A Dynamic Landslide Hazard Assessment System for Central America and Hispaniola

D. B Kirschbaum[1], T. Stanley,[1,2], J. Simmons[3]

[1]{Hydrological Sciences Laboratory, Goddard Space Flight Center, Greenbelt, Maryland}

[2]{Universities Space Research Association, Columbia, Maryland}

[3]{Columbia University, New York, New York}

Correspondence to: D. B. Kirschbaum (dalia.b.kirschbaum@nasa.gov)

Abstract

Landslides pose a serious threat to life and property in Central America and the Caribbean Islands. In order to allow regionally coordinated situational awareness and disaster response, an online decision support system was created. At its core is a new flexible framework for evaluating potential landslide activity in near real-time: Landslide Hazard Assessment for Situational Awareness. This framework was implemented in Central America and the Caribbean by integrating a regional susceptibility map and satellite-based rainfall estimates into a binary decision tree, considering both daily and antecedent rainfall. Using a regionally distributed, percentile-based threshold approach, the model outputs a pixel-by-pixel nowcast in near real-time at a resolution of 30 ~~arcseconds~~arc-seconds to identify areas of moderate and high landslide hazard. The daily and antecedent rainfall thresholds in the model are calibrated using a subset of the Global Landslide Catalog in Central America available for 2007-2013. The model was then evaluated with data for 2014. Results suggest reasonable model skill over Central America and poorer performance over Hispaniola, due primarily to the limited availability of calibration and validation data. The landslide model framework presented here demonstrates the capability to utilize globally available satellite products for regional landslide hazard assessment. It also provides a flexible framework to interchange the indiviual model components and adjust or calibrate thresholds based on access to new data and calibration sources. The availability of free, satellite-based near real-time rainfall data allows the creation of similar models for any study area with a spatiotemporal record of landslide

events. This method may also incorporate other hydrological or atmospheric variables such as numerical weather forecasts or satellite-based soil moisture estimates within this decision tree approach for improved hazard analysis.

1 Introduction

The ability to estimate or forecast landslide ~~hazard~~ activity is largely dependent on the scale at which the analysis is undertaken as well as the availability of geomorphologic, atmospheric, and landslide data ~~accessible~~ for the study region. Physically based models focusing on the local hillslope scale require a broad range of geotechnical and hydromechanical in situ variables for accurate modelling of individual slope failures (Baum et al., 2010; Liao et al., 2012; Montgomery and Dietrich, 1994; Montrasio et al., 2011). Empirical studies can focus on local to regional scales but are constrained by the availability of landslide information and surface products that can be used to create a homogenous picture of landslide hazard over the region. The timing of rainfall-triggered landslides is challenging to predict due to the scarcity of real-time precipitation measurements, in situ landslide inventories and information about local ground conditions. Satellite rainfall products provide the opportunity to approximate the conditions that lead to rainfall-triggered landslides over regional scales, especially where rain gauge networks are sparse. The Tropical Rainfall Measuring Mission (TRMM) and its successor, the Global Precipitation Measurement (GPM) mission, provide a multi-decadal record of precipitation estimates that can be used to systematically evaluate rainfall and estimate landslide triggering relationships over multiple spatial and temporal scales (Kirschbaum et al., 2012a; Ray and Jacobs, 2007; Rossi et al., 2012).

Previous work has used rainfall intensity-duration (I-D) thresholds to estimate the landslide hazard over time at a variety of spatial scales (Berti et al., 2012; Caine, 1980; Godt et al., 2006; Guzzetti et al., 2008; Li et al., 2011; Mathew et al., 2014; Saito et al., 2010; Terlien, 1998). Landslide susceptibility zonation studies have examined the components of landslide hazard using a range of heuristic (Farahmand and AghaKouchak, 2013; Hong et al., 2007; Mora and Vahrson, 1994; Nadim et al., 2006) and statistical (Eeckhaut et al., 2009; Lee and Pradhan, 2007; Pradhan and Lee, 2010) models at diverse spatial scales. Recent work has also combined both rainfall accumulation thresholds and susceptibility information to provide early warning for landslides at a sub-national level (Lagomarsino et al., 2013; Martelloni et al., 2012; Segoni et al., 2014). Kirschbaum et al. (2012b) adapted a prototype global system

developed by Hong et al. (2006) into a dynamic landslide model at the regional scale that applied a single I-D threshold to TRMM Multi-satellite Precipitation Analysis (TMPA; Huffman et al., (2007, 2010) precipitation data and a susceptibility map for four countries in Central America: Honduras, Nicaragua, El Salvador and Guatemala. The Landslide Hazard Assessment for Situational Awareness (LHASA) model presented here builds upon ~~the~~ this previous work to develop a dynamic regional framework that provides estimations of ~~regional~~ landslide hazard in near real-time across multiple countries in Central America. The model incorporates a new landslide susceptibility map developed for Central America and the Caribbean region (Kirschbaum et al., 2015a) with local percentile-based rainfall and antecedent rainfall thresholds. LHASA has been incorporated into a prototype regional natural hazard website: http://ojo-streamer.herokuapp.com/meso. The public is now able to view a daily map identifying moderate and high landslide hazard areas, static landslide susceptibility, precipitation and antecedent rainfall over the study domain and download the model's major data inputs.

This study proposes one method for ~~approx~~estimating ~~addresses the challenge of providing a near real-time estimation of~~ potential landslide activity across broad regions ~~using when there are~~with sparse ~~in situ~~ landslide inventories and other in situ information. Ideally, an empirical relationship between precipitation and landslide occurrence would be based on a long historical landslide record with many events and corresponding gauge-based rainfall at the local scale (Frattini et al., 2009; Guzzetti et al., 2007). However, due to the dearth of both landslide information and rainfall gauges over this region, extracting local I-D thresholds was not possible. The approach presented here leverages the long-term TMPA precipitation record to relate landslide events from the Global Landslide Catalog (GLC) to a statistical distribution of rainfall from 2001~~0~~-2013. While intense rainstorms are the most important trigger of landslides in the Caribbean region (Larsen and Simon, 1993), landslides are often exacerbated by prior soil moisture conditions (Nadim et al., 2009). Using antecedent daily rainfall has been shown to help predict landslides, especially those cases where the triggering precipitation event is small (Cepeda et al., 2009). The LHASA model incorporates an antecedent rainfall index (ARI) to represent the conditions prior to the day of the triggering event. Since the relationship between rainfall, antecedent rainfall, susceptibility and landslide triggering is not linear, we employ a binary decision tree approach to test the feasibility of the model to accurately resolve landslide nowcasts while minimizing the overall number of alerts issued. This paper first presents an overview of the regional application of the model and data

used for this study. Next the LHASA threshold-based decision tree model framework is presented and calibration and validation procedures for parameterizing the model over Central America~~n~~ and the Caribbean are outlined. Lastly, the paper concludes with a discussion outlining the applicability of this model framework across a range of spatiotemporal scales and the possibility of using different hydrometeorological and in situ data products.

2 Data

2.1 Regional Setting

The LHASA model provides a flexible architecture that can be applied over a variety of spatiotemporal scales by leveraging regional environmental, climatic and landslide data. The model was initially developed to serve two regions: Central America and Hispaniola. The Central America study area ranges from 93 to 76^0 west longitude and from 6 to 19^0 north latitude. In addition to the nations of Central America, the analysis includes Jamaica and small portions of Mexico and Colombia. Central America has a tropical climate and a wide range of terrain dominated by the Central American Volcanic Arc along the western coast and active geologic faults throughout the region. The region experiences intermittent spring rains and a long rainy season from July through November, marked by landfalling tropical cyclones from the Caribbean Sea and eastern Pacific Ocean. These heavy rains combine with tectonic activity to make Central America a hotspot for landslide activity (Nadim et al., 2006). The Hispaniola study area encompasses Haiti, the Dominican Republic and Puerto Rico. The timing and intensity of the rainy season varies across Hispaniola due to five major mountain ranges across the island, with elevations varying from 3,000 meters to below sea level. Due to the associated rain shadow, annual precipitation varies from over 2,500 mm in the elevated northern regions to as little as 500 mm in the semi-arid southern regions (Alpert, 1941). In Haiti, the impact of earthquakes and tropical cyclones on deforested slopes with poorly constructed dwellings makes this area exceptionally vulnerable to landslides (Mora, 1995). All of the countries in this region have experienced significant losses from landslides as a result of previous disasters such as Hurricane Mitch in 1998 and the 2010 Haiti earthquake, among others (Guha-Sapir et al., 2014). This region is also very likely to suffer significant losses from landslides as a result of changing precipitation and tropical cyclone patterns in a changing climate (IPCC, 2007).

2.2 Landslide Catalogs

~~consistently~~ There are several different landslide inventories available within Central America that have varying geographic extents, compilation methodologies, temporal information and accuracies. We used four landslide inventories to develop and test the regional landslide susceptibility map, which are outlined in (Kirschbaum et al., 2015a). These inventories include: 1) landslides triggered by Hurricane Mitch in 1998, compiled by USGS and others (Bucknam et al., 2001; Cannon et al., 2001; Crone et al., 2001; Harp et al., 2004); 2) a historical Nicaragua database compiled by (Devoli et al., 2006, 2007); 3) a historical landslide database from El Salvador (Gerencia de Geología, 2012); and 4) the Global Landslide Catalog. While each of these inventories were useful in various ways to compute the regional static susceptibility map, the GLC had the most relevant spatial and temporal information for calibrating and evaluating the LHASA model. As a result, ~~In order to understand the conditions under which landslides occur and evaluate model performance,~~ a record of historical landslides was selected from the GLC (Kirschbaum et al., 2010). We also selected 24 landslides from the El Salvador inventory compiled by the Ministry of the Environment and Natural Resources (MARN) (Gerencia de Geología, 2012). No times of occurrence were available for these points, nor were spatial accuracies defined. The combined landslide data covered the years 2007-2013.

Despite its limitations ~~of this catalog~~, the GLC is a key resource in ~~systematically~~ evaluating landslide patterns and represents the only event-based landslide database available across all countries in Central America and the Caribbean Region. The GLC is populated primarily from media reports, but it also incorporates online disaster databases, and personal communication in some instances. Due to the compilation methodology of the GLC, there are several types of ~~biases~~ error that impact the accuracy of the catalog, including regional reporting biases, variations in cataloging methodology, and report accuracy. Kirschbaum et al. (2015b) outlines these biases in more detail. Another uncertainty stems from the landslide typ~~olog~~ies presented in this catalog. The GLC includes ~~all~~ mass movements that are reported to have been triggered directly by rainfall (including debris flows, mudslides, rock falls, etc.), all of which we herein refer to ~~all~~ as landslides. While it is often impossible to differentiate between landslide types from ~~only~~ a media report unless detailed descriptions or a photo is included, we believe that the majority of landslides that are used to calibrate and evaluate the LHASA model are rapid, shallow ~~in nature with compositions~~movements of soil, rock, and

other debris ~~that occur rapidly (opposed to creeping events)~~. The size of each landslide is often even more difficult to determine in most cases, but ~~based on reporting many of~~ the reported landslides often occur above roads and tend to be narrow, long runout debris flows. These assertions are based on review of GLC event entries as well as previous work in this region (Bucknam et al., 2001; Cepeda et al., 2010a; Devoli et al., 2006, 2008). ~~Despite the limitations, this database is a key resource in systematically evaluating landslide patterns and represents the only event-based landslide database available across all countries in Central America and the Caribbean Region.~~ In many cases, there is uncertainty on exactly where and when the landslide took place due to limited information. ~~As a result~~To limit the effects of this problem, only rain-triggered landslides with a spatial accuracy of 25 kilometres or better and a known date of occurrence were chosen. From this selection, 99 landslides from Central America and 24 landslide events from Hispaniola were used for this study (Figure 1). The exact time of occurrence was only known for 17 of these reports. Because the landslides occurred in multiple time zones, it was necessary to correct a few dates. However, no time zone correction could be made for the vast majority of events; thus, in addition to any errors present in the original report, it is likely that some landslides may have actually occurred on a different UTC date than the date of record. ~~In addition to the GLC, 24 landslides were selected from a catalog compiled by El Salvador's Ministry of the Environment and Natural Resources (MARN) (Gerencia de Geología, 2012). No times of occurrence were available for these points, nor were spatial accuracies defined. The combined catalog covered the years 2007-2013.~~

In 2014, 877 new landslides were added to the GLC. These were not available during the development of the dynamic landslide model and represent an independent dataset of the same type as the 2007-2013 catalog. 79 landslides were located within the study areas described above, accounting for 49 deaths and 30 injuries. Of these, 56 were known to be triggered by rain and had a spatial accuracy better than 25 kilometres. Due to the submission of a single detailed report, the exact location of 14 landslides was known. However, these points represent a single cluster of landslides occurring on June 23rd, 2014 near El Ayote, Nicaragua (INETER, 2014). In order to reduce the weight placed on this cluster, 6 closely spaced landslides were pruned from the GLC. The resulting 2014 catalog used in the analysis includes 42 landslides that occurred in Central America, 1 in Jamaica, and 7 that occurred in the Hispaniola study area.

2.3 Susceptibility Map

A susceptibility map was created for all of Central America and the Caribbean Islands at a resolution of 30 arc-seconds with the goal of discriminating susceptible from non-susceptible regions (Figure 1) (Kirschbaum et al., 2015a). In order to achieve a consistent output across the region, one regional ~~dataset~~ (faults) and three global ~~datasets~~ (slope, soils, and roads) geographic datasets were combined (Table 1). These ~~datasets~~ variables were selected ~~from a modified sensitivity analysis conducted over parts of the region~~on the basis of geographic extent, consistency, expert opinion, and empirical relevance. Several ~~different~~other surface variables, such as ~~percent~~ forest cover and geology, were also tested within the susceptibility model framework, but did not enhance predictions. In some cases, variables that were largely redundant (e.g. cation exchange capacity) were ~~.~~eliminated, despite good validation results.

The distance to the nearest major fault was derived from a geologic map of the Caribbean (French and Schenk, 2004). The United States Geological Survey (USGS) produced a global product highlighting statistical properties of slope and elevation (*e.g.*, median, maximum, 70th percentile, etc.) where Shuttle Radar Topography Mission (SRTM) data was aggregated from the nominal resolution (3 arc-seconds) to 30 arc-seconds (Verdin et al., 2007). The 70th ~~quantile~~ percentile slope gradient was considered for this evaluation because the slope distribution most closely correlated to the 3 arc-second SRTM data for reported landslide locations. The distribution of major soil types was obtained in raster format from the Harmonized World Soil Database (FAO/IIASA/ISRIC/ISSCAS/JRC, 2012). The variable "percent clay" was selected to represent the regional soil properties. Road locations were obtained from the Global Roads Open Access Data Set, Version 1 (CIESIN and ITOS 2013).

These 4 layers were overlaid in ArcGIS through the use of fuzzy operators. First, each variable was transformed into a "possibility" between 0 (representing low landslide hazard) and 1 (representing high hazard) through the use of a fuzzy membership function. Next, the non-topographic variables were combined with a "fuzzy gamma" function, in which gamma was set to 0.4. Finally, the output was overlaid with the transformed slope values with the "fuzzy product" operator, a simple function chosen to prevent the identification of flat ground as hazardous.

Using the standard deviation classification scheme, the susceptibility map was divided into 5 categories that represented relative ~~hazard~~susceptibility: very low, low, medium, high, and very high. Four historical landslide catalogs—varying greatly in temporal and~~,~~ spatial scale,

size, and completeness—were used to evaluate the susceptibility map. Figure 2 compares the distribution of the 5 susceptibility categories at recorded landslides to the distribution over the total study area. Few landslides occurred in locations rated as having "very low" landslide susceptibility (SI = 1), despite the fact that this is the largest category by area. These locations were not considered susceptible to landslides for the purposes of the binary decision tree model.

2.4 Rainfall Estimates

Satellite precipitation estimates from the Tropical Rainfall Measuring Mission (TRMM) Real-Time Multi-Satellite Precipitation Analysis (TMPA-RT) are available at a resolution of 0.25 x 0.25 degrees (Huffman et al., 2007, 2010). This product provides a snapshot of precipitation rates utilizing TRMM and other satellites to provide a precipitation map every 3 hours from 50^0N-S. TMPA-RT data is available from March 2000 to the present. For this analysis, ~~the 3-hourly files were merged to compute~~ daily rainfall totals were used~~for this region~~. The Global Precipitation Measurement (GPM) mission was launched in February, 2014 and is a global successor to TRMM. GPM's multi-satellite product called IMERG (Integrated Multi-Satellite Retrievals for GPM) is already providing data, although TMPA-RT will continue to be processed into the near future.

3 Methods

The model inputs to LHASA required several processing and calibration steps before they could be directly applied. Unless otherwise noted, all calculations were performed in the statistical programming language, R v3.1.2 (R Core Team 2013). Raster operations were performed with ~~the~~ "raster", a third-party package for R (Hijmans and van Etten, 2014). All raster files were resampled to a resolution of 30 arc-seconds, the same as the susceptibility map, using the nearest neighbour method, and clipped to the two study areas. Thus, each pixel represents a data point of approximately 1 square kilometre (30 ~~arcseconds~~arc-seconds) and the model is run every 24 hours. Where possible, the dates of landslides were adjusted to match the time zone of the daily rainfall data. The landslide catalog was converted from a shapefile to a series of daily presence/absence landslide rasters over the same extent as the rainfall files. This format enabled landslide events to be directly compared to the daily maps

of landslide hazard outputs of the LHASA model. Most raster operations were performed in parallel on a Linux server.

3.1 Landslide Catalog

The landslide catalogs used to calibrate this model are known to contain both spatial and temporal errors (Kirschbaum et al., 2015b). The spatial accuracy of each landslide point was estimated at the time of recording. Although the least accurate reports were not used in this study, some landslides may have occurred as much as 25 kilometres from their reported locations. In these cases, as well as with more accurate reports, the terrain and rainfall where the landslide occurred may differ from the conditions at the reported coordinates. In addition, the exact date and time of an event is often unknown. To evaluate the extent to which landslide reports were incorrectly evaluated as false negatives due to spatiotemporal uncertainty, landslide points were spatially and temporally buffered. Temporal uncertainties were accounted for by considering 1, 3 and 7-day windows surrounding the reported landslide date. Spatial uncertainties were evaluated by considering the exact location of the reported landslide, 1 and 5 km circular buffers surrounding the location, as well as a variable buffer based on the spatial accuracy denoted in each landslide entry. For this analysis, if a nowcast was generated anywhere within the buffer spatiotemporal window, we considered this nowcast to be a success (true positive). There are challenges inherent with this assumption, which are outlined in the Discussion section. True ~~Positive~~ positive ~~Rates~~ rates were calculated for each combination of windows.

3.2 Daily Rainfall

Owing to the diverse topography, coastal zones and prevailing wind patterns, rainfall is unevenly distributed over ~~the~~ Central America and Hispaniola ~~study regions~~. There have been many different approaches to representing the intensity-duration rainfall triggering relationships at various scales including critical rainfall (Aleotti, 2004; Li et al., 2011; Saito et al., 2010; Tiranti and Rabuffetti, 2010), normalized rainfall or return periods~~using mean annual precipitation~~ (Dahal and Hasegawa, 2008; Hromadka II et al., 2010; Terlien, 1998), ~~return period (Hromadka II et al., 2010)~~ and the combination of multiple thresholds (Brunetti et al., 2010; Cepeda et al., 2010b; Chleborad et al., 2006), among others. ~~Many of these approaches are summarized in Guzzetti et al. (2008)~~. One of the challenges with approaching rainfall-triggering relationships at a regional scale is that the prevailing precipitation regimes

are regionally heterogeneous. As a result, developing a single regional intensity-duration threshold to represent the landslide triggering relationships across this region was not optimal.

To address this problem, we considered the statistical distribution of daily rainfall over a 13-year record, using percentiles to create a precipitation metric that could be compared across morphologies and landslides events. A daily precipitation time series from January 1st, 2001 to December 31st 2013 was prepared for each pixel over the study region and days with no rainfall were removed from the calculations. Then, every 5th percentile was calculated from the distribution of non-zero values using the "quantile" function's default method. The resulting series of raster files identify the local precipitation distribution at each pixel and provide a more localized way to address regional landslide triggering (Figure 3). The rainfall thresholds were then calibrated with the landslide data to assign a separate rainfall threshold for each 0.25-degree pixel. The calibration procedure is described below.

3.3 Antecedent Rainfall

Real-time measurements of subsurface pore pressure are not available at most locations within the study area. Satellite-based soil moisture retrievals are often biased or limited over complex terrain, particularly with dense vegetation (Jackson and Schmugge, 1991; Njoku et al., 2003). Therefore, remotely sensed rainfall was chosen as a proxy for this variable. Time is required for rain to infiltrate soil and rock and generate higher pore pressures that lead to slope instabilities, as well as for pore pressure to dissipate. High levels of precipitation often trigger landslides, but the amount required to do so is usually dependent upon the volume of prior rain and the permeability of the soils and rocks. To describe this phenomenon, an antecedent rainfall index (ARI) was created from the TMPA-RT daily rainfall estimates using a time-weighted average of the previous 60 days:

$$\frac{\sum_{t=1}^{60} w_t * p_t}{\sum_{t=1}^{60} w_t} \quad (1)$$

Where t = the number of days before the present, p_t = the precipitation at time t, and $w_t = t^{-0.5}$. This closely resembles the antecedent precipitation index introduced by Kohler and Linsley (1951). The decay exponent and the 60-day window were chosen by calibrating a preliminary decision tree model using landslide reports and 500 randomly selected locations.

No rainfall or ARI threshold can serve as a perfect classifier to differentiate landslide and non-landslide rainfall events because the distributions of these two datasets overlap (Figure

4). However, Figure 5 shows that the relative frequency of landslides increases at higher precipitation levels, allowing thresholds for current and antecedent rainfall to be combined in such a way that the resulting classification is reasonably effective.

3.4 Decision Tree Framework

At the most simplistic level, higher soil moisture values prior to a landslide occurrence can be a key predisposing factor in future landslide triggering (Wieczorek, 1987). ~~Past d~~Previous decision tree models considering precipitation and antecedent values have been derived at ~~mostly~~ the city level and apply the trade-off between rainfall and past rainfall infiltration to create an alert framework (Aleotti, 2004; Godt et al., 2006) or at a regional (sub-national) scale considering accumulated precipitation and specifying a critical rainfall threshold (Lagomarsino et al., 2013; Martelloni et al., 2012; Segoni et al., 2014). In this study, antecedent rainfall was incorporated into a 3-level binary decision tree structure (Figure 6). At the first level, those pixels believed to have "very low" susceptibility to landslides (Susceptibility Index of 1 or 0) are excluded from further analysis. All other pixels are considered as having a non-negligible chance of slope failure. An SI ≥ 2 (low) was chosen to exclude a large portion of Central America without losing the ability to predict most landslide events. ~~We selected this binary threshold due to the XXXX.~~ This categorization means that ~~over XX~~less than 60% of the study area is considered to be susceptible ~~within the model, which we feel is reasonable because XXX.~~ , while over 90% of the landslides occurred in susceptible pixels. ~~easysusceptibility ,~~

At the second level of analysis, the antecedent rainfall value for each pixel is compared to the 50th percentile value. This antecedent rainfall threshold roughly corresponds to the division between the Central American wet season, during which most landslides occur, and the dry season (Figure 7b). At the third level of the decision tree, the current daily rainfall accumulation is compared to the daily rainfall threshold and a moderate or high landslide hazard nowcast is issued. If the 50th percentile rainfall is exceeded and the soils are considered to be wet (ARI > 50th percentile), a moderate hazard level is assigned. If the rainfall exceeds the 95th percentile, a high hazard is assigned. In dry conditions (ARI < 50th percentile), a moderate hazard level is assigned if the daily rainfall exceeds the 90th percentile and a high hazard nowcast is generated if rainfall exceeds the 95th percentile. The "high hazard" nowcast is design~~at~~ed to represent the extreme triggering conditions under which landslides have a higher p~~oss~~roba~~i~~bility of occurrence; whereas the "moderate hazard"

nowcasts represent a lower possirobability of potential landslide activity but still represent the possibility of landslides occurring in the designated areas. With additional data, future work will seek to assign more probabilistic values onarrowly quantify the probability of landslides potential tofor each of these hazard classes.

The moderate hazard nowcast thresholds were determinedwas calibrated by comparing varying the ARI and daily rainfall thresholds, then determining the model's success for the 123 events from the Central American catalog and optimizing the predominant thresholds for these instances. Due to the computational burden, Thethe calibration process involved did not include every possible set of thresholds due to the compute time, but provided a representative sample of the percentiles thresholds between the 50th and 95th percentiles for both ARI and daily rainfall records, not every possible set of thresholds. The model was not calibrated for the 24 landslides that occurred in Hispaniola because preliminary analysis of the reported landslide locations indicated that no combination of rainfall thresholds would provide a good fit to the landslide observations. The high- landslide hazard classification nowcast was determined based on the desirecreated to provide a representation of extreme rainfall at any time over the study region. The 95th percentile was chosen based on past research and qualitative analysis of the rainfall distributions over this area (Kirschbaum et al., 2015b). Further discussion of this topic can be found in the Results section.

Given the triggering variables, surface information and landslide catalogs considered within LHASA, we posit that the LHASA model is more successful in resolving the potential conditions for landslides with a mix of soil, rock and other debris, ranging from moderate to shallow depths and occurring at moderate to high velocities. This assertion is mostly due to the main types of landslides observed within the study area as well as from the fact that currently we do not consider other triggering variables such as earthquake occurrence, anthropogenic triggers (mining, construction, etc.), etc.

3.5 Evaluation

The predictive success of the dynamic landslide model was evaluated with the "distance to perfect classification" metric, r_j (Cepeda et al., 2010a), which combines the true positive rate (TPR) with the false positive rate (FPR):

$$r_j = \sqrt{(FPR^2 + (1 - TPR)^2)} \quad (2)$$

Where TPR = true positives / (true positives + false negatives) and FPR = false positives / (false positives + true negatives). These metrics are often used to compare binary classifiers (Fawcett, 2006). In this case, the true positives are pixels where a landslide occurred on the same date that the model issued a nowcast, the false negatives are pixels where a landslide occurred on the same date that the model failed to identify the potential for slope failure, the false positives are pixels where a landslide did not occur on the same date that the model identified the potential for slope failure, and the true negatives are pixels where the model did not identify the potential for slope failure and no landslide was reported. The confusion matrix was calculated for each day in the study period, ~~and~~ then results were summed before calculating the FPR and TPR. This process was repeated each time the model ran during calibration. For the 2007-2013 dataset, the TPR varied more rapidly with threshold changes than the FPR, so changes in r_j largely reflected changes in the model's ability to predict the occurrence of a relatively small number of landslides (Figure 8). To quantify how the predictions respond to different levels of accuracy ~~of~~in the GLC, we varied the spatial area from 0 to 25~~+~~ km around each of the landslide points ~~as well as~~and varied the temporal window ~~considered~~ around each landslide reported date/time by 1, 3 and 7 days. This provided a way to ~~better quantify~~calculate the probability of detection more realistically since the uncertainty in both the location and the date of the validation landslides was variable.~~The predictions were also evaluated across a variety of modest spatial and temporal windows.~~ After calibration, the same metrics were calculated for an independent 2014 dataset. ~~Thresholds ranging from XXX – XXX were tested and results with r_j statistics and FPR values are shown in Table 3.~~

4 Results

The LHASA model was evaluated over Central America and Hispaniola for 2007-2013 and then compared with an independent validation dataset for 2014. The best performance, an r_j of 0.38, was observed with an antecedent rainfall threshold of the 50th percentile and current daily rainfall thresholds of the 50th and 90th percentiles. The high hazard threshold (95^{th} percentile) was not calibrated with the available landslide data. The model results are summarized in Table 2 and Table 3. The TPR varied depending on the spatiotemporal window considered for identifying the landslide. As introduced above, we accounted for the uncertainty in the reported date and location of the landslide by applying spatial and temporal

buffers around the reported latitude and longitude and date of the event. We considered TPR results for 1, 3, and 7-day windows surrounding the date of the landslide and 0, 1 and 5 km buffers, as well as a~~nother~~ buffer ~~field that directly applies~~equal to the qualitative location accuracy value reported for the event (ranging from a 0 to 25 km radius surrounding the landslide location). For each of these ~~buffered~~ windows, if a nowcast were issued anywhere within the buffer~~area considered~~, we considered it to be a successful prediction. The same approach was taken for both moderate and high hazard categories.

4.1 Central America (2007-2013)

Results for all regions and time spans considered indicate that generally as the buffer temporal and spatial window increase, the TPR increases as well. This result is not surprising as it provides more opportunities for a nowcast to be successful. The temporal window (or spatial buffer) has more variability in the TPR values at the most conservative tolerances assigned, ranging from 64% to 83% for the temporal windows and from 64% to 81% for the spatial windows over Central America. For the high hazard nowcasts, results are considerably lower, with only a TPR of 26 to 40% based on the spatial buffers and 26 to 47% TPR when a 1, 3 and 7-day window was considered. The FPR for Central America was 11% for the moderate hazard nowcasts and 1% for the high hazard nowcasts. Table 2 also shows the percentage of landslides that fall above the threshold (SI $\geq$ 2) specified in the decision tree structure according to the spatial buffer considered (maximum susceptibility value was considered within each buffered area). For Central America, 12 (10%) landslides are located within the lowest hazard zone~~ of the susceptible pixels~~. This may be due to the map's 1-km resolution, where generally "safe" pixels may contain small areas that are prone to landslides, or due to the location of the reported landslide (*e.g.*, the landslide was reported in the ~~landslide~~ runout zone where slopes are very gradual vs. higher up on the slope where ~~the~~ initiation likely occurred). ~~Implications for this spatiotemporal buffering are presented below.~~

Even with the largest spatiotemporal buffers applied to the reports, not every landslide was predicted. For example, one group of landslides was reported in western Costa Rica along the highway between Balsa de Atenas and Orotina (Figure ~~8~~9). This report represents at least 5 different rockfalls that occurred over this section of the road, denoted in the figure as a single point. As the specific location was unknown, a spatial accuracy of 10 kilometres was assigned. The date of the event was reported as April 14th, 2010, but the exact time is unknown. The model issued a moderate hazard nowcast for the area east of the reported

landslide on both the 14th and 15th of April, 2010, but not to the west where the rockfalls occurred (Figure ~~8a~~9a). Because the exact timing of these events was unknown, there was a possible temporal error of 1 day for this event. A second event occurred in the same area on May 22nd and a high hazard nowcast was issued in the area of the reported rockfall event (Figure ~~8b~~9b). Figure 7 plots the rainfall and antecedent rainfall for the same area shown in Figure ~~8~~9, where four landslides occurred in 2010: April 14th, May 22nd, July 30th, and November 5th. As described above, a nowcast was not issued for the April 14th event because it occurred before antecedent rainfall exceeded the 50th percentile, despite the daily rainfall exceeding the 50th percentile (Figure ~~8a~~9a). This specific incident might also be attributed to slope destabilization associated with recent highway construction because relatively little rain was required to move these steep slopes out of equilibrium. The May 22nd event generated a high hazard nowcast (Figure ~~8b~~9b). Rainfall totals for July 30th were not substantial enough to trigger a nowcast. Lastly, a moderate landslide hazard nowcast was issued for the November 5th event, which occurred the day after a very intense precipitation event.

4.2 Hispaniola (2007-2013)

Because of the limited number of data points in Hispaniola, no calibration was performed in this study area. Instead, the thresholds calibrated from the Central American catalog were applied to the rainfall distributions over these islands. A FPR of 9% was observed for the moderate hazard threshold and 1% for the high hazard threshold. The TPR for the exact location and date of the reported landslide was 21% (i.e. the model predicted 5 of the 24 landslide events). The high hazard model performed poorly, with a TPR of 17-21% for the spatial buffers on the day of the landslide. The highest TPR value at the most liberal spatial and temporal window was 54%. This rate is largely due to the fact that half of the landslide events are recorded in locations not considered to be susceptible to landslides, as shown in the far right column of Table 2. The low susceptibility values corresponding to the locations of the reported events may be due to spatial errors in the GLC, since only 13 of the 24 reports were assigned an accuracy better than 5 kilometres. As the buffer size increases, the TPR significantly improves. In addition, 11 of the 24 landslides reportedly occurred on days without any rainfall. It is likely that some degree of temporal error in the catalog explains this ~~issue~~fact. The limited data inventory for Hispaniola may also affect the susceptibility map calculations over this area.

4.3 Global Landslide Catalog (2014)

The landslides recorded in the GLC during 2014 make up an independent dataset with which to evaluate the performance of the thresholds developed for the previous years. The 50 points in the GLC were separated by study area. The TPR for the Central American study area for the exact location and date was 58%, while the FPR was 9% for the moderate hazard nowcasts. By contrast, the TPR for the 7 landslides in the Hispaniola study area was only 43%. If the spatial and temporal tolerance is increased, the TPR ~~rates~~ ranges from 63-91% for Central America and 57-86% for Hispaniola. Overall, these results are similar to the results for the period 2007-2013. This suggests that the model is not overfitted to a single dataset. On the other hand, the fact that the model does not perform as well in Hispaniola over both time periods suggests that the use of percentile thresholds may be limited to the geographic regions for which they were developed.

The high hazard nowcasts had a FPR of 1% and consistently gave a TPR of 43% for the 1 and 3-day windows irrespective of the buffer considered, while the 7-day window produced a TPR of 71%. The TPR values were consistent across all spatial buffers because this very small dataset (7 events) did not happen to contain any events in which a storm occurred near~~by~~ but not at the reported location of landslide. Three of the reported events had rainfall that exceeded the 95th percentile, while 2 of the events occurred on fairly dry days.

The use of validation data with substantial spatial and temporal errors makes the evaluation of model performance difficult. One approach to this issue is to find times and locations near recorded points that may be closer to the true location of the landslides. Considering model outputs within 1 day and 1 kilometre (less than the spatial uncertainty of most reports) of reported landslides raised the TPR of the Central American catalog to 79% and the TPR of the Hispaniola catalog to 71%. The performance of the model against a perfectly complete and precise landslide catalog is, unfortunately, unknown. However, when the cluster of 14 landslides with exact locations in Nicaragua was compared to the model output for June 23rd, 2014, only 1 landslide was not predicted (Figure ~~9~~10). This cluster only represents a single event, but supports an optimistic interpretation of the results for the larger catalogs. Another potential approach could consider dividing the study area into geomorphologically similar regions and re-calibrating the rainfall and ARI thresholds at sub-regional scales, allowing the rainfall thresholds and even susceptibility bins to be adjusted. However, this approach requires a robust landslide inventory ~~data~~ for calibration ~~and would need to assume that all~~

areas had consistent and sufficient data points. We may consider this approach should new datasets become available or we apply this model over a different study area.

5 Discussion

The objectives of this LHASA system are is to estimate potential landslide activity over a very broad area in near real-time using input data that both has very few points (relative to the area considered) as well as variable accuracy. These ambitious challenges difficulties require restrict the usage of this model to be approached in the appropriate context: a situational awareness tool for further investigation of that flags potentially affected areas for further investigation rather than a direct tool for issuing warnings or declaring impacts. Results of the evaluation suggest that when the finest spatial buffers and temporal windows are considered for the 2014 validation dataset, the true positive rate for the moderate hazard model is between 43 and 58% for the Hispaniola and Central America datasets respectively (Table 2). However, as the search criteria are expanded, even slightly, results are more promising. The high hazard model continues to have poorer performance has a relatively low probability of predicting landslides, which most likely results from the few nowcasts made as well as the due to the fact that very limited number of many landslide reports in the GLC that are not recorded on the same day as coincide with these extreme rainfall periods events. Given the limitations of the data available for evaluating this system as well as for calibration of its components, we feel that results of the LHASA model are encouraging for serving nevertheless encourage its use as a regional situational awareness tool for potential landslide activity.

The LHASA model is currently implemented in a multi-hazard website servicing Central America and Hispaniola. While the model is currently parameterized for this region, it could be adapted to serve other landslide-prone locations. This flexible binary decision tree framework enables different forcing variables (precipitation, soil moisture antecedent precipitation) and susceptibility variables to be considered dynamically. The process of transferring the model to another location requires a susceptibility map and an event-based landslide catalog such as the GLC. The simplicity of the model, combined with the fact that rainfall data are made available without cost by NASA for every location between 50 degrees north to 50 degrees south latitude, means that it should be possible for web developers, students, and other programmers to implement the LHASA system rapidly.

~~There are several limitations in this methodology owing to the landslide catalog available for both calibrating and validating the LHASA system.~~ F~~A f~~uture ~~step in this~~ work could ~~consider~~include calibrating the LHASA model in an area with an extensive ~~, extended~~ and accurate landslide inventory to fully assess the performance of this system. ~~However, t~~ To our knowledge, no landslide inventories of this type exist over the current study region. ~~and therefore~~Therefore, ~~we~~it would ~~need~~be necessary to parameterize the model over a new domain. This is outside the scope of the existing work but may be feasible as we continue to test this system in other regions.

The thresholds chosen by calibration in this study represent a compromise between identifying landslide hazard and limiting the number of days on which an alert is issued. As a result, a large number of days are identified as having moderate hazard across the study area, especially during the rainy season. The frequency of high-hazard nowcasts is ~~reduced when considering the high hazard category using the 95th percentile of the 13-year TRMM record~~significantly lower. ~~–~~Ultimately, optimal thresholds could be determined from information on the relative cost to model users of false positives and false negatives. Gathering this economic data is beyond the scope of this study, as it would require extensive consultation with current and potential users of the system.

This model relies on TRMM's TMPA-RT data inputs but GPM's IMERG data will be incorporated into this system in the near future, which will extend the latitudinal boundaries of the precipitation information to 65 degrees N-S and increase the spatiotemporal resolution to ~~30-~~30-minute sampling at a 0.~~1-~~1-degree spatial resolution. If rain gauge or forecasted rainfall data is available for a region, this may also be applied to create a more accurate real-time hazard assessment system. The LHASA regional system is currently run on the Heroku Cloud Application Platform (Heroku, 2015) with limited computational resources required for generating regional, daily nowcast products. The real-time IMERG product was made available in March, 2015 while TMPA-RT continues to provide data. Once the IMERG algorithm has been running routinely for a period of time, the entire TRMM archive will be reprocessed to encompass one consistent precipitation dataset for both TRMM and GPM era. We will begin testing the application of IMERG in the near future but will only be able to recompute the percentiles in the analysis and fully recalibrate the model once the reprocessing takes place (tentatively scheduled for 2017). Additional testing is required to determine the effectiveness of satellite products from the Soil Moisture Active Passive (SMAP;

http://nasa.gov/smap) or modelled soil moisture products within this area of complex terrain and dense tropical vegetation. This is a topic of future study. It may be possible to determine the relationship between the water content of surficial soils and deeper soils (Swenson et al., 2008), resulting in an estimate of pore pressure at critical depths below the ground surface. A different approach would be to separate the geologic and topographic properties currently embodied in the susceptibility map, then use them directly in the decision tree structure. Finally, other triggering variables such as seismicity ~~or temperature~~ may also be considered in a future version of this model by adding another branch to the decision tree. In some regions, temperature has been shown to drive landslide triggering during freeze/thaw episodes or spring snowmelt (do Amaral Vargas Jr. et al., 2013; Chleborad, 1997; Li et al., 2013; Tatard et al., 2010); however, in the Central America region, this triggering variable is less relevant given the predominant tropical or subtropical temperatures. ~~—~~The inclusion of other susceptibility or triggering variables within the LHASA framework is both feasible and fairly straightforward to implement.

Fundamentally, the model calibration for both the susceptibility map and rainfall thresholds can be significantly improved with a more robust, event-based landslide archive. The GLC used here provides a global source for investigating rainfall-triggered landslides. However, due to topographic, linguistic and other reporting biases, there remain large gaps in the landslide inventory for many areas. An additional capability of the prototype regional natural hazard website is the ability to access, share, edit and accept volunteered geographic information on landslide events in multiple languages ~~on landslide events~~. Future dynamic landslide models are likely to benefit from the improved accuracy and completeness of event-based landslide catalogs compiled and edited through citizen science efforts by large numbers of local end users.

6 Conclusion

The LHASA model provides a simple, flexible framework that can be easily calibrated and transferred to other regions. This model is meant to provide a regional near real-time perspective of moderate to high landslide hazard potential and currently lacks the spatial resolution and accuracy to be considered over smaller (e.g. city, municipality) scales. The ultimate goal of this paper is to present the LHASA model framework with a set of calibrated thresholds for Central America and the Caribbean Region. With improved landslide inventory

information we feel that the model calibration could be significantly improved. The availability of satellite-based rainfall, susceptibility products and the GLC provide the opportunity to expand this analysis to other susceptible regions. With TRMM and GPM data it is possible to produce rainfall thresholds that incorporate the unique climate of each site, even in locations where no rain gauges exist. Susceptibility maps can benefit from the many landslide catalogs that lack date or time attributes. Thus, dynamic models that incorporate susceptibility maps derived from these long-term catalogs may estimate long-term hazard in ways that models derived solely from recent event-based catalogs cannot.

Given the regional scope of this system, the LHASA model correctly identifies the potential for a majority of the landslide events recorded in the GLC. Although a large number of days without recorded landslides were identified as moderately hazardous, many of these data points may have had slope failures that went unrecorded due to lack of observations, economic impact or other factors. This dynamic landslide model made use of the best available real-time rainfall data. However, future inclusion of GPM's IMERG will enable landslide modelling at finer spatial and temporal resolutions. In addition, improved soil moisture estimates from SMAP may help to better quantify the ground conditions prior to extreme rainfall events. Precipitation forecast data may also be considered within this framework to provide landslide forecasts, rather than near-real-time nowcasts. This is an area of active research. While more sophisticated landslide hazard assessment and prediction models are feasible when considering smaller spatial scales and more landslide information, the LHASA model serves as a template for rapid adaptation of remote sensing datasets for dynamic situational awareness of landslide hazards at the regional scale.

Acknowledgements

The authors would like to acknowledge the individuals who helped to develop the GLC, including Stephanie Hill, Lynne Shupp, Teddy Allen, Pradeep Adhikari, Lauren Redmond, David Adler, Kimberly Rodgers, Lee Sanders, Benjamin Hall, and Melanie Franchek. We would also like to thank the individuals who provided expertise in developing the regional susceptibility map used in this study, including Soni Yatheendradas and José Cepeda. A special thank you to Pat Cappelaere for developing the online prototype of this model and the

landslide catalog interface. This work was supported by the NASA SERVIR Program, NNH11ZDA001N-SERVIR.

7 References

Aleotti, P.: A warning system for rainfall-induced shallow failures, Eng. Geol., 73, 247 – 265, doi:10.1016/j.enggeo.2004.01.007, 2004.

Alpert, L.: THE AREAL DISTRIBUTION OF MEAN ANNUAL RAINFALL OVER THE ISLAND OF HISPANIOLA, Mon. Weather Rev., 69(7), 201–204, 1941.

Do Amaral Vargas Jr., E., Velloso, R., Chávez, L., Gusmão, L. and do Amaral, C.: On the Effect of Thermally Induced Stresses in Failures of Some Rock Slopes in Rio de Janeiro, Brazil, Rock Mech. Rock Eng., 46(1), 123–134, doi:10.1007/s00603-012-0247-9, 2013.

Baum, R. L., Godt, J. W. and Savage, W. Z.: Estimating the timing and location of shallow rainfall induced landslides using a model for transient, unsaturated infiltration, J. Geophys. Res., 115(F03013), doi:10.1029/2009JF001321, 2010.

Berti, M., Martina, M. L. V, Franceschini, S., Pignone, S., Simoni, A. and Pizziolo, M.: Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach, J. Geophys. Res. Earth Surf., 117(F4), n/a–n/a, doi:10.1029/2012JF002367, 2012.

Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D. and Guzzetti, F.: Rainfall thresholds for the possible occurrence of landslides in Italy, Nat. Hazards Earth Syst. Sci., 10, 447–458, 2010.

Bucknam, R. C., Coe, J. A., Chavarria, M. M., Godt, J. W., Tarr, A. C., Bradley, L., Rafferty, S., Hancock, D., Dart, R. L. and Johnson, M. L.: Landslides triggered by Hurricane Mitch in Guatemala - Inventory and Discussion, U.S. Geol. Surv. Open-File Rep., 01-443, 1–40, 2001.

Caine, N.: The Rainfall Intensity: Duration Control of Shallow Landslides and Debris Flows, Geogr. Ann. Phys. Geogr., 62(1/2), 23–27, 1980.

Cannon, S. H., Haller, K. M., Ekstrom, I., Schweig III, E. S., Devoli, G., Moore, D. W., Rafferty, S. A. and Tarr, A. C.: Landslide Response to Hurricane Mitch Rainfall in Seven Study Areas in Nicaragua, Open-File Report 01-412-A, , 21 [online] Available from: http://pubs.usgs.gov/of/2001/ofr-01-0412-a/OFR01-412-A.access.pdf, 2001.

CIESIN (Center for International Earth Science Information Network ~~(CIESIN)~~)/Columbia University and ITOS ~~Georgia, I. T. O. S. (ITOS)/~~(University of Georgia): Global Roads Open Access Data Set, Version 1 (gROADSv1), [online]

Available from: http://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1, 2013.

Cepeda, J., Chávez, J. A. and Cruz Martínez, C.: Procedure for the selection of runout model parameters from landslide back-analyses: application to the Metropolitan Area of San Salvador, El Salvador, Landslides, 7(2), 105–116, doi:10.1007/s10346-010-0197-9, 2010a.

Cepeda, J., Díaz, M. R., Nadim, F., Høeg, K. and Elverhøi, A.: Generalised form of a power law threshold function for rainfall-induced landslides, in EGU General Assembly, p. 8256, Vienna, Austria. [online] Available from: http://adsabs.harvard.edu/abs/2010EGUGA..12.8256C, 2010b.

Cepeda, J., Hoeg, K. and Nadim, F.: Landslide-triggering rainfall thresholds: a conceptual framework, Q. J. Eng. Geol. Hydrogeol., 43(1), 69–84 [online] Available from: http://qjegh.geoscienceworld.org/content/43/1/69.short?rss=1&ssource=mfr, 2009.

Chleborad, A. F.: Temperature, Snowmelt, and the Onset of Spring Season Landslides in the Central Rocky Mountains. [online] Available from: http://landslides.usgs.gov/recent/archives/1997rockymtn.php, 1997.

Chleborad, A. F., Baum, R. L. and Godt, J. W.: Rainfall Thresholds for Forecasting Landslides in the Seattle, Washington, Area—Exceedance and Probability, U.S. Geol. Surv. Open-File Rep., 2006-1064, 2006.

Crone, A. J., Baum, R. L., Lidke, D. J., Sather, D. N., Bradley, L.-A. and Tarr, A. C.: Landslides induced by Hurricane Mitch in El Salvador -- an inventory and descriptions of selected features, Open File Rep. 01-444, 28 [online] Available from: http://pubs.usgs.gov/of/2001/ofr-01-0444/OF01444_EN.pdf, 2001.

Dahal, R. K. and Hasegawa, S.: Representative rainfall thresholds for landslides in the Nepal Himalaya, Geomorphology, 100, 429 – 443, doi:10.1016/j.geomorph.2008.01.014, 2008.

Devoli, G., De Blasio, F. V, Elverhoi, A. and Hoeg, K.: Statistical Analysis of Landslide Events in Central America and their Run-out Distance, Geotech. Geol. Eng., doi:10.1007/s10706-008-9209-0, 2008.

Devoli, G., Morales, A. and Hoeg, K.: Historical landslides in Nicaragua—collection and analysis of data, Landslides, ADD, doi:10.1007/s10346-006-0048-x, 2006.

Devoli, G., Strauch, W., Chavez, G. and Hoeg, K.: A landslide database for Nicaragua: a tool for landslide-hazard management, Landslides, 4, 163–176, doi:10.1007/s10346-006-0074-8, 2007.

Van den Eeckhaut, M. ~~Van Den~~, Reichenbach, P., Guzzetti, F., Rossi, M. and Poesen, J.: Combined landslide inventory and susceptibility assessment based on

different mapping units: an example from the Flemish Ardennes, Belgium, Nat. Hazards Earth Syst. Sci., 9, 507–521, 2009.

FAO/IIASA/ISRIC/ISSCAS/JRC: Harmonized World Soil Database (Version 1.2), [online] Available from: http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/, 2012.

Farahmand, ~~a~~A. and AghaKouchak, a.: A satellite-based global landslide model, Nat. Hazards Earth Syst. Sci., 13(5), 1259–1267, doi:10.5194/nhess-13-1259-2013, 2013.

Fawcett, T.: An introduction to ROC analysis, Pattern Recognit. Lett., 27, 861–874, doi:10.1016/j.patrec.2005.10.010, 2006.

Frattini, P., Crosta, G. and Sosio, R.: Approaches for defining thresholds and return periods for rainfall-triggered shallow landslides, Hydrol. Process., 1460(March), 1444–1460, doi:10.1002/hyp, 2009.

French, C. D. and Schenk, C. J.: Map Showing Geology, Oil and Gas Fields, and Geologic Provinces of the Caribbean Region. Open File Report 97-470-K, Denver, CO., 2004.

Grencia de Geología~~, G. de~~: Landslide inventory of El Salvador, El Salvador. [online] Available from: http://www.marn.gob.sv/, 2012.

Godt, J. W., Baum, R. L. and Chleborad, A. F.: Rainfall characteristics for shallow landsliding in Seattle, Washington, USA, Earth Surf. Process. Landforms, 31, 97–110, doi:10.1002/esp.1237, 2006.

Guha-Sapir, D., Below, R. and Hoyois, P.: EM-DAT: International Disaster Database, Univ. Cathol. Louvain - Brussels - Belgium [online] Available from: www.em-dat.net (Accessed 20 March 2014), 2014.

Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, Meteorol Atmos Phy, 98(3-4), 239–267, 2007.

Guzzetti, F., Peruccacci, S., Rossi, M. and Stark, C. P.: The rainfall intensity–duration control of shallow landslides and debris flows: an update, Landslides, 5, 3–17, doi:10.1007/s10346-007-0112-1, 2008.

Harp, E. L., Reid, M. E. and Michael, J. A.: Hazard Analysis of Landslides Triggered by Typhoon Chata'an on July 2, 2002, in Chuuk State, Federated States of Micronesia, U.S. Geol. Surv. Open-File Rep., 2004-1348, 1–24, 2004.

Heroku: Heroku Cloud Application Platform, [online] Available from: https://aws.amazon.com/marketplace/pp/B008DJG1TY/ref=sp_mpg_product_title/178-1902068-2217163?ie=UTF8&sr=0-2 (Accessed 8 March 2015), 2015.

Hijmans, R. and van Etten, J.: raster: raster: Geographic data analysis and modeling, R Packag. version, 2, 2014.

Hong, Y., Adler, R. and Huffman, G.: Evaluation of the potential of NASA multi-satellite precipitation analysis in global landslide hazard assessment, Geophys. Res. Lett., 33(L22402), 1–5, doi:10.1029/2006GL028010, 2006.

Hong, Y., Adler, R. and Huffman, G.: Use of satellite remote sensing data in the mapping of global landslide susceptibility, Nat. Hazards, 43(2), 245–256, doi:10.1007/s11069-006-9104-z, 2007.

Hromadka II, T. V, Hromadka III, T. V and Phillips, M.: Use of Rainfall Statistical Return Periods to Determine Threshold for Mass Wasting Events, Environ. Eng. Geosci., XVI(4), 343–356, 2010.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F. and Wolff, D. B.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, J. Hydrometeorol., 8, 38–55, doi:10.1175/JHM560.1, 2007.

Huffman, G. J., Adler, R. F., Bolvin, D. T. and Nelkin, E. J.: The TRMM Multi-satellite Precipitation Analysis (TMPA), in Satellite Rainfall Applications for Surface Hydrology, edited by F. Hossain and M. Gebremichael, pp. 3–22, Springer Verlag., 2010.

INETER: Localización de deslizamientos en las comunidades Cerro Azul y La Gongolona, municipo El Ayote, 2014.

IPCC: Climate Change 2007: Working Group I: The Physical Science Basis, IPCC Fourt., edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Cambridge University Press, Cambridge and New York. [online] Available from: http://www.ipcc.ch/publications_and_data/ar4/wg1/en/contents.html, 2007.

Jackson, T. J. and Schmugge, T. J.: Vegetation effects on the microwave emission of soils, Remote Sens. Environ., 36(3), 203–212, doi:http://dx.doi.org/10.1016/0034-4257(91)90057-D, 1991.

Kirschbaum, D. B., Adler, R., Adler, D., Peters-Lidard, C. and Huffman, G.: Global Distribution of Extreme Precipitation and High-Impact Landslides in 2010 Relative to Previous Years, J. Hydrometeorol., 13(5), 1536–1551, doi:10.1175/JHM-D-12-02.1, 2012a.

Kirschbaum, D. B., Adler, R., Hong, Y., Hill, S. and Lerner-Lam, A.: A global landslide catalog for hazard applications: method, results, and limitations, Nat. Hazards, 52(3), 561–575, doi:10.1007/s11069-009-9401-4, 2010.

Kirschbaum, D. B., Adler, R., Hong, Y., Kumar, S., Peters-Lidard, C. and Lerner-Lam, A.: Advances in landslide nowcasting: evaluation of a global and regional modeling

approach, Environ. Earth Sci., 66(6), 1683–1696, doi:10.1007/s12665-011-0990-3, 2012b.

Kirschbaum, D. B., Stanley, T. and Yatheendradas, S.: Modeling Landslide Susceptibility over Large Regions with Fuzzy Overlay, Landslides, doi:10.1007/s10346-015-0577-2, 2015a.

Kirschbaum, D. B., Stanley, T. and Zhou, Y.: Spatial and Temporal Analysis of a Global Landslide Catalog, Geomorphology, doi:10.1016/j.geomorph.2015.03.016, 2015b.

Kohler, M. . and Linsley, R. .: Predicting the runoff from storm rainfall, Weather Bur. Res. Pap. No. 34, (September), 10 [online] Available from: http://pbadupws.nrc.gov/docs/ML0819/ML081900279.pdf, 1951.

Lagomarsino, D., Segoni, S., Fanti, R. and Catani, F.: Updating and tuning a regional-scale landslide early warning system, Landslides, 10(1), 91–97, doi:10.1007/s10346-012-0376-y, 2013.

Larsen, M. C. and Simon, A.: A Rainfall Intensity-Duration Threshold for Landslides in a Humid-Tropical Environment, Puerto Rico, Geogr. Ann. Phys. Geogr., 75(1/2), 13–23, 1993.

Lee, S. and Pradhan, B.: Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models, Landslides, 4, 33–41, doi:10.1007/s10346-006-0047-y, 2007.

Li, C., Ma, T., Zhu, X. and Li, W.: The power – law relationship between landslide occurrence and rainfall level, Geomorphology, 130(3-4), 221–229, doi:10.1016/j.geomorph.2011.03.018, 2011.

Li, T., Li, P. and Wang, H.: Forming Mechanisms of Landslides in the Seasonal Frozen Loess Region in China, in Landslides in Cold Regions in the Context of Climate Change, edited by W. Shan, Y. Guo, F. Wang, H. Marui, and A. Strom, pp. 41–52, Springer., 2013.

Liao, Z., Hong, Y., Kirschbaum, D. and Liu, C.: Assessment of shallow landslides from Hurricane Mitch in central America using a physically based model, Environ. Earth Sci., 66(6), 1697–1705, doi:10.1007/s12665-011-0997-9, 2012.

Martelloni, G., Segoni, S., Fanti, R. and Catani, F.: Rainfall thresholds for the forecasting of landslide occurrence at regional scale, Landslides, 9(4), 485–495, doi:10.1007/s10346-011-0308-2, 2012.

Mathew, J., Babu, D. G., Kundu, S., Kumar, K. V. and Pant, C. C.: Integrating intensity–duration-based rainfall threshold and antecedent rainfall-based probability estimate towards generating early warning for rainfall-induced landslides in parts of the Garhwal Himalaya, India, Landslides, 11(4), 575–588, doi:10.1007/s10346-013-0408-2, 2014.

Montgomery, D. R. and Dietrich, W. E.: A physically based model for the topographic control on shallow landsliding, Water Resour. Res., 30(4), 1153, doi:10.1029/93WR02979, 1994.

Montrasio, L., Valentino, R. and Losi, G. L.: Towards a real-time susceptibility assessment of rainfall-induced shallow landslides on a regional scale, Nat. Hazards Earth Syst. Sci., 11(7), 1927–1947, doi:10.5194/nhess-11-1927-2011, 2011.

Mora, S. C.: Extent and Socio-Economic Significance of Slope-Instability on the Island of Hispaniola (Haiti and Dominican Republic), in Energy and Mineral Potential of Central America-Caribbean Regions, edited by R. L. Miller, G. Escalante, J. A. Reinemund, and M. J. Bergin, Springer-Verlag, Berlin, Heidelberg., 1995.

Mora, S. C. and Vahrson, W.-G.: Macrozonation Methodology for Landslide Hazard Determination, Bull. Assoc. Eng. Geol., XXXI(1), 49–58, 1994.

Nadim, F., Cepeda, J., Sandersen, F., Jaedicke, C. and Heyerdahl, H.: Prediction of Rainfall-Induced Landslides through Empirical and Numerical Models, in Rainfall-Induced Landslides: mechanisms, monitoring techniques and nowcasting models for early-warning systems, pp. 1–10, First Italian Workshop on Landslides, Naples., 2009.

Nadim, F., Kjekstad, O., Peduzzi, P., Herold, C. and Jaedicke, C.: Global landslide and avalanche hotspots, Landslides, 3, 159–173, doi:10.1007/s10346-006-0036-1, 2006.

Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K. and Nghiem, S. V: Soil moisture retrieval from AMSR-E, Geosci. Remote Sensing, IEEE Trans., 41(2), 215–229, 2003.

Pradhan, B. and Lee, S.: Landslide susceptiblity assessment and factor effect analysis: backpropagation artifical neural networks and their comparison with frequency ratio and bivariate logistic regression modelling, Environ. Model. Softw., 25, 747–759, doi:10.1016/j.envsoft.2009.10.016, 2010.

Ray, R. L. and Jacobs, J. M.: Relationships among remotely sensed soil moisture, precipitation and landslide events, Nat. Hazards, 43, 211–222, doi:10.1007/s11069-006-9095-9, 2007.

Rossi, M., Kirschbaum, D., Luciani, S. and Guzzetti, F.: Comparison of TRMM satellite rainfall estimates with rain gauge data and landslide empirical rainfall thresholds under different morphological and climatological conditions in Italy, in EGU General Assembly, p. 9354, 22-27 April, Vienna, Austria. [online] Available from: http://adsabs.harvard.edu/abs/2012EGUGA..14.9354R, 2012.

Saito, H., Nakayama, D. and Matsuyama, H.: Relationship between the initiation of a shallow landslide and rainfall intensity — duration thresholds in Japan, Geomorphology, 118(1-2), 167–175, doi:10.1016/j.geomorph.2009.12.016, 2010.

Segoni, S., Lagomarsino, D., Fanti, R., Moretti, S. and Casagli, N.: Integration of rainfall thresholds and susceptibility maps in the Emilia Romagna (Italy) regional-scale landslide warning system, Landslides, 1–13, doi:10.1007/s10346-014-0502-0, 2014.

Swenson, S., Famiglietti, J., Basara, J. and Wahr, J.: Estimating profile soil moisture and groundwater variations using GRACE and Oklahoma Mesonet soil moisture data, Water Resour. Res., 44(1), W01413, doi:10.1029/2007WR006057, 2008.

Tatard, L., Grasso, J. R., Helmstetter, A. and Garambois, S.: Characterization and comparison of landslide triggering in different tectonic and climatic settings, J. Geophys. Res., 115(F04040), 1–18, doi:10.1029/2009JF001624, 2010.

R Core Team~~, R. C.~~: R: A language and environment for statistical computing, R Found. Stat. Comput., Version 3.0.1 [online] Available from: http://www.r-project.org/ (Accessed 16 May 2013), 2013.

Terlien, M. T. J.: The determination of statistical and deterministic hydrological landslide-triggering thresholds, Environ. Geol., 35(2-3), 124–130, 1998.

Tiranti, D. and Rabuffetti, D.: Estimation of rainfall thresholds triggering shallow landslides for an operational warning system implementation, Landslides, 7(4), 471–481, doi:10.1007/s10346-010-0198-8, 2010.

Verdin, K. L., Godt, J., Funk, C., Pedreros, D., Worstell, B. and Verdin, J.: Development of a Global Slope Dataset for Estimation of Landslide Occurrence Resulting from Earthquakes, U.S. Geol. Surv., Open-File , 1–29, 2007.

Wieczorek, G. F.: Effect of rainfall intensity and duration on debris flows in central Santa Cruz Mountains, in Debris Flows/Avalanches: Process, Recognition, and Mitigation, Reviews in Engineering Geology, vol. 7, edited by J. E. Costa and G. F. Wieczorek, pp. 93–104, Geological Society of America, Boulder, CO., 1987.

Table 1. Data sources used in the LHASA model.

Data Type	Data Set	Resolution	Extent	Source and Details
Slope	70th ~~quantile~~ percentile slope, USGS	30 arc-seconds (~1 km)	65 degrees N-S	(Verdin et al., 2007) derived from 3-arc-second SRTM DEM
Soils	Harmonized World Soil Database	30 arc-seconds, nominal scale (1:5,000,000)	Global	(FAO/IIASA/ISRIC/ISSCAS/JRC, 2012)
Roads	Global Roads Open Access Data Set, Version 1	Multiple sources (accuracy ranges from 30m to 1265m)	Global	~~(~~(CIESIN and ITOS ~~Center for International Earth Science Information Network (CIESIN)/Columbia University and Georgia,~~ 2013)~~CIESIN and ITOS, 2013)~~
Fault zones	Map Showing Geology, Oil and Gas Fields, and Geologic Provinces of the Caribbean Region	1:2,500,000	Central America and Caribbean	(French and Schenk, 2004)
Rainfall	TRMM Multi-satellite Precipitation Analysis Version 7, Real-Time	0.25 degree, 3-hourly resolution	50 degrees N-S	http://pmm.nasa.gov
Landslide Inventory	GLC, 322 points in region	Accuracy is defined on a point-by-point basis	Global	(Kirschbaum et al., 2010)
	MARN, 297 points	Various mapping scales and survey types	El Salvador	(Gerencia de Geología, 2012)

Table 2. True Positive Rate considered over 0, 1, 5 km buffers and a variable buffer based on the reported location accuracy. Temporal windows of 1, 3 and 7-days were also evaluated. Results are shown in percentages, with high hazard TPR percentages shown in parentheses. The susceptible pixels column (far right) shows the percentage of reported landslides that have a Susceptibility Index of 2 or greater within the spatial buffer considered, indicating the maximum TPR that could be generated based on the rainfall and antecedent thresholds of the model.

	Spatial Buffer Distance	1-Day Window	3-Day Window	7-Day Window	Susceptible pixels
Central America 2007-2013	0 km	64 (26)	77 (37)	83 (47)	90
	1 km	67 (28)	81 (39)	87 (50)	92
	5 km	72 (34)	85 (48)	93 (59)	100
	Variable	81 (40)	89 (57)	94 (65)	100
Hispaniola 2007-2013	0 km	21 (17)	29 (21)	46 (21)	50
	1 km	33 (21)	67 (29)	67 (29)	75
	5 km	46 (21)	71 (46)	88 (50)	100
	Variable	54 (21)	71 (50)	88 (54)	96
Central America 2014	0 km	58 (12)	74 (33)	79 (47)	86
	1 km	63 (12)	79 (35)	84 (49)	93
	5 km	72 (14)	86 (44)	91 (56)	100
	Variable	67 (12)	84 (37)	91 (51)	91
Hispaniola 2014	0 km	43 (43)	57 (43)	71 (71)	86
	1 km	57 (43)	71 (43)	86 (71)	100
	5 km	71 (43)	86 (43)	86 (71)	100
	Variable	71 (43)	86 (43)	86 (71)	100

Table 3. False Positive Rate ~~considered over 0, 1, 5 km buffers and a variable buffer based on the reported location accuracy~~ and Distance to Perfect Classification (r_j) for both high-hazard and moderate-hazard nowcasts.~~. Temporal windows of 1, 3 and 7-days were also evaluated.~~ The moderate-hazard model was calibrated for the period 2007-2013 in Central America, so it is not surprising that this location has the lowest r_j value. Because the high-hazard nowcast is a heuristic intended to identify only the most ~~dangerous~~ extreme conditions while minimizing false alarms, r_j values for this model are expected to be large.~~Results are shown in percentages, with high hazard FPR percentages shown in parentheses. The susceptible pixels column (far right) shows the percentage of reported landslides that have a Susceptibility Index of 2 or greater within the spatial buffer considered, indicating the maximum FPR that could be generated based on the rainfall and antecedent thresholds of the model.:~~

	Hazard Level	FPR	r_j
Central America 2007-2013	Moderate	11%	0.38
	High	1%	0.74
Hispaniola 2007-2013	Moderate	9%	0.80
	High	1%	0.83
Central America 2014	Moderate	9%	0.43
	High	1%	0.88
Hispaniola 2014	Moderate	7%	0.57
	High	1%	0.57

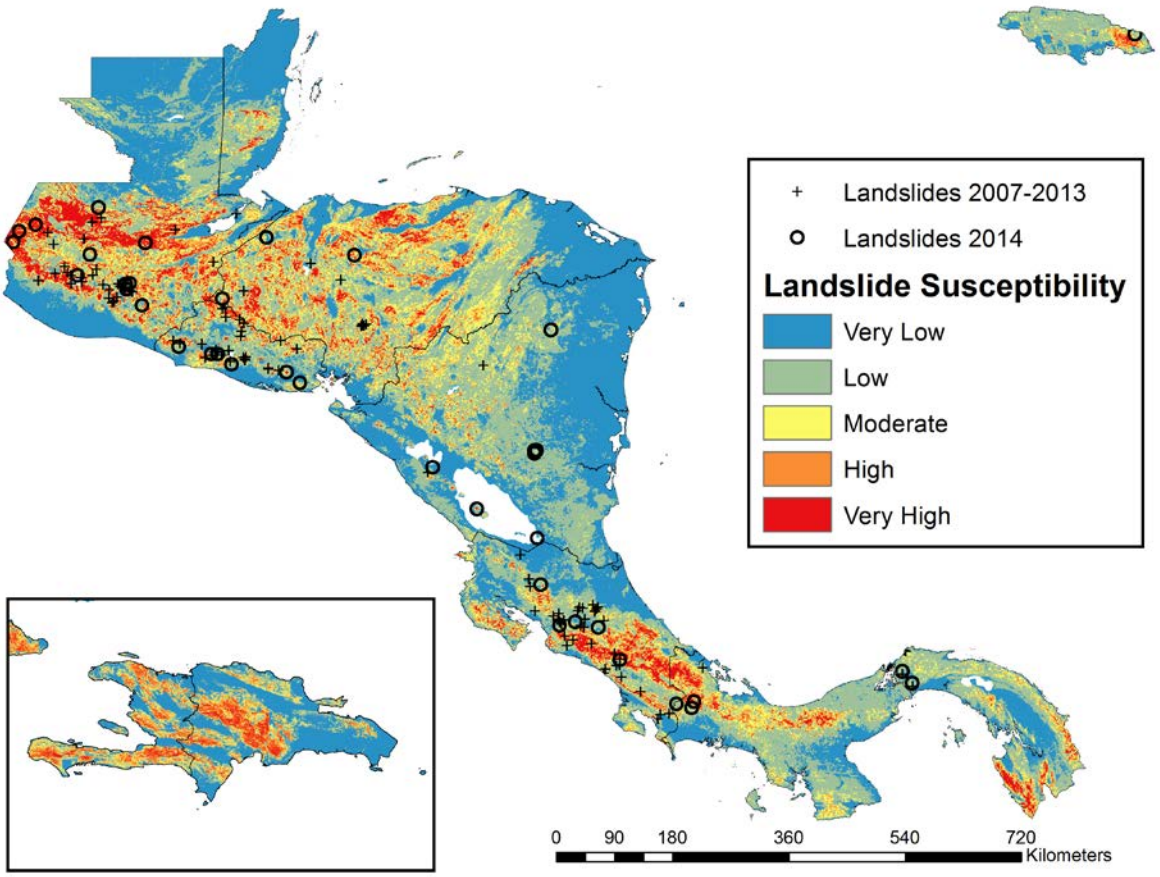

Figure 1. Regional landslide susceptibility map created using a fuzzy overlay methodology from global slope, soil, and road databases, as well as a regional map of faults at a resolution of 30 ~~arcseconds~~arc-seconds (approximately 1-kilometre) ~~resolution~~ (Kirschbaum et al., 2015a). Landslides recorded in the study areas are shown for the years 2007-2013 (+) and 2014 (o). 166 landslide reports are from the Global Landslide Catalog and 24 landslides in El Salvador were selected from the MARN catalog. Large numbers of landslides are located near capital cities, implying some degree of reporting bias. Inset map shows landslide susceptibility map for Hispaniola with 31 reported landslides.

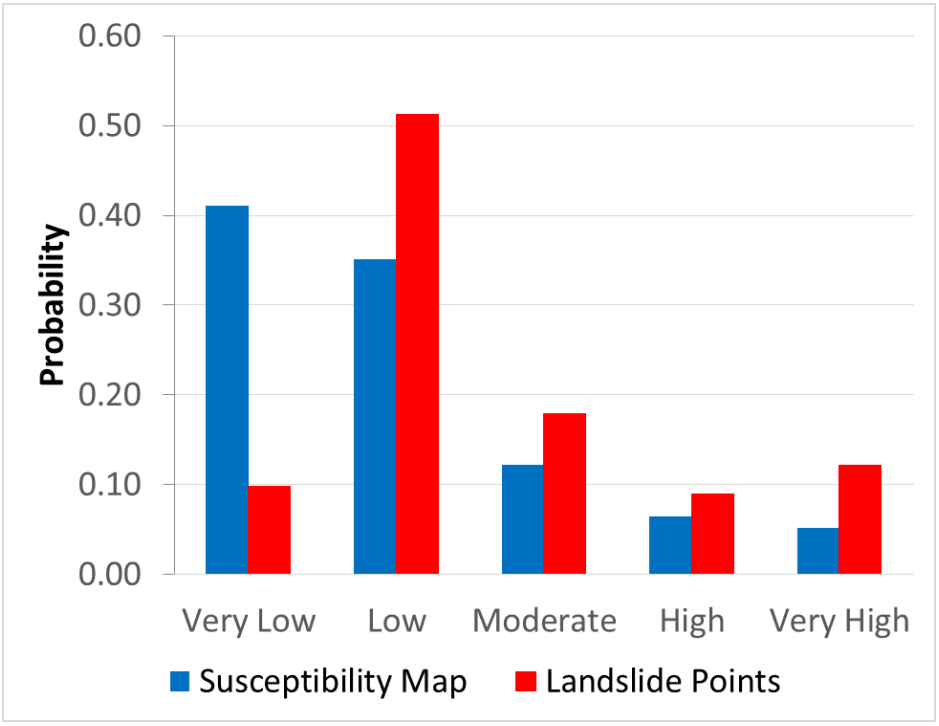

Figure 2: Probablity distribution of susceptibility values at landslide locations reported from 2007 to 2013 compared to landslide susceptibility values for all of Central America. Landslides occurred in all susceptibility categories, but few landslides (<10%) occurred in the lowest category. The LHASA model used a threshold of "low" susceptibility or greater (SI ≥ 2) with rainfall and antecedent rainfall thresholds within the decision tree framework (Figure 6). ~~An~~ SI ≥ 2 (low) was chosen to exclude a large portion of Central America without losing the ability to predict most landslide events.

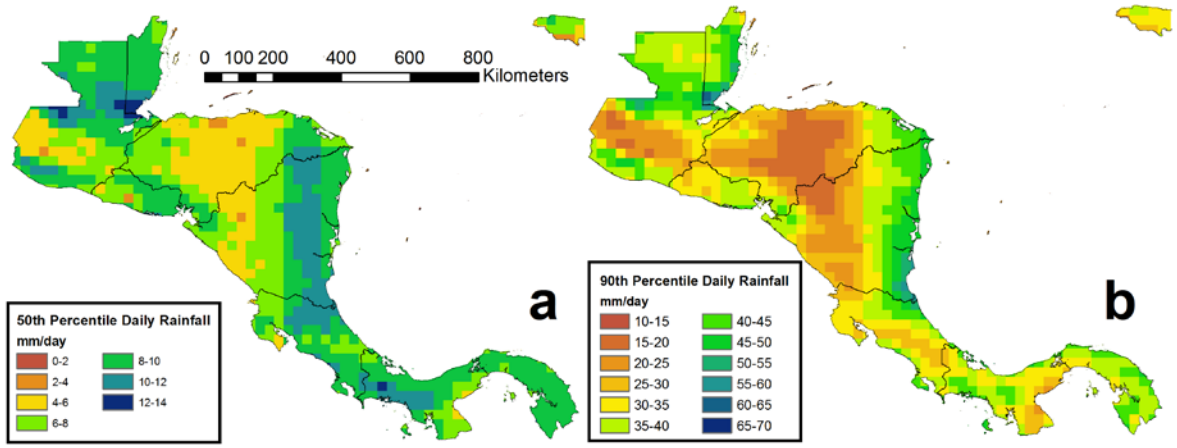

Figure 3. Maps showing rainfall values in mm/day for a) 50th and b) 90th percentiles calculated from TMPA daily rainfall estimates covering the years 2001-2013. These two percentile maps were used within the decision tree framework to produce moderate landslide hazard nowcasts.

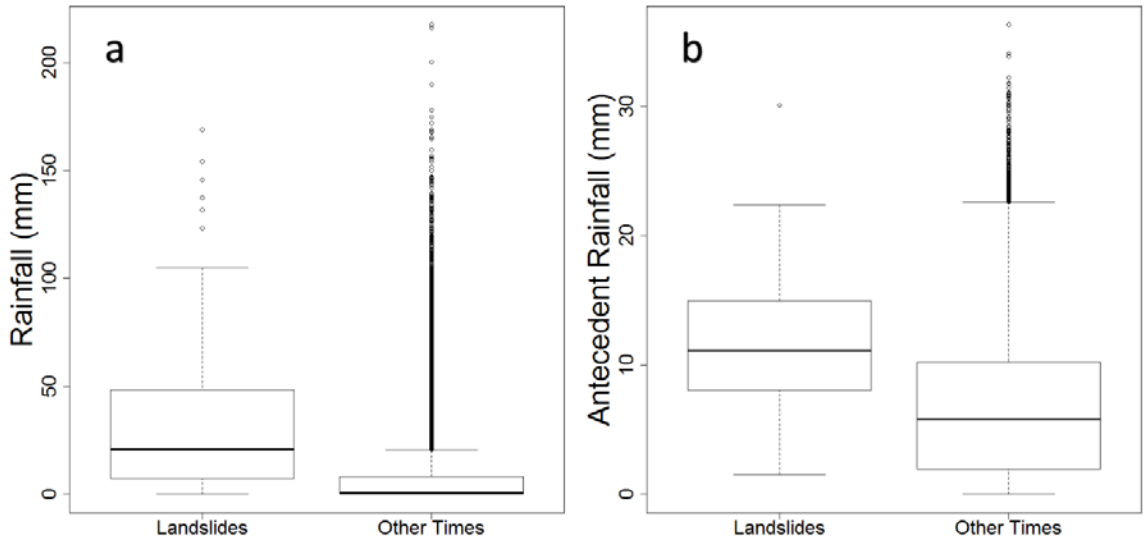

Figure 4. Comparison of a) daily rainfall and b) antecedent rainfall distributions (in mm) for days with and without reported landslides. Results are plotted only for locations ~~the~~ where landslides were historically recorded and include 123 landslides from Central America and 316,697 data points from ~~other~~ times landslides were not reported. Due to the limitations in the GLC, it is likely that unreported landslides may have occured in the category we are classifying as "other times". The distributions substantially overlap, suggesting that daily and antecedent rainfall thresholds cannot classify the data perfectly into landslide events and non-events.

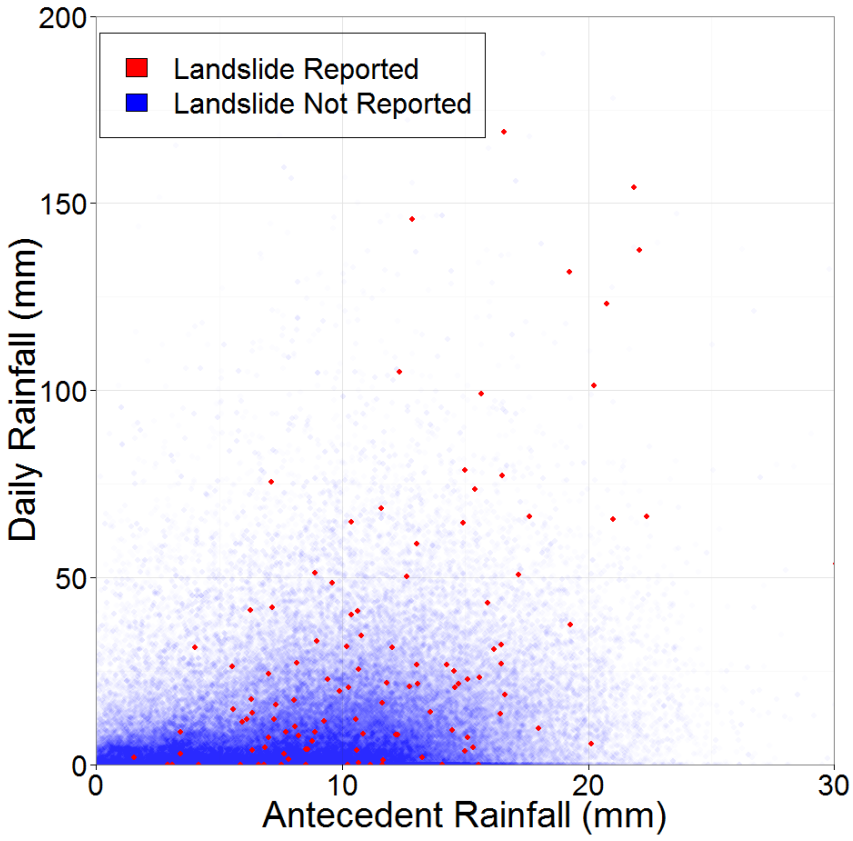

Figure 5. Scatter plot showing the distribution of landslides (red) and dates without recorded landslides (blue) comparing antecedent rainfall and daily rainfall.

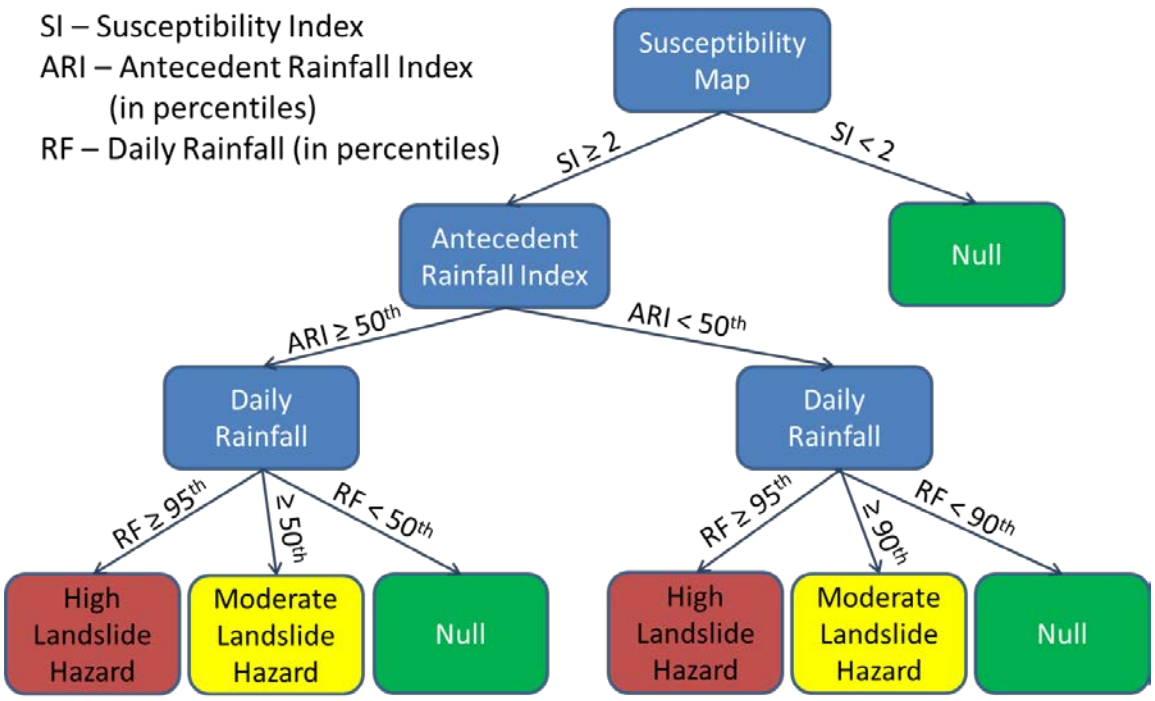

Figure 6. Decision tree structure highlighting the three tiers of decisions made in this model on a pixel-by-pixel basis across the study area, which is computed daily. First, if a pixel has an SI ≥ 2), the antecedent rainfall index (ARI) is considered using the 50th percentile value. If the ARI ≥ 50th percentile, a nowcast is issued if the daily rainfall exceeds the 50th percentile (moderate hazard) or 95th percentile (high hazard). If the ARI < 50th percentile, a nowcast is issued if the daily rainfall exceeds the 90th percentile (moderate hazard) or 95th percentile (high hazard). In all other cases, there is no nowcast issued (Null).

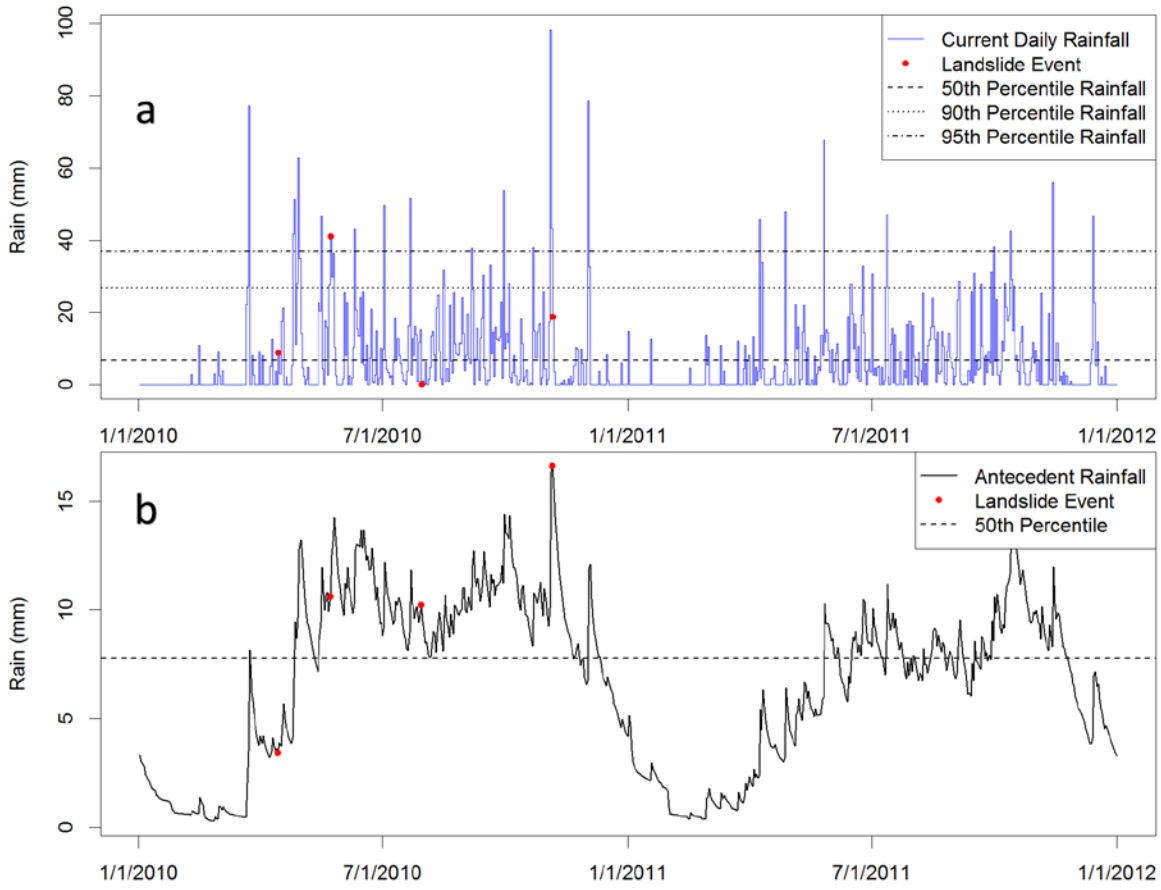

Figure 7: Example time series of a) rainfall and b) antecedent rainfall for a pixel west of San Jose, Costa Rica. The time series highlight 4 landslides that were reported in this area in 2010: April 14th, May 22nd, July 30th, and November 5th (shown as red ~~crosses~~circles). The first landslide (Figure ~~8a~~9a) occurred very early in the year, before the peak rainy season and a nowcast was not issued. The series indicates that a moderate hazard nowcast was triggered for the November 5th event and a high hazard nowcast was triggered for the May 22nd event

(Figure ~~8b~~9b). However, due to the antecedent rainfall for April 14th event and low rainfall totals for the July 30th event, no nowcasts were issued.

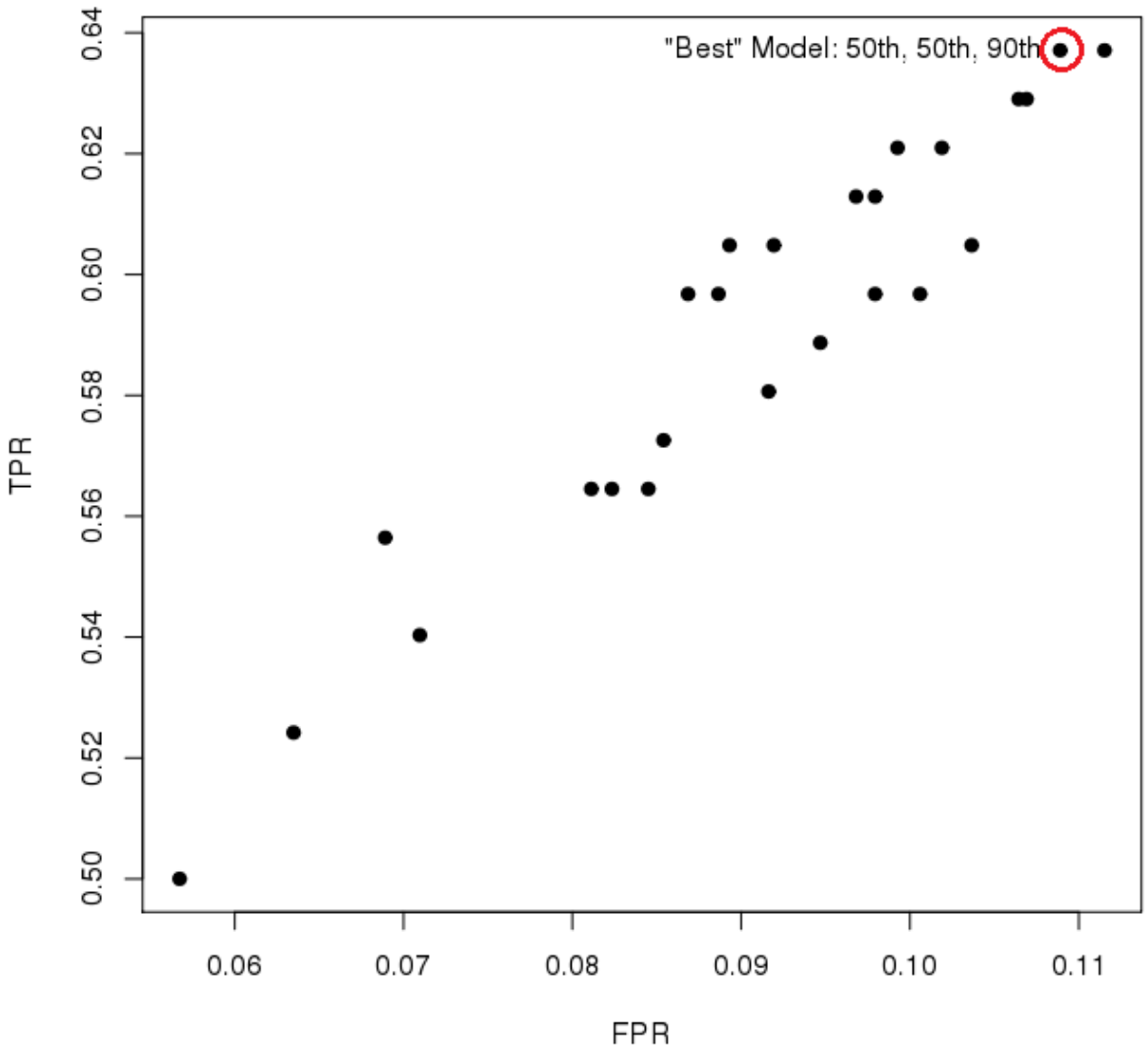

Figure 8. The moderate-hazard nowcast was calibrated by calculating r_j, the distance to perfect classification, for a variety of rainfall and ARI thresholds that ranged from the 50th to the 95th percentile values. The lowest (best) r_j value was observed for the 50th percentile ARI, 50th percentile rainfall over moist-condition ground, and 90th percentile rainfall over ~~moist~~dry-condition ground. These thresholds are represented by the second dot from the right in the red circle. Ideally, the choice of the "best" rainfall thresholds depends upon the intended use of the model and the economic costs of different errors. Since these details are not yet available, r_j was selected to provide a balanced and easily interpreted measure of model success.

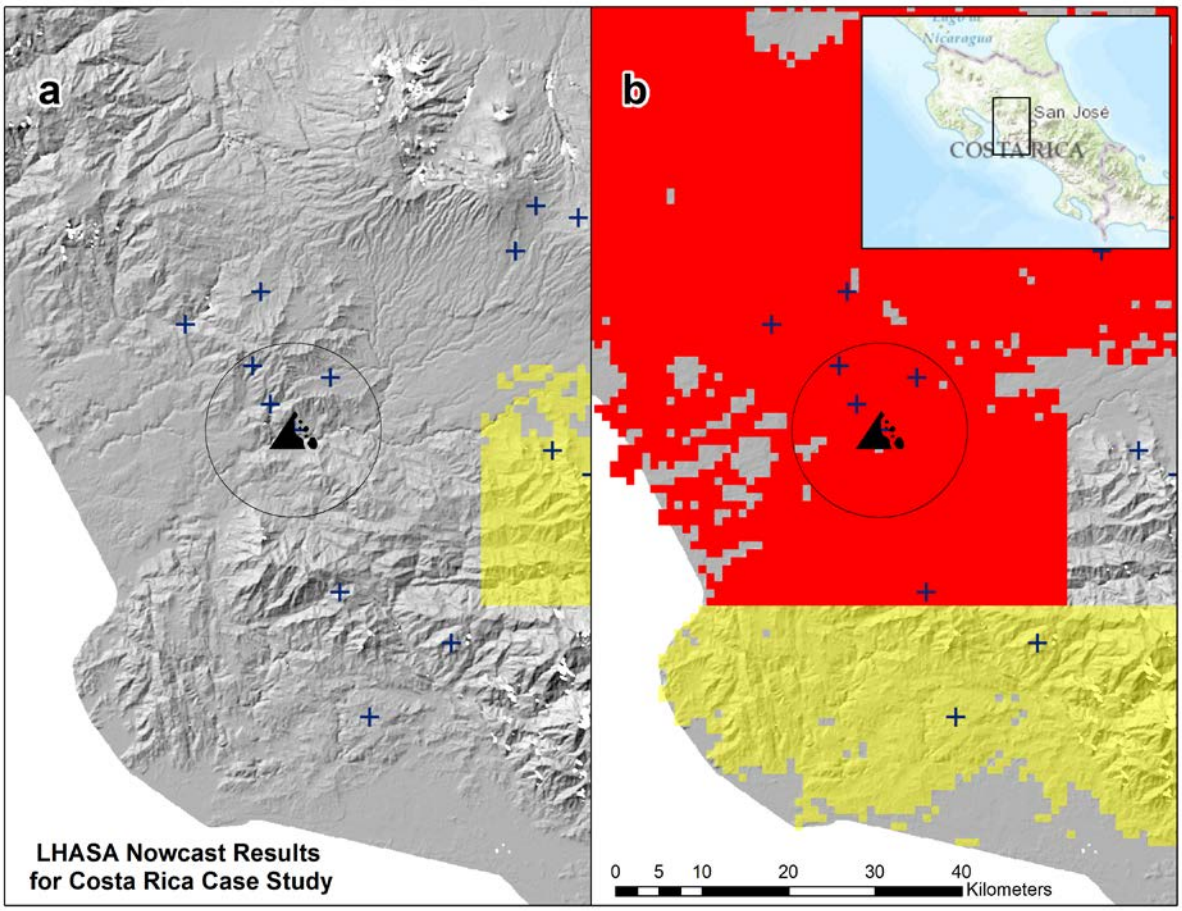

Figure ~~8~~9. This shows the results of the model nowcasts for a cluster of rockfalls (denoted by a single black rockfall icon) that occurred within the same area of the highway near Caldera, Costa Rica for two dates in 2010: a) April 14th and b) May 22nd. The black circle indicates the estimated maximum spatial error of these landslide reports (10 km), suggesting that the landslides could have occurred anywhere within that area. Blue crosses indicate the locations of other landslides in the GLC from 2007-2013. The April 14th event did not generate a moderate hazard nowcast (yellow), but enough rainfall was observed to the east of the landslide location to trigger a moderate hazard nowcast. Comparatively, the May 22nd event (b) shows that many high and moderate hazard nowcasts were generated within the proximity of the reported event.

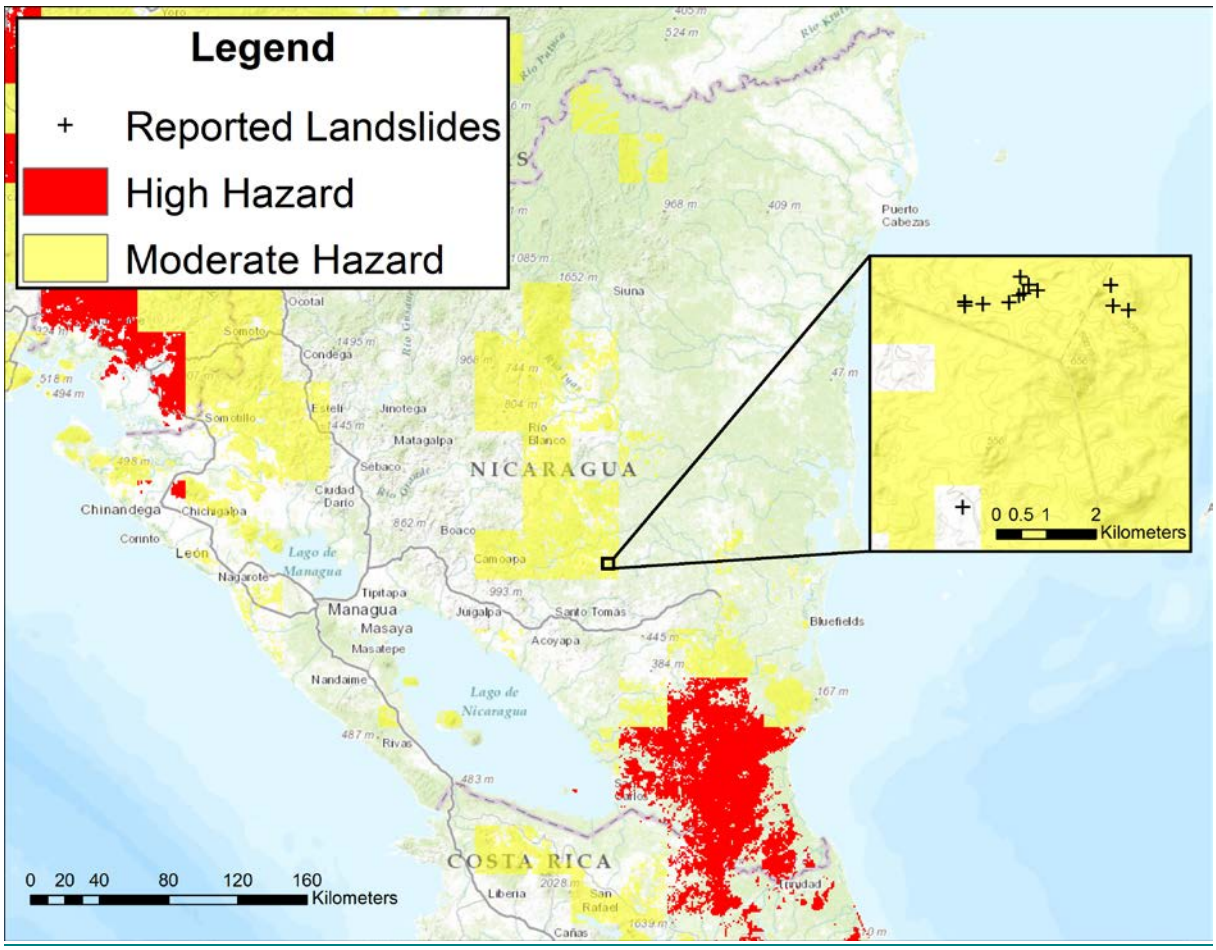

Figure ~~9~~10: Landslide hazard potential for 6/23/2014. Black crosses indicate locations where a cluster of landslides occurred near El Ayote, Nicaragua. Yellow pixels (moderate hazard) and red pixels (high hazard) are shown for that day. 13 of 14 landslides were predicted by the moderate hazard category. The southernmost landslide was located in a relatively flat location that had been mapped as having "very low" landslide susceptibility, so it was not predicted, despite daily rainfall exceeding the 50th percentile threshold.