Landslide Hazard and Exposure Modelling in Data-Poor Regions: The Example of the Rohingya Refugee Camps in Bangladesh

R. A. Emberson1,2,3, D. B. Kirschbaum1, and T. Stanley1,2,3

1Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA, 2Universities Space Research Association, Columbia, MD, USA, 3Goddard Earth Sciences Technology and Research, Columbia, MD, USA

Abstract
Landslide hazards significantly affect economies and populations around the world, but locations where the greatest proportional losses occur are in data-poor regions where capacity to estimate and prepare for these hazards is most limited. Earth observation (EO) data can fill key knowledge gaps, and can be rapidly used in settings with lower analytical capacity. In this study, we describe a novel series of methods designed to analyze landslide susceptibility, hazard and exposure in the region in and around the Rohingya refugee camps in Bangladesh, where limited data is juxtaposed with a major humanitarian crisis. We demonstrate that a high degree of accuracy is possible even when estimating susceptibility of relatively small landslides. In the context of this example, we also explore how estimates of landslide hazard and exposure are most beneficial to decisions made by humanitarian stakeholders relevant to natural hazards and risk. The unique opportunity to work alongside humanitarian end-users has allowed us to produce focused products that can be tested while in development. In particular, we stress the importance of communicating the difference between a landslide “early warning system”—for which satellite data may be unsuitable at local scales—and a model that provides relative hazard estimates, where EO may be valuable. The toolbox of methods presented here could be used to generate landslide hazard and exposure maps in other data-poor regions around the globe.

1. Introduction

Human settlements and infrastructure affected by landslide hazards are spread widely around the globe, in essentially every region where high topographic relief intersects human societal footprints. The net result is that landslides cause hundreds to thousands of fatalities each year, as well as billions of dollars of damage (Froude & Petley, 2018; Kjekstad & Highland, 2009; Petley, 2012). While landslides are distributed globally, their impact is felt locally; only in extreme cases does the footprint of a single landslide extend beyond more than a few kilometers from its point of origin (Dade & Huppert, 1998; Davies, 1982; Roback et al., 2018). In well-instrumented regions where resources have been allocated to mitigate landslide exposure, vulnerability and risk, local geotechnical information has proven to be invaluable for such efforts (Michel et al., 2014; Von Ruette et al., 2013). However, the value of such detailed measurements must be weighed against the cost of obtaining them, leading to data gaps in many vulnerable areas around the world. Where local high-resolution data are lacking, researchers have looked to satellite-based Earth observations (EOs) for comprehensive mapping. Satellite data have been used for mapping landslides after catastrophic earthquakes and tropical cyclones (e.g., Marc et al., 2018; Roback et al., 2018; Tanyas et al., 2017). Relative to in situ data some limitations exist in using satellite data as inputs for local assessment of landslide hazard. Lower spatial resolution of publicly available datasets, such as topography, place a limit on how finely terrain features can be resolved. Shuttle Radar Topography Mission (SRTM; Slater et al., 2006) data can resolve terrain features greater than 30 m, far lower than field-based estimates from Light Detection and Ranging (<8 m), for example. High resolution Digital Elevation Models (DEMs) derived from commercial stereo pairs (e.g., Shean et al., 2016) and DEMs derived from Synthetic Aperture Radar (SAR) (e.g., TanDEM-X DEM, Wessel et al., 2018) can provide sub-8 m resolution imagery; however this data often comes at a high cost and/or is limited in its coverage. However, EO data have a handful of key advantages; repeated and widespread coverage means no region is excluded, and the larger-scale topographic signatures of landsliding can be observed.
Additionally, many parts of the developing world may lack sufficient local expertise or financial resources to conduct field scale mapping of slope, soil and lithology characteristics, and hydrological parameters to a sufficient extent for in situ landslide hazard mapping. This is especially significant given that the deaths associated with landslides disproportionately affect many countries with lower levels of economic development (D. B. Kirschbaum et al., 2015). The dearth of this important data can be even more pronounced in areas where humanitarian crises have led to the displacement of populations. For example, refugee camps are often established in previously uninhabited areas where there may be no existing analysis of landslide hazard. The Rohingya refugee camps in Southern Bangladesh, which have grown exponentially since 2017 (UNDRR, 2019), have been built on steep hillslopes where a combination of cyclonic and monsoonal rainfall has historically caused extensive landsliding (Khan et al. 2012; Petley, 2012). In locations similar to these, satellite data may be the only available resource from which a landslide susceptibility and hazard analysis can be conducted.

A recent review of published studies (Reichenbach et al., 2018) highlights the wide range of methods used to build maps of landslide susceptibility. Reichenbach et al. (2018) also point out that only a fraction of these studies exploit EO data, and the bulk of these studies estimate susceptibility for regions smaller than 10,000 km², limiting the applicability of any given study to a wider region. This further highlights the potential for data-poor areas to remain unmapped or void of detailed hazard information.

As such, we suggest that it is important to establish a rapid, reproducible and robust method to use EO and other freely available data to produce landslide hazard and exposure assessments where little to no other data exist. This work focuses on an area in Southern Bangladesh with high-landslide hazard potential and a highly vulnerable population within and proximate to refugee camps associated with the Rohingya refugee crisis. In response to needs identified in discussion with in-country humanitarian stakeholders, we have developed a toolbox of methods to estimate landslide hazard and exposure using EO and other global, open-access data. While the initial application of these tools is to the areas in and around the refugee camps in Bangladesh, the toolbox of methods incorporates machine-learning techniques that can incorporate diverse datasets for application in other settings. This toolbox presents methodologies that can be used by a range of stakeholders with differing data availability situations and decision making needs in other data-poor regions.

A fundamental and important aspect of this work is engagement with regional and local stakeholders to ensure that the most effective and actionable products are created. From our experience, we found it crucial to have open discussions with humanitarian stakeholders to define the nature of the model outputs and their utility for hazard-relevant decisions. We adopt this approach in designing experimental landslide products produced within the Rohingya camps by United Nations agencies, which has provided a useful testing ground while also assisting with a humanitarian crisis. We stress the importance of this collaborative aspect to ensure that scientific data can be used most effectively in vulnerable locations.

2. Context

As a result of social and ethnic strife in Myanmar, hundreds of thousands of people from the Rohingya ethnic group have fled their homes in Myanmar since 2016. At the time of writing, these groups are residents of refugee camps in southern Bangladesh. The large refugee camps in southern Bangladesh are situated in an area where landsliding is prevalent (D. B. Kirschbaum et al., 2010). Intense rainfall from both the Indian Summer Monsoon and tropical storms act as potent triggers for landslides in areas with steep slopes and poorly consolidated (Islam & Uyeda, 2007; Shahid, 2011).

The addition of hundreds of thousands of refugees has not only increased susceptibility to landslides through extensive deforestation of local hillslopes (UNDP Bangladesh & UN WOMEN, Bangladesh, 2018), but also placed a large and vulnerable population in harm’s way. Through collaboration with partners at Columbia International Research Institute for Climate and Society we have built connections with stakeholders including UNDP, UNHCR, and IOM, who are involved with management of the refugee camps and disaster mitigation therein. Multilateral discussions with stakeholders have enabled a clearer picture of the needs in data-poor humanitarian crisis zones as related to landslide hazards, and have informed the
development of our model. Continual testing and iterative co-development of our products alongside humanitarian end-users allowed us to clearly identify the limitations and potential of an EO-based modeling approach for landslide hazard and exposure. In Section 4, we discuss the importance of the scientist-end-user connection in the co-development of our models.

It is crucial in any humanitarian context to clarify the terminology surrounding natural hazards and risks to ensure that any products are appropriately communicated and used. This is especially important in the context of EO data as the key determinants of risk—hazard, exposure and vulnerability—differ greatly in how easily they can be distinguished from satellite-derived data. In the following section, we briefly discuss these parameters and describe existing efforts to document them using EO data for landslide applications.

2.1. Hazard

Hazard is the physical manifestation of the landslide event, before consideration of human elements such as infrastructure or local population. Burton et al. (1978) define hazard as: “the threat potential posed to man or nature by events originating in, or transmitted by, the natural or built environment.” Thus, for landslides, this should consider the extent (both in terms of area and volumetrically) of the landslide, as well as how fast the onset of landsliding is e.g., (Fell et al., 2008). At a very local scale, many of these parameters can be estimated and linked to the magnitude of a given triggering event (seismic or rainfall-induced) using geophysical methods (G. B. Crosta & Prattini, 2003; Iverson, 2000; Marc et al., 2016), but when assessment time is limited or spatial scales are larger these methods become impractical.

As a result, the probability and severity of landslide is often lumped into a singular static metric of “susceptibility,” which describes the propensity of a given hillside to fail due to a triggering event. In general, metrics for susceptibility incorporate a range of input variables that affect slope stability and landslide occurrence, and apply either subjectively or empirically defined weighting to combine them into a single proxy for landslide probability (Reichenbach et al., 2018).

On a physical basis, landslides occur when gravitational forces acting on a mass of hillslope material exceed the frictional forces retaining the material in place. Geophysical or engineering methods aim to measure the parameters that determine these forces; slope angle, soil cohesiveness, soil saturation, material strength, among others. EO data must necessarily approximate some of these variables using proxy data, especially if the resolution of the available data exceeds the size of the hillslopes. Prior research has used EO-derived digital elevation models or digital surface models (DEMs/DSMs) to estimate slope angle, slope roughness, altitude or relief as key landslide susceptibility factors (Costanzo et al., 2012; Reichenbach et al., 2018; Stanley & Kirschbaum, 2017).

Other inputs that EO have yet to measure can be gleaned from global open-source data. Hillslope material parameters are often approximated using global or local maps of subsurface lithology (Stanley & Kirschbaum, 2017; van Westen et al., 2003) and/or soil and regolith (Nandi & Shakoor, 2010). Other data include distance to major geologic faults, distance to roads (Stanley & Kirschbaum, 2017) and land use categorization (Lee & Talib, 2005). The relative importance of each of these parameters is estimated heuristically based on expert assessment (Stanley & Kirschbaum, 2017), or using statistical methods, with a set of landslides mapped within the study area to calibrate and estimate accuracy (e.g., van Westen et al., 2003).

The input variables derived from EO or freely available data discussed so far are generally static on the timescales over which landslides are triggered—i.e., there is not repeat mapping or they do not change on timescales of hours to weeks. As such, susceptibility derived from such data should be contextualized as “susceptibility to a given trigger for landsliding.” In most cases, this is either seismic shaking or intense rainfall. To transition from a susceptibility model to estimates of hazard, such triggering factors need to be incorporated. Earthquakes can trigger widespread landsliding in many parts of the world with active seismicity; however, predicting these sudden events is challenging. For this study, we consider rainfall-triggered landslides since we can approximate the conditions using EO data.

Some regional models incorporate weather data from ground-based radar or other precipitation estimates to obtain a more complete picture of landslide hazard, where landslide hazard is elevated based upon both
short term triggering rainfall and longer term antecedent rainfall (Guzzetti et al., 2008), both of which affect the internal frictional state of a hillslope depending on hydraulic parameters. Many versions of antecedent rainfall and immediate rainfall thresholds exist (G. Crosta, 1998; Glade et al., 2000; Guzzetti et al., 2007; Jakob & Weatherly, 2003; Martelloni et al., 2012), which can make it difficult to generalize local models to other regions.

Using the vantage point of space, EO data can provide consistent global estimates of surface and atmospheric conditions that trigger landslides. NASA’s Global Precipitation Measurement (GPM) satellite precipitation product IMERG (Integrated Multi-satellite Retrievals for GPM) is available globally every 30 min (Huffman et al., 2013). EO data from moderate resolution optical imaging satellites like Landsat or Sentinel-2 have been used to map landslides (Marc et al., 2018), and higher resolution data from commercial optical imaging satellite products offer promise to provide more refined landslide maps.

Global, real-time landslide hazard estimates provided by the LHASA model represent the most expansive and rapid-response use of satellite data in landslide assessment to date, but the relatively coarse scale of the LHASA output product (∼1 km grid cells) limits the range of applications to more regional considerations. Moreover, because the output of the model is landslide hazard, it may not directly provide the risk or exposure information required in humanitarian scenarios. Satellite-based methods to estimate the exposure and vulnerability of infrastructure are generally less developed than those to estimate hazard, and we briefly summarize some of these below.

2.2. Exposure and Vulnerability

To ascertain risk to structures, infrastructure and populations, risk analysts often delineate where the geographic footprints of such elements intersect with a defined hazard footprint, and multiply that by a defined metric for vulnerability—representing the proportion of the value (dollars, lives, or other metric) of the element expected to be lost as a result of the hazard—to obtain the estimated risk (UNDRR, 2019). While exposure is generally a mapping exercise, and therefore more amenable to satellite based or other freely available methods, vulnerability incorporates a whole range of factors depending on the element in question, and is thus more challenging and problematic to determine with satellite-derived products (Corominas et al., 2014).

Satellite-based estimates of land use and urban extent (Friedl et al., 2010; Schneider et al., 2009) can be used as proxies for infrastructure or population density, and have been used by some studies to estimate risk for large scale hazards (Philpott et al., 2008). However, these estimates may be relatively coarse in terms of spatial resolution for landslide risk assessment, as the landslide hazard footprints are generally small.

Freely available datasets of infrastructure and building footprints are available from OpenStreetMap (OSM) and partners (OpenStreetMap contributors 2015), and while there are data gaps in some parts of the world, in many other locations the data completeness are high. Combined with the publicly available nature of the data, OSM data often represents the best available maps for exposure analysis, and have been used in several studies of landslide hazard (Jaedicke et al., 2014; Stanley & Kirschbaum, 2017; Van Den Eeckhaut et al., 2012). The global nature of EO data and the OSM library provides the opportunity for methods based upon these observations to be globally applicable.

Vulnerability data, on the other hand, is challenging to effectively extrapolate and generalize at a global scale. The dimensions of vulnerability vary widely depending on the type of element under consideration; if buildings, this could be physical parameters like construction style and type of building material, while if population is the element considered then social parameters like national healthcare and emergency planning metrics could also be part of vulnerability. This is not an exhaustive list but even these few metrics demonstrate that vulnerability is either highly local or harder to quantify than many other physical parameters, which explains why EO data are rarely if ever used to quantify vulnerability (Corominas et al., 2014).

As such, while we mention it here, we do not incorporate metrics for vulnerability into our modeling; we focus primarily on hazard and exposure. We suggest that providing a product that defines exposure to landslide hazards in real time for developing humanitarian scenarios is still valuable, and allows end-users to incorporate heuristic or expert-assigned metrics for vulnerability themselves.
While we have focused on the specific case study of the Rohingya refugee camps and surrounding areas as a way to build and test our model, from the outset our intention has been to develop a methodology that would be applicable globally in similar situations and deployable within short time spans. This maximizes the utility of global datasets from EO or other public sources, while providing a workable system for unsupervised use by other end-users. An added benefit of using a globally applicable method is that it allows comparison of model robustness in a range of climatological regimes and with varying degrees of data availability.

3. Methodology and Approach

We have divided our methodology broadly into three parts; the first part constructs a landslide susceptibility map from EO data, the second part incorporates satellite-based precipitation to estimate landslide hazard in near-real time, and the third part estimates the exposure of infrastructure or population to this hazard. Where possible, we have used freely available data and have scripted all the computational methods within Python to allow for complete reproducibility of our methodology. Below, we describe the method as we have applied to the Rohingya refugee camps and surrounding areas, which represents our case study. In the discussion section, we explain how this method can be applied elsewhere.

3.1. Landslide Susceptibility

A plethora of parameters have been demonstrated to be important in controlling landslide susceptibility (Costanzo et al., 2012; van Westen et al., 2003, 2008). These range from topographic factors such as slope, altitude, or roughness, to substrate criteria such as soil type and thickness, bedrock lithology, forest cover, vegetation type, and extent (De Rose et al., 1993; García-ruiz et al., 2017; Glade, 2003). These factors can also be related to anthropogenic disturbance, such as extent of impermeable surfaces (Gill & Malamud, 2017). The relative importance of these parameters can vary across geographic regions, and their utility in modeling will depend on the resolution at which observations are made (van Westen et al., 2008). To account for these differences, we suggest it is essential to use a methodology that does not arbitrarily assign weights a priori to the input parameters.

A number of other studies have used machine learning techniques to avoid such arbitrary parameterization (Brenning, 2005; Bui et al. 2016). Methods such as support vector machines (SVMs), decision trees, or random forests (Catani et al. 2013) avoid this tendency, and have already been extensively used for landslide susceptibility assessment (Bui et al., 2016; Catani et al., 2013; Vorpahl et al. 2012). We have selected the random forest model as it generally performs well for landslide susceptibility modeling (Chen et al., 2017; Trigila et al., 2015; Vorpahl et al., 2012) and also provides a good balance between computational efficiency and accuracy in these cases. Some studies have shown that random forests (Breiman, 2001) provide marginally better performance in terms of predictive skill in comparison with other susceptibility models (Goetz et al. 2015).

The first step in the process is to obtain data sources that could approximate the physical parameters that control landslide susceptibility (see Section 2); these should be of as high resolution as possible. The data sources we have used in the Rohingya camp area are detailed in Table 1. The availability of high-resolution data such as the Vricon 10 m DEM in this region allows us to build a higher resolution model, but this data set is not freely available elsewhere. In the discussion below, we describe the basic subset of EO data that is globally available to allow this method to be used elsewhere. We also suggest limiting data to those that are regularly updated, such as satellite-based land-use estimates, since in transient humanitarian crises land use may change rapidly in the matter of one season. As a result, it is vital to use data that is as relevant to the current situation as possible. However, some parameters (like soil thickness or bedrock lithology) may not vary on such short timescales.

It is possible to derive multiple topographic variables from a DEM, including slope, roughness, aspect, topographic wetness index (Sørensen et al., 2006), or local relief. Susceptibility models developed by other authors have included a diverse range of such parameters (Reichenbach et al., 2018), and it is possible to
include all of these parameters in a Random Forest model. However, some input data associated with landsliding may be strongly correlated; slope and topographic roughness, for example, are highly correlated in the study region (0.964 Pearson R2 coefficient). Co-linearity of input variables in random forest models can make interpreting feature importance problematic, (e.g., determining whether slope or roughness is more important may be challenging if both are included) but unless new inputs are perfectly correlated with existing data inputs they still provide some informational gain to the model (e.g., including roughness may explain marginally more of the variability in landslide susceptibility). The choice of parameters to include will depend upon the location in question and expert judgment of the end-users. For this model, we have included a limited set of parameters that are known to be relevant for landslide susceptibility (slope, relief, forest loss, soil thickness, and land use classification). While we note that other factors may be relevant in other settings, these factors are likely to be applicable in a broad range of settings.

The second step in the process is to construct the dependent variable data (i.e., landslide polygons) to train the random forest classifier. In this case, the dependent variable constitutes the occurrence (or not) of landslides. We map a subset of the region in question using Google Earth imagery. While higher resolution imagery is available from commercial sources, the freely available imagery on Google Earth combined with the historical catalog of imagery means that not only are our results reproducible by any stakeholder with internet access, but our methods can be used widely in other applications without commercial data.

Methods for landslide mapping using EO or aerial photographic data have been applied by a wide range of authors and include hand-mapping (Hovius et al., 1997; Stark & Hovius, 2001) to more automated methods (Li et al., 2016; Martha et al., 2010; Mondini et al., 2011). While hand-mapping requires some specialized training, and can be liable to human error, algorithmic mapping methods suffer from other biases, and may not be universally applicable depending on input imagery resolution (Marc & Hovius, 2015). Algorithmic methods may be more computationally intensive, and may therefore prevent some stakeholders from following the method. In addition, algorithmic methods generally require manually mapped landslides as

| Variable                        | Data source                                                                 | Spatial resolution/Extent | Details                                                                 |
|---------------------------------|-----------------------------------------------------------------------------|----------------------------|------------------------------------------------------------------------|
| Susceptibility map              | Vricon DSM                                                                  | 10 m; covers Study region  | End-User License Agreement (EULA) Restrictions sharing of data; could alternatively use freely available SRTM DEM (30 m resolution) |
| Slope (degrees)                 | Vricon DSM                                                                  | 10 m; covers Study region  | Local relief is more generalizable to other regions compared to absolute elevation. |
| Local Relief in 100 m radius window | Global 1-km Gridded Thickness of Soil, Regolith, and Sedimentary Deposit Layers | 1 km, global                         | Available from ORNL DAAC: https://daac.ornl.gov/Soils/Guides/Global_Soil_Regolith_Sediment.html |
| Soil thickness                  | MCD12Q1: MODIS/Terra and Aqua Combined Land Cover Type Yearly Global 500 m SIN Grid V006 | 500 m, global                         | Available from NASA/USGS: https://lasp.usgs.gov/products/mcd12q1v006/Note: using 2018 data, using Annual International Geosphere-Biosphere Programme (IGBP) classification |
| Land cover                      | MCD12Q1: MODIS/Terra and Aqua Combined Land Cover Type Yearly Global 500 m SIN Grid V006 | 500 m, global                         | Available from NASA/USGS: https://lasp.usgs.gov/products/mcd12q1v006/Note: using 2018 data, using Annual International Geosphere-Biosphere Programme (IGBP) classification |
| Forest loss                     | Global Forest Change 2000–2018 (Hansen et al., 2013)                         | 30 m, global                         | Available here: https://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.5.html |
| Rainfall                        | NASA IMERG 30 min precipitation Near-Real Time product                     | 0.1°, globally between 65N/65S    | Accessible from NASA FTP server: ftp://jsjimsonpps.eosdis.nasa.gov/Available from year 2000-present.Requires login to access data. |
training data. One important consideration is that the area within which landslides are mapped should be as representative of the wider region to be modeled as possible; otherwise, systematic biases may result due to extrapolation.

For the case study area, we mapped landslides by hand in an area around the Rohingya refugee camp, focusing on catchments within Bangladesh that intersected areas of humanitarian concern. The mapped area is ∼600 square kilometers, and using current and historic Google Earth imagery we mapped 1750 individual landslides. Imagery from Google Earth was primarily sourced from Maxar Technologies Digital Globe Satellite products, with dates 16/1/18, 14/2/17, 26/1/16, 19/4/13, 26/11/12, 11/11/09, and 5/11/02. One image from CNES/Airbus satellites was used, with acquisition date 19/12/15. The high cloud cover during summer monsoon periods means that imagery is generally available during the drier season, which improves comparability between images. The minimum size of these landslides is limited by the resolution of the imagery, with the smallest landslide around 50 square meters in area. This minimum size is still larger than the smallest events occurring within the refugee camp itself (meter scale), but is still representative of regional landslide occurrence and is based solely on satellite imagery. We were generously provided with data on landslide events within the Kutupalong mega-camp by United Nations Development Program (UNDP) partners, and we have used those landslides in conjunction with others mapped regionally to calibrate our models. The combined landslide inventory contains >2,300 landslides. The UNDP data was also mapped using high-resolution satellite imagery. For both inventories, the landslide polygons included the entire landslide area (both initiation area and deposit), ensuring they are comparable. We did not attempt to separate initiation areas from deposit, as the UNDP data was provided without separation of the different areas. The number of landslides, and the completeness of the inventory more generally, that is needed to effectively calibrate the model is an important question, and we discuss it in detail below.

Once the mapped landslides are obtained as a polygon layer, they are converted into a raster layer of landslide presence/absence at the same extent and resolution as the base map. Similarly, we reproject and resample each raster data set from the other variables to the base map projection and resolution. The resolution should be the finest of any of the raster data sources to maximize preservation of input information—in our case, the high resolution DSM (10 m resolution). We then extract the values for each of the variables for each pixel within the area in which landslides were mapped. While random forest models can theoretically incorporate categorical data without encoding (Breiman, 2001), the Scikit-learn algorithm we use must encode the categories as numerical values (Pedregosa et al., 2011). Where data sources are categorical (in this case, the MODIS land use categorization), we encode the categories to allow for evaluation by the Scikit-learn random forest classifier. We use the “leave-one-out” encoding method (McGinnis, 2016), which reduces the number of features associated with binary encoding of variables with high cardinality. This method calculates the mean of the target variable (landslide presence) for each of the data points of the same category as the one in question, but leaves out the variable in question to avoid overfitting. This encoding is applied to the test data. This method allows categorical data to be considered as a single variable, rather than multiple binary variables for each category. This reduces the dimensionality of the model, allowing for simpler understanding of the outputs.

To train and test the machine learning model, the data must be split into a training and test data set. We use a training/testing split proportion of 70% and 30% of total data respectively. While we use pixel values for landslides, we split the data before pixelating the landslides. We generate two shapefiles of landslides, a training and test shapefile, that are then each pixelated. The remaining nonlandslide pixel areas from within the mapped area are also split by the same proportions and combined with the respective landslide pixel areas to provide the training and testing datasets. It is important to split at the polygon, rather than pixel level, since otherwise model performance estimates will be over-optimistic (Peña & Brenning, 2015).

The model chosen for prediction is a random forest (Breiman, 2001). This is an ensemble model of multiple decision trees, where each tree is fit to a random bootstrap sample of the training data (random sample with replacement). The ensemble model is designed to reduce the effect of overfitting inherent in decision trees, with the output the modal value of the individual trees. A schematic of this type of model is shown in Figure 1.
The random forest algorithm is run on the training data; the detailed parameters for the random forest model are shown in Table 2. The parameters are selected to avoid over-fitting and to reduce the computational load. Once fitted, the model is applied to the “test” subset of the data to assess the accuracy, false positive rate (FPR), and true positive rate (TPR) (at a pixel basis). We use the Scikit-learn package (Pedregosa et al., 2011) for Python with the inbuilt Random Forest Classifier function.

We have conducted sensitivity tests of the parameters described in Table 2 to determine the appropriate values for the model to maximize computational efficiency without sacrificing model performance. We test predictive accuracy of the model on the training data set, which for the binary classification of landslides is calculated as the Jaccard similarity coefficient. We also calculate out-of-bag error (OOB error) for the training data. OOB error is the average error for each subset of training observations calculated using predictions from trees that do not contain the subset in their training sample, and it represents a measure of how over-fitted the model is (Hastie et al., 2009). Description of the sensitivity analysis carried out on the different model iterations is found in the supplementary material. It is worth noting that while these parameters are fit for the data associated with the area around the Rohingya refugee camps, parameters may differ elsewhere. To test the importance of each parameter, we use the permutation-based feature importance estimates (Pedregosa et al., 2011; Strobl et al., 2007).

We assess the accuracy of the susceptibility model using a Receiver Operating Characteristic (ROC) curve, applied on the testing data set. For a given susceptibility value, we plot the TPR versus FPR, which then gives a curve indicating the performance of the model (Figure 3). These TPR/FPR combinations also allows us to define susceptibility thresholds; for example, susceptibility of 0.8 has TPR of 0.46, and an FPR of 0.02, while susceptibility of 0.4 has TPR of 0.95 and FPR of 0.25.

An important parameter is the area under the curve (AUC value) which tells us how effectively the model predicts 0 values as 0s, and 1s as 1s. In the case of the model, this value is 0.94, demonstrating that the model...
Earth’s Future

is a good classifier of the testing data. As pointed out by Reichenbach et al. (2018), this kind of test can be problematic, since although the model training and testing datasets are split, they are still derived from the same initial landslide inventory, and as such we are not truly validating the model against a new set of data. We discuss this in more detail in Section 4.4.

Using this random forest fit for the data, we use the model to estimate the probability of each pixel for the entire area containing a landslide. We treat this probability as a quantitative estimate for landslide susceptibility, and once the entire set of probabilities have been rescaled from 0 to 1, these values are plotted as pixels within a new raster of susceptibility. Since our intention is to provide a method that can inform landslide susceptibility mapping where other data may be lacking, this inherently limits the degree to which we can use external data sources to validate our model.

### 3.2. Dynamic Landslide Hazard

For the case in Bangladesh, once we have a final static raster representing susceptibility, we are able to incorporate the dynamic changes due to rainfall to generate a dynamic hazard map. Following the approach of D. Kirschbaum and Stanley (2018), we incorporate GPM-derived satellite precipitation data to help describe the likelihood of landslides occurring. In the absence of seismic activity, immediate, intense rainfall is frequently the key trigger for landsliding (Burtin et al., 2013) but preconditions such as hillslope water saturation due to antecedent rainfall may also matter (Gabet et al., 2004; Guzzetti et al., 2008). Many methods exist to constrain both the preconditions and the immediate rainfall, from intensity-duration thresholds (Baum & Godt, 2010; Guzzetti et al., 2008; Nikolopoulos et al., 2017) to consideration of soil moisture. We focus on satellite-based methods to allow for global applicability in near-real time.

We have assessed characteristics for triggering rain for a data set of local landslides with known date. Humanitarian stakeholders in the Rohingya camps have collected incident data on landslides in the camp, including day of occurrence and GPS location. These data were gathered using local reports from stakeholders and does not include the polygon perimeters of the landslides; however, the location of each event is recorded to an accuracy of a few meters. These data were gathered separately to the mapped landslides within the camps discussed above. We have utilized a part of these data collected between the beginning of May 2018 and December 31, 2018 that has been generously shared with us to help calibrate the relationship between satellite-estimated rainfall and landslide occurrence. Since the timings of these landslides are not recorded at a degree of accuracy greater than the day on which they occurred, we are unable to define the exact immediate precursory rainfall conditions. However, IMERG rainfall data were extracted for the 7 days prior to the event to calculate the associated Antecedent Rainfall Index (ARI) value.

#### Table 2

| Model Parameters and the Justification for Those Values |
|--------------------------------------------------------|
| Random forest model parameter | Value | Justification |
|--------------------------------|-------|---------------|
| Number of estimators (number of decision trees in forest) | 100 | Sensitivity analysis of models suggests limited increase in accuracy and decrease in out-of-bag error above this point |
| Maximum number of leaf nodes in each tree | 100 | Smaller values reduce computation time and overfitting, and sensitivity analysis of models suggests limited increase in accuracy and decrease in out-of-bag error above this point |
| Maximum number of features to consider when looking for best split of data at each node | All features | Sensitivity analysis of models suggests this has best accuracy and lowest out-of-bag error, compared to Square root of total features (which would reduce computational time). |

We have assessed characteristics for triggering rain for a data set of local landslides with known date. Humanitarian stakeholders in the Rohingya camps have collected incident data on landslides in the camp, including day of occurrence and GPS location. These data were gathered using local reports from stakeholders and does not include the polygon perimeters of the landslides; however, the location of each event is recorded to an accuracy of a few meters. These data were gathered separately to the mapped landslides within the camps discussed above. We have utilized a part of these data collected between the beginning of May 2018 and December 31, 2018 that has been generously shared with us to help calibrate the relationship between satellite-estimated rainfall and landslide occurrence. Since the timings of these landslides are not recorded at a degree of accuracy greater than the day on which they occurred, we are unable to define the exact immediate precursory rainfall conditions. However, IMERG rainfall data were extracted for the 7 days prior to the event to calculate the associated Antecedent Rainfall Index (ARI) value.

![Number of landslide events per day compared with ARI for that day. ARI, Antecedent Rainfall Index.](image-url)
D. Kirschbaum and Stanley (2018) define a single rainfall metric for the LHASA model—ARI to estimate the rainfall impact on landslide hazard, and we follow the same approach. This method uses two parameters to define the antecedent rainfall index—a time window, and a weighting exponent to strongly favor the more recent rainfall. Plotting the number of events per day against this ARI value shows a strong positive correlation—Figure 2 (Pearson’s $R^2$ value of 0.865). A sensitivity analysis of exponent and number of days considered found only minor differences in the residuals in the fit between ARI and number of daily events, and so we continue to use the parameters defined by D. Kirschbaum and Stanley (2018)—window length 7 days, and exponent of 2.

The local model developed here also follows the example of the LHASA model in providing landslide hazard “nowcasts” when rainfall and susceptibility exceed specific thresholds. However, the selection of thresholds for hazard “nowcasts” for both susceptibility and hazard will be dependent on the local characteristics of the area. Hazard thresholds will depend upon the tolerance for uncertainty of a given end-user, but here we provide a generally applicable approach based on analysis of the input landslides. To obtain susceptibility thresholds, we use the derived ROC curve for the testing landslide data, shown in Figure 3, below.

The susceptibility thresholds for a TPR of 0.5 is 0.85, and for 0.9 it is 0.407. The associated FPRs are 0.04 and 0.24. While arbitrary, it is valuable to assign thresholds in some way to help provide thresholds for action taken by humanitarian actors in settings like the refugee camps, based on discussion with stakeholders there. Assigning thresholds based on true positive or FPRs allows for quantification of the meaning of these thresholds. We suggest that TPRs of 0.5 and 0.9 are reasonable estimates of high and moderate hazard, respectively.

To assign rainfall thresholds in a similar fashion, we can compare the cumulative distributions of ARI over the observation period with the ARI associated with landslide events, as shown in Figure 4.

The cumulative distributions of ARI shown in Figure 4 allow us to provide thresholds associated with TPRs of landslides while simultaneously examining the sensitivity of such thresholds relative to the distribution of the overall ARI values. For a TPR of 0.5, an ARI value of 34 mm is calculated, while for 0.9 the value of ARI is 80 mm. It is important to note that the sensitivity of the lower threshold is relatively good, since only around 8% of days have an ARI above that value in the observation period. Based on discussion with stakeholders, it is valuable to have a threshold above which action can be taken, and we suggest that a TPR of 0.5, while arbitrary, provides appropriate context for humanitarian stakeholders.

It is reasonable to ask whether these thresholds may vary depending on the underlying susceptibility. Physically based models such as the SHALSTAB model (Montgomery & Dietrich, 1994) show such dependence between rainfall and susceptibility. We have assessed the average susceptibility in a 10 m radius area around the landslide recorded points to establish whether a relationship exists between the ARI triggering an event and the susceptibility.

Figure 5 shows the relationship between susceptibility and ARI for the landslide events in question. A clear relationship does not emerge, which does not provide support for changing rainfall thresholds based upon susceptibility. As such, while we acknowledge that a physical basis for
variable rainfall thresholds may exist, it cannot be inferred from the available data.

The dynamic hazard map produced via this method has a 2-level hazard prediction, with a high hazard prediction issued for pixels where susceptibility exceeds 0.85 and antecedent rainfall exceeds the 34 mm, and medium hazard where the susceptibility exceeds 0.407. We note that landslides with known dates may not be available in every setting, and in the discussion section below we suggest alternatives where this method may not be possible.

We did not observe a strong correlation between surface soil moisture, as measured by the Soil Moisture Active Passive (SMAP) level 3 satellite product (Entekhabi et al., 2010), and landslide events per day. The location of the refugee camps is close to the ocean, and since the SMAP tile in this area overlaps with the ocean, SMAP data from this location cannot be reliably used to determine soil moisture. The relationship between ARI and the number of events per day observed does not change before or after the onset of the monsoon (Figure 2), suggesting either that soil moisture is not an important precondition for the observed landslides in this region, or that ARI captures the variability associated with soil moisture changes. As a result, we did not incorporate soil moisture into our further analyses.

3.3. Exposure Modeling

To address a need for information about the exposure of specific infrastructure to landslide hazard, we have used open datasets of infrastructure to locate the exposed elements in the study area.

For specific critical infrastructure elements, such as road segments, medical centers, power substations, or food distribution centers, we need to estimate the hazard in the vicinity. We do this by summing the nowcasts in the area around each point based on a given buffer radius. In this setting, we used a radius of 50 m. By leveraging experience of end-users on the ground in this setting, we identified this distance as a reasonable length scale within which landslide impacts may be felt. The fraction of the area around each building that contains hazard nowcasts can be used to rank the elements in order of their likely impacts. Since high hazard or moderate hazard nowcasts are not an ordinal scale (they are arbitrary categories) we separately calculate the fraction of surrounding areas occupied by high nowcasts, as well as all nowcasts, to provide estimates of buildings exposed to higher hazard, and all exposed buildings. The results of this analysis can be provided in a tabular format (e.g., identifying the percentage of hospitals over the affected area in high exposure categories) or by a map (e.g., highlighting road segments that may be exposed along key transportation routes, or specific exposed buildings).

For the region in question, we utilized publicly available data from OSM—OSM contributors 2015), which provides global coverage and relatively detailed data on infrastructure, even in rapidly changing humanitarian settings; in this case, the data are also supplemented with infrastructure mapped by the Humanitarian OpenStreetMap Team (HOTOSM). For example, the infrastructure associated with the Rohingya camps is well documented within the OSM databases despite its recent development. Below we show specific examples of exposed elements drawn from OSM data using this method. In the first example, we illustrate the buffering of individual buildings to assess the local hazard to which the building may be exposed. The second example shows the exposure of roads to hazard. An important point that was raised in discussion with end-users is the value of aggregating hazard, for example to administrative district level. Since landslide hazard, as flagged by a nowcast, does not guarantee a landslide occurrence, aggregating at a larger spatial level allows end-users to assess the overall hazard level and assign resources more effectively to mitigate potential landslide impacts. To compare across different regions, we can calculate the average proportion
of surrounding areas exposed to hazard for buildings across an entire district. These two example outputs fill explicit needs highlighted by discussions with end-users, but could differ depending on the specific decisions to be addressed. The methodology can incorporate any type of point or line shapefile. We outline how other datasets can easily be incorporated and discuss the transferability and wider use of the method in more detail below.

4. Results and Discussion

Based on the methodology outlined in Section 3 and the inputs for the study region around the Rohingya refugee camps, we have generated a 10 m resolution landslide susceptibility map, a dynamic hazard map, and near-real time risk assessments. In the section below, we describe the results for the study region, then explain how to transfer the system to other regions and discuss the caveats and limitations associated with the model.

4.1. Susceptibility and Landslide Mapping

Figure 6 shows the susceptibility map generated using the random forest model trained with >2,300 hand-mapped landslides. The model is generated only for the mapped area, to avoid extrapolation. The raster from which this sample image is taken can be found in the data uploaded as part of the supplemental material. Shown in Figure 7 is a smaller area from within this region around the Kutupalong mega-camp, to demonstrate the high resolution nature of the model.

To calculate the importance of the different input variables to the final random forest classifier, we use the permutation feature importance method of Scikit-learn v 0.23.1 (Pedregosa et al., 2011). This algorithm defines the importance of a parameter as the decrease in the classification accuracy of the model when a single feature value is randomly shuffled (Breiman, 2001). This reduces bias toward high cardinality features that results from assessing parameter importance using impurity-based algorithms (Altmann et al. 2010).

As shown in Figure 8, the model determines that the most important factor is relief, followed by soil thickness, land use type, slope, the presence of the refugee camp, and finally forest loss. The importance of these factors is in line with prior studies of landslide susceptibility (Catani et al., 2013; Trigila et al., 2015). Perhaps surprisingly, areas where forest loss has been observed since 2000 (Hansen et al., 2013) do not exhibit a major increase in landslide propensity, which is somewhat contradictory to what has been reported by previous studies (De Rose et al., 1993; García-ruiz et al., 2017; Glade, 2003). This may be due to the training data set since there are few landslides reported within these areas and they are difficult to map due to difficulties in resolving failures when there is no vegetation or other feature changes to indicate a failure from satellite imagery.

As discussed above, some variables may be correlated to a degree. Because the model identifies the variable that best explains the landslide patterns at each step in the decision tree, the relative importance of correlated variables will differ from a simple bivariate assessment of importance. Local relief and slope are correlated, so while relief is a more important predictor for the model in this setting, slope remains important. We can examine the importance of key variables by calculating the “area under the curve” for the receiver operating characteristic curve (AUC-ROC value). A higher value represents a better predictor of the output. When applied to the test data, the ROC value is 0.766 for slope, and 0.757 for relief. The two variables have similar predictive strength when considered individually, but the random forest model assigns greater weight to relief. The same method allows us to look at the predictive strength of the model-derived susceptibility model; the AUC-ROC value estimated for the susceptibility model calculated on the testing data is 0.906.

As has been highlighted by Goetz et al. (2015), assessment of susceptibility methodologies should be linked to the goals of the study. We suggest that this methodology is appropriate for use in settings that are a relatively data-limited, with few constraints on landslide triggering conditions. In such settings, several predictor variables may be relevant that could be co-linear, such as roughness, slope, and relief, and the use of Random Forest methods are appropriate for such scenarios (Goetz et al., 2015). The methodology
Outlined here can quickly provide an initial assessment of landslide susceptibility using openly available data that can be supplemented by more accurate local scale assessment of hazards as improved data becomes available.

With the susceptibility map in hand, we can incorporate the dynamic precipitation estimates to assess changing landslide hazard in near-real time.

**Figure 6.** Sample of susceptibility map generated using random forest methodology. The susceptibility values are scaled between 0 and 1, where one is the highest possible likelihood of a landslide in the area based on the model, and 0 the lowest. Note that the extent of the susceptibility model is the extent of the area mapped for landslides, not extrapolated across the DEM area. DEM, Digital Elevation Models.
4.2. Dynamic Hazard

Applying the methodology from the LHASA model to this high-resolution susceptibility map provides a near-real time estimate of the relative landslide hazard in this region. This is provided to end-users via a web interface and is updated daily with a spatial resolution of 10 m. An example output of the hazard model is shown in Figure 9.

Although our method defines thresholds for rainfall based on historical records of landslides with known dates, such data may not be available in every setting to the extent it is here. In some settings, the NASA Global Landslide Catalog (D. B. Kirschbaum et al., 2015) may be used to provide specific landslides with
known dates and times that could be used to calibrate a true-positive rate threshold, as shown in Figure 4. If such data are still not available, we suggest caution is necessary in defining thresholds. Based on global relationships compiled by Guzzetti et al. (2008), it may be possible to identify minimum rainfall intensity-duration thresholds for landslide initiation, below which landslide triggering is unlikely. These simplified minimum thresholds may still benefit humanitarian stakeholders in some settings.

It is important to note that daily hazard output should be considered explicitly as situational awareness, rather than a warning system. Since rainfall from the most recent 24 h is weighted equally, the ARI value does not completely capture any short event with high instantaneous rainfall that might trigger a landslide. Additionally, the satellite precipitation used to define the ARI has a latency of 4 h after acquisition and as such the hazard model should not be considered as an early warning system.

Instead, we consider it a model that highlights areas where landslide hazard is elevated, and where greater attention should be paid by end-users to rainfall triggers.

This differs from other recent studies that have developed “early warning systems” for landslides in this setting (Ahmed et al., 2020). Based upon discussion with end-users and stakeholders in the refugee camps, the risk of “crying wolf” in humanitarian settings has the potential for negative effects including loss of trust in model outputs, and we do not recommend the use of our model in such a capacity.

4.3. Exposure Modeling

A schematic example of building exposure as well as a map of exposure for roads within the study region are presented below (Figures 10 and 11).

While the exposure estimates rely primarily on the hazard estimates, there are some specific considerations and assumptions associated with the exposure modeling that may influence the product outcome.

First, when considering the hazard over a specific location, it is likely that the degree and severity of impact from a landslide on an infrastructural element (e.g., building or road segment) will differ depending on whether the element is up-slope or down-slope of the landslide. We do not explicitly consider differences in the location of the element compared to the landslide hazard. This may lead to errors in exposure estimates, but we note that landslide hazards can still impact elements in up-slope positions, for example by creating unstable hillslopes (Samia et al., 2017).

In the specific example of the Rohingya refugee camps and the vicinity around them, humanitarian stakeholders highlighted the value of this kind of exposure mapping to help make logistical decisions about where to position supplies given the exposure of different road sections to landslide hazard. As discussed above, this information should not be considered an early warning for specific locations, but instead as a tool to help prioritize resource allocation as part of hazard-relevant decisions.

Since the decisions made by humanitarian stakeholders with respect to infrastructure exposure are sensitive to factors specific to each setting, like capacity to rapidly inform exposed people and evacuate them if necessary, and relative risks associated with “crying wolf,” we stress that researchers should consult with end-users and humanitarian stakeholders when generating exposure estimates. While it is possible to estimate semi-quantitative exposure to landslide hazard for each individual building for a given area, this may not be a useful product. Through discussion with UN partner agencies, we ascertained that an aggregated output at an administrative district level (e.g., Figure 8) may provide more actionable information. This further highlights the importance of scientist—stakeholder communication and co-development. Moreover, some humanitarian stakeholders may prefer exposed elements only to be highlighted when their exposure exceeds a certain threshold (as in Figure 9), while others may prefer a graduated scaling. The model output is a continuous variable, and as such to set thresholds it is important to discuss with end-users what
constitutes an acceptably high or low threshold. We note that the combination of hazard information with exposure provides an added level of information relevant to disaster response stakeholders that is not generally combined with landslide hazard models (Ahmed et al., 2020; Reichenbach et al., 2018).

4.4. Uncertainties and Limitations

One of the key aims of our study is to provide a series of methods that can be used in diverse locations around the world, with no existing prior local data in that region. While this supports decision making for end-users working in these locations who may lack significant capacity to supplement EO data with ground based data, it also presents challenges to validation and calibration. There are a number of caveats to these
methods that should be considered in each of the susceptibility, hazard, and exposure aspects, and we consider those in detail below.

The machine-learning based nature of the susceptibility model introduces some issues that need to be considered before implementation over a region. First, a shortcoming of random forest models is that extrapolating a model to datasets with variables that differ in range from their training data set can be problematic. For example, a model trained with landslide data on hillslopes ranging from 0 to 30° would not perform well in calculating susceptibility for hillslopes with slope angles above 30°. Similarly, if the landslide inventory used to calibrate a model does not contain landslides mapped within a specific land use type (e.g., savannah), then no landslides will be predicted on this land use elsewhere; in the case of water bodies, this would be accurate. Extrapolation is thus a challenging issue to address, as the range of variables within the mapped landslide area will not necessarily match the overall area mapped. As a first step to avoid issues with extrapolation, we suggest not extending the analysis outside of the area covered by the landslide inventory.

This also emphasizes the importance of the input landslide inventory. Since the model outputs represent areas that are similar to areas where the landslides used for calibration have occurred before, the input landslide inventory strongly controls the output. If the input inventory includes only landslides from a specific event that affected only a small part of the mapped region, biases would be introduced. Similarly, if the inventory was systematically missing certain types of landslides (e.g., below a certain size), the model outputs would also fail to effectively capture the susceptibility associated with those. It is possible to produce statistically complete inventories for specific triggering events (Guzzetti et al., 2012) using remote sensing data, but in some locations the rapid rate of revegetation may limit how well preserved landslide scars are from historic events. As such, we suggest using Google Earth’s historical imagery catalog as a way to capture the recent history of landslides in a given location. While not guaranteed to be comprehensive, this can capture multiple years of landslide events and avoid bias to one singular event. In addition, given that we have mapped the entire area of landslides, rather than just the initiation area, the model outputs areas that correspond to the entire impacted area of a landslide.

Another important consideration is how large the landslide catalog needs to be. For the case study in the area around the Rohingya refugee camps, we used an inventory of more than 2,300 landslides mapped over ~597 square km, but this number may not be possible to attain in other regions. To test the importance of inventory size, we re-ran the susceptibility model using random subsets of the landslides as input, reducing the subset size each time. In Figure 12 we plot the proportion of the inventory used as input against the AUC value for the output model (where AUC is measured against the test landslides—i.e., those not used to train the model), as a measure for the relative accuracy of the output. We also plot the AUC-ROC values for slope and relief calculated for the test data set at each point. We observe that with only 10% of the inventory used (~230 landslides), the AUC value declines, but still performs better than the model only considering slope or relief. This suggests that inventories of 100s, rather than 1000s, of landslides are likely sufficient to generate this kind of model output. However, it is key that the landslides be mapped...
An important issue that many landslide susceptibility models fail to address is validation, as few statistically based studies validate their outputs against data other than the data used to calibrate their models (Reichenbach et al., 2018). Since our intention is to provide a method for data-poor areas, we will inevitably run into this issue if no prior data exists. We have taken some steps to address this. First, in the analysis performed using the random subsets of the inventory as input and validating against the test data (Figure 12), the input and validation datasets are not fully coupled. In addition, calculating the susceptibility associated with each of the events recorded in the camp by UN stakeholders (a separate data set) allows us to test model performance. Median susceptibility values for those landslides is 0.70, and higher susceptibility shows an increase in number of landslides (Figure S2). This is a spatially disjoint set of data that provides a useful independent test of model performance, which is important for susceptibility model evaluation (Reichenbach et al., 2018).

Uncertainty in input data remains a challenge. The soil thickness data is at a relatively coarse resolution (250 m), limiting how effectively small features can be distinguished. In addition, the IMERG rainfall data used to generate the dynamic hazard estimates is a very coarse spatial resolution (0.1°) and temporal resolution, while good for satellite data, is not perfect (every 30 min with 4-h latency). This limits how effectively highly localized, short-term rainfall can be detected (Bookhagen & Burbank, 2006), and thus guides the use of the model—as situational awareness, not as early warning.

4.5. Interaction With End-Users

One of the most important and unique parts of this landslide modeling study area and humanitarian setting is the opportunity to work alongside and learn from humanitarian end-users. As discussed above, there have been many points in our model development process where input on specific parameters has been guided by input from those on the ground. While the product development was driven by research scientists, the involvement of end-users whose expertise included refugee camp management, geo-engineering and resilience presented both challenges in communication and the chance to develop a product that explicitly filled needs identified at the earliest stages of model development.

The UN and other NGO user base for our EO-based product in the Rohingya setting is led by a cross-agency working group of humanitarian staff with expertise in natural hazards. This working group facilitated communication of scientific terminology and concepts to users working directly within the camp who may have less familiarity with remote sensing science. Since the overlap of expertise between scientists and those on the ground may be limited, we suggest that each stakeholder must act to some degree as a translator of their needs or products to those stakeholders representing other links in the chain.

While iterative co-development of model products increased the time taken to produce an initial working prototype, it reduced the chances that we developed products that would not be fit-for-purpose, in some ways acting as a live peer-review process. By focusing on the specific decisions made by end-users relating to landslide hazards and how they initially incorporated geological data, we were able to isolate where EO data and products could be most beneficial.

While at present this collaborative effort between scientists and humanitarian stakeholders represents a test case for using EO data, we suggest that formalizing the methodology for collaboration could lead to more effective use of EO products in future in other, similar settings.
4.6. Transferability

Each aspect of the model can be applied to the majority of global settings. With the exception of the commercially obtained 10 m DEM, all of the input data for the susceptibility model used around the Rohingya refugee camps can be freely obtained for any other part of the world (see Table 1 for details). The 30 m resolution SRTM DEM is a globally available topographic data set that can be used to estimate slope and altitude and replace the 10 m DEM in other settings. Landslides can be mapped using Google Earth anywhere, although this may be limited by availability of imagery and intervals at which it is provided, and reproducibility may be affected if Google Earth changes the availability of imagery. IMERG rainfall data is available from 65°N to 65°S, encompassing the majority of the planet. Finally, OSM data is available across the globe, and while the level of completeness differs (Barrington-Leigh & Millard-Ball, 2017), it can serve as an initiation point for the exposure elements.

Furthermore, any of these input data can be exchanged with better quality or higher resolution inputs if the end-users see fit. The method is agnostic to type and resolution of data, which allows for a flexible approach; however, it is important to note that the quality of the susceptibility maps will be dependent on the input, so end-users should take care when interpreting susceptibility to consider it side-by-side with inputs to ensure a more complete picture is obtained. Crucially, we note that the same random forest classifier cannot be applied to other regions; it is necessary to derive a new random forest model for each location this model is applied using new data.

5. Conclusions

We have presented a new methodology for satellite-based estimation of landslide hazard and exposure in data-poor regions that can be accomplished with generally available datasets. We have outlined a protocol for choosing publicly available data sets as well as examples that we have applied to our case study in the Rohingya refugee camps in South Bangladesh. This methodology does not require arbitrary weighting of input variables, and is agnostic to the type and resolution of the data input. A key factor in determining the accuracy of the susceptibility model is the representativeness of the landslide data used to train and test the model. We stress that data should be both spatially and temporally extensive. This supports the application of comprehensive landslide mapping using remote sensing imagery in humanitarian settings. For robust hazard modeling, data are needed not only on the location of landslide occurrence, but also landslide timing. Data on timings tends to be more complex to collect, but we emphasize the value of such information.

The derived susceptibility model is combined with satellite-derived precipitation to estimate near-real time hazard. While estimating triggering rainfall for landslides may be problematic if dates of landslide events are unknown, we suggest that a susceptibility map for landslides provides an initial tool to aid in hazard identification in data-poor regions; minimum thresholds for landslide triggering defined from global relationships (e.g., Guzzetti et al., 2008) could be used in combination with susceptibility maps to help approximate hazard. When landslide data are available that allows for suitable definition of hazard thresholds, the derived hazard map can then be linked to infrastructure information to provide either itemized or mapped estimates of relative exposure to landslide hazard in near real time.

While local methods based on detailed field observation will likely out-perform satellite based methods, we suggest that the global applicability, low or zero overhead costs, and ease of data access associated with our modeling approach makes it a useful addition to the inventory of models available to end-users in data-poor or rapidly changing regions of the world. End-users must be aware of the caveats associated with the model, namely that it cannot serve as an early warning or immediate warning system, that the hazard and exposure are not fully quantitative, and that extrapolation beyond areas where landslides were mapped should be avoided. Despite these caveats, within our case study region the model performs well and meets the needs of regional and local stakeholders who will use the products. As with any machine-learning model, increased accuracy and relevance of input data will likely improve the model performance, and as such end-users can use their own in situ products, which are likely more detailed within the same modeling framework, to obtain better results. This model and approach with end users demonstrates the feasibility of characterizing hazard severity and establishing models that work across many geographies and humanitarian instances. It
may also be transferable to assess other hazards, in particular those associated with extreme weather, where global near real-time satellite data could be most useful.

Data Availability Statement

While all data and methods necessary to generate the model outputs are detailed in the text or supplementary material, a packaged version of the code will be publicly available once it has been cleared by NASA's internal technology review process.

Acknowledgments

All authors acknowledge no conflict of interest, financial, or otherwise. RE's research was supported by an appointment to the NASA Postdoctoral Program at NASA Goddard Space Flight Center, administered by Universities Space Research Association under contract with NASA. The authors greatly appreciate the input and expertise of the members of the COMPASS project, supported by NASA’s Rapid Response Grant (18-RNRES18-0008), and from end-users and stakeholders from UNHCR, UNDP, IOM, and ISCG in Cox’s Bazar, Bangladesh. Pukar Amatya provided useful insight during model development and manuscript review.

References

Ahmed, B., Rahman, S., Sammonds, P., & Islam, R. (2020). Application of geospatial technologies in developing a dynamic landslide early warning system in a humanitarian context: The Rohingya refuge crisis in Cox’s Bazar, Bangladesh. Geomatics, Natural Hazards and Risk, 11(1), 446–468. https://doi.org/10.1080/19475705.2020.1730988

Altmann, A., Tolo, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. Bioinformatics, 26, 1340–1347.

Barrington-Leigh, C., & Millard-Ball, A. (2017). The world’s user-generated road map is more than 80% complete. PLoS One, 12(8), 1–20. https://doi.org/10.1371/journal.pone.0180698

Baum, R. L., & Godt, J. W. (2010). Early warning of rainfall-induced shallow landslides and debris flows in the USA. Landslides, 7(3), 259–272. https://doi.org/10.1007/s10346-009-0177-0

Bookhagen, B., & Burbank, D. W. (2006). Topography, relief, and TRMM-derived rainfall variations along the Himalaya. Geophysical Research Letters, 33, 1–5. https://doi.org/10.1029/2006GL026037

Breiman, L. (2001). Random Forests. Machine Learning, 45, 5–32.

Brenning, A. (2005). Spatial prediction models for landslide hazards: Review, comparison and evaluation. Natural Hazards and Earth System Sciences, 5, 853–862.

Bui, D. T., Tuan, T. A., Klempe, H., Pradhan, B., & Revhaug, I. (2016). Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides, 13, 361–378. https://doi.org/10.1007/s10346-015-0557-6

Burtin, A., Hovius, N., Milodowski, D. T., Chen, Y. G., Wu, Y. M., Lin, C. W., et al. (2013). Continuous catchment-scale monitoring of geomorphic processes with a 2-D seismological array. Journal of Geophysical Research: Earth Surface, 118(3), 1956–1974. https://doi.org/10.1002/jgrf.20137

Burton, J., Kates, R. W., & White, G. F. (1978). The environment as hazard. New York: Oxford University Press.

Catani, F., Lagomarsino, D., Segoni, S., & Tofani, V. (2013). Landslide susceptibility estimation by random forests technique: Sensitivity and scaling issues. Natural Hazards and Earth System Sciences, 13, 2851–2853. https://doi.org/10.5194/nhess-13-2815-2013

Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D. T., et al. (2017). A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena, 151, 147–160.

Corominas, J., van Westen, C., Frattini, P., Cascini, L., Malet, J. P., Fotopoulou, S., et al. (2014). Recommendations for the quantitative analysis of landslide risk. Bulletin of Engineering Geology and the Environment, 73(2), 209–263. https://doi.org/10.1007/s10064-013-0538-8

Costanzo, D., Rotigliano, E., Irigaray, C., Jiménez-Perálvarez, I. D., & Chacón, J. (2012). Factors selection in landslide susceptibility modelling on large scale following the gis matrix method: Application to the river Beiro basin (Spain). Natural Hazards and Earth System Sciences, 12(2), 327–340. https://doi.org/10.5194/nhess-12-327-2012

Crosa, G. (1998). Regionalization of rainfall thresholds: An aid to landslide hazard evaluation. Environmental Geology, 35(2–3), 131–145.

Crosa, G. B., & Frattini, P. (2003). Distributed modelling of shallow landslides triggered by intense rainfall. Natural Hazards and Earth System Sciences, 3, 81–93.

Dade, W. B., & Huppert, H. E. (1998). Rock runout. Geology, 26(9), 803–806.

Davies, T. R. H. (1982). Spreading of rock avalanche debris by mechanical fluidization. Rock Mechanics, 15, 9–24.

De Rose, R. C., Trustnum, N. A., & Blaschke, P. M. (1993). Post-deforestation soil loss from steepland hillslopes in Taranaki, New Zealand. Earth Surface Processes and Landforms, 18, 131–144.

De la Haba, D., Nogueira, C., O’Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The soil moisture active passive (SMAP) mission. Proceedings of the IEEE, 98(5), 704–716. https://doi.org/10.1109/JPROC.2010.2047750

Dell, M., Roetman, J., Bonnard, J. C., Cascini, L., Lelou, E., & Savage, W. Z. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning. Engineering Geology, 102(3–4), 99–111. https://doi.org/10.1016/j.enggeo.2008.03.014

Friedl, M. A., Sulla-menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). Remote sensing of environment MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114(4), 168–182. https://doi.org/10.1016/j.rse.2009.08.016

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. Natural Hazards and Earth System Sciences, 18, 2161–2181.

Gabert, E. J., Burbank, D. W., Putkonen, J. K., Pratt-Sitaula, B. A., & Obja, T. (2004). Rainfall thresholds for landsliding in the Himalayas of Nepal. Geomorphology, 63(3–4), 131–143. https://doi.org/10.1016/j.geomorph.2004.03.011

García-ruiz, J. M., Beguería, S., Arnáez, J., Sanjuán, Y., & Lana-renault, N. (2017). Deforestation induces shallow landsliding in the montane and subalpine belts of the Urbión Mountains, Iberian Range, Northern Spain. Geomorphology, 296, 31–44. https://doi.org/10.1016/j.geomorph.2017.08.016

Gill, J. C., & Malamud, B. D. (2017). Anthropogenic processes, natural hazards, and interactions in a multi-hazard framework. Earth-Science Reviews, 166, 246–269. https://doi.org/10.1016/j.earscirev.2017.01.002

Glade, T. (2003). Landslide occurrence as a response to land use change: A review of evidence from New Zealand. Catena, 51, 297–314.

Glade, T., Crozier, M., & Smith, P. (2000). Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical “Antecedent daily rainfall model”. Pure and Applied Geophysics, 157, 1059–1079.
Samia, J., Temme, A., Bregt, A., Wallinga, J., Guzzetti, F., Ardizzone, F., & Rossi, M. (2017). Do landslides follow landslides? Insights in path dependency from a multi-temporal landslide inventory. Landslides, 14(2), 547–558. https://doi.org/10.1007/s10346-016-0739-x

Schneider, A., Friedl, M. A., & Potere, D. (2009). A new map of global urban extent from MODIS satellite data. Environmental Research Letters, 4, 1–11. https://doi.org/10.1088/1748-9326/4/4/044003

Shahid, S. (2011). Trends in extreme rainfall events of Bangladesh. Theoretical and Applied Climatology, 104, 489–499. https://doi.org/10.1007/s00704-010-0363-y

Shean, D. E., Alexandrov, O., Moratto, Z. M., Smith, B. E., Joughin, I. R., Porter, C., & Morin, P. (2016). An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 116, 101–117. https://doi.org/10.1016/j.isprsjprs.2016.03.012

Staker, C., & Hovius, N. (2001). The characterization of landslide size distributions. Geophysical Research Letters, 28(6), 1091–1094.

Strobl, C., Boulesteix, A., Zeileis, A. & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics 8, 1–21.

Tanyas, H., van Westen, C., Allstadt, K., Nowicki Jessee, A., Görüm, T., Jibson, R. W., et al. (2017). Presentation and analysis of earthquake-induced landslide inventories. Journal of Geophysical Research: Earth Surface, 122, 1991–2015. https://doi.org/10.1002/2017JF004236

Van Den Eeckhaut, M., Hervás, J., Jaedicke, C., Malet, J.-P., Montanarella, L., & Nadim, F. (2012). Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data. Landslides, 9, 357–369. https://doi.org/10.1007/s10346-011-0299-z

Van Westen, C. J., Castellanos, E., & Kuriakose, S. L. (2008). Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. Engineering Geology, 102(3–4), 112–131. https://doi.org/10.1016/j.enggeo.2008.03.010

Van Westen, C. J., Rengers, N., & Soeters, R. (2003). Use of geomorphological information in indirect landslide susceptibility assessment. Natural Hazards, 30, 399-419.

Vorholt, P., Ptacek, J., & Or, D. (2013). Rainfall-triggered shallow landslides at catchment scale: Threshold mechanics-based modelling for abruptness and localization. Water Resources Research, 49(10), 6266–6285. https://doi.org/10.1002/wrcr.20418

Vorpahl, P., Elsenbeer, H., Marker, M., & Schroder, B. (2012). How can statistical models help to determine driving factors of landslides? Ecological Modelling, 239, 27–39.

Wessel, B., Huber, M., Wohlhart, C., Marschalk, U., Kosmann, D., & Roth, A. (2018). Accuracy assessment of the global TanDEM-X Digital Elevation Model with GPS data. ISPRS Journal of Photogrammetry and Remote Sensing, 139, 171–182. https://doi.org/10.1016/j.isprsjprs.2018.02.017