AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Earth Surface

RESEARCH ARTICLE
10.1002/2014JF003137

Key Points:
- A general deterministic search algorithm predicts discrete shallow landslides
- Size, shape, and location of synthetic and observed landslides are reproduced
- Many failure shapes possible at a site, local variations define which occurs

Correspondence to:
D. Bellugi,
dinob@mit.edu

Citation:
Bellugi, D., D. G. Milledge, W. E. Dietrich, J. A. McKean, J. T. Perron, E. B. Sudderth, and B. Kazian (2015), A spectral clustering search algorithm for predicting shallow landslide size and location, *J. Geophys. Res. Earth Surf.*, *120*, doi:10.1002/2014JF003137.

Received 24 FEB 2014
Accepted 27 DEC 2014
Accepted article online 6 JAN 2015

A spectral clustering search algorithm for predicting shallow landslide size and location

Dino Bellugi[1], David G. Milledge[2], William E. Dietrich[3], Jim A. McKean[4], J. Taylor Perron[1], Erik B. Sudderth[5], and Brian Kazian[6]

[1]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, [2]Department of Geography, Durham University, Durham, UK, [3]Department of Earth and Planetary Science, University of California, Berkeley, California, USA, [4]U.S. Department of Agriculture, U. S. Forest Service, Rocky Mountain Research Station, Boise, Idaho, USA, [5]Department of Computer Science, Brown University, Providence, Rhode Island, USA, [6]Palantir Technologies, Palo Alto, California, USA

Abstract The potential hazard and geomorphic significance of shallow landslides depend on their location and size. Commonly applied one-dimensional stability models do not include lateral resistances and cannot predict landslide size. Multidimensional models must be applied to specific geometries, which are not known a priori, and testing all possible geometries is computationally prohibitive. We present an efficient deterministic search algorithm based on spectral graph theory and couple it with a multidimensional stability model to predict discrete landslides in applications at scales broader than a single hillslope using gridded spatial data. The algorithm is general, assuming only that instability results when driving forces acting on a cluster of cells exceed the resisting forces on its margins and that clusters behave as rigid blocks with a failure plane at the soil-bedrock interface. This algorithm recovers predefined clusters of unstable cells of varying shape and size on a synthetic landscape, predicts the size, location, and shape of an observed shallow landslide using field-measured physical parameters, and is robust to modest changes in input parameters. The search algorithm identifies patches of potential instability within large areas of stable landscape. Within these patches will be many different combinations of cells with a Factor of Safety less than one, suggesting that subtle variations in local conditions (e.g., pore pressure and root strength) may determine the ultimate form and exact location at a specific site. Nonetheless, the tests presented here suggest that the search algorithm enables the prediction of shallow landslide size as well as location across landscapes.

1. Introduction

Shallow landslides generally involve only the colluvial soil mantle (less than a few meters deep) and are often translational, failing along a quasi-planar surface [e.g., *Rogers and Selby*, 1980; *Lehre*, 1982; *Glade*, 1998; *Robison et al*., 1999; *Guimaraes et al*., 2003; *Baum et al*., 2005; *Chigira and Yagi*, 2006]. Often triggered by extreme precipitation events, they can be the primary sources of debris flows, which sweep downslope with the capability to destroy property and cause loss of life (see review in *Sidle and Ochiai* [2006]). Shallow landslides also play an important role in landscape evolution. They are a primary cause of erosion in steep landscapes, and when mobilized as debris flows can carve valley networks [e.g., *Dietrich and Dunne*, 1978; *Benda and Dunne*, 1997; *Stock and Dietrich*, 2006]. The linkages between shallow landslides and debris flow initiation are complex, as a landslide may (a) mobilize as a debris flow, (b) trigger a debris flow upon entering a headwater channel, (c) transport sediment into channel heads and channels and only subsequently cause a debris flow, or (d) make sediment available on the hillslope to future hydrogeomorphic events [*Sidle and Ochiai*, 2006]. While the details of these processes may not be completely understood, two key characteristics that determine the importance of a shallow landslide, both in terms of hazard and geomorphic significance, are its location and size. Location and size affect the amount of sediment that is mobilized, the distance that the landslide debris then travels, the potential for mass gain (bulking up) or loss in transit, and the scale of local morphological change [*Benda and Cundy*, 1990; *Fannin and Wise*, 2001]. Moreover, *Hungr et al*. [2008] suggest that the volume of a shallow landslide controls the extent of the hazard area, the intensity of impact within it and the vulnerability of elements at risk. Location and size together thus determine the downslope effects, particularly the potential hazard for people and property.

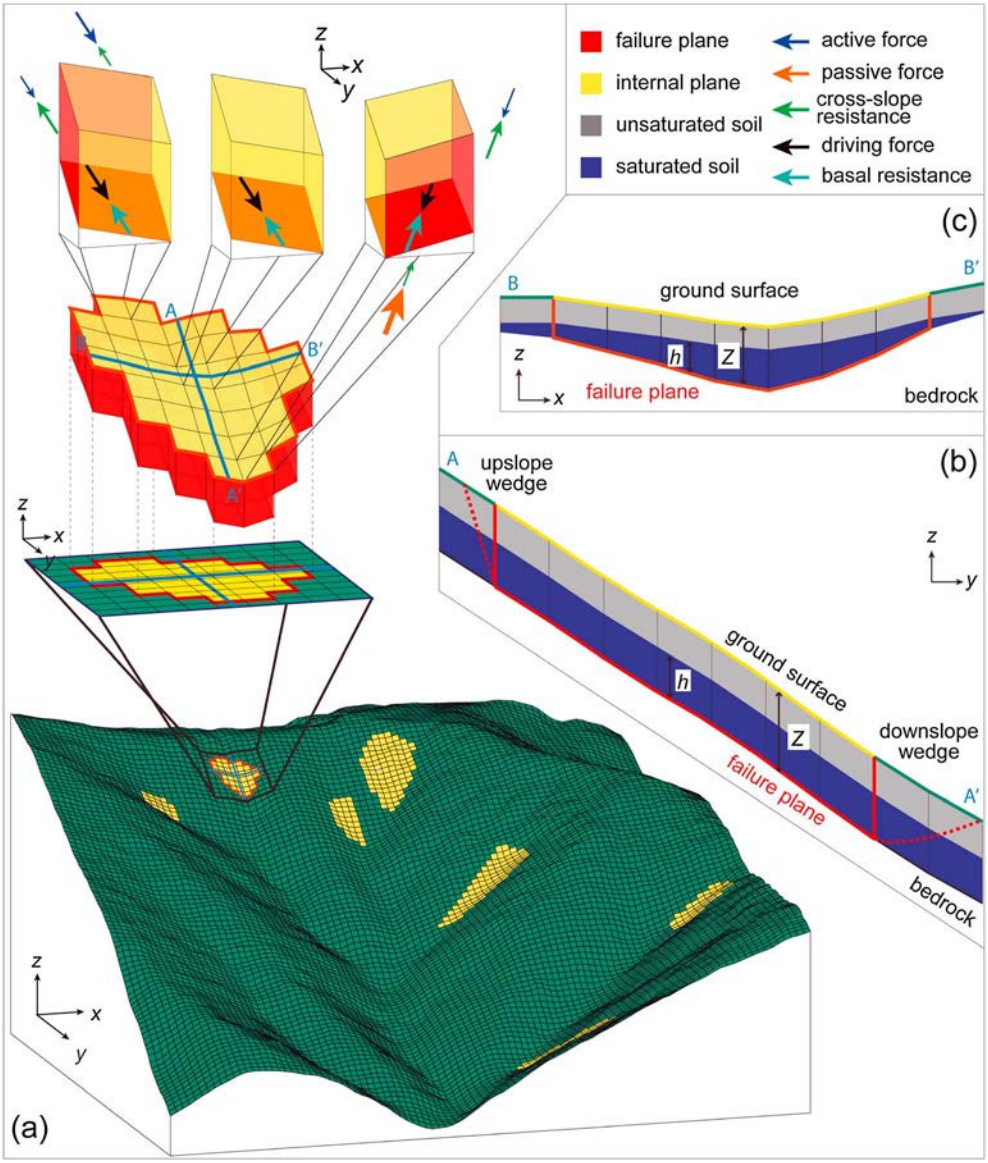

Figure 1. (a) Oblique view of a discretized landscape (Coos Bay, Oregon) showing steep convergent topography with discretized landslides in yellow. The inset shows the CB-1 landslide discretized into columns with landslide margins in red. The collective FS for these columns is the ratio of the total resisting force to the total driving force. Colored arrows show the driving and resistive forces acting on three example columns: an upslope boundary column, an internal column, and a downslope boundary column. Each landslide margin has two forces acting on it because grid cells are not oriented parallel to the slope (see Figure 2b). (b) Profile view of the landslide from A to A′ indicating the upslope and downslope wedges. (c) Cross section from B to B′ where *Z* is the soil depth and *h* is the water table height above the failure plane.

The advent of high-resolution topographic data has improved the detection of various types of landslides (e.g., *McKean and Roering* [2004], *Chigira and Yagi* [2006], and review in *Jaboyedoff et al.* [2012]), as well as the application of slope stability models to determine locations with high landslide susceptibility (see reviews in *Casadei et al.* [2003a] and *Simoni et al.* [2008]). However, we currently lack mechanistic models for specifically predicting the location and size of individual shallow landslides across landscapes, thus reducing the effectiveness of landslide hazard delineation and inhibiting our ability to formulate and apply mechanistic models for landslide-derived sediment flux and surface erosion due to landslides [*Dietrich et al.*, 2003].

Figure 1 presents an example of a landscape discretized into grid cells, containing information about their physical properties (e.g., elevation, slope, soil depth, contributing drainage area, pore water pressure, and

root strength). These cells extend vertically from the top of the bedrock to the surface, forming columns that represent the soil mantle. Although some of the landscape properties are often underconstrained, such a representation of a landscape is what is typically available as input to physically based models for predicting landslides, such as those shown in Figure 1. A shallow landslide prediction procedure thus needs to identify the collections of cells that together would be unstable considering the assigned properties for each cell. Common models that adopt a one-dimensional representation of slope stability define a landslide as either a single cell or a set of cells that will fail when pore water pressure is above a critical threshold [e.g., *O'Loughlin and Pearce*, 1976; *Wu et al*., 1979; *Montgomery and Dietrich*, 1994; *Wu and Sidle*, 1995; *Pack et al*., 1998; *Dietrich et al*., 1995, 2001; *Borga et al*., 2002; *Casadei et al*., 2003a; *Dhakal and Sidle*, 2003; *Iida*, 2004; *Rosso et al*., 2006; *Baum et al*., 2008; *Tarolli et al*., 2008; *Simoni et al*., 2008; *Lanni et al*., 2012]. In these models, landslides such as those shown in Figure 1 can be identified only if each of the cells that they contain is individually unstable and surrounding cells are not. Furthermore, such one-dimensional approaches cannot include lateral effects (notably root strength and soil friction), which are known to be important in defining instability [e.g., *Arellano and Stark*, 2000; *Schmidt et al*., 2001].

In order to account for lateral effects a multidimensional slope stability analysis is required. Many three-dimensional slope stability models have been proposed [e.g., *Hovland*, 1977; *Chen*, 1981; *Burroughs*, 1985; *Dietrich et al*., 2008]. These models perform a limit-equilibrium analysis for a defined failure surface, assuming that the soil mass behaves as one or more rigid blocks. While they usually require similar input parameters as the infinite-slope methods, their application has been limited as they require the treatment of discrete shapes that are not known a priori. Because the number of possible shapes grows exponentially with the number of grid cells, an exhaustive search that tests all possible shapes is effectively intractable at any relevant scale [*Dietrich et al*., 2008].

Few attempts have been made to apply three-dimensional methods at the watershed scale; all involve the introduction of constraints that effectively reduce the computational complexity of the problem. *Montgomery et al*. [2000] incorporated the effects of lateral root strength by using a predefined single rectangular landslide shape with fixed size, which results in the same limitations of the infinite-slope approaches with respect to landslide size (i.e., size is defined a priori). To examine the controls of lateral root reinforcement on landslide size, *Gabet and Dunne* [2002] and *Casadei et al*. [2003b] assumed landslides are rectangular and have a fixed length to width ratio but did not apply their model to a landscape. *Okimura* [1994] also assumed landslides to be rectangular but relaxed the restriction of a single length to width ratio. In this model landslide size is determined by computing a least stable cell using an infinite-slope stability model then exploring a fixed number of potential rectangular slide masses (constrained to be oriented downslope) that include the least stable cell, resulting in good agreement between predicted locations and shapes and of observed failures but over a very limited area. *Qiu et al*. [2007] instead tested potential ellipsoidal slip surfaces, using Hovland's method [*Hovland*, 1977] to compute their three-dimensional stability. They did not account for additional resistance provided by roots on the margins of the unstable block. To reduce the computational complexity, they took a nondeterministic approach in which 100 random potential failure surfaces centered on each grid cell were tested (thus limiting the number of slope stability tests to at most 100 times the number of grid cells).

Lehmann and Or [2012] proposed an alternative approach that relaxes the assumption of a regular landslide shape but requires landslides to originate at a single cell to reduce its computational complexity. In their model a hillslope is discretized into soil columns interconnected by frictional and tensile mechanical bonds represented as fiber bundles [e.g., *Schwarz et al*., 2010]. If a failure threshold is reached based on the forces acting on the base of a column, its load is redistributed to its neighbors via the fiber bundles which in turn can gradually fail, allowing the failure to progress in both the upslope and downslope directions. This model allows for irregular shapes to develop but is computationally very intensive and as a result has only been applied to small synthetic landscapes. *Ruette et al*. [2013] developed a simplified version of this model that relies on prescribed brittle mechanical bonds between adjacent columns rather than the fiber bundle representation of lateral resistance used by *Lehmann and Or* [2012] and applied it to two catchments in the foothills of the Swiss Alps. While they generally yielded reasonable predictions of landslide events, they overpredicted the number and volume of observed landslides and greatly overpredicted the slopes on which landsides initiated. Both these approaches can capture the progressive nature of failure observed in some shallow landslides but have a dependence on grid resolution, as the likelihood of initial failure will vary with

cell size. These models also implicitly assume that deformation is large enough to allow failure at the base of a soil column and subsequent detachment upslope, yet small enough to allow the transfer of the loads by columns leaning on their downslope neighbors. However, because these models assume failure initiates at individual grid cells, and test the stability of groups of cells failing in series rather than failing simultaneously, they do not require a search algorithm.

Here we develop an efficient search algorithm that can be applied at the catchment scale to identify unstable clusters of adjacent cells without constraints on their size or shape, provided that the grid resolution is sufficiently fine (1–2 m) to represent a landslide as a collection of grid cells. Our approach is different from prior work in that it is deterministic rather than involving a random sample of possible shapes and is less dependent on grid spatial resolution in the sense that it does not require instability of individual cells to locate possible search locations. To test the search algorithm, we couple it to the limit-equilibrium multidimensional slope stability model MD-STAB [*Milledge et al*., 2015], which uses a traditional Mohr-Coulomb formulation of strength to calculate the driving and resistive forces on each cell. Other three-dimensional limit-equilibrium models could also be used with the search algorithm that identifies landslide location and size. With this definition of slope stability, we apply the search algorithm to a synthetic landscape and show that it can recover predefined clusters of unstable cells having a variety of regular and irregular shapes; and that it offers better performance and greater flexibility when compared to constrained rectangular or elliptical exhaustive searches using the same slope stability model. We then apply the search algorithm to a field setting and find that it recovers the size and location of a rainfall-triggered shallow landslide at a research site near Coos Bay, Oregon (CB-1), where all the relevant physical parameters such as hydrological conditions, soil depth, and root strength were field measured [*Montgomery et al*., 2009]. We do not perform model calibration to minimize misfits between predictions and observations. Rather, we use field measurements to estimate the landslide-relevant parameters (e.g., soil depth, root strength, and pore water pressure). In contrast to *Milledge et al*. [2015], who test the stability of predefined shapes, we let the search algorithm find the unstable (as defined by the slope stability model) shapes. We also assess the robustness of the procedure by varying each field-measured parameter by ±5% and ±10%, finding that it is not very sensitive to modest changes in the relevant parameters. Finally, in Appendix A we show that our procedure is efficient, with the number of operations growing quadratically with the number of grid cells, a dramatic reduction from the prohibitive exponential number of operations required to test every combination of grid cells. In a companion paper (*D. Bellugi*, *Milledge*, *D., Dietrich*, *W.E., McKean*, *J., Perron*, *J.T*., Predicting shallow landslide size and location across a natural landscape: Application of a spectral clustering search algorithm, submitted to *JGR-Earth Surface*, 2014, hereafter referred to as Bellugi et al., submitted), we use process-based submodels to estimate soil depth, root strength, and pore water pressure, to test the coupled search algorithm and slope stability model across a larger landscape where the parameters have not been directly measured at all locations.

2. A Three-Dimensional Slope Stability Model

In a landscape during times of elevated pore pressure, patches of unstable ground (represented as a collection of neighboring grid cells), where the collective driving forces exceed the resisting forces, may develop. While our search algorithm does not depend on the specific slope stability method adopted to characterize these forces, here we adopt the multidimensional slope stability model MD-STAB [*Milledge et al*., 2015]. This model is fully three-dimensional and incorporates the effects of root strength and soil friction on sloping boundaries and is thus appropriate for application to a natural landscape. *Milledge et al*. [2015] provide an extensive derivation and discussion of this model, and we summarize the approach here. The landscape is discretized into columns defined in the vertical by the soil layer bounded by the ground surface and the bedrock interface (Figure 1). To compute the stability of a group of adjacent columns, a framework similar to *Hovland*'s [1977] method to calculate the limit equilibrium stability is utilized. The Factor of Safety is computed from the ratio of total resisting force to total driving force along the failure surface [*Hovland*, 1977]. Resisting forces (forces that act against the driving force to maintain equilibrium) due to friction and root cohesion exist on the base, cross slope, upslope (head), and downslope (toe) margins of the group of columns (Figure 1). Driving forces are the downslope component of the block's weight and the force exerted from the soil mass upslope. MD-STAB assumes that failure occurs by simultaneous shear on the boundary of the landslide, without internal deformation and with the soil columns constituting the body of the landslide not

moving relative to each other. This implies that any intercolumn shear forces within the unstable block do not exceed the shear strength at the time of failure. As a result, resistance to movement by internal columns is only by basal shear resistance, and resistance to movement by boundary columns is by basal shear resistance, shear resistance on the vertical boundary between the failure and surrounding soil columns, and active or passive forces (forces exerted under active or passive conditions, i.e., under soil tension or compression, respectively) from soil columns upslope or downslope. No progressive failure, strain weakening, or strain-induced pore water pressure dynamics are considered. It is also assumed that (1) slope instability occurs in drained conditions, in which fluids can freely drain through the pores, and effective stresses apply throughout; (2) the base of the failure surface is at the interface between the soil and the bedrock; (3) the water table surface, ground surface, and soil-bedrock interface are parallel within each column; and (4) groundwater conditions are characterized by steady, slope-parallel subsurface flow. For simplicity, any infiltration, suction, or capillary rise effects in the unsaturated zone that might cause negative pore pressures are ignored. Instead, the block is partitioned into saturated and unsaturated zones. These can be combined as additive terms and simplified using a ratio of water table height (h (m)) to the depth of the failure plane (z (m)) to define a saturation ratio $m = h/z$ (Figure 1).

The formulation for basal shear resistance force (R_b (Pa)) is similar to that of the infinite-slope method [*Taylor*, 1948], except that the force is calculated over the basal area of each column ($wl \sec \theta$, where l and w are the length and width of each cell (m) and θ is the slope angle (deg), measured parallel to the slope (Figure 1)). The reinforcing effect of roots on the base of the block is quantified using an effective root cohesion term (C_b (Pa)). The basal shear resistance force R_b (Pa) is defined as

$$R_b = (\sec(\theta)C_b + \cos(\theta)z(\gamma_s - \gamma_w m)\tan(\phi))lw, \tag{1}$$

where ϕ is the soil's effective friction angle (deg), γ_s is the unit weight of the soil defined as $\gamma_s = g\rho_s$ (N/m^3), γ_w is the unit weight of water defined as $\gamma_w = g\rho_w$ (N/m^3), g (m/s^2) is acceleration due to gravity, ρ_s (kg/m^3) is the soil bulk density, and ρ_w (kg/m^3) is the density of water. A single soil density independent of soil moisture content is assumed [*Milledge et al.*, 2015]. It is assumed that the soils are normally consolidated, cohesionless, and have isotropic frictional properties. Soil cohesion can be included in the model when appropriate.

Deformation on the cross-slope lateral margins is assumed to be by pure shear along the vertical faces. The shear resistance due to earth pressure on the cross-slope sides can be calculated from the horizontal stress at a point and the soil friction angle [*Burroughs*, 1985; *Stark and Eid*, 1998], integrated over the area of the block's cross-slope sides. An effective cohesion term (C_l) is added to account for the effect of roots on lateral resistance. Horizontal stresses are calculated from vertical stresses using earth pressure coefficients [*Lambe and Whitman*, 1969]. The cross-slope shear resistance forces R_l and R_r are defined as

$$R_l = R_r = \frac{1}{2}K_o z^2(\gamma_s - \gamma_w m^2)\tan(\phi)l + C_l z l. \tag{2}$$

The at-rest earth pressure coefficient (K_0) is calculated from the soil friction angle (ϕ) using Jacky's empirical relation: $K_o = 1 - \sin\phi$ [*Jacky*, 1944].

The active (σ_a) and passive (σ_p) stresses at a given depth (z_i) on the upslope or downslope margin of the central block can be calculated from vertical stress (σ_z) using active (K_a) or passive (K_p) earth pressure coefficients. Classical soil mechanics theory defines several methods for earth pressure prediction, described in standard soil mechanics textbooks [e.g., *Das*, 2010]. Here we adopt the Coulomb method [*Coulomb*, 1776], on the upslope margin, and the log-spiral method [*Caquot and Kerisel*, 1948] on the downslope margin. The former assumes a planar failure surface on the upslope margin, while the latter accounts for the curvature in the failure surface that typically develops on the downslope margin. Both are upper bound plasticity solutions based on statically admissible stress fields [*Das*, 2010]. Both methods assume that the soil is incompressible and that its properties are homogeneous and isotropic [*Chen*, 1975]. These methods have been adapted from their original formulations to allow a sloping ground surface and cohesive soil by *Chugh and Smart* [1981], in the Coulomb case, and by *Soubra and Macuh* [2002], in the log-spiral case.

The total passive force (R_d) on the downslope margin is calculated by integrating the passive stress over the area of the block's downslope side. This force is the resultant of both the normal and shear forces (due to friction) on the downslope margin of the block and is inclined at the soil friction angle (ϕ). The passive force is

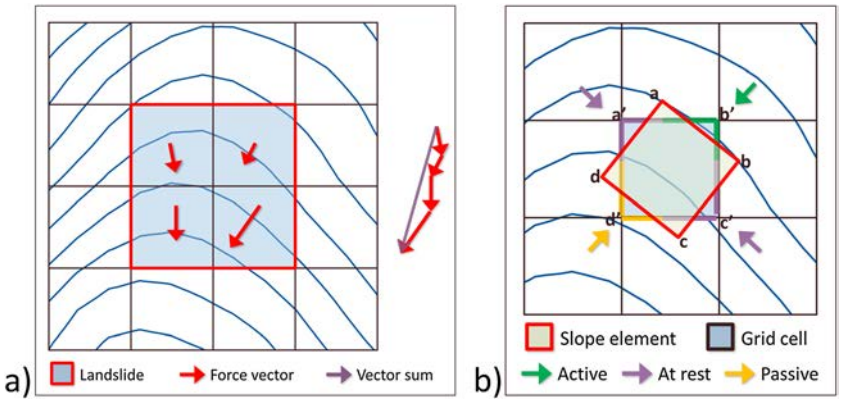

Figure 2. An example grid with elevation contours (blue lines) illustrating force partitioning and reorientation. (a) A four-cell potential landslide. The total driving force is the vector sum of the individual forces contributed by each cell. Note that the length of the resulting vector (the vector sum) is less than or equal to the sum of the individual lengths (the arithmetic sum). (b) The decomposition of active, passive, and at-rest earth pressure forces accounting for grid orientation. In this example, edges a–b and a–d experience active and at-rest forces, respectively. The grid cell edge a′–b′ (which intersects both a–b and a–d) experiences a combination of active and at-rest forces proportional to the amount of rotation between slope direction and the grid cell.

decomposed into its slope-parallel component, which acts as a resisting force on the downslope margin, and its slope-normal component, which modifies the normal (and thus resisting) force on the base. These forces are combined to give the net resisting force on the passive downslope margin R_d defined as

$$R_d = \frac{1}{2}K_p z^2\left(\gamma_s - \gamma_w m^2\right)w(\cos(\phi-\theta) - \sin(\phi-\theta)\tan(\phi)). \quad (3)$$

The active earth pressure conditions on the upslope margin can provide a net driving or resisting force, depending on the local parameters. This force is calculated as in equation (3), but with the active earth pressure coefficient K_a. In this case the slope-parallel component acts as a driving force, while the slope-normal component modifies the normal stress on the basal failure plane and thus the basal shear resistance. These can be combined to define the net driving force on the active upslope margin R_u as

$$R_u = \frac{1}{2}K_a z^2\left(\gamma_s - \gamma_w m^2\right)w(\cos(\phi-\theta) - \sin(\phi-\theta)\tan(\phi)), \quad (4)$$

In soils with a strong cohesive component, the force on the upslope margin is negative since the resisting forces due to cohesion exceed the driving force of the upslope soil wedge (Figure 1). In this case R_u represents a net resisting force on the upslope margin of the central block.

The downslope driving force F_d (the downslope component of the weight force on a column) is defined as

$$F_d = \sin(\theta)z\gamma_s lw. \quad (5)$$

The cell edges defined by the gridded topography are typically not aligned with the downslope direction at a given location (Figure 2a), so the total driving and resisting forces acting on a cluster of columns that represents a landslide are computed with vector sums. The driving force (equation (5)) is calculated for each column based on its aspect, and then these vectors are summed over the entire landslide area to obtain the total driving force. The vector sum is used for the Factor of Safety calculations, while the arithmetic sum is used to guide the search algorithm (Figure 2a). The sum of the magnitudes of forces is an upper bound on the true force (i.e., as computed by a vector sum) and provides an approximate criterion for identifying candidate landslides. The resisting force of each column is the sum of its basal (equation (1)) and lateral components (equations 2–4). Lateral components are only included for column margins that represent the edge of the landslide. Since the grid is not oriented slope parallel, most cells have more than one-force component (upslope, downslope, and cross slope) acting on them. The lateral resistance on each column margin is decomposed into its components by assigning a fraction of the edge length to each resistance component (Figure 2b). The dependence of lateral resistance on perimeter length introduces grid resolution dependence to the stability calculations. Its impact on stability predictions is complicated by the resolution dependence

of other driving parameters (e.g., slope) and is not examined in detail here. In a related study, Bellugi et al. (submitted) find that, as a result of grid discretization, the perimeter length of 34 landslides mapped over a 10 year period in the area surrounding our study site can increase up to 140%. Because discretized area instead does not vary significantly, the estimation of driving forces is not affected. As a result, the stability of these mapped landslides tends to be overestimated (Bellugi et al., submitted).

The Factor of Safety (FS) for the group of columns is defined as the ratio of its total resisting to driving forces,

$$\mathrm{FS} = \frac{\sum_{\text{all}} R_b + \sum_{\text{left}} R_l + \sum_{\text{right}} R_r + \sum_{\text{down}} R_d - \sum_{\text{up}} R_u}{\sum_{\text{all}} F_d}, \tag{6}$$

where the summation subscripts indicate whether the sums are performed over all the columns, or those on the left, right, downslope, or upslope margins only.

3. Search Algorithm

3.1. General Framework

In a landscape such as that illustrated in Figure 1, a shallow landslide can be conceptualized as a collection of surface elements (e.g., grid cells), which behave coherently as a solid block, mobilizing together in accordance to their physical properties. We assume that the resolution of the grid is sufficiently fine to represent a landslide as a collection of grid cells, i.e., that it is finer than the smallest landslide of interest. While the effect of grid resolution is not explored in this manuscript, it is apparent that a resolution that is too coarse will result in smaller landslides not being predicted, and a resolution that is too fine will result in unnecessary computational burden. In practice, a grid resolution of 1–2 m results in a good balance between these two end-member cases. Our aim is to find all unstable collections of grid cells, but the exact solution to this problem requires testing the stability of every possible combination of grid cells. The number of possible combinations of cells grows exponentially with the number of cells n in the grid, resulting in up to 2^n different combinations of cells. For example, exhaustively testing every combination of grid cells across a small 1 km^2 landscape discretized into a square grid composed of 1000 by 1000 cells (as one would obtain using modern lidar data) could result in exploring up to $2^{1,000,000}$ combinations of cells, a number so vast that the task would be unfeasible even using the world's most powerful computers [*Dasgupta et al.*, 2006].

Alternatively, the collection of cells that fails when its cumulative driving forces exceed the cumulative resistive forces (the landslide) could be thought of as the result of an optimization process that minimizes the ratio of resistive to driving forces (the FS). While no algorithm currently exists for finding exact solutions for this class of optimization problems in a polynomial number of iterations, clustering algorithms based on graph theory can provide good approximations. Graphs are combinatorial mathematical structures used to model pairwise relations between objects, consisting of a set of vertices (or nodes) that represent the objects, and a set of edges (or lines) that represent their relation or connection. Vertices and edges may have additional attributes that describe objects and their connections. If connections are asymmetric (e.g., the flow between two locations on a river), edges have directions and the graph is directed. If instead the connections are symmetrical (e.g., the line of sight between two mountaintops), edges have no direction and the graph is undirected. For more formal definitions we refer to common graph theory textbooks [e.g., *Diestel*, 2005; *Gross and Yellen*, 2005]. Graphs can also be represented using matrices, in which entries (i, i) encode the attributes of vertex i and entries (i, j) encode the attributes of the edge connecting vertex i and vertex j. In recent years spectral graph theory, the study of the properties of a graph in relationship to its associated matrices and their eigenvectors and eigenvalues (the graph's spectrum) [*Chung*, 1997], has enabled the use of modern linear algebra techniques in graph theory applications. In particular, this has allowed the development of algorithms that, by making use of graph spectra, can efficiently solve clustering problems [*Von Luxburg*, 2007]. Clustering is the task of grouping a set of objects such that those in the same group (a cluster) are more similar (in some sense or another) to each other than to those in other groups [*Kaufman and Rousseeuw*, 2005]. A gridded landscape can be depicted as a graph, with vertices corresponding to soil columns and edges to the forces that can develop between them. A cluster (e.g., of unstable cells) can be obtained by cutting all the edges that link those vertices (cells) inside the cluster (the landslide) with those outside (the stable landscape). For example, the four cells outlined in red in Figure 2a correspond to the four blue vertices of the graph

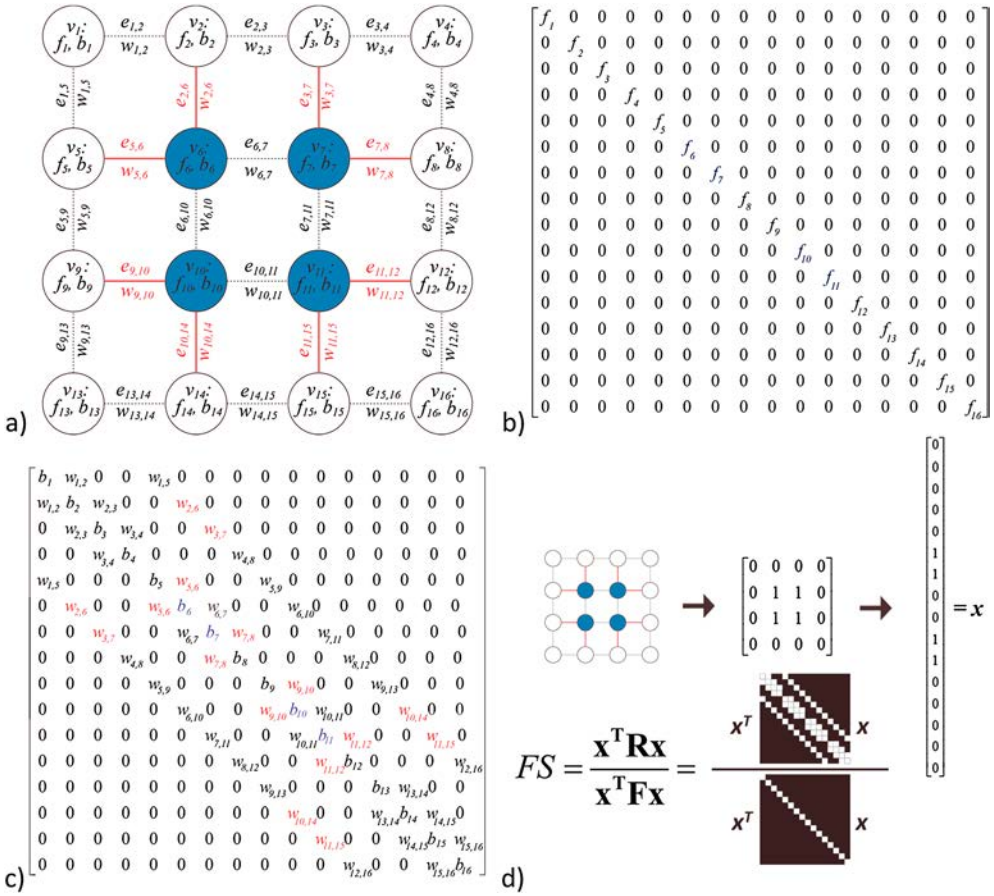

Figure 3. (a) Graph $G = (V,E)$ of a discretized landscape, consisting of 16 grid cells. The vertices, labeled v_1–v_{16}, are associated with the mobilizing forces f_i, as well as the frictional and cohesion forces b_i, contributed by each grid cell. The edges represent the resistive forces between grid cells: the edge labeled $e_{i,j}$, encodes the symmetrized weights $w_{i,j}$ between vertices v_i and v_j. For the partition S (the blue vertices) to fail as a landslide, their mobilizing forces must exceed the resistive forces of the red edges combined with their frictional forces. (b) The driving force matrix **F** for the graph G; the f_i, values are computed using equation (7). (c) The resistive force matrix **R** for the graph G; the r_i, values are computed using equations (8) and (9). Red, blue, and black colors of entries in Figures 3b and 3c correspond to the colors in Figure 3a. (d) The indicator vector **x** for partition S of G and its FS in matrix form.

shown in Figure 3a. For this group of cells to fail it must overcome the resistances that develop at its boundary, which is equivalent to cutting the red edges shown in Figure 3a.

Choosing which vertices to include in a cluster, and therefore which edges to cut, depends on the objective, in our case to minimize the stability of the cluster. This is a combinatorial problem involving a global optimization (i.e., applied to all vertices and edges). In the context of image segmentation, *Shi and Malik* [2000] associate each image pixel with a vertex of the graph, and a measure of pixel similarity to the edges of the graph. To identify clusters of similar pixels, they introduced an objective function defined by the normalized cut, which computes the cut cost (the sum of the edges that must be cut to form the cluster) as a fraction of the total edge connections to all the vertices in the graph. They then use a spectral method to find partitions that minimize the normalized cut. The advantage of the normalized cut is that it minimizes cost by removing weak (low similarity) edges between clusters and maximizes benefit by including more similar vertices in a cluster. In contrast to other graph cut definitions (e.g., the minimum cut [see *Papadimitriou and Steiglitz*, 1998]) that mostly focus on removing weak edges, the normalized cut combines global and local information resulting in more balanced graph partitions.

In the case of landslides costs and benefits derive from the resistive and driving forces experienced by a cluster of unstable cells, and the objective function is the Factor of Safety of that cluster. Thus, we can identify

landslides using a spectral clustering approach similar to that of *Shi and Malik* [2000], but with the advantage that the objective function is physically based. Using a slope stability model such as that in section 2, we determine the contribution of each soil column to the driving force of a potential landslide that includes that cell, the resistance on the base of each soil column, and the resistance between each soil column and its neighbors. We encode the driving and basal resisting forces contributed by each column in the vertices of the graph, and the potential lateral resisting forces that can develop between columns in the edges of the graph. Each vertex thus corresponds to a discretized soil column (i.e., a grid cell), and each edge corresponds to a boundary between columns (Figure 1). We define the objective function as the ratio of all the resistances acting on a cluster of cells and the total driving force contributed by those cells. We develop a spectral method to find clusters (i.e., landslides) that minimize this function, guiding the search for the unstable areas of the landscape.

3.2. Spectral Clustering Derivation

The landscape and all its spatial attributes are discretized into a regular grid (i.e., with square grid cells). The attributes of the landscape relevant to shallow landslides are represented as an undirected weighted graph $G = (V,E)$, where the vertices $v \in V$ of the graph represent the vertical soil columns corresponding to the original discretized grid cells. An edge $e_{ij} \in E$ is formed between every pair of neighboring vertices v_i and v_j and represents the forces (which are encoded in the edge weights) acting between two adjacent columns (Figure 3a). The discretization of the landscape into soil columns with four sides results in each vertex having four neighbors. The edges of G are thus associated with the four lateral faces of the soil columns, on which the forces acting between adjacent grid cells are defined. This graph G can be represented as a weighted adjacency matrix $\mathbf{W}$, with entries w_{ij} set to 0 when v_i and v_j are not connected in G and to the magnitude of the resistive forces on edge e_{ij} otherwise. G is four regular as each vertex has four neighbors (upslope, downslope, left, and right). For a grid of size n by m, $\mathbf{W}$ will be of size nm by nm but will contain mostly zero values, and thus will be extremely sparse [*Cormen et al*., 2001]. This is important computationally, as the use of sparse matrices introduces little memory overhead in comparison to the original elevation grid.

The vertices of this graph are annotated with the forces contributed by each grid cell: the driving forces f and the resistive forces b due to friction and cohesion acting on the base of the column (Figure 3a). Similarly, the edges of the graph are annotated with the forces acting between a grid cell and its neighbors: these resistive forces, which arise from root strength and earth pressure, are mostly positive, but may be negative to represent the active "push" from upslope neighbors. Spectral clustering methods associate a scalar weight with each node [*Von Luxburg*, 2007], and thus, we associate each grid cell with the magnitude of its driving force, rather than magnitude and direction. The total driving force is thus the arithmetic sum rather than the vector sum, resulting in an over estimation of the force (Figure 2a). As a result, some stable areas may be unnecessarily examined by the search algorithm. However, when candidate landslides are examined by the algorithm, the FS is correctly computed using the vector sum (Figure 2a). The driving and resistive forces are encoded in two nm by nm matrices $\mathbf{F}$ and $\mathbf{R}$. The force matrix $\mathbf{F}$ contains the driving forces associated with the vertices of G along the diagonal (Figure 3b):

$$\mathbf{F_{ii}} = F_{d_i}, \tag{7}$$

where F_{d_i}, is the force contributed by vertex v_i. The resistance matrix $\mathbf{R}$ contains the resistive forces associated with both the vertices (basal resistance) and the edges (lateral resistance) of G (Figure 3c). The diagonal of $\mathbf{R}$ is

$$\mathbf{R_{ii}} = R_{b_i} + \sum_{j \neq i}^{m*n} w_{ij}, \tag{8}$$

where R_{b_i} is the basal resistance contributed by the soil column represented by vertex v_i and the w_{ij} term is the sum of all the resistances between v_i and its neighbors. This definition differs from the general formulation of the Laplacian matrix typically used in spectral clustering methods [e.g., *Shi and Malik* [2000]], as it also includes the R_{b_i}, terms (which are associated with the vertices of G). A central assumption of spectral clustering approaches is that the relationship between neighbors is symmetric (i.e., the graph is undirected) [*Von Luxburg*, 2007]. The pixel similarity used by *Shi and Malik* [2000] is a symmetrical measure. In their graph representation, v_i similar to v_j implies v_j similar to v_i and that $w_{ij} = w_{ji}$, resulting in an undirected graph. This is not the case in this application, as the resistive forces between grid cells are not always symmetrical. In particular, the active and passive earth pressures across a vertical boundary separating an

upslope cell from its downslope neighbor can be different, as the passive resistance of the downslope cell generally is greater than the active force from the upslope cell. To obtain an undirected graph, the boundary forces are first symmetrized: the off-diagonal entries of **R**, visualized in Figure 3c, contain the negated average of the weights along a given edge:

$$\mathbf{R_{ij}} = \mathbf{R_{ji}} = -\frac{w_{ij} + w_{ji}}{2}, i \neq j. \tag{9}$$

While this formulation is rather different from the usual spectral clustering definitions [e.g., *Von Luxburg*, 2007], it allows the exact computation of the numerator of equation (6) while retaining a symmetric matrix. The matrices **R** and **F** have a similar form to the weighted adjacency (**A**) and degree (**D**) matrices of a graph, and $\mathbf{F}^{-1/2}\mathrm{RF}^{-1/2}$ is thus similar in form to the normalized Laplacian ($\mathbf{L} = \mathbf{D} - \mathbf{A}$) matrix of a weighted graph [*Chung*, 1997], which has the form $\mathbf{D}^{-1/2}\mathrm{LD}^{-1/2}$ and is commonly used in spectral clustering algorithms [*Von Luxburg*, 2007].

The aim is to partition G to delineate unstable clusters of cells $S \in G$ (Figure 3a), defined by a binary indicator vector **x** of length nm. Each component $\mathbf{x}_k \in \{0,1\}$ indicates whether vertex v_k is part of S: $\mathbf{x}_k = 1$ if $v_k \in S$, and $\mathbf{x}_k = 0$ if $v_k \notin S$. The index k corresponds to the position of a grid cell in a linearized representation of the grid (Figure 3d). We define the cost function $C(\mathbf{x})$ of the partition S as the FS of S expressed in terms of **x**, **R**, and **F**:

$$C(\mathbf{x}) = \frac{\mathbf{x^T R x}}{\mathbf{x^T F x}}, \tag{10}$$

where $\mathbf{x^T}$ is the transpose of **x**. Consistent with the assumption of a rigid block with resistive forces acting on its margins, entries $\mathbf{R_{ji}}$ from equation (9), corresponding to the resistance between a pair of adjacent vertices v_i and v_j, are considered only if one of v_i and v_j corresponds to a nonzero entry of **x**. This is equivalent to assuming that lateral resistance acts between adjacent grid cells i and j if and only if they straddle the margin of the block. Similarly, entries $\mathbf{R_{ii}}$ from equation (8), which include the resistance on the base of v_i, are only considered if v_i corresponds to a nonzero entry of **x**. This is consistent with grid cell i contributing to the resistance on the base of the block if and only if it is part of the block. With this construction, averaging of the weights w_{ij} and w_{ji} in equation (9) does not change the cost function: $C(\mathbf{x})$ in equation (10) is exactly equal to FS in equation (6) for any partition defined by an indicator vector **x**.

The optimal partition S^* (the one having the lowest FS) is determined by the indicator vector $\mathbf{x}^*$ which minimizes $C(\mathbf{x}^*)$. Finding $\mathbf{x}^*$ in the discrete case (i.e., $\mathbf{x}^* \in \{0,1\}$) is an intractable problem [see *Shi and Malik*, 2000, Appendix A], and thus, an approximation must be used. Following an approach similar to *Shi and Malik* [2000], we find an approximate solution by relaxing the condition that $\mathbf{x}^* \in \{0,1\}$, and let $\mathbf{x}^*$ take real values (i.e., $\mathbf{x}^* \in \mathbb{R}$) (commonly referred to as a spectral relaxation). While this allows a tractable minimization of $\mathbf{x}^*$, the solution is no longer discrete (i.e., landslide or not landslide). Instead, $\mathbf{x}^*$ can be interpreted as a continuous indicator vector, representing the fraction of each node that contributes to the optimal partition S^*. If we let $\boldsymbol{y} = \mathbf{F}^{1/2}\mathrm{x}$, and substitute into equation (10), the cost function becomes

$$C(\mathbf{y}) = \frac{\mathbf{y^T F^{-1/2} R F^{-1/2} y}}{\mathbf{y^T y}}. \tag{11}$$

The cost function defined by equation (11) has the same form as the Rayleigh quotient of a positive semidefinite matrix **M** and a nonzero vector **x**, defined as $(\mathbf{x^T M x})/(\mathbf{x^T x})$ [*Horn and Johnson*, 1985]: as **R** is a real-valued symmetric square matrix and **F** is a diagonal matrix with positive entries, the inner product of $\mathbf{F}^{-1/2}\mathrm{RF}^{-1/2}$ is also symmetric and positive semidefinite. The minimizing solution to the Raleigh quotient is the eigenvector $\mathbf{y}^*$ pointed to by the smallest nonzero eigenvalue of the linear system defined by

$$\mathbf{F}^{-1/2}\mathbf{R}\mathbf{F}^{-1/2}\mathbf{y} = \lambda\mathbf{y} \tag{12}$$

[*Horn and Johnson*, 1985]. The optimal partition S^* is thus defined by $\mathbf{x}^* = \mathbf{F}^{-1/2}\mathrm{y}^*$, after reversing the change of variable $\mathbf{y} = \mathbf{F}^{1/2}\mathrm{x}$ used to obtain the Rayleigh quotient. As $\mathbf{x}^*$ is only an approximation of the original discrete **x**, we do not limit ourselves to the first eigenvector, rather we take the first k eigenvectors, where k can be limited by a maximal value of $C(\mathbf{x}^*)$ or, in practice, by computational constraints. These eigenvectors effectively guide the search algorithm in finding patches of instability in the landscape. In this study we set $k = 164$, a compromise between a value that results in fast running times and one that results in a more exhaustive search.

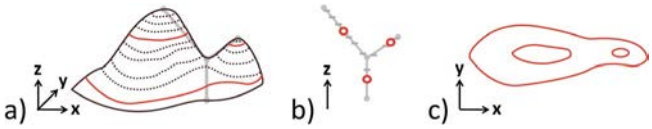

Figure 4. (a) Side view of the 3-D surface defined by the continuous vector $\mathbf{x}^*$. Dashed contour lines define regions r_i associated with threshold t_i, solid contour lines indicate the regions with the minimal FS for each gray edge in the contour tree T_c, with solid gray circles representing the local minima and maxima of $\mathbf{x}^*$. (b) The contour tree T_c arising from $\mathbf{x}^*$. Solid gray circles are the vertices of T_c, corresponding to the birth, and merging of the regions r_i; solid gray lines correspond to the edges of T_c in which individual regions grow; dashes represent the thresholds t_i; unfilled circles show the thresholds that minimize the FS for each edge. (c) Overhead view of Figure 4a showing only the three overlapping regions r_i from thresholds t_i that minimize the FS for each edge of T_c, corresponding to the solid red lines in Figure 4a and the unfilled red circles in Figure 4b.

3.3. Recovering Discrete Landslides

An approximation to the discrete solution $\mathbf{x}$ of equation (10) can be obtained by thresholding the continuous vector $\mathbf{x}^*$. As $\mathbf{x}^*$ is also of length nm, all $nm-1$ possible interesting threshold values t_i are examined. All $\mathbf{x_i}$ for which the corresponding $\mathbf{x_i}^*$ are greater than t_i are set to 1, and the rest are set to 0. In other words, $\mathbf{x}^*$ can be visualized as a surface with each threshold defining a contour line on the surface, and the regions at threshold t_i are the projection on the Cartesian plane of the points with values higher than t_i (Figure 4a). These regions are extracted using connected-component analysis [*Haralick and Shapiro*, 1992], by which any threshold t_i may give rise to disconnected discrete regions r_i in which $r_i = 1$ if $\mathbf{x_i}^* > t_i$ and $r_i = 0$ otherwise. In particular, a region r_i at threshold t_i must either have not existed at threshold t_{i-1}, or be an expansion of a region r_{i-1}' at threshold t_{i-1}, or be the result of the merging of regions r_{i-1}' and r_{i-1}'' at threshold t_{i-1}. The evolution of the regions as the threshold is varied gives rise to a topological contour tree T_c (Figure 4b) in which vertices represent the birth and merging of regions, while edges represent the expansion or contraction of an individual region [*Freeman and Morse*, 1967; *Carr et al.*, 2000].

As we are interested not only in the least stable cluster of cells in a landscape but in all the unstable clusters, we examine the complete contour trees resulting from each eigenvector. Each edge in T_c is traversed, and at each threshold the FS of r_i is computed for the corresponding collection of cells. This allows the use of a vector sum to correctly compute the driving forces. On each edge, multiple thresholds may result in shapes that have a FS below unity (Figures 4a and 4b). These shapes will overlap one another (Figure 4c), but only one will actually fail at a specific location. There are many ways to choose among these unstable shapes. Here we choose the unstable shape with the lowest unstable FS, (referred to as FS_{min}) as it is most consistent with the optimization and is commonly assumed to be the most probable [e.g., *Montgomery and Dietrich*, 1994; *Rosso et al.*, 2006; *Stark and Guzzetti*, 2009]. However, it is also the case that under evolving conditions the first cluster of cells to cross the stability threshold can be the one that fails [*Casadei et al.*, 2003a]. Thus, the algorithm offers the option of retaining the overlapping shape with FS closest to but less than unity (i.e., the highest unstable FS, referred to as FS_{max}).

The thresholding process is then repeated for the complement of $\mathbf{x}^*$, $\mathbf{x}^{*\prime}$ ($\mathbf{x}^{*\prime} = -\mathbf{x}^*$), which is also an eigenvector of equation (12), after the change of variables. This is equivalent to inverting the $\mathbf{x}^*$ surface and generating a new contour tree T_c'. Thresholding $\mathbf{x}^{*\prime}$ results in the complements of the discrete regions arising from $\mathbf{x}^*$ (i.e., $r_i = 0$ if $\mathbf{x_i}^* > t_i$, and 1 otherwise), thus allowing the exploration of all the regions initiating from the local maxima and minima of $\mathbf{x}^*$ (the local maxima of $\mathbf{x}^*$ are the local minima of $\mathbf{x}^{*\prime}$). The set of all regions extracted from T_c and T_c' represent the predicted landslides from the eigenvector $\mathbf{y}^*$. This process is repeated for each of the k eigenvectors. While only one landslide is retained per edge of T_c and T_c', multiple landslide predictions arise from the many thresholds applied to many eigenvectors, and these landslides may overlap (Figure 4 illustrates the case of overlapping predictions from a single eigenvector). As only one of these overlapping landslides will fail for a particular storm event, they are pruned a posteriori using the same criterion (FS_{min} or FS_{max}) as in the thresholding process. In the FS_{min} case, predicted shapes are sorted by their FS in ascending order into a list. The first element of the list is pairwise compared to all the other elements in sorted order, and whenever another element of the list overlaps with the first it is removed from the list. The process is then repeated for the next element of the list until the list is fully traversed. The pruning process is similar in the FS_{max} case, but with the predicted shapes sorted in descending order.

4. Testing the Shallow Landslide Prediction Procedure

We test the ability of our procedure to identify clusters of unstable cells on small synthetic and natural landscapes. The modest size of these landscapes enables the comparison of our search algorithm to exhaustive search methods in which the number of possible shapes is constrained by assuming landslides to be rectangular or elliptical. The synthetic landscape consists of a stable planar slope where soil slope, depth, pore water pressure, and root strength are constant. On this slope we introduce patches with regular and irregular shapes and sufficiently reduced root strength to make them the only unstable parts of the landscape. We then test the ability of the procedure to recover landslides with size and location similar to the pre-defined patches, and contrast to the alternative search methods using the same slope stability model. In the natural landscape, where soil slope, depth, pore water pressure, and root strength are spatially varying, we test the ability of the procedure to recover a known landslide using field measurements to constrain all the relevant parameters at the time of the failure. In this case we can compare size, location, and shape of predictions to those of the known landslide. We also contrast to the alternative search methods and assess the robustness of the procedure by varying each field-measured parameter by ±5%.

4.1. Application to a Synthetic Landscape

The synthetic hillslope consists of a 51 m by 51 m planar surface inclined at 35°, discretized into 1 m^2 grid cells. We set the soil depth z to 1 m and assume the soil has a friction angle of 35°, a bulk density of 1600 kg m^{-3}, and is fully saturated ($m = 1$). We set the basal root strength (the parameter C_b in equation (1)) to 1.5 kPa and the depth-averaged lateral root strength (the parameter C_l in equation (2)) to 60 kPa. These values are chosen so that no part of the slope can be unstable and are representative of natural forests [*Schmidt et al*., 2001]. We then define regular and irregular patches of reduced strength centered on the planar slope. The four regular patches consist of a 25 m^2 square, a 45 m^2 rectangle, a 49 m^2 circle, and an 87 m^2 ellipse (Figure 5a). The four irregular patches are generated using the method of *Robidoux et al*. [2011] that produces random, smooth, simply-connected (i.e., with no holes) shapes, by connecting eight random points drawn from a uniform distribution on a disk, applying a Gaussian blur and retaining the pixels above a 5% intensity threshold [*Robidoux et al*., 2011, Figure 2]. To make these patches similar in size to the regular ones, they are then rescaled to 76, 53, 73, and 51 m^2, respectively, and translated to the center of the planar slope (Figure 6a). We then ensure that each of the patches is just unstable by lowering the root strength parameters C_b and C_l of the patch until it becomes unstable, but the same patch is stable when eroded or dilated by one grid cell. The C_b and C_l values selected for each patch are listed in Table 1 and are in the range reported by *Schmidt et al*. [2001] for clear-cut conditions.

This reduced-strength patch generation method generally results in unstable areas confined to a ±1 grid cell buffer zone around the reduced strength patch boundary. However, it does not guarantee that there is a single unstable area within any reduced-strength patch or that the least stable area has the same shape as the reduced-strength patch. For rectangular reduced-strength shapes generated with this method, the only possible unstable shapes are in the ±1 pixel buffer around the perimeter. While this results in many millions of possible unstable combinations, the least stable shape can still be found by an exhaustive search of all possible shapes. For nonrectangular shapes the least stable area cannot be constrained in a ±1 pixel buffer, so the number of possible combinations grows exponentially larger, and testing all the combinations is prohibitive. As a result no "ground truth" can be defined, and thus, we define success as (1) the ability to identify a landslide where we know that one should exist (i.e., some configuration of cells in the low strength patch is known to be unstable); and (2) the similarity between the predicted landslide and the defined low strength patch, in terms of size and shape.

For each test, the force matrix **F** and Laplacian-like resistance matrix **R** are assembled (equation (7) and equations (8) and (9), respectively). The eigenvectors associated with the 164 smallest eigenvalues of equation (12) are examined. Thresholding these 51^2-element eigenvectors is analogous to taking 51^2-1 slices through these surfaces parallel to the x-y plane and setting the value of each cell to be one if the surface is above the slice at that cell or zero if not. The eigenvectors resulting in the shapes with the lowest FS out of all the unstable shapes found are shown in Figures 5b (for regular shapes) and 6b (for irregular shapes). The corresponding shapes with the lowest FS are shown in Figures 5c and 6c. The theoretical maximum number of shapes examined by the search algorithm in a 51 by 51 grid is approximately 20 million, a dramatic

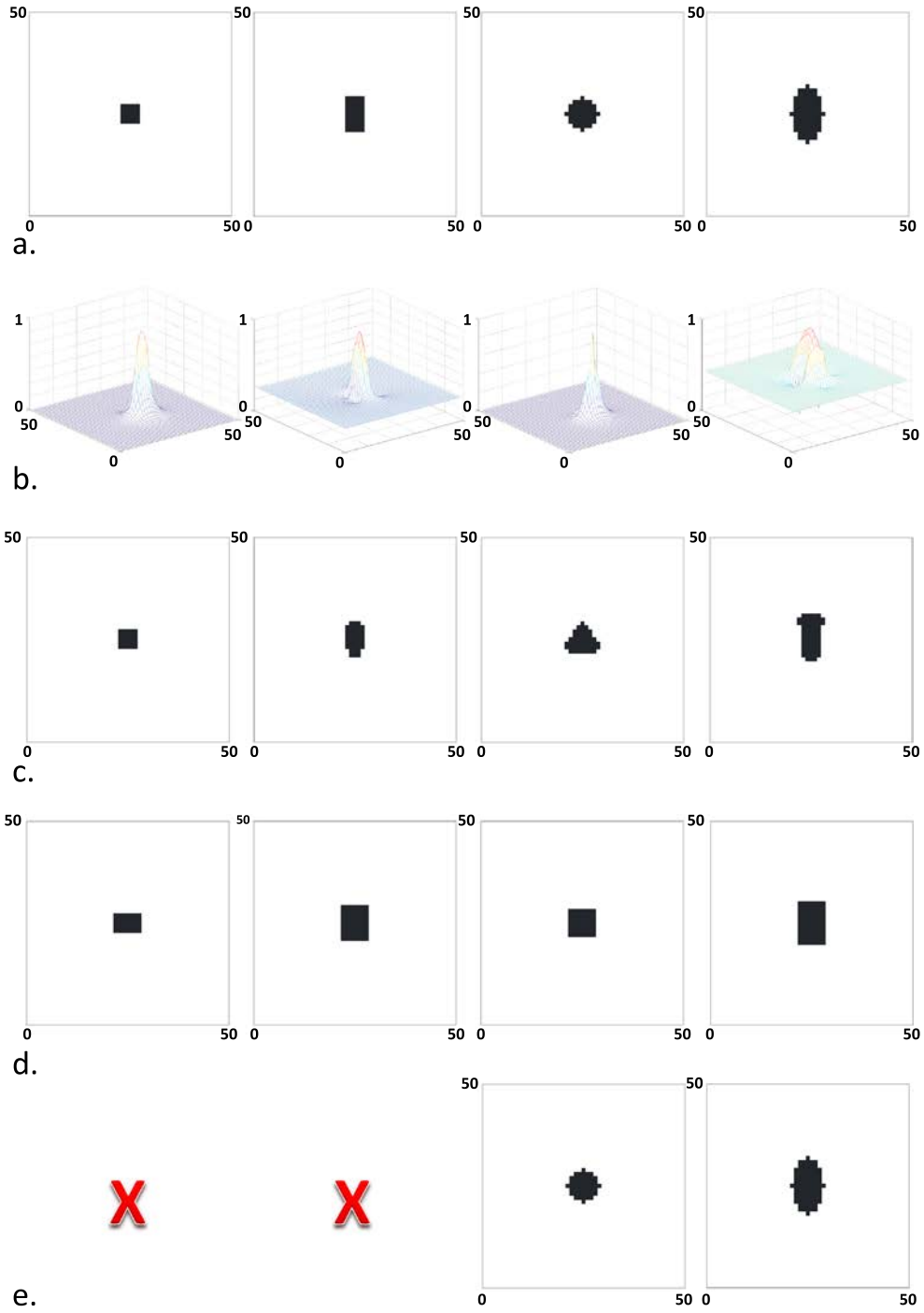

Figure 5. (a) Synthetic landscape with square, rectangular, circular, and elliptical low root strength patches. (b) The eigenvectors of equation (12) resulting in the shapes with the lowest FS. (c) Predictions resulting from thresholding the eigenvectors in Figure 5b. (d) Predictions resulting from the exhaustive rectangular search. (e) Predictions resulting from the exhaustive elliptical search; X indicates that an unstable shape was not found in the corresponding low-strength patch.

reduction from the theoretical maximum of $2^{51\times51}$ required by an exhaustive search. As the number of grid cells is very small in these test cases, constrained exhaustive searches are possible. If test shapes are assumed to be rectangular [e.g., *Okimura*, 1994] or elliptical [e.g., *Xie et al*., 2006], the number of shapes to test scales with the square of the number of grid cells, similarly to our search algorithm. We can thus compare our

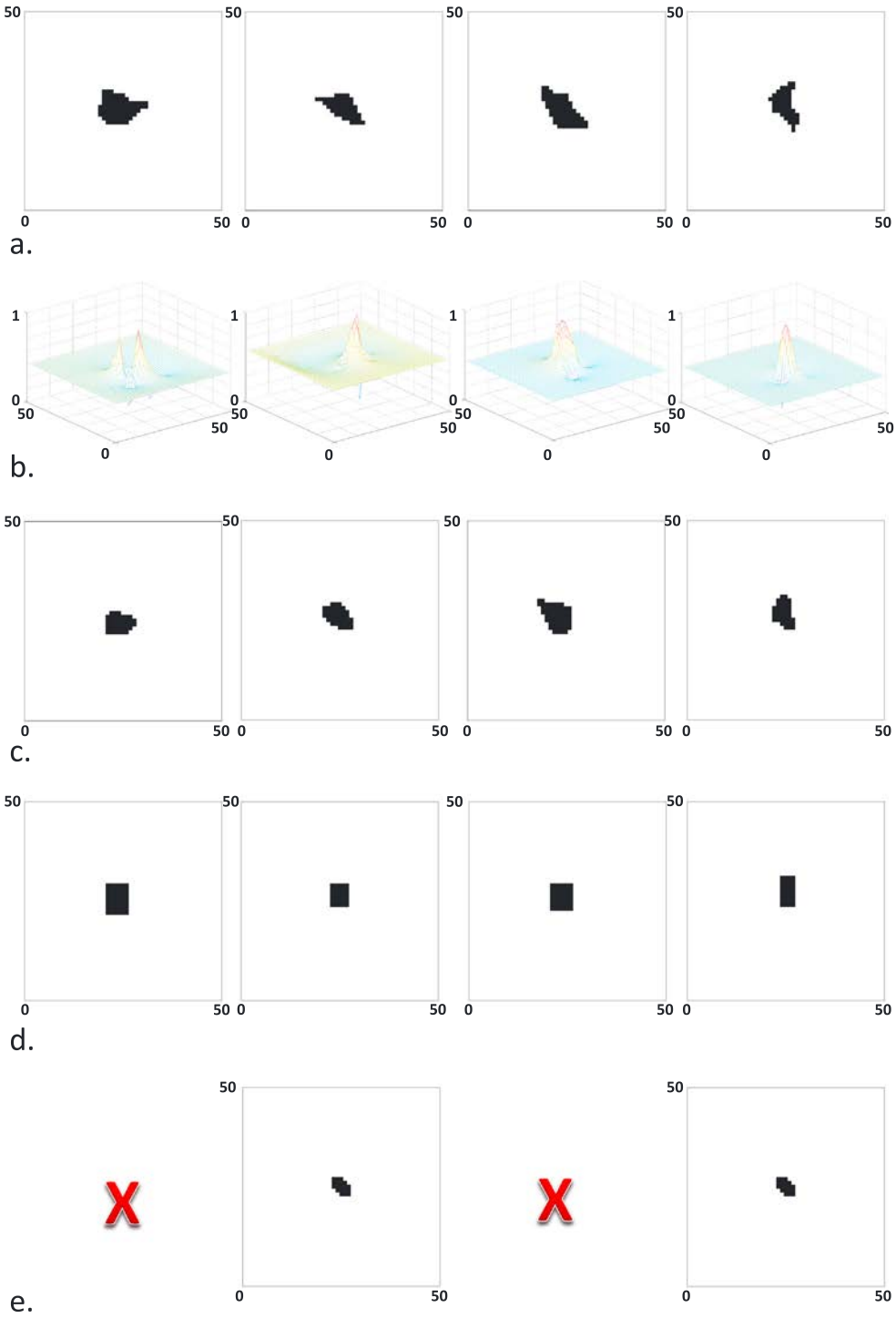

Figure 6. (a) Synthetic landscape with randomly generated irregular low root strength patches. (b) The eigenvectors of equation (12) resulting in the shapes with the lowest FS. (c) Predictions resulting from thresholding the eigenvectors in Figure 6b. (d) Predictions resulting from the exhaustive rectangular search. (e) Predictions resulting from the exhaustive elliptical search; X indicates that an unstable shape was not found in the corresponding low-strength patch.

search algorithm to an exhaustive rectangular search (with rectangle orientation constrained to be axis parallel), and to an exhaustive elliptical search (with ellipse orientation constrained to the eight principal grid directions) by applying the same slope stability model to all candidate rectangular and elliptical shapes. The least stable shapes found by these alternate search methods are shown in Figures 5e, 5d, 6e, and 6d.

Table 1. Parameters and Results for the Synthetic Landscape Tests

	Original			Search Algorithm		Rectangle Search		Ellipse Search	
Patch Type	Size (m^2)	C_l (kPa)	C_b (kPa)	Size (m^2)	Difference (|%|)	Size (m^2)	Difference (|%|)	Size (m^2)	Difference (|%|)
Square	25.00	3.97	0.14	25.00	0.00	35.00	40.00	Not found	100
Rectangle	45.00	4.97	0.17	39.00	13.33	63.00	40.00	Not found	100
Circle	49.00	4.17	0.14	46.00	6.12	49.00	0.00	49	0.00
Ellipse	87.00	5.96	0.20	62.00	28.74	77.00	11.49	87	0.00
Irregular 1	76.00	4.67	0.16	39.00	48.68	48.00	36.84	Not found	100
Irregular 2	53.00	2.78	0.09	40.00	24.53	30.00	43.40	19	64.15
Irregular 3	73.00	5.36	0.18	55.00	24.66	42.00	42.47	Not found	100
Irregular 4	51.00	2.78	0.09	38.00	25.49	32.00	37.25	19	62.75
Average (regular)	51.50	4.77	0.16	43.00	12.05	56.00	22.87	68.00	50.00
Average (irregular)	63.25	3.90	0.13	43.00	30.84	38.00	39.99	19.00	81.72
Average (all)	57.38	4.33	0.15	43.00	21.44	47.00	31.43	43.50	65.86

In all cases our procedure recovers landslides within the reduced-strength patches (Figures 5c and 6c). The predictions are always contained within the patches and have similar shapes. This is because the eigenvectors resulting from equation (12), shown in Figures 5b and 6b, have very localized peaks over these patches, and thus provide very good constraints to the predicted unstable areas. The exhaustive rectangular search also predicts landslides in every case (Figures 5d and 6d), while the exhaustive eight-direction ellipse search predicts a landslide in only half the cases (Figures 5e and 6e). The difference in size between our search algorithm predictions and the reduced-strength patches ranges from 0% (for the square) to almost 50% (for the first irregular shape), with an average difference of ~20% (Table 1). In comparison, the difference in the exhaustive rectangular and elliptical search is ~30% and 65%, respectively (Table 1). These results illustrate that our procedure is not only effective in finding unstable areas but also quite flexible with respect to their shapes. Furthermore, to the degree that it may be considered a measure of success, our procedure reproduces the shapes of the low-strength patches more faithfully than the alternate search methods tested here.

4.2. Application to Field Measurements of a Known Landslide

In this section we apply the shallow landslide prediction procedure to a watershed in southwestern Oregon, USA. To investigate the mechanisms responsible for pore water pressure development and shallow subsurface runoff generation, an experimental monitoring site, known as CB-1, was set up in a small north facing drainage basin about 15 km north of Coos Bay, Oregon [*Montgomery et al*., 1997]. Unlike many neighboring locations, CB-1 had not experienced landsliding after being clear cut in 1987. The experimental site was equipped with rain gages, piezometers, sprinklers, wells, tensiometers, weirs, lysimeters, and a weather station [*Ebel et al*., 2007]. The CB-1 catchment area is 860 m^2, and the average slope is 43°. Measured soil depths throughout the CB-1 catchment range from a few centimeters on the divergent side slopes to a little less than 2 m in the central portion of the hollow [*Schmidt*, 1999].

The largest recorded 24 h rainfall event in the Oregon Coast Range occurred in November 1996. It triggered widespread landsliding in many areas, causing extensive damage and several deaths [*Robison et al*., 1999; *Beaulieu and Olmstead*, 1999]. Over the period of 16–18 November, the CB-1 rain gages measured 225 mm of rain with a maximum daily intensity of 145 mm/d and a 48 h average intensity of 85 mm/d [*Montgomery et al*., 2009]. As recorded by the CB-1 instrumentation, a landslide occurred at the site approximately 1 h after the peak rainfall. The colluvium from the axis of the CB-1 hollow mobilized as a debris flow. The debris flow left a 156 m^2 scar that was on average 5 m wide and 26 m long, with depths ranging from 0.3 to 1 m. A detailed forensic investigation led to the conclusion that the initial landslide only comprised the upper part of the scar, with an area of approximately 58 m^2 and that the remainder of the colluvium was likely removed by surface erosion and debris flows later in the event [*Montgomery et al*., 2009]. In the upper part of the scar, the colluvial soil was excavated to the bedrock. Very few roots were found on the base of the scar, and those exposed in the head scarp and lateral margins were snapped, suggesting full engagement of their tensile strength during initial failure [*Montgomery et al*., 2009]. The CB-1 measurements possibly

represent one of the most complete data sets available for any catchment that underwent slope failure [*Ebel et al*., 2007].

The data used for this application consist of 1 m resolution grids of elevation, soil depth, and pore pressure values used in the reanalysis of the failure performed by *Montgomery et al*. [2009], without performing any reinterpolation or adjustments (Figure 7c). The fundamental difference between the *Montgomery et al*. [2009] reanalysis and the analysis presented here is that we do not prescribe the locations of the landslide boundaries, instead allowing the search algorithm to find the least stable configuration. Triaxial tests on CB-1 colluvium samples performed by *Montgomery et al*. [2009] indicate that the soil friction angle ranges from 39.5° to 41° and that soil cohesion ranges from 0 to 1.8 kPa, consistent with values previously reported for the Oregon Coast Range [e.g., *Yee and Harr*, 1977; *Schroeder and Alto*, 1983]. Based on these measurements we set the soil friction angle to 40° and assume that the soil is cohesionless with all its effective cohesion provided by roots [*Schmidt et al*., 2001]. We set the soil density ρ_s to 1600 kg/m^3, and water density ρ_w to 1000 kg/m^3 [*Montgomery et al*., 2009]. Following *Schmidt et al*. [2001], *Montgomery et al*. [2009] measured the spatial distribution of root type, root diameter, root/area ratio, and root depth along the perimeter of the scarp of the CB-1 landslide. They provide data on root cohesion as a function of depth [*Montgomery et al*., 2009, Figure 8c], and report a depth-averaged, spatially weighted apparent cohesion value of 4600 Pa around the CB-1 scar. Although numerous studies model basal and lateral root cohesion as invariant with depth [e.g., *Montgomery et al*., 2000; *Schmidt et al*., 2001; *Casadei et al*., 2003a, 2003b; *Montgomery et al*., 2009], these data suggest that the relationship between root cohesion and depth at this site is better represented as an exponential function [e.g., *Dunne*, 1991; *Benda and Dunne*, 1997; *Roering*, 2008] of the form

$$C_b = C_{r0} e^{-zj}, \tag{14}$$

where C_b (Pa) is the root cohesion at the base of a soil column of thickness z (m), C_{r0} (Pa) is a coefficient representing the maximum root cohesion value at the surface, and j (m^{-1}) is a coefficient that controls the rate at which C_b approaches zero with increasing depth z. Root cohesion can be averaged over the thickness of the soil column z (in the vertical coordinate z_c) to obtain the average lateral root cohesion C_l per unit perimeter area (Pa):

$$C_l = \frac{1}{z}\int_0^z C_{r0}{}^{-z_c j} \mathrm{d}z_c = \frac{C_{r0}}{jz}\pi\left(1 - e^{-zj}\right). \tag{15}$$

By fitting equations (14) and (15) to the *Montgomery et al*. [2009] data, we obtain values of $C_{r0} = 22$ kPa and $j = 5$ m^{-1}. While the exponential representation at this site leads to an overestimation of root strength in the top 10 to 20 cm of the soil column, it is a better fit to the data than a constant average value through the soil profile [*Montgomery et al*., 2009, Figure 8c]. Although in nature root strength will vary according to vegetation type and spacing, we assume the parameters C_{r0} and j to be spatially uniform but note that root strength varies spatially in this application because soil depth varies spatially (C_l varies inversely and C_b declines exponentially with soil depth).

We apply the search algorithm as in the previous section, assembling the force matrix **F** and Laplacian-like resistance matrix **R** (equation (7) and equations (8) and (9), respectively), and examining the eigenvectors associated with the 164 smallest eigenvalues of equation (12). Because the grid is small (56 m by 30 m), we can also apply the exhaustive rectangular and elliptical searches defined in the previous section. The continuous solution that minimizes equation (11), the first eigenvector of the linear system defined by equation (12), is shown in Figure 7a. Eigenvectors 2–5 (not shown) also define areas of instability. Thresholding these eigenvectors (see section 3.3) results in a set of 475 overlapping shapes, represented in Figure 7b. The final prediction, consisting of the shape having the lowest FS (i.e., FS_{min} method), is shown in Figure 7c. The predicted landslide has a similar size and is in a similar location to the observed landslide. They differ only in that our prediction extends approximately 2 m further upslope, resulting in an area 8 m^2 larger than the observed failure (a 13% size increase). In contrast, the predictions resulting from the FS_{max} method and from the rectangular and elliptical searches, also shown in Figure 7b, result in larger landslides (+18%, +132%, and +140%, respectively). We also compare the aspect ratio (the ratio of landslide length to width) as a proxy for shape, the median topographic index (defined as $\ln(A/(b \times \sin\theta))$), where A is drainage area (m^2), b is the cell size (m), and θ is the slope angle [*Dietrich et al*., 2001]) as a proxy for location, and the FS of the predicted and observed landslides, all listed in Table 2. The predicted aspect ratio is 1% lower than observed, indicating that

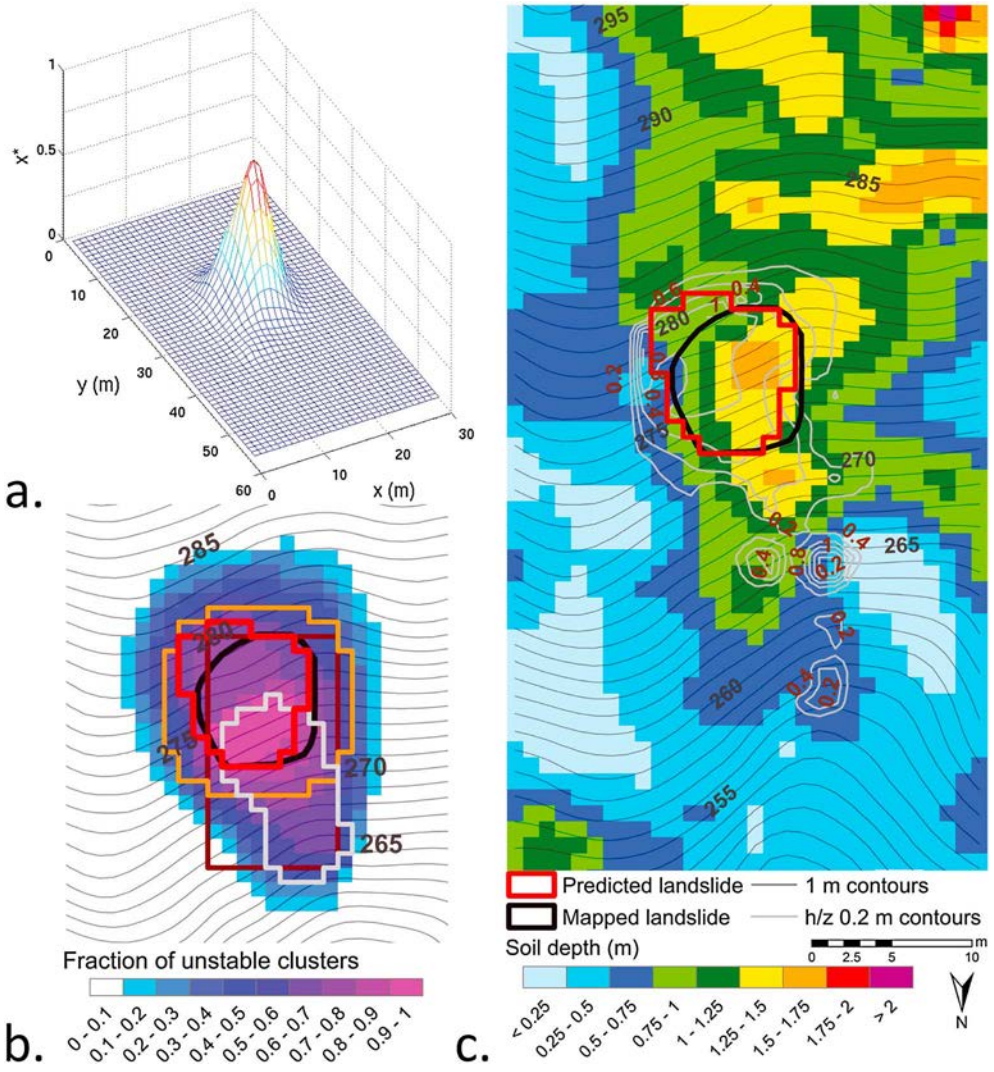

Figure 7. Application of the model to the CB-1 experimental catchment. (a) The first eigenvector **x***, which corresponds to the smallest eigenvalue of the linear system defined by equation (12). (b) The outlines of the search algorithm predictions with the minimum FS (red), the maximum FS below unity (gray); those from an exhaustive rectangular search (dark brown), and from an exhaustive elliptical search (light brown); the observed CB-1 landslide (black). The landslides are overlaid on a map showing 1 m elevation contours (thin black lines) and the fraction of the 475 predicted landslide shapes that include each grid cell. (c) The outlines of the search algorithm prediction with the minimum FS (red) and of the observed CB-1 landslide (black). The landslides are overlaid on a map showing 1 m elevation contours, 0.2 m/m saturation ratio (*h*/*z*) contours, and the soil depth grid. Data from *Montgomery et al.* [2009].

the predicted landslide is slightly rounder; the topographic index is 6% lower, indicating that the predicted landslide is slightly shifted up the hollow axis; and the FS is 1% lower, indicating that the predicted landslide is slightly more unstable (Table 2). We apply the same measures to the predictions resulting from the FS_{max} methods, as well as those resulting from the application of the exhaustive rectangular and elliptical search methods (Figure 7b). The FS_{min} method vastly outperforms the other methods in this application, based on the relative change between the observations and the predictions (Table 2).

As errors in field measurements used to determine the input parameters in this application may impact the results, we test the sensitivity of our procedure to modest changes in the input parameters by varying soil depth, pore water pressure, root strength, and soil friction angle by ±5%. While uncertainty in these parameters is difficult to assess from published data, the values reported suggest that at this site uncertainty may in fact be relatively small. Friction angle measurements for CB-1 range from 39.5° to 41°

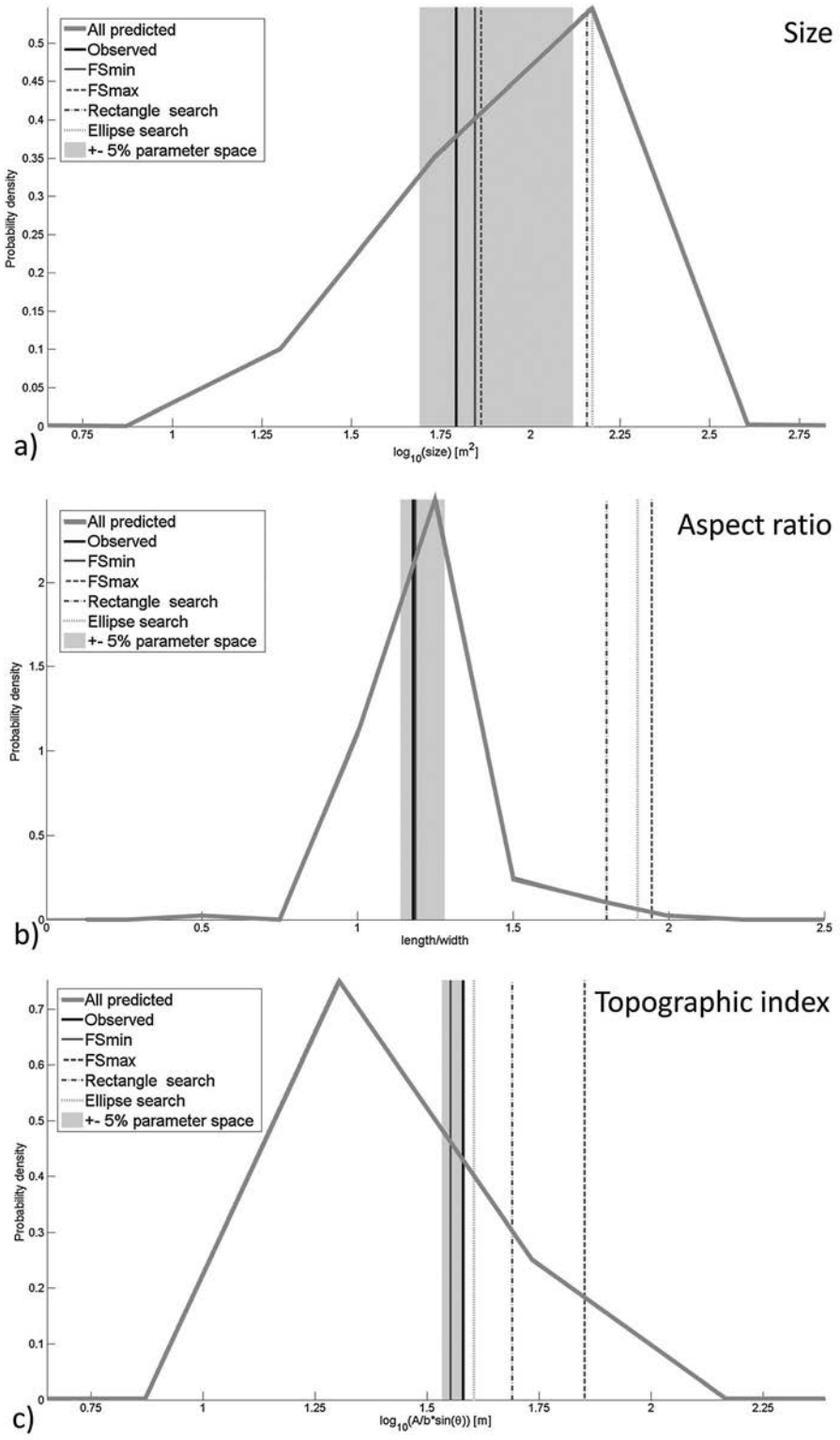

Figure 8. Distributions of (a) size, (b) aspect ratio, and (c) topographic index of all the 475 unstable grid cell combinations returned by the search algorithm. Vertical lines indicate the attributes (size, aspect ratio, and topographic index) of the observed landslide, of the least stable shape (FS_{min}) and of that with FS closest to 1 (FS_{max}), as well as the attributes resulting from the alternate search methods (rectangle search, ellipse search). Gray-shaded area shows the range of the attributes of the least stable (FS_{min}) predictions when individual parameters are varied by ±5%. Aspect ratio and topographic index are defined as length/width and $\ln(A/(b \times \sin\theta))$, respectively.

Table 2. Size, Aspect Ratio, Topographic Index, and FS of the CB-1 Observed and Predicted Landslides[a]

	CB-1 Landslide	Fs_{min} Search	Fs_{max} Search	Rectangle Search	Ellipse Search
Size (m^2)	62	70 (+13%)	73 (+18%)	144 (+132%)	149 (+140%)
Aspect ratio []	1.20	1.19 (−1%)	1.95 (+62%)	1.80 (+50%)	1 (−17%)
Topographic index (m)	38.0	35.7 (−6%)	71.2 (+87%)	49 (+29%)	40.2 (+6%)
FS []	0.88	0.89 (+1%)	0.99 (+12%)	0.85 (−3%)	0.9 (+2%)

[a]The FS of the observed landslide is computed using MD-STAB. Values in parentheses indicate the relative change between the predicted shapes and the observed.

[*Montgomery et al*., 2009]. Fitting equations (14) and (15) to data reported by *Montgomery et al*. [2009] and *Schmidt et al*. [2001] resulted in standard error for parameters C_{r0} and j of 0.5 kPa and 0.7 m^{-1}, respectively [*Milledge et al*., 2015]. Accounting for error in pore pressure and soil depth fields is more difficult since they vary spatially and are the result of interpolation. While ±5% is a small fluctuation at individual grid cells, it is a large variation when applied uniformly over the entire site. The effect of these parameter variations on the size, aspect, ratio, and topographic index of the predicted least stable landslide is shown by the shaded areas in Figure 8. While predicted aspect ratio and topographic index remain stable under modest variation in input parameters, predicted size can vary from 30% smaller to 87% larger than the observed landslide, with the largest variations resulting from changes in the pore pressures. A ±10% variation in the same parameters results in minimal fluctuations in predicted aspect ratio and topographic index, but the predicted size range becomes larger (from −53% to +94% of the observed size). Nevertheless, the resulting range in predicted least stable shape size is only a small fraction of the size distribution of all the 475 unstable areas found, which range from 7 m^2 to 400 m^2 (Figure 8). We compare the distributions of size, aspect ratio, and topographic index of all the 475 unstable combinations with the predictions resulting from the FS_{min}, FS_{max}, and the rectangular and elliptical search methods. We find that FS_{min} predictions are most similar to the observations in every case (Figure 8).

Together these results illustrate that our procedure is effective in finding unstable areas on a natural landscape with continuously varying parameters, that it can faithfully recover the size, shape, and location of an observed landslide and that the attributes of the least stable predicted shape are relatively stable under modest changes in the relevant parameters.

5. Discussion

5.1. Assumptions

One of the strengths of our approach is that it allows an efficient search (see Appendix A) of unstable patches of the landscape without assuming predefined shapes or that instability must originate from a single unstable cell [e.g., *Okimura*, 1994; *Xie et al*., 2006; *Qiu et al*., 2007; *Lehmann and Or*, 2012; *Ruette et al*., 2013]. However, to make this possible we assume that the forces can be represented as an undirected planar graph, relax the solution of the FS-based objective function from binary (equation (10)) to continuous (equation (11)), and perform an instantaneous global optimization. Each of these terms is explained below.

A planar graph implies that the search is not performed in the vertical domain, which corresponds to our assumption of the base of the failure surface at the soil-bedrock interface. An undirected graph does not allow the encoding of the force vectors but rather only their magnitude (see section 3.2). As a result the ratio of resistance over force $C(\mathbf{x})$ in equation (10) may be underestimated (i.e., the values in the matrix $\mathbf{F}$ are upper bounds), and thus the eigenvectors of equation (12) may show peaks over areas that are in fact stable. While this could lead to unnecessarily examining stable shapes, it does not affect the FS of the predictions, which is correctly calculated using vector sums (see section 3.3). However, because the eigenvectors of equation (12) are also computed using the matrix $\mathbf{F}$, their shape could be affected and some potentially unstable clusters of cells may never be examined. Testing how many clusters are missed is prohibitive since it would require an exhaustive search.

Solving the force balance instantaneously and globally is consistent with limit equilibrium approaches that assume that the potential failure mass behaves coherently as a rigid entity [*Duncan*, 1996], that the tensile strength of individual root fibers is fully mobilized, and that all roots fail simultaneously.

Failure may also occur in a progressive fashion, both at the scale of individual roots [*Schwarz et al*., 2010] and at the hillslope scale [*Petley et al*., 2005]. Alternative stability models that account for internal deformation explicitly [e.g., *Griffiths and Lane*, 1999; *Itasca, F. L. A. C*., 2000; *Borja and White*, 2010] or implicitly [e.g., *Lehmann and Or*, 2012; *Ruette et al*., 2013] can capture some of these effects. However, since the mechanics of progressive failure remain poorly understood, we opt to use a limit-equilibrium framework in which the stability problem is well defined, resulting in a suite of robust approximations [e.g., *Fellenius*, 1927; *Bishop*, 1955; *Spencer*, 1967; *Hovland*, 1977]. In its current form, the algorithm could be used to represent progressive failure with an iterative approach in which the unstable shapes are removed or redistributed, in a manner similar to *Lehmann and Or* [2012].

While the search algorithm can use any three-dimensional slope stability model within the limit-equilibrium framework, specific models make different assumptions. The model MD-STAB [*Milledge et al*., 2015] used here assumes that the central block is rigid with failure by shear on a surface parallel to the ground surface at a prescribed depth. Additionally, the model assumes (1) that failure occurs in drained conditions under steady slope-parallel groundwater flow without suction in the unsaturated zone; (2) that the soil is normally consolidated with isotropic frictional properties and its density is independent of moisture content; (3) that the cross-slope boundaries are vertical, and that earth pressure on these boundaries is in an at-rest condition; (4) that earth pressure on the upslope and downslope margins is characterized by active and passive conditions, respectively; and (5) that intercolumn shear forces do not exceed the shear strength within the unstable block.

5.2. Multiple Overlapping Predictions

Applying this search algorithm even to very small landscapes introduces a new problem: the search identifies multiple overlapping clusters of unstable cells (e.g., Figure 7), but any cell can only fail once. This problem is not specific to our search algorithm but to any search that calculates the stability of many candidate clusters (e.g., an exhaustive search). In our search algorithm overlapping predictions may arise when each edge of the contour tree is traversed (Figure 4c), when contour tree edges merge (i.e., within the same eigenvector), and when combining results from many eigenvectors or from overlapping spatial windows. Here we propose to retain the least stable shape (FS_{min}), which performed best in our tests and is most consistent with the optimization; and also test an alternate end-member method that retains the unstable shape with FS closest to unity (the FS_{max} method). Selecting which model output is most likely to occur in reality is not a well-studied problem, particularly given the levels of approximation that go into all models. In our tests we found that the FS_{min} method performed consistently well, but that while most of the 475 unstable clusters found in the CB-1 test overlap with the observed landslide (Figure 7b), the choice of which overlapping prediction to select can have an impact on predicted landslide location and size. For example, the FS_{max} method predicted a larger and more elongated landslide, located further downslope (gray line in Figure 7b). While retaining all the possible unstable configurations in a catchment-scale application is impractical, retaining the size distribution of all landslides found at each location would not require significant additional computational costs and would provide probabilistic constraints on the possible outcomes. However, sampling from this distribution is a nontrivial problem. Figure 8 shows that the observed landslide did not correspond to the modes of the size, location, and shape distributions, suggesting the need for a weighted sampling based on prior knowledge. This weighting could be based on the FS, as suggested by *Stark and Guzzetti* [2009], whereby the least stable overlapping shape would be assigned a higher probability of being selected than more stable shapes. Alternatively, as more accurate and diverse landslide data sets become available, data-driven methods could be used to make an informed choice on which of the overlapping predictions is more likely under a specific set of conditions (e.g., using observed landslide size and shape data for a particular region to constrain which landslide to pick from the modeled distribution).

The presence of many overlapping predictions does indicate that many outcomes are possible for any scenario and suggests that subtle variations in local conditions not captured by the model could determine which outcome is ultimately realized. In most natural landscapes the details of the local conditions are not known, and may be unknowable at a scale relevant to shallow landslides. This implies that data quality may place an unsurpassable constraint on our ability to forecast the exact size of a landslide at a specific location and suggests that instead calibration and testing of the model should be based on the frequency

distribution of sizes and locations of the predicted and observed landslides rather than on prediction of exact location of individual slides. Nevertheless, Figure 8 suggests that estimating field parameters (e.g., soil depth, root strength, and pore pressure) to a few percent of their actual value can lead to satisfactory approximate results, highlighting the importance of characterizing these parameters in the field as accurately as possible.

5.3. Performance

In the synthetic tests we planted landslides on a landscape, applied a search algorithm to find them, and recovered their approximate size and location, irrespective of the shape of the low-strength patch. However, in many cases all the search methods found different landslides within the same low-strength patch. In absence of any other information, each of these predictions is equally likely. While the set of possible landslides is tightly constrained to a ±1 grid cell from the boundary of the low-strength patches, making these tests nontrivial, multiple configurations of grid cells can fail implying that there is not a single true answer. This prevents a strict quantitative comparison between the search methods and the planted landslides. Nevertheless, the tests demonstrate that our search algorithm always found the landslides and that the predictions had similar size and shape to the low-strength patches.

Application of the search algorithm to a real landslide at CB-1 is a somewhat stricter test because (1) there is a ground truth (i.e., we know which of the possible landslides actually failed) and (2) all the landslide-relevant parameters vary continuously across the landscape rather than changing abruptly as in the synthetic landscape. The search algorithm applied to a real landscape at CB-1 predicted failure at the same location and of similar size to the landslide that occurred during the November 1996 storm (Figure 7c), without any parameter tuning (all parameters are defined by field measurements). The least stable predicted landslide precisely straddles the area of high pore pressure and thicker soils, while the observed head scarp intersects the pore pressure peak (Figure 7c). As a result, it has an area that is 9 m^2 larger than the observed failure, extending approximately 2 m further upslope. This is likely due to the small-scale parameter variations that exist in natural landscapes which cannot be captured even in a site as well studied as CB-1. In particular, we note that the spatial arrangement of the piezometer nests used to measure the pore pressure values (maps in *Montgomery et al.* [2009]) will inevitably result in the highest pressure values being extended upslope when a continuous field is interpolated, because no other piezometers were in the immediate vicinity. Moreover, the use of a raster framework will inevitably result in slight differences between predictions and observations that could become more significant with coarser resolutions.

Any failure at this site is likely to be centered around the maximal values of pore pressure and soil depth, so that location alone is not a definitive test. However, considering the wide distributions of the possible unstable configurations, their size, shape, and location (Figure 8), the similarity of the least stable predicted landslide to the observed is very strong. In contrast, the predictions resulting from the alternative exhaustive search methods deviate considerably from the observations (Table 2 and Figure 8).

The rectangular search finds a landslide with a FS that is lower than both that for our predicted landslide and that for the observed landslide. This is likely due to the fact that, in situations where the grid is oriented slope parallel, rectangles will minimize perimeter length for a given area. Furthermore, the calculated FS of the observed landslide is also affected by grid discretization. The reason that our search identifies a rounder landslide (despite the higher perimeter to area ratio for this shape) is very likely related to the balance between global and local components in the optimization, resulting in a smooth eigenvector surface (Figure 7a) guiding the search algorithm. This points to the fact that while the mathematical properties of Laplacian matrices are extremely useful (see overview in *Mohar* [1991]), the link between those matrices and real-world clustering problems such as ours is somewhat heuristic.

From a methodological point of view this application (searching for potential landslides) is particularly well suited to spectral clustering. The lack of a rigorous definition of what is a good measure of similarity in the input data results in a number of spectral clustering algorithm variants, and no agreement on which one is best (see review in *Von Luxburg* [2007]). In contrast, we were able to take a physically based model and use a well-defined Factor of Safety formulation to define the objective function for the spectral clustering. While we ultimately use similar linear-algebraic techniques as in other spectral clustering methods [e.g., *Shi and Malik*, 2000], we suggest that their justification is in this case much stronger.

6. Conclusion

To calculate the size and location of shallow landslides, we propose conceiving of landscapes as clusters of small cells with properties (e.g., soil depth, root strength, and pore pressure) that influence the forces driving their failure and the forces resisting failure that they exert on one another. The potential hazard and geomorphic significance of shallow landslides depend on their location and size. Commonly applied one-dimensional stability models do not include lateral effects and cannot predict landslide size. Multidimensional models must be applied to specific geometries that are typically not known a priori, and testing all possible geometries is computationally prohibitive. We present an efficient deterministic search algorithm based on spectral graph theory and couple it with a multidimensional stability model to predict discrete landslides in catchment-scale applications using gridded spatial data. The algorithm is general, assuming only that instability results when driving forces acting on a cluster of cells exceed the resisting forces on its margins and that clusters behave as rigid blocks with a failure plane at the soil-bedrock interface. When applied to a synthetic landscape with predefined regularly and irregularly shaped reduced-strength unstable patches, the algorithm recovers landslides with shape and size similar to these patches. When applied to an intensely investigated field site near Coos Bay, Oregon (CB-1), the algorithm predicts the size and location of an observed shallow landslide using field-measured physical parameters. While predictions of location and shape are robust to modest variations in input parameters, size is more sensitive, particularly to pore pressure variations. In these applications, the search algorithm identifies patches of potential instability within large areas of stable landscape. Within these patches will be many different combinations of cells with Factor of Safety less than one, suggesting that subtle variations in local conditions may determine the ultimate form and exact location at a specific site. Nonetheless, the preliminary tests presented here suggest that the search algorithm enables the predictions of shallow landslide locations, sizes, and shapes across landscapes.

Our results suggest that in natural landscapes prone to landsliding, the procedure identifies large areas of stable landscape and patches of potential instability. Within these patches, many different combinations of cells are found with FS less than unity, implying that subtle variations in local conditions determine the form and location of failure at a specific site. This implies that, in the absence of high-resolution strength and resistance data, calibration and testing of the model should be based on the frequency distribution of sizes and locations of the predicted and observed landslides rather than on prediction of exact location of individual slides. This new search algorithm enables the prediction of shallow landslide size and locations across landscapes. However, application of our method to a larger landscape will involve making choices on how to parameterize landslide-relevant spatial properties at sufficiently fine resolution. This is the subject of subsequent research presented in D. Bellugi et al. (submitted).

Appendix A: Computational Complexity and Implementation

A naïve approach that tests every combination of grid cells requires a number of operations that grows exponentially with the number of grid cells. In contrast, here we show that the number of operations required by our procedure grows quadratically with the number of grid cells and linearly with the number of eigenvectors examined. This appendix also presents implementation details, including the parallelization of the search algorithm.

Iterative algorithms for numerical computation of eigenvectors and eigenvalues, typically based on the Lanczos algorithm [*Lanczos*, 1950], proceed by multiplying the target matrix by a series of vectors and are thus dominated by matrix-vector multiplication [*Demmel et al*., 2007]. Due to the sparsity of our Laplacian-like resistance matrices which have at most five nonzero entries in each row (Figure 3c), the cost of one sparse matrix-vector multiplication scales linearly with the number of grid cells. The number of Lanczos iterations is upper bounded by the matrix dimension n, so in the worst case the number of operations required to compute k eigenvectors is kn^2. In practice the Lanczos method often converges to numerical precision after many fewer than n iterations, and thus, running time is significantly lower than the kn^2 bound.

The cost of finding connected components in a graph is linear in the number of nodes and edges [*Tarjan*, 1972], but the process is repeated for n thresholds of k eigenvectors, resulting in an upper bound of kn^2 operations. Similarly, comparing at each threshold the current and previous connected components to determine which regions overlap can be accomplished via a loop which examines each spatial cell once, also

resulting in a kn^2 upper bound on the number of operations. The a posteriori pruning of regions is also bounded by the number of regions, which is always less than n.

The use of more sophisticated algorithms and data structures that exploit the very small and controlled change in the regions after each thresholding step would further reduce the maximum number of operations to $kn\log(n)$. Nevertheless, the current kn^2 upper bound is a huge reduction from the intractable brute-force method. In particular, kn^2 operations are not prohibitive when the problem size is small (i.e., small n). Thus, the input layers are partitioned into N windows, each with a small n. In order to avoid edge effects, these windows must have significant overlap in both the x and y directions. In practice, nN is approximately 25% greater than the number of cells of the landscape grid. A typical window size should be significantly larger than the size of shallow landslides in the region of interest, in the range of 40,000 m^2 and upward (100 by 100 pixels for a 2 m grid resolution).

Noise on the surface of the eigenvectors of equation (12) $\mathbf{x}^*$ or $\mathbf{x}^{*\prime}$ (see section 3.3) could result in a very large number of insignificant regions, which impose an unnecessary computational cost. To further reduce the computational cost of the procedure, we include a maximum number of p edges in the contour trees T_c and $T_c{}'$ (see section 3.3), corresponding to the p tallest peaks in $\mathbf{x}^*$ and $\mathbf{x}^{*\prime}$. Examining both $\mathbf{x}$ and $\mathbf{x}^{*\prime}$ ensures that all the regions associated with the p most significant peaks are tested as potential landslides. The choice of p should reflect a compromise between redundancy and computational efficiency. For a typical window size of 100 by 100 pixels, $p = 50$ guarantees that the tallest 5% of all possible peaks will be examined in the worst case of the eigenvector surface consisting only of noise. In this study the window size is significantly smaller (51 by 51 and 30 by 56) so our choice $p = 50$ results in approximately 35% to 50% most significant of all possible peaks. The maximum number of regions examined is thus given by $(nm - 1)pk$, the product of the maximum number of thresholds, the maximum number of contour tree edges, and the maximum number of eigenvectors.

There are two trivial levels of parallelization that are compatible with the procedure: both the order in which windows are examined within the landscape and the order in which eigenvectors are examined within each window are irrelevant, making this an "embarrassingly parallel" problem [*Foster*, 1995]. The implementation is also well suited to modern parallel computing architecture, which typically consists of many nodes each with multiple processors and shared memory. The landscape can be partitioned across nodes that do not require shared memory, while the processing of eigenvectors, although independent, uses the same information and thus is most efficient on a shared memory platform.

The software is written in MATLAB R2014a (v. 8.3.0.532) and runs similarly under Linux, Windows, and Mac OS operating platforms. Two versions have been developed and tested: a standard serial version and a parallel implementation. The parallel version uses the MATLAB Parallel Computing toolbox and can take advantage of up to 12 processor cores with a nearly linear speedup.

Acknowledgments
This research was financially supported by grants from the United States Forest Service (09-JV-11221634-233), NASA (ROSES 09-IDS09-0049), NSF (NSF-EAR-0828047), and NERC (NER/S/A/2004/12248). All data used to evaluate the model are derived from previously published sources. Thanks to Kevin Schmidt for providing root cohesion and soil depth data and to Dave Montgomery for providing landslide dimension data. All data used in this study have been previously published, please see citations in the text.

References

Arellano, D., and T. D. Stark (2000), Importance of three-dimensional slope stability analyses in practice, *Geotech. Spec. Publ.*, 18–32, doi:10.1061/40512(289)2.

Baum, R. L., et al. (2005), Regional landslide-hazard assessment for Seattle, Washington, USA, *Landslides*, *2*(4), 266–279, doi:10.1007/s10346-005-0023-y.

Baum, R. L., W. Z. Savage, and J. W. Godt (2008), TRIGRS—A Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis, Version 2.0. US Geological Survey Open-File Report, 2008–1159.

Beaulieu, J. D., and D. L. Olmstead (1999), *Mitigating Geologic Hazards in Oregon: A Technical Reference Manual*, Dep. of Geology and Mineral Industries, State of Oregon.

Benda, L. E., and T. W. Cundy (1990), Predicting deposition of debris flows in mountain channels, *Can. Geotech. J.*, *27*, 209–417, doi:10.1139/t90-057.

Benda, L. E., and T. Dunne (1997), Stochastic forcing of sediment supply to the channel network form landsliding and debris flow, *Water Resour. Res.*, *33*, 2849–2863, doi:10.1029/97WR02387.

Bishop, A. W. (1955), The use of the slip circle in the stability analysis of slopes, *Geotechnique*, *5*(1), 7–17.

Borga, M., G. D. Fontana, C. Gregoretti, and L. Marchi (2002), Assessment of shallow landsliding by using a physically based model of hillslope stability, *Hydrol. Process.*, *16*, 2833–2851, doi:10.1002/hyp.1074.

Borja, R. I., and J. A. White (2010), Continuum deformation and stability analyses of a steep hillside slope under rainfall infiltration, *Acta Geotech.*, *5*, 1–14, doi:10.1007/s11440-009-0108-1.

Burroughs, E. R. J. (1985), Landslide hazard rating for portions of the Oregon Coast range, paper presented at Symposium on effects of forest land use on erosion and slope stability, Univ. of Hawaii, Honolulu.

Caquot, A., and J. Kerisel (1948), *Tables de Poussee et de Butee*, Gauthier-Villars, Paris.

Carr, H., J. Snoeyink, and U. Axen (2000), Computing contour trees in all dimensions: Proceedings of the eleventh annual ACM-SIAM symposium on discrete algorithms, pp. 918–926.

Casadei, M., W. E. Dietrich, and N. L. Miller (2003a), Testing a model for predicting the timing and location of shallow landslide initiation in soil mantled landscapes, *Earth Surf. Processes Landforms*, *28*, 925–950, doi:10.1002/esp.470.
Casadei, M., W. E. Dietrich, and N. L. Miller (2003b), *Controls on shallow landslide size*, in *Proceedings of the 3rd International Conference on Debris-Flow Hazards Mitigation: Mechanics, Prediction, and Assessment*, Davos, Swizerland, Millpress, Rotterdam.
Chen, R. J. (1981), *Three-Dimensional Slope Stability Analysis, Rep. FHWA/IN/JHRP-81/17*, Joint Highway Research Project, Indiana Dep. of Transportation and Purdue University, West Lafayette, Indiana, doi:10.5703/1288284314038.
Chen, W. F. (1975), *Limit Analysis and Soil Plasticity*, Elsevier Science, Amsterdam.
Chigira, M., and H. Yagi (2006), Geological and geomorphological characteristics of landslides triggered by the 2004 Mid Niigata prefecture earthquake in Japan, *Eng. Geol.*, *82*(4), 202–221, doi:10.1016/j.enggeo.2005.10.006.
Chugh, A. K., and J. D. Smart (1981), Suggestions for slope stability calculations, *Comput. Struct.*, *14*(1), 43-50.
Chung, F. R. K. (1997), *Spectral Graph Theory*, AMS Press, Providence, R. I.
Cormen, T. H., C. E. Leiserson, R. L. Rivest, and C. Stein (2001), *Introduction to algorithms (Vol. 2)*, Cambridge: MIT press.
Coulomb, C. A. (1776), Essai sur une application des regles des maximis et minimis a quelques problemes de statique relatifs a l'architecture, Memoires de l'Achademie Royale pres Divers Savants, 7.
Das, B. M. (2010), *Principles of Geotechnical Engineering*, 7th ed., PWS-KENT, Boston, Ma.
Dasgupta, S., C. H. Papadimitriou, and U. Vazirani (2006), *Algorithms*, McGraw-Hill, New York.
Demmel, J., I. Dumitriu, and O. Holtz (2007), Fast linear algebra is stable, *Numer. Math.*, *108*, 59–91, doi:10.1007/s00211-007-0114-x.
Dhakal, A. S., and R. C. Sidle (2003), Long-term modelling of landslides for different forest management practices, *Earth Surf. Processes Landforms*, *28*(8), 853-868, doi: 10.1002/esp.499.
Diestel, R. (2005), *Graph Theory, Volume 173 of Graduate Texts in Mathematics*, 3rd ed., Springer, Heidelberg.
Dietrich, W. E., and T. Dunne (1978), Sediment budget for a small catchment in mountainous terrain, *Z. Geomorphol. Suppl.*, *29*, 191–206.
Dietrich, W. E., R. Reiss, M. Hsu, and D. R. Montgomery (1995), A process-based model for colluvial soil depth and shallow landsliding using digital elevation data, *Hydrol. Processes*, *9*, 383–400, doi:10.1002/hyp.3360090311.
Dietrich W. E., D. Bellugi, and R. Real de Asua (2001), Validation of the shallow landslide model, SHALSTAB, for forest management, in *Land use and Watersheds: Human influence on Hydrology and Geomorphology in Urban and Forest areas*, Water Sci. Appl., vol. 2, edited by M. S. Wigmosta and S. J. Burges, pp. 195–227, AGU, Washington, D. C., doi:10.1029/WS002p0195.
Dietrich, W. E., D. Bellugi, A. M. Heimsath, J. J. Roering, L. Sklar, and J. D. Stock (2003), Geomorphic transport laws for predicting landscape form and dynamics, in *Prediction in Geomorphology*, *Monogr. Ser.*, vol. 135, edited by P. Wilcock and R. Iverson, pp. 103-132, AGU, Washington, D. C., doi:10.1029/135GM09.
Dietrich, W. E., J. McKean, D. Bellugi, and J. T. Perron (2008), The prediction of shallow landslide location and size using a multidimensional landslide analysis in a digital terrain model, Proceedings of the Fourth International Conference on Debris-Flow Hazards Mitigation.
Duncan, J. M. (1996), State of the art: Limit equilibrium and finite-element analysis of slopes, *J. Geotech. Eng.-ASCE*, *122*(7), 577–596.
Dunne, T. (1991), Stochastic aspects of the relations between climate, hydrology and landform evolution, *Trans. Jpn. Geomorphol. Union*, *12*, 1–24.
Ebel, B. A., K. Loague, W. E. Dietrich, D. R. Montgomery, R. Torres, S. P. Anderson, and T. W. Giambelluca (2007), Near-surface hydrologic response for a steep, unchanneled catchment near Coos Bay, Oregon: 1. Sprinkling experiments, *Am. J. Sci.*, *307*, 678–708.
Fannin, R. J., and M. P. Wise (2001), An empirical-statistical model for debris flow travel distance, *Can. Geotech. J.*, *38*, 982–994, doi:10.1139/t01-030.
Fellenius, W. (1927), *Erdstatische Berechnungen mit Reibung und Kohasion* [in German], Ernst, Berlin.
Foster, I. (1995), *Designing and Building Parallel Programs*, pp. 83–135, Addison Wesley, Boston.
Freeman, H., and S. P. Morse (1967), On searching a contour map for a given terrain elevation profile, *J. Franklin Inst.*, *284*(1), 1–25, doi:10.1016/0016-0032(67)90568-6.
Gabet, E. J., and T. Dunne (2002), Landslides on coastal sage-scrub and grassland hillslopes in a severe El Nino winter: The effects of vegetation conversion on sediment delivery, *Geol. Soc. Am. Bull.*, *114*(8), 983–990.
Glade, T. (1998), Establishing the frequency and magnitude of landslide-triggering rainstorm events in New Zealand, *Environ. Geol.*, *35*(2–3), 160–174.
Griffiths, D. V., and P. A. Lane (1999), Slope stability analysis by finite elements, *Geotechnique*, *49*(3), 387–403, doi:10.1680/geot.1999.49.3.387.
Gross, J. L., and J. Yellen (2005), *Graph Theory and Its Applications*, 2nd ed., CRC Press, Boca Raton, Fla.
Guimaraes, R. F., D. R. Montgomery, H. M. Greenberg, N. F. Fernandes, R. A. Trancoso Gomes, and O. A. de Carvalho Júnior (2003), Parameterization of soil properties for a model of topographic controls on shallow landsliding: Application to Rio de Janeiro, *Eng. Geol.*, *69*(1), 99–108, doi:10.1016/S0013-7952(02)00263-6.
Haralick, R. M., and L. G. Shapiro (1992), *Computer and Robot Vision*, Addison-Wesley, Reading, Mass.
Horn, R. A., and C. A. Johnson (1985), *Matrix Analysis*, pp. 176–180 , Cambridge Univ. Press, London.
Hovland, H. J. (1977), Three-dimensional slope stability analysis method, *ASCE J. Geotech. Eng. Div.*, *103*(GT9), 971–986.
Hungr, O., S. McDougall, M. Wise, and M. Cullen (2008), Magnitude-frequency relationships of debris flows and debris avalanches in relation to slope relief, *Geomorphology*, *96*(3–4), 355–365, doi:10.1016/j.geomorph.2007.03.020.
Iida, T. (2004), Theoretical research on the relationship between return period of rainfall and shallow landslides, *Hydrol. Process.*, *18*, 739–756, doi:10.1002/hyp.1264.
Itasca, F. L. A. C. (2000), Fast Lagrangian analysis of continua, Itasca Consulting Group Inc., Minneapolis, Minn.
Jaboyedoff, M., T. Oppikofer, A. Abellán, M. H. Derron, A. Loye, R. Metzger, and A. Pedrazzini (2012), Use of LIDAR in landslide investigations: A review, *Nat. Hazards*, *61*(1), 5–28, doi:10.1007/s11069-010-9634-2.
Jacky, J. (1944), The coefficient of earth pressure at rest, *J Soc Hungarian Architects Eng*, *7*, 355–358.
Kaufman, L., and P. J. Rousseeuw (2005), *Finding Groups in Data: An Introduction to Cluster Analysis*, pp. 342, John Wiley, Hoboken, N. J.
Lambe, T. W., and R. V. Whitman (1969), *Soil Mechanics*, John Wiley, New York.
Lanczos, C. (1950), An iteration method for the solution of eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Stand.*, *45*, 255.
Lanni, C., M. Borga, R. Rigon, and P. Tarolli (2012), Modelling shallow landslide susceptibility by means of a subsurface flow path connectivity index and estimates of soil depth spatial distribution, *Hydrol. Earth Syst. Sci.*, *16*(11), 3959–3971, doi:10.5194/hess-16-3959-2012.
Lehmann, P., and D. Or (2012), Hydromechanical triggering of landslides: From progressive local failures to mass release, *Water Resour. Res.*, *48*, W03535, doi:10.1029/2011WR010947.
Lehre, A. K. (1982), Sediment mobilization and production from a small mountain catchment: Lone Tree Creek, Marin County, California, PhD dissertation, Univ. of California, Berkeley.

McKean, J., and J. Roering (2004), Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry, *Geomorphology*, *57*(3), 331–351, doi:10.1016/S0169-555X(03)00164-8.
Milledge, D. G., D. Bellugi, J. A. McKean, A. L. Densmore, and W. E. Dietrich (2015), A multi-dimensional stability model for predicting shallow landslide size and shape across landscapes, *J. Geophys. Res. Earth Surf.*, doi:10.1002/2014JF003135.
Mohar, B. (1991), The Laplacian spectrum of graphs, in *Graph Theory, Combinatorics, and Applications, Kalamazoo, MI, 1988*, vol. 2, pp. 871–898, Wiley, New York.
Montgomery, D. R., and W. E. Dietrich (1994), A physically-based model for topographic control on shallow landsliding, *Water Resour. Res.*, *30*, 1153–1171, doi:10.1029/93WR02979.
Montgomery, D. R., W. E. Dietrich, R. Torres, S. P. Anderson, J. T. Heffner, and K. Loague (1997), Hydrologic response of a steep unchanneled valley to natural and applied rainfall, *Water Resour. Res.*, *33*, 91–109.
Montgomery, D. R., K. M. Schmidt, W. E. Dietrich, and H. M. Greenberg (2000), Forest clearing and regional landsliding in the Pacific Northwest, *Geology*, *28*, 311–314, doi:10.1130/0091-7613(2000)28<311:FCARL>2.0.CO;2.
Montgomery, D., K. M. Schmidt, W. E. Dietrich, and J. McKean (2009), Instrumental record of debris flow initiation during natural rainfall: Implications for modeling slope stability, *J. Geophys. Res.*, *114*, F01031, doi:10.1029/2008JF001078.
Okimura, T. (1994), Prediction of the shape of a shallow failure on a mountain slope: The three-dimensional multi-planar sliding surface method, *Geomorphology*, *9*, 223–233, doi:10.1016/0169-555X(94)90064-7.
O'loughlin, C. L., and A. J. Pearce (1976), Influence of Cenozoic geology on mass movement and sediment yield response to forest removal, North Westland, New Zealand, *Bull. Int. Assoc. Eng. Geol.*, *13*(1), 41–46.
Pack R. T., D. G. Tarboton, and C. N. Goodwin (1998), Terrain stability mapping with SINMAP, technical description and users guide for version 1.00, *Rep. 4114-0*, Terratech Consulting Ltd, Salmon Arm, Canada.
Papadimitriou, C. H., and K. Steiglitz (1998), "6.1 The Max-Flow, Min-Cut Theorem." Combinatorial optimization: Algorithms and complexity, Dover, pp. 120–128.
Petley, D. N., T. Higuchi, D. J. Petley, M. H. Bulmer, and J. Carey (2005), Development of progressive landslide failure in cohesive materials, *Geology*, *33*, 201–204, doi:10.1130/G21147.1.
Qiu, C., T. Esaki, M. Xie, Y. Mitani, and C. Wang (2007), Spatio-temporal estimation of shallow landslide hazard triggered by rainfall using a three-dimensional model, *Environ. Geol.*, *52*(8), 1569–1579, doi:10.1007/s00254-006-0601-x.
Robidoux, N., P. Stelldinger, and J. Cupitt (2011), Simple random generation of smooth connected irregular shapes for cognitive studies, in *Proceedings of The Fourth International C* Conference on Computer Science and Software Engineering*, pp. 17-24, ACM, Montreal, Canada, doi:10.1145/1992896.1992899.
Robison, E. G., K. Mills, J. Paul, L. Dent, and A. Skaugset (1999), Oregon Department of Forestry storm impacts and landslides of 1996: Final report, Oregon Department of Forestry, Forest Practices Technical Report 4.
Roering, J. J. (2008), How well can hillslope evolution models "explain" topography? Simulating soil transport and production with high resolution topographic data, *Geol. Soc. Am. Bull.*, *120*, 1248–1262, doi:10.1130/B26283.1.
Rogers, N. W., and M. J. Selby (1980), Mechanisms of shallow translational landsliding during summer rainstorms: North Island, New Zealand, *Geogr. Ann. Ser. A. Phys. Geogr.*, *62*, 11–21.
Rosso, R., M. C. Rulli, and G. Vannucchi (2006), A physically based model for the hydrologic control on shallow landsliding, *Water Resour. Res.*, *42*, W06410, doi:10.1029/2005WR004369.
Ruette, J., P. Lehmann, and D. Or (2013), Rainfall-triggered shallow landslides at catchment scale: Threshold mechanics-based modeling for abruptness and localization, *Water Resour. Res.*, *49*(10), 6266–6285, doi:10.1002/wrcr.20418.
Schmidt, K. M. (1999), Root strength, colluvial soil depth, and colluvial transport on landslide-prone hillslopes, PhD thesis, Dep. of Earth and Space Sciences, Univ. of Wash., Seattle.
Schmidt, K. M., J. J. Roering, J. D. Stock, W. E. Dietrich, D. R. Montgomery, and T. Schaub (2001), The variability of root cohesion as an influence on shallow landslide susceptibility in the Oregon Coast Range, *Can. Geotech. J.*, *38*(1), 995–1024, doi:10.1139/t01-031.
Schroeder, W. L., and J. V. Alto (1983), Soil properties for slope stability analysis; Oregon and Washington Coastal Mountains, *For. Sci.*, *29*, 823–833.
Schwarz, M., P. Lehmann, and D. Or (2010), Quantifying lateral root reinforcement in steep slopes—From a bundle of roots to tree stand, *Earth Surf. Processes Landforms*, *35*, 354–367, doi:10.1002/esp.1927.
Shi, J., and J. Malik (2000), Normalized cuts and image segmentation, *IEEE Trans. Pattern Anal. Machine Intelligence*, *22*(8), 888–905, doi:10.1109/34.868688.
Sidle, R. C., and H. Ochiai (2006), *Landslides: Processes, Prediction, and Land Use*, *Water Resour. Monogr.*, vol. 18, AGU, Washington, D. C.
Simoni, S., F. Zanotti, G. Bertoldi, and R. Rigon (2008), Modelling the probability of occurrence of shallow landslides and channelized debris flows using GEOtop-FS, *Hydrol. Processes*, *22*(4), 532–545, doi:10.1002/hyp.6886.
Soubra, A. H., and B. Macuh (2002), Active and passive earth pressure coefficients by a kinematical approach, *Proc. ICE-Geotech. Eng.*, *155*(2), 119–131.
Spencer, E. (1967), A method of analysis of the stability of embankments assuming parallel interslice forces, *Geotechnique*, *17*(1), 11–26, doi:10.1680/geot.1967.17.1.11.
Stark, T. D., and H. T. Eid (1998), Performance of three-dimensional slope stability methods in practice, *J. Geotech. Geoenviron. Eng.*, *124*(11), 1049–1060.
Stark, C. P., and F. Guzzetti (2009), Landslide rupture and the probability distribution of mobilized debris volumes, *J. Geophys. Res.*, *114*, F00A02, doi:10.1029/2008JF001008.
Stock, J. D., and W. E. Dietrich (2006), Erosion of steepland valleys by debris flows, *Geol. Soc. Am. Bull.*, *118*(9–10), 1125–1148.
Tarjan, R. E. (1972), Depth first search and linear graph algorithms, *SIAM J. Comput.*, *1*(2), 146–160, doi:10.1137/0201010.
Tarolli, P., M. Borga, and G. D. Fontana (2008), Analysing the influence of upslope bedrock outcrops on shallow landsliding, *Geomorphology*, *93*(3), 186–200, doi:10.1016/j.geomorph.2007.02.017.
Taylor, D. W. (1948), Fundamentals of soil mechanics, *Soil Sci.*, *66*(2), 161.
Von Luxburg, U. (2007), A tutorial on spectral clustering, *Stat. Comput.*, *17*(4), 395–416, doi:10.1007/s11222-007-9033-z.
Wu, T. H., W. R. McKinnel, and D. N. Swanston (1979), Strength of tree roots and landslides on Prince of Wales Island, Alaska, *Can. Geotech. J.*, *16*, 19–33, doi:10.1139/t79-003.
Wu, W., and R. C. Sidle (1995), A distributed slope stability model for steep forested hillslopes, *Water Resour. Res.*, *3*, 2097–2110, doi:10.1029/95WR01136.
Xie, M., T. Esaki, C. Qiu, and C. Wang (2006), Geographical information system-based computational implementation and application of spatial three-dimensional slope stability analysis, *Comput. Geotech.*, *33*, 260–274, doi:10.1016/j.compgeo.2006.07.003.
Yee, C. S., and D. R. Harr (1977), Influence of soil aggregation on slope stability in the Oregon Coast Range, *Environ. Geol.*, *1*, 367–377.