Assessing the Impact of Changes in Extreme Precipitation Events on Shallow Landslide Abundance, Location, and Size

Dino Bellugi[1] [dinob@mit.edu], Catherine Slesnick[2], Erin Leidy[2], Natasha Markuzon[2], Paul O'Gorman[1], J. Taylor Perron[1], John Regan[2], Adam Schlosser[1], John West[2]

[1]Department of Earth, Atmospheric, and Planetary Sciences, MIT, Cambridge, MA

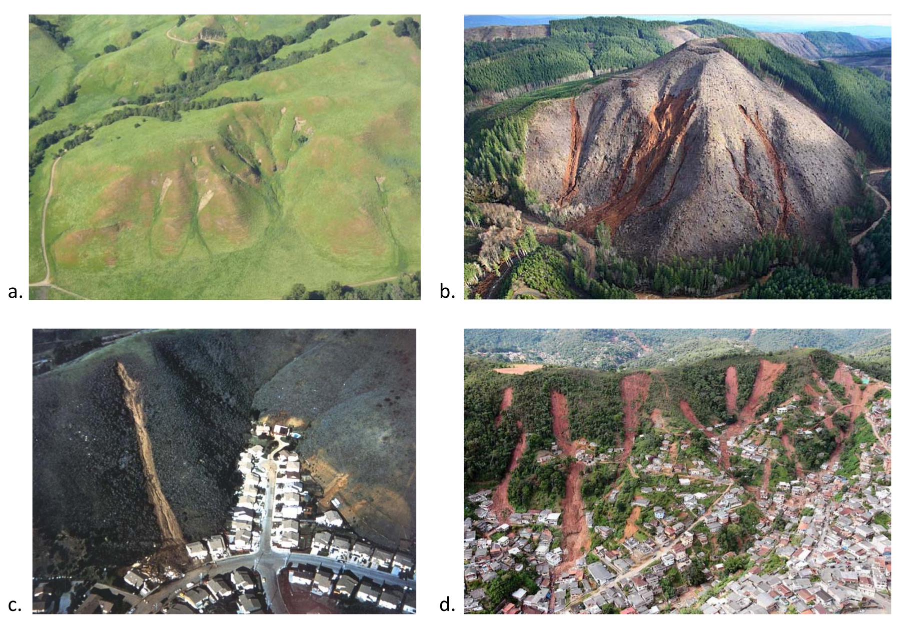
[2]Charles Stark Draper Laboratory, Cambridge, MA

Abstract

Shallow landslides are a widespread phenomenon in the United States and the world. Often triggered by extreme precipitation events, they can be the primary sources of debris flows, and are generally a threatening source of hazards, causing loss of life, destruction of property and infrastructure, and affecting communities all across the nation. It is crucial to accurately assess such hazards, particularly in light of expected climate and land use changes. The overall goal of this NASA-funded project is to assess how climate change will impact extreme precipitation and landslide hazards, and what risks those events will pose for natural and human systems in the future.

Here we explore the landsliding response of a prototype landscape located in the Oregon Coast Range (OCR) to hypothetical changes in intensity, duration, and frequency of extreme rainfall events. We adopt a mechanistic landslide prediction procedure which couples a three-dimensional slope stability model with an efficient search algorithm to predict discrete shallow landslides. We use a landslide inventory collected by repeat field mapping over a 10-year period in an area with constraints on soil, vegetation, hydrological, and rainfall characteristics. In hind-cast mode, the procedure reproduces the distribution of sizes and locations of the landslide inventory under a suite of rainfall and moisture characteristics representative of the observation period. We use projections of precipitation extremes under different climate change scenarios to generate landslide forecasts and explore the sensitivity of landslide abundance, size and location to the intensity, duration, and frequency of rainfall events, as well as to antecedent moisture conditions, resulting from the different scenarios.

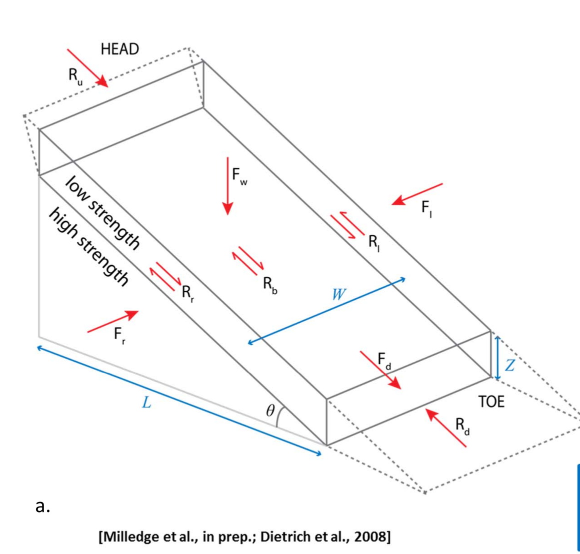
We also present progress in the development of a data-driven approach to understanding landslide activity and the response to changes in extreme precipitation in an evolving climate. Resultant models forecast landslides based on a combination of remote sensing data and historical surface observations including weather patterns, landcover and lithology, and topographic attributes. We present spatially explicit results from the application of a non-linear classifier (a support vector machine constructed using topographic, soil, and vegetation attributes) to the OCR dataset and compare the results to the existing landslide inventory. Finally, we address the need for eco-atmo-geo-hydrological models to capture the linkages between climate, vegetation, and the landscape.



Examples of size, location, and impacts of shallow landslides: a) grasslands near Briones, CA; b) de-forested slopes near Chehalis, WA, 2007, $57 million in property damage, no drinking water for months; c) Pacifica, CA, 1982, 3 fatalities ($66 million in damage and 25 fatalities over the San Francisco Bay Area); d) Brazil, Rio de Janeiro region, 2011, 1000+ fatalities.

Shallow landslide prediction procedure

We adopt and test a novel procedure which couples a three-dimensional slope stability model that captures the basic physics of shallow landsliding, with a new and efficient deterministic search algorithm that can predict discrete shallow landslides. The slope stability model is fully mechanistic, and considers resistances acting on all sides of a discretized slope element. It is applicable to gridded data, is parsimonious in its parameterization, and thus is easily applicable to the regional scale. Sub-models are also defined to produce spatial data of pore water pressure, soil depth, and root strength. The search algorithm is based on spectral graph theory and is fully deterministic. The landscape and all the landslide-relevant spatial attributes are represented as a weighted undirected graph. Spectral clustering is then used to aggregate grid cells which together minimize a factor of safety for an exponential reduction in complexity.



Frictional and cohesive forces on 5 boundaries

Limit-equilibrium analysis

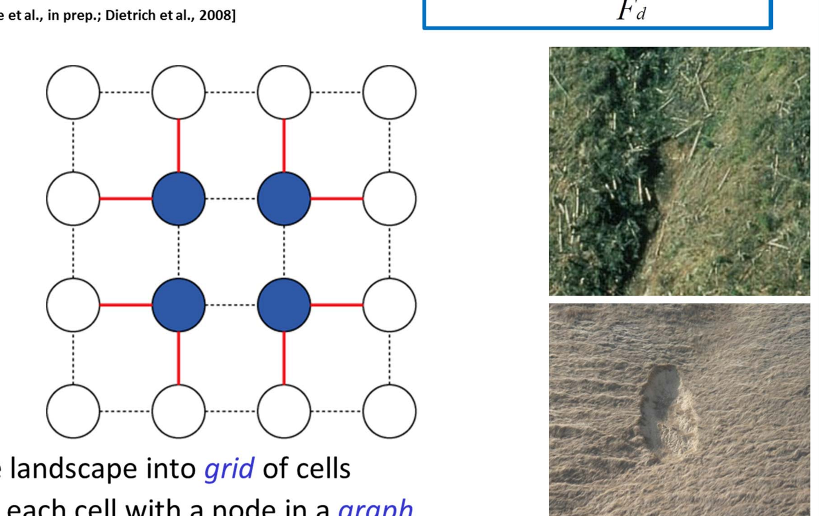
Spatial Data:
- Elevation
- Soil depth
- Pore water pressure
- Root strength field

Material Properties:
- Friction angle
- Bulk density

Implicit assumption: Rigid block

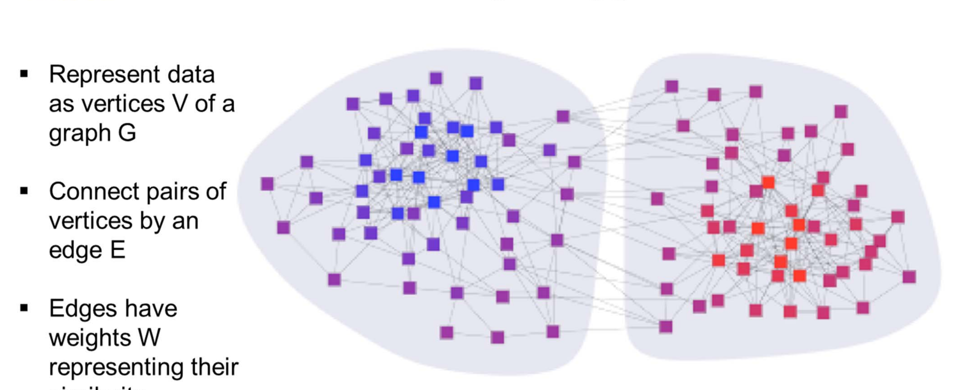
Explicit assumption: Failure plane

$$FS = \frac{R_u + R_d + R_l + R_r + R_b}{F_d}$$



b.

- Discretize landscape into *grid* of cells
- Associate each cell with a node in a *graph*
- *Nodes:* landscape cells annotated by driving forces
- *Edges:* resistive forces between neighboring grid cells



- Represent data as vertices V of a graph G
- Connect pairs of vertices by an edge E
- Edges have weights W representing their similarity
- The graph construction depends on the application

Loss function: inspired by *Normalized Cut*

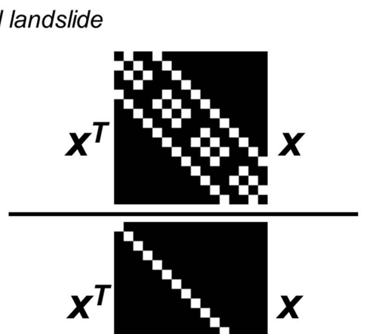
$$nCut(A,B) = \frac{cut(A,B)}{\sum_{i \in A, j} W_{ij}} + \frac{cut(B,A)}{\sum_{i \in B, j} W_{ij}}$$

c.

R = *Resistance matrix*
F = *Force matrix*
x = *Indicator vector for hypothesized landslide*

Factor of Safety:

$$FS = \frac{x^T R x}{x^T F x}$$



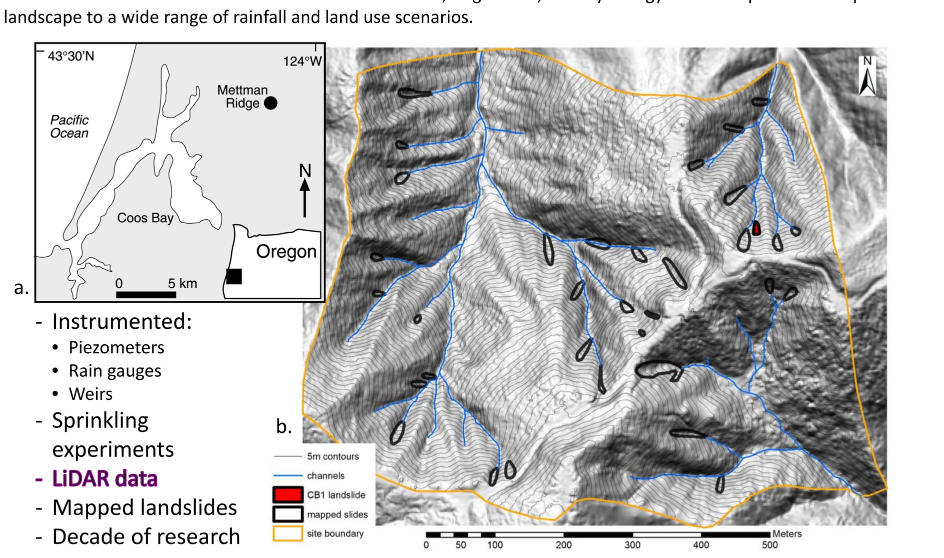
- We want to find the landslide indicators x which have a factor of safety < 1
- Extremely well studied in *numerical linear algebra*

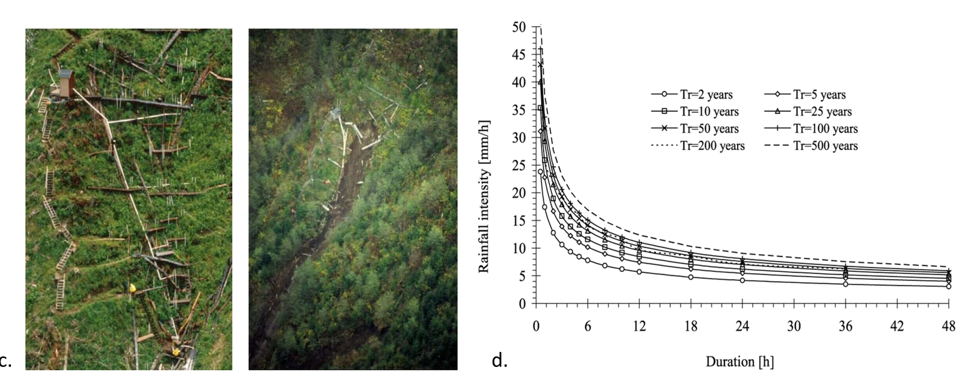
d.

The mechanistic, deterministic landslide prediction procedure: a) three-dimensional slope stability model; b) graph representation; c) graph partitioning example; d) matrix-based factor of safety calculation and optimization.

Study area

The Mettman Ridge test site is in a landslide-prone area in the Oregon Coast Range, near Coos Bay, OR. It was selected because of two unique datasets deriving from over a decade of research at this site. An instrumental record of a rainfall-triggered shallow landslide that occurred in a small catchment allowed testing the slope stability model and the search algorithm using field-measured physical parameters such as hydrological conditions, soil depth, and root strength. Across a larger area repeat field mapping provides an inventory of all the shallow landslides that occurred over a 10-year period. Also during this 10-year period intensive research was conducted in the area, providing detailed information on soil, vegetation, hydrological, and rainfall characteristics. This dataset thus presents a unique opportunity to apply all the sub-models which estimate the local characteristics of soil, vegetation, and hydrology and to explore the response of the landscape to a wide range of rainfall and land use scenarios.



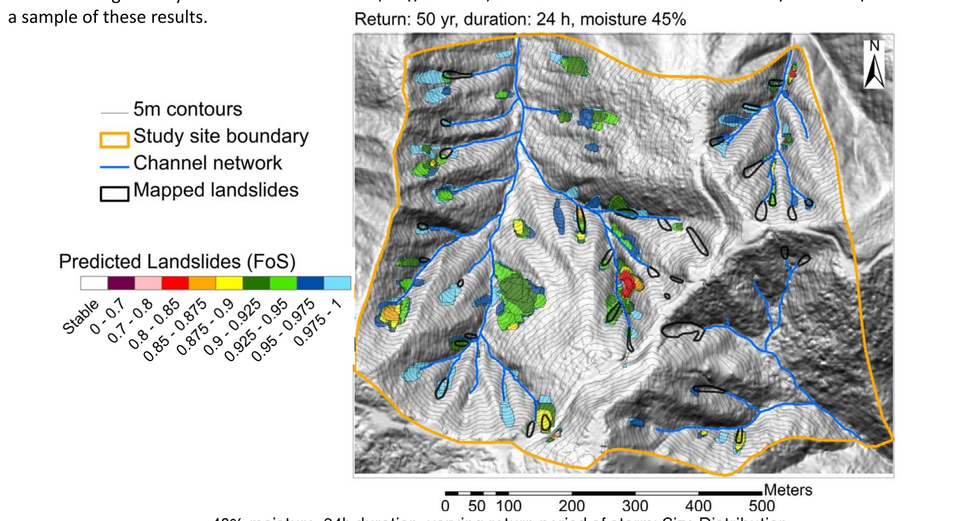


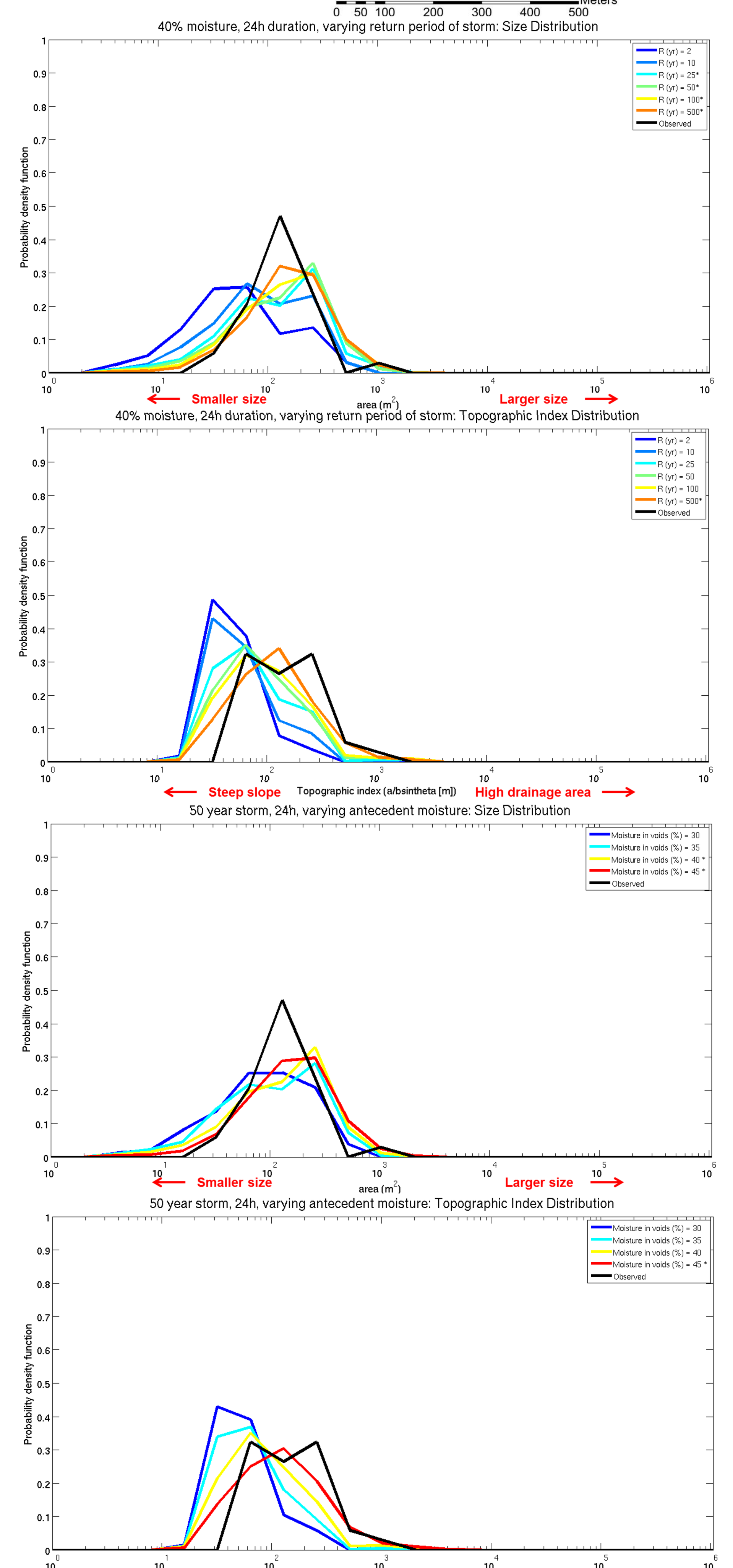
The test area in the Oregon Coast Range: a) location of the Metman Ridge; b) landslide inventory from ten years of observations mapped onto high-resolution Lidar-derived topographic data; c) the instrumented CB-1 experimental first order catchment before and after the November 1996 storm, which triggered a shallow landslide that destroyed the site; d) Intensity duration frequency (IDF) curves, derived from historical rainfall data from the nearby Allegany, OR, rain gauge.

Results

We sample the rainfall intensity duration frequency (IDF) curves and run the landslide prediction procedure with current soil and vegetation conditions. We then plot the landslide size distributions and the topographic index distributions for varying storm durations, return periods, and antecedent moisture conditions.

Results indicate that increasing the duration, the return period, or the antecedent moisture conditions, causes landslides to generally become more abundant, larger in size, and to move further down the valley axis. We present here a sample of these results.





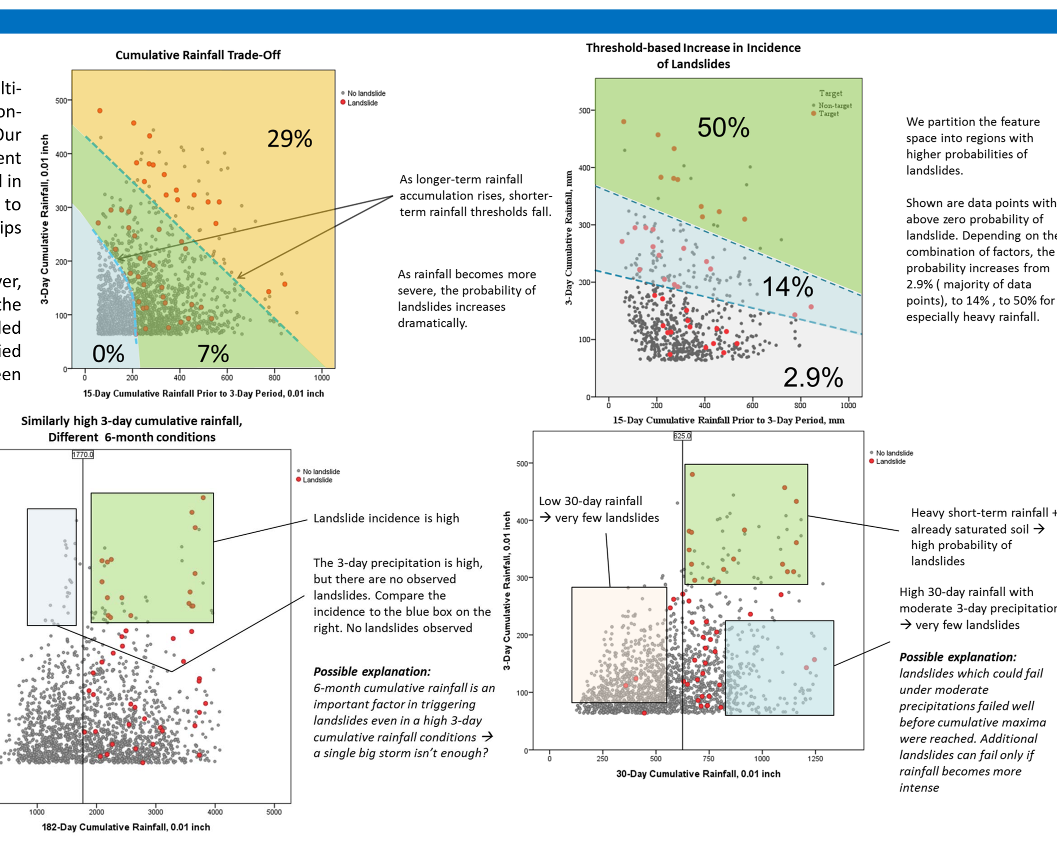
* Denotes that predicted distributions are statistically indistinguishable from the observations according to 2-sample K-S test

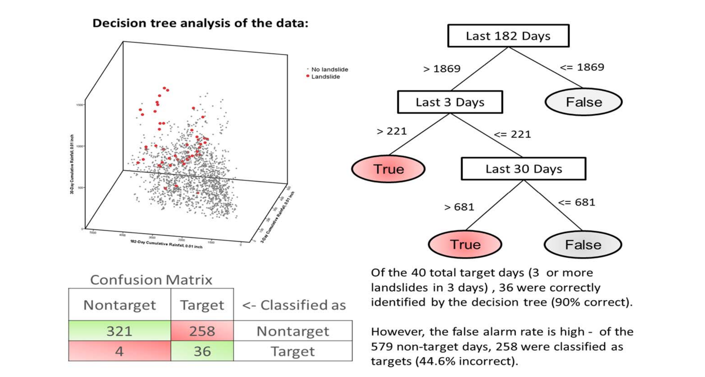
Data-driven models

We are developing data driven models aimed at predicting landslide activity. The models learn multi-dimensional weather and geophysical patterns associated with historical landslides and estimate location-dependent probabilities for landslides under current or future weather and geophysical conditions. Our approach uses machine learning algorithms capable of determining non-linear associations between dependent variables and landslide occurrence without requiring detailed knowledge of geomorphology. Our primary goal in this phase of the project is to evaluate the predictive capabilities of data mining models in application to landslide activity, and to analyze if the approach will discover previously unknown variables and/or relationships important to landslide occurrence, frequency or severity.

The models include remote sensing and ground-based data, including rainfall, soils and land-cover, topographic (e.g. slope, elevation, and drainage area) information as well as urbanization data. In addition to the Coos Bay, OR dataset, the historical landslide dataset we used to build our preliminary models was compiled from City of Seattle landslide files, United States Geological Survey reports, newspaper articles, and a verified subset of the Seattle Landslide Database that consists of all reported landslides within Seattle, WA, between 1948 and 1999. Most of the landslides analyzed to-date are shallow.

Using statistical analysis and unsupervised clustering methods we have thus far identified subsets of weather conditions that lead to a significantly higher landslide probability in the Seattle region, and have developed statistically predictive models for individual storms.





Confusion Matrix

Nontarget	Target	<- Classified as
321	258	Nontarget
4	36	Target

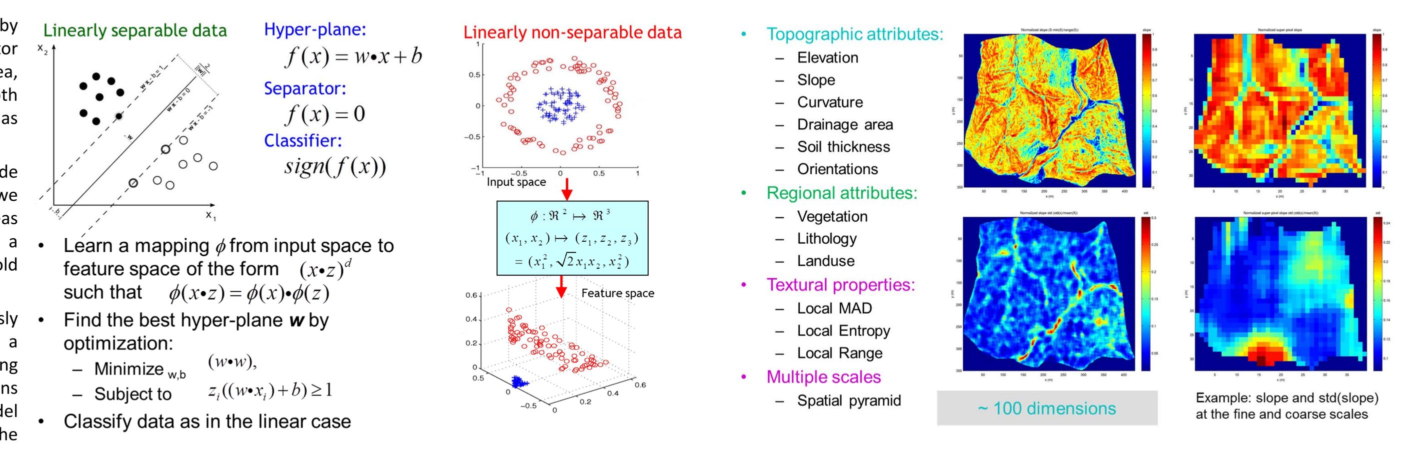
Of the 40 total target days (3 or more landslides in 3 days), 36 were correctly identified by the decision tree (90% correct).

However, the false alarm rate is high - of the 579 non-target days, 258 were classified as targets (44.6% incorrect).

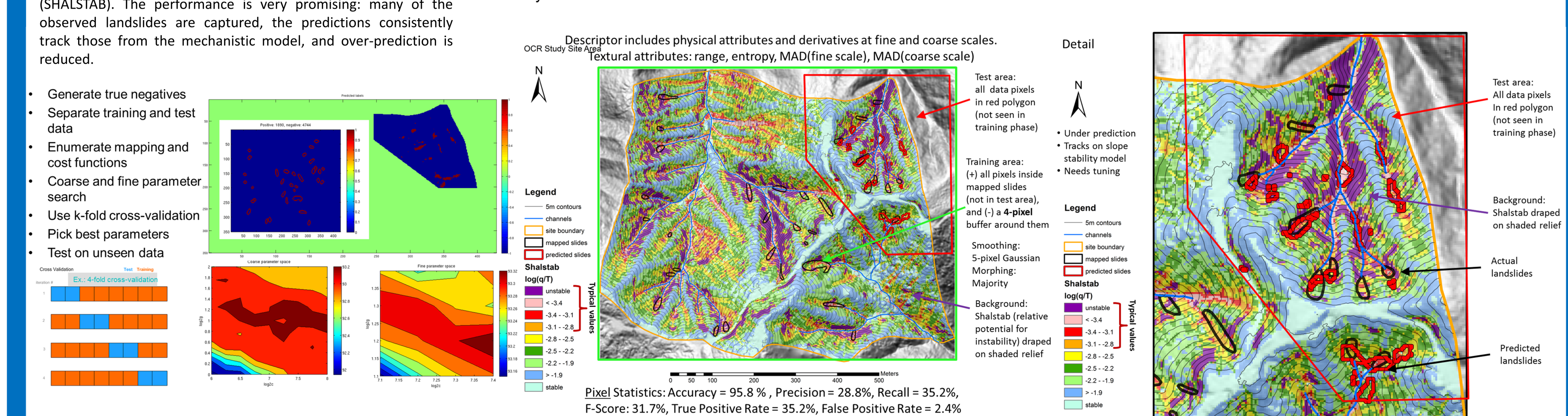
We implement a spatial semi-supervised nonlinear classifier by using a Support Vector Machine (SVM). The feature vector consists of topographic attributes such as slope, drainage area, curvature, topographic index, soil depth, root strength, at both fine and course scales. Also included are textural attributes such as local entropy, standard deviation, and range.

We train the classifier on a subset of the Coos Bay, OR, landslide dataset. Because of the limited temporal span of observations, we define the observed landslides as positive examples and areas neighboring these landslides as negative examples. We search a parameter space for a radial-basis function, and use a 10-fold cross-validation scheme.

The classifier is then applied to the remaining (previously unseen) landscape where each pixel is then assigned to be a landslide or a non-landslide cell. We present the results overlaying a map showing the observed landslides, as well as the predictions from a common mechanistic relative landslide potential model (SHALSTAB). The performance is very promising: many of the observed landslides are captured, the predictions consistently track those from the mechanistic model, and over-prediction is reduced.

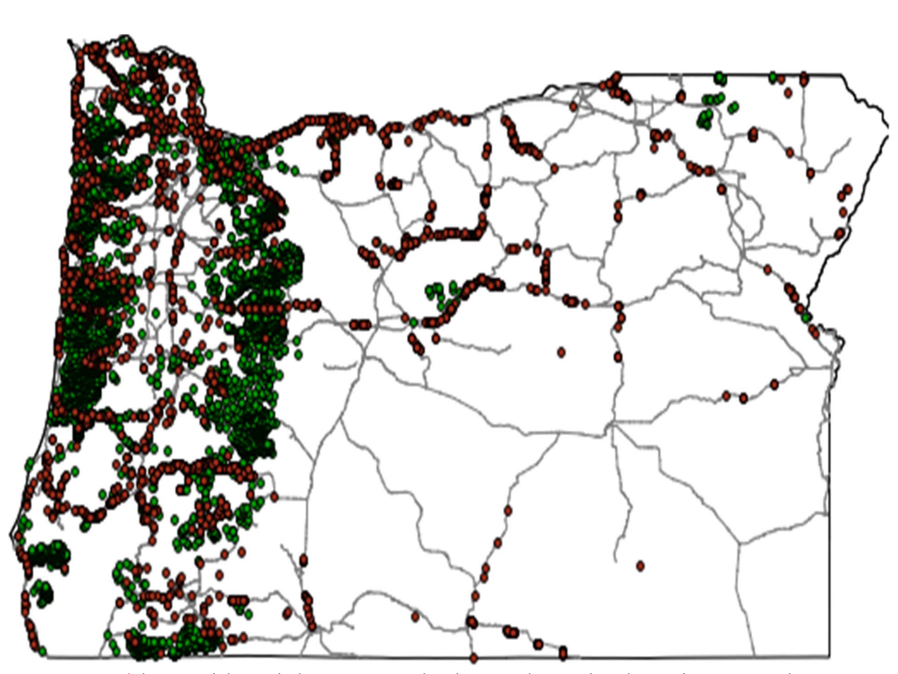


- Generate true negatives
- Separate training and test data
- Enumerate mapping and cost functions
- Coarse and fine parameter search
- Use k-fold cross-validation
- Pick best parameters
- Test on unseen data

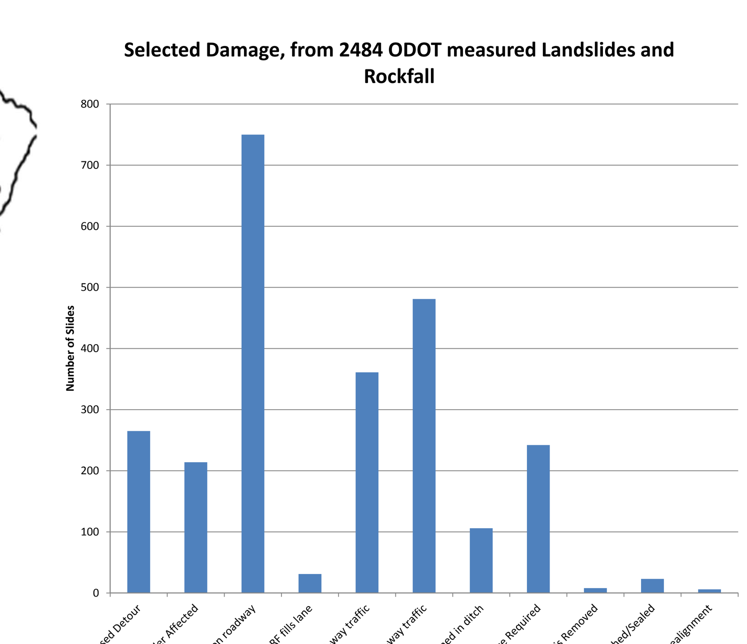


Socio-economic impacts

Oregon's SLIDO-2 database contains the location of ~3400 historical landslides with information on the socio-economic impacts and damage caused. For each landslide, we have identified the characteristics of the population and infrastructure surrounding the landslides (population density, roads, railroads, land cover). These characteristics, together with the physical characteristics of the landslide locations, will be correlated to the amount or damage that a landslide or weather event has caused. We will look for an association between damage and event severity and duration. At the same time, we will be able to identify regions that are particularly at risk. The two images shown here illustrate some of the contents of this dataset. The map show the pattern of where recorded landslides have occurred and the bar graph shows the types of damage caused by landslides along roads.

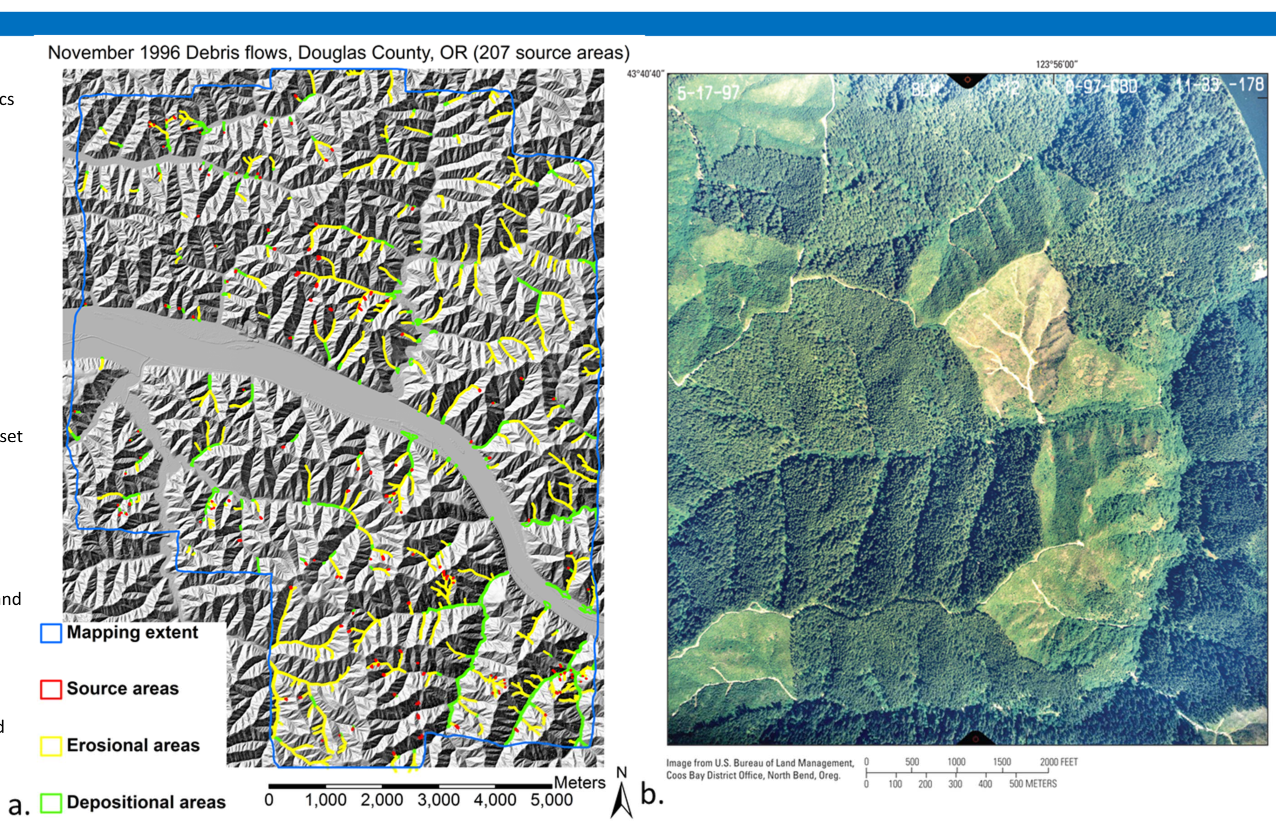


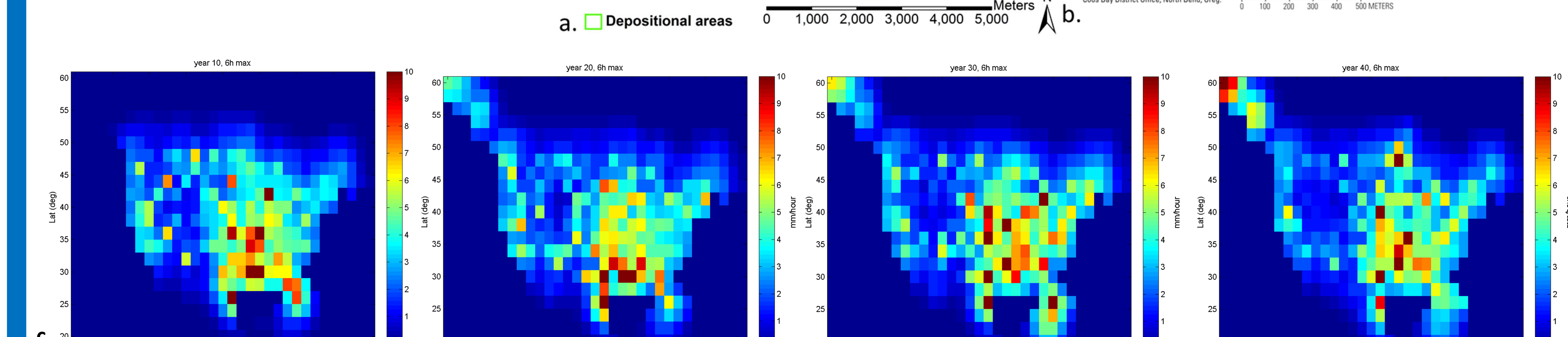
Red dots are slides with damage, Green dots have no damage listed, Grey lines are Roads



Next steps

- Use regional climate models (e.g. NARCCAP) to compile precipitation statistics
- Compare across models and to observations
- Repeat for areas where orographic precipitation is dominant
- Understand how IDF curves may change in both cases
- Test on prototype landscape to see landslide response
- Identify regional scale test area with high-resolution topographic data, landslide inventory, and well constrained soil and vegetation characteristics
- Generate precipitation statistics for specific climate scenarios (e.g. NASA NEWS)
- Test at regional scale
- Continue data-driven spatial classification approach with regional-scale dataset
- Continue developing weather-based models, focusing on developing probabilistic predictions for landslide activity for different weather patterns
- Integrate data-driven spatial and temporal classifiers
- Quantifying qualitative damage descriptions on a numerical intensity scale.
- Develop a model that can estimate the amount of damage from a weather and landslide prediction
- Explore models which include vegetation, evapotranspiration, soil moisture
- Explore models which simulate the effects of climate change on vegetation
- Explore new remote sensing techniques to better quantify soil moisture, and vegetation type and age





Examples of planned developments:
a) Shallow landslide inventory from November 1996 in Douglas County, OR, mapped onto 1-m Lidar-derived topographic data;
b) Detail of aerial imagery for Douglas County, OR, showing contrasting vegetation type and age.
c) Historical observed 6-hour block yearly maxima from the NOAA CPC gridded hourly dataset spanning the years 1948-2002

Acknowledgements

This work is funded by NASA ROSES grant number 09-IDS09-0049.