaluation. Chin J Rock Mech Eng. 6:969–976.

Nampak H, Pradhan B, Manap MA. 2014. Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. J HydroI. 513:283–300.

Neuhausser B, Terhorst B. 2007. Landslide susceptibility assessment using ‘weights-of-evidence’ applied to a study area at the Jurassic escarpment (SW-Germany). Geomorphology. 86:12–24.

Noguchi S, Tsuboyama Y, Siddle RC, Hosoda I. 1999. Morphological characteristics of macropores and the distribution of preferential flow pathways in a forested slope segment. Soil Sci Soc Am J. 63:1413–1423.

North WB. 1964. Coastal landslides in northern Oregon [dissertation]. Corvallis: Oregon State University.

Palmer L. 1977. Large landslides of the Columbia River gorge, Oregon and Washington. Rev Eng Geol. 3:69–84.

Paudel PP, Pradhan B, Xu C, Gokceoglu C. 2013. Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran. J Earth Syst Sci. 122:349–369.
Umar Z, Pradhan B, Ahmad A, Jebur MN, Tehrany MS. 2014. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. Catena. 118:124–135.

Van Den Eeckhaut M, Reichenbach P, Guzzetti F, Rossi M, Poesen J. 2009. Combined landslide inventory and susceptibility assessment based on different mapping units: an example from the Flemish Ardennes, Belgium. Nat Hazard Earth Syst Sci. 9:507–521.

Van Westen CJ, Rengers N, Terlien MTJ, Soeters R. 1997. Prediction of the occurrence of slope instability phenomena through GIS-based hazard zonation. Geol Rund. 86:404–414.

Ventura G, Vilardo G, Terranova C, Sessa EB. 2011. Tracking and evolution of complex active landslides by multi-temporal airborne LiDAR data: the Montaguto landslide (Southern Italy). Remote Sens Environ. 115:3237–3248.

Vijith H, Krishnakumar KN, Pradeep GS, Ninu Krishnan MV, Madhu G. 2013. Shallow landslide initiation susceptibility mapping by GIS-based weights-of-evidence analysis of multi-class spatial data-sets: a case study from the natural sloping terrain of Western Ghats, India. Georisk: Assess Manag Risk Eng Syst Geohazards. 8:48–62.

Wang Y, Summers RD, Hofmeister RJ. 2002. Landslide loss estimation pilot project in Oregon. Portland (OR): Department of Geology and Mineral Industries. Watershed Sciences. 2009. Lidar remote sensing data collection. Portland, OR: Department of Geology and Mineral Industries, Oregon Department of Forestry, and Puget Sound Lidar Consortium. Available from: http://www.oregongeology.org/sub/projects/olc/watershedsci/PDX/Hood_Lidar_Report_2007

Westen CJv, Terlien MJT. 1996. An approach towards deterministic landslide hazard analysis in GIS. A case study from Manizales (Colombia). Earth Surf Process Landforms. 21:853–868.

Xie M, Esaki T, Zhou G. 2004. GIS-based probabilistic mapping of landslide hazard using a three-dimensional deterministic model. Nat Hazards. 33:265–282.

Xu C, Dai F, Xu X, Lee YH. 2012. GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. Geomorphology. 145:70–80.

Yalcin A, Reis S, Aydinoglu AC, Yomralioglu T. 2011. A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. Catena. 85:274–287.

Yao X, Tham LG, Dai FC. 2008. Landslide susceptibility mapping based on support vector machine: a case study on natural slopes of Hong Kong, China. Geomorphology. 101:572–582.

Yao X, Zhang Y, Zhou N, Guo C, Yu K, Li LJ. 2013. Application of two-class SVM applied in landslide susceptibility mapping. Information Technology and Computer Application Engineering: Proceedings of the International Conference on Information Technology and Computer Application Engineering (ITCAE 2013), Hong Kong. p. 203.

Yilmaz C, Topal T, S üzen ML. 2012. GIS-based landslide susceptibility mapping using bivariate statistical analysis in Devrek (Zonguldak-Turkey). Environ Earth Sci. 65:2161–2178.

Yilmaz I. 2010. Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. Environ Earth Sci. 61:821–836.

Yilmaz I, Nefeslioglu HA, Marschalko M, Bednarik M. 2013. A. Blais-Stevens, P. Behnia, M. Kremer, A. Page, R. Kung, G. Bonham-Carter: Comments on Landslide susceptibility mapping of the Sea to Sky transportation corridor, British Columbia, Canada: comparison of two methods, Bulletin of Engineering Geology and the Environment, 2012, 71(3):447–466. doi: 10.1007/s10064-012-0421-z. Bull Eng Geol Environ. 72:625–627.

Yu FC, Chen TC, Lin ML, Chen CY, Yu WH. 2006. Landslides and rainfall characteristics analysis in Taipei City during the Typhoon Nari event. Nat Hazards. 37:153–167.

Zare M, Pourghasemi HR, Vafakhah M, Pradhan B. 2012. Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: a comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. Arab J Geosci. 6:2873–2888.