Accounting for Satellite Precipitation Uncertainty: The Development of a Probabilistic Landslide Hazard Nowcasting System

by

Samantha Hartke

Master of Science

(Civil and Environmental Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2019

Table of Contents

Acknowledgements

I would like to thank my advisor, Dr. Daniel Wright, for his support and advice over the past two years and my committee members, Dr. Paul Block and Dr. Nimish Pujara, for their guidance and feedback on my MS thesis. I thank the Hydrologic Extremes Research Group for their companionship and participation in many presentation rehearsals. I would also like to thank the water resources engineering graduate student group and my friends and family.

I would like to acknowledge the generous funding support from the Arthur H. Frazier Fellowship and the NASA Earth and Space Science Fellowship, which has made my research possible over the past two years.

1. Introduction

Satellite multisensor precipitation products (SMPPs) provide valuable precipitation information for hydrological modeling, water resources management, natural disaster assessment, and other applications on a global scale and can be especially beneficial to data-limited regions. Uncertainty in SMPPs presents a challenge to fully utilizing these near real time sources of global precipitation. As the key forcing variable in hydrological processes, precipitation and any associated variability will impact model output. While various models have been developed to characterize SMPP uncertainty, methods of modeling SMPP uncertainty for applied use in models have been limited. This study explores the challenges of modeling SMPP uncertainty for applied use and develops a probabilistic landslide hazard nowcasting system that demonstrates one method of incorporating SMPP uncertainty in a decision tree framework. The improved performance of a probabilistic landslide hazard model which accounts for SMPP uncertainty confirms the value of converting SMPP data from an unknown source of uncertainty to a known source of uncertainty. Even without knowing the true antecedent rainfall, representing the known uncertainty of SMPP input within the model improves model performance.

Chapter 2 will review the accuracy of SMPPs, discuss common metrics of SMPP error, and explore SMPP uncertainty models that have been developed over the past two decades. The limitations of current error models and needs for applied use in environmental models is discussed and methods of accounting for these limitations are proposed using a CSGD-based error model. Chapter 3 details the development of a probabilistic landslide hazard model that incorporates SMPP uncertainty using a CSGD-based error model. The advantages of probabilistically representing SMPP data as input to the model are demonstrated by increased true positive rate and more realistic variability in model output. A summary of findings and discussion of future work is included in Chapter 4.

2. SMPP Uncertainty

This chapter will discuss two types of studies addressing SMPP uncertainty: studies which evaluate the accuracy of SMPPs (Section 2.1) and studies which model the uncertainty associated with SMPP retrievals (Section 2.2). Challenges to incorporating SMPP uncertainty into modeling applications are discussed in Section 2.3.

2.1 Characterizing SMPP Error

A variety of SMPPs provide global precipitation data at a range of spatial and temporal resolutions. The most commonly-used are described here. The 3-hourly, 0.25˚ TRMM Multisatellite Precipitation Analysis (TMPA; Huffman et al., 2007, 2010) has been produced by NASA and the Japan Aerospace Exploration Agency since 1998 through the Tropical Rainfall Measuring Mission (TRMM). With the launch of the Global Precipitation Mission (GPM) Core Observatory satellite in February 2014, NASA released TMPA's successor, the 30-min, 0.1˚ Integrated MultisatellitE Retrievals for GPM (IMERG; Huffman et al., 2015). The 30-min, 8-km Climate Prediction Center (CPC) morphing technique (CMORPH; Joyce et al., 2004) was developed by NOAA/NCEP and the hourly, 4-km Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN; Sorooshian et al., 2000) is available through the University of California, Irvine. SMPPs generally have a near real-time version as well as a late version which incorporates ground-reference precipitation data. TMPA and IMERG are used for the analyses in this work.

SMPP algorithms combine passive microwave (PMW) and infrared (IR) retrievals to derive precipitation. IR sensors detect precipitation using cloud top brightness temperature (Levizzani et al., 2002) while PMW sensors retrieve microwave emission and scattering data at specific frequencies to estimate precipitation rates (Kummerow et al., 1998). Although PMW retrievals are a more accurate source of precipitation data, they are mounted on low Earth orbit satellites and have infrequent temporal sampling (Gebregiorgis et al., 2018; Hossain & Anagnostou, 2004). IR sensors, on the other hand, are geostationary and can continuously monitor precipitation at a given time-space resolution. SMPPs contain uncertainty in the form of systematic bias and random error, which often varies with season and location (Abebe S. Gebregiorgis & Hossain, 2013; Tian & Peters-Lidard, 2010; Tian et al., 2009). Sampling errors (infrequent satellite passes), measurement

issues (cloud cover and reflective land surfaces), and imperfect retrieval algorithms all contribute to SMPP error. Spatial differences between SMPP and ground-reference precipitation contribute to errors in rain detection.

Even in newer, higher resolution, and more accurate SMPPs like IMERG, near real-time versions of the product, IMERG-Early and IMERG-Late, still contain errors (Tan & Santo, 2018; Wang et al., 2017). Many studies have evaluated SMPP accuracy across a range of spatial and temporal scales (Ebert et al., 2007; Gebregiorgis & Hossain, 2014; Gebregiorgis et al., 2017; Sapiano & Arkin, 2008; Tian et al., 2007; among many others). While many studies directly compare SMPPs to ground reference estimates and quantify the error between the two records, some studies evaluate SMPPs based on their ability to predict streamflow in a hydrological model (e.g. Bitew & Gebremichael, 2011; Gebregiorgis & Hossain, 2013). Uncertainty is often quantified in terms of bias, root mean square error (RMSE), probability of detection (POD), correlation, and false alarm ration (FAR).

Uncertainty is greater in high latitudes due to fewer sensors as well as ice and water coverage (especially in the northern hemisphere) which is difficult for sensors to distinguish from clouds (Tian & Peters-Lidard, 2010). Areas with high elevations and complex terrain such as the Tibetan Plateau, the Rockies, and the Andes exhibit consistently higher error (Gebregiorgis & Hossain, 2014). SMPP uncertainty in areas with mountainous topography can be partially attributed to SMPP algorithms' difficulty in detecting orographically induced precipitation (Dinku et al., 2008). SMPP retrievals during extreme precipitation events and over coastlines and inland water bodies also exhibit higher error (Aghakouchak et al., 2011; Tian & Peters-Lidard, 2007). SMPPs have been found to be most reliable over areas with strong convective precipitation and uniform surfaces such as oceans (Tian & Peters-Lidard, 2010). Certain error metrics are scale dependent, such as the probability of rain detection (Hossain & Huffman, 2008) and, in general, SMPP errors diminish with temporal and spatial aggregation (Artan et al., 2007). In our study area in the southeastern U.S. (see Section 3.2), SMPPs IMERG and TMPA consistently exhibit errors relative to NEXRAD Stage IV radar data, especially for high magnitudes of precipitation (Figure 1).

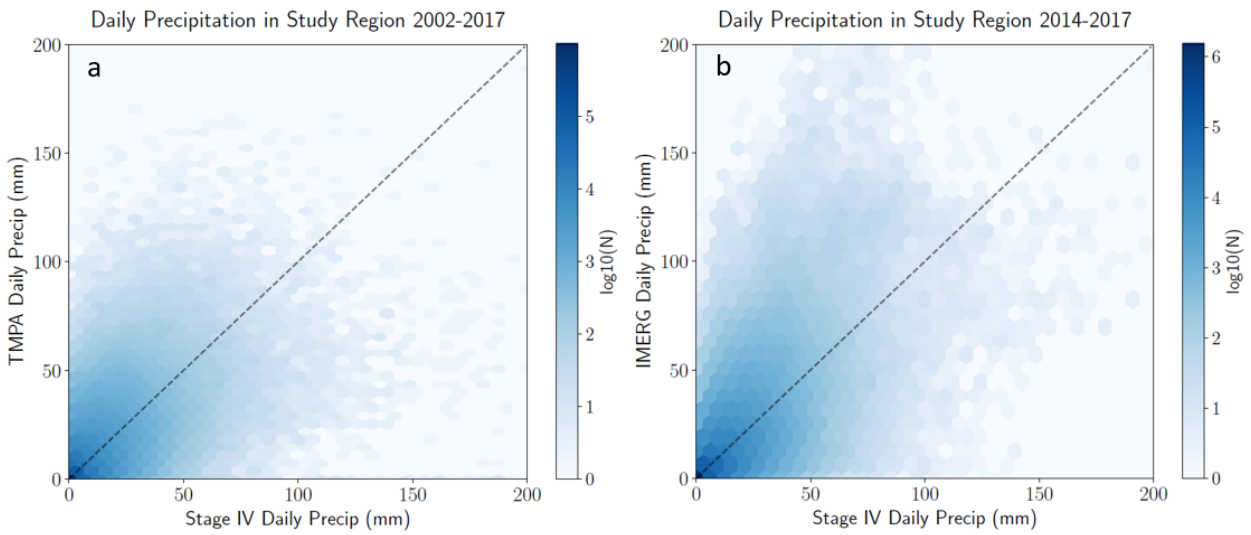

Figure 1*: Comparison of daily precipitation for every SMPP grid cell in study area for (a) TMPA product 2002-2017 and (b) IMERG product 2014-2017 compared with corresponding Stage IV daily precipitation.*

The next section will discuss how SMPP uncertainty has been modeled with the intent of 'correcting' SMPP data or simulating possible realizations (i.e. ensembles) of precipitation conditional on given SMPP retrievals.

2.2 Modeling SMPP Uncertainty

While quantifying SMPP accuracy in terms of RMSE, bias, and other validation statistics may be useful for judging the reliability of SMPP retrievals, modeling SMPP uncertainty is required to simulate and estimate SMPP errors. SMPP uncertainty has been modeled as Gaussian in several studies (Sarachi et al., 2015), but has been shown to better fit gamma, Weibull, and lognormal distributions (Alemohammad et al., 2015; Hossain & Anagnostou, 2006; Wright et al., 2017). Error distributions vary with spatial and temporal resolution of SMPP data, but one consistent characteristic of random error is that it is multiplicative, not additive, such that it increases with the value of SMPP estimates (Tian et al., 2013). SMPP error models should model three types of error cases: false alarms, when an SMPP records precipitation but the ground-reference product does not; missed cases, when reference data records precipitation but the SMPP observes none; and hits, when both SMPP and ground-reference data record precipitation but differ in magnitude.

SMPP error models have been developed for SMPPs at monthly scales (Yan & Gebremichael, 2009) and for SMPPs over specific regions and resolutions (AghaKouchak et al. , 2009; Maggioni et al., 2014; Tian et al., 2013; Villarini et al., 2009). The general goal of most of these studies has

been to develop methods to provide relatively simple uncertainty estimates (in the form of distributions) to SMPP users. Error models are calibrated using SMPP and reference precipitation, which limits their applicability to data-limited regions without sufficient ground-reference records. Though radar and rain-gauge precipitation data is acknowledged to contain errors, this field of research assumes that such errors are small relative to SMPP errors and that reference rainfall provides a sufficiently accurate approximation of actual rainfall (Tian et al., 2013).

The two-dimensional satellite rainfall error model (SREM2D) was developed by Hossain and Anagnostou (2006) to generate ensembles of simulated satellite precipitation based on reference rain fields. Simulated satellite rainfall fields preserve the spatial autocorrelation of rain/no-rain detection and temporal autocorrelation in the SMPP which is used to calibrate the SREM2D model (Hossain & Anagnostou, 2006). The gridded satellite precipitation ensembles which SREM2D models can be used to investigate the propagation of SMPP uncertainty through hydrologic models. The SREM2D model can be reversed to generate an ensemble of possible ground reference rainfall observations given satellite rainfall input. This method is used in Maggioni et al. (2013) to investigate rainfall-to-runoff error propagation of three SMPPs over the southeastern U.S. The SREM2D model 'perturbs' SMPP data to create a random realization of precipitation which estimates reference rainfall by replicating the multidimensional structure of SMPP error. This process is repeated 50 times to obtain a 50-member ensemble of rainfall and, through the National Weather Service (NWS) Hydrology Laboratory-Research Distributed Hydrological Model (HL-RDHM), a 50-member ensemble of basin runoff (Maggioni et al., 2013). This study found that SMPP random error is enhanced at smaller basin scales by the streamflow hydrological model, but dampened at larger basin scales. Relative bias doubled from rainfall to runoff for all basin scales and all three SMPPs, confirming the potential of SMPP uncertainty to disproportionately affect hydrological model output.

Studies using SREM2D have described it as "elaborate" and "complex": a comprehensive, but not concise model of SMPP uncertainty. Training the model requires numerous parameters: probability of rain detection as a function of rainfall magnitude, probability of no-rain detection, first- and second-order moments of a probability distribution for false alarms, correlation lengths for both the detection of rain and no-rain, mean field bias, error variance, correlation length for retrieval error, and lag-one autocorrelation of the mean field bias. Only one ensemble member is

generated when the SREM2D model randomly 'perturbs' SMPP data so a 50-member ensemble of rainfall requires running SREM2D 50 times, which is computationally resource-intensive. In Maggioni et al. (2013), SREM2D-generated runoff ensembles consistently underestimate actual error variability.

The Precipitation Uncertainties for Satellite Hydrology (PUSH) framework quantifies SMPP error by explicitly modeling each combination of rain and no-rain cases for SMPP and reference rainfall (Maggioni et al., 2014). For each of four cases (correct no-rain detection, missed case, false alarm, and hit), the framework generates a PDF of the actual precipitation for a given day based on the satellite observation. Similar to the SREM-2D approach, PUSH requires several different distributions to model SMPP error and separately models the discrete and continuous portions of SMPP uncertainty. Unlike SREM-2D, PUSH doesn't produce spatially or temporally autocorrelation realizations of possible true precipitation. Neither SREM-2D nor PUSH have had success in modeling SMPP uncertainty at extreme precipitation rates (Maggioni et al., 2013; Wright et al., 2017).

Various other models have been developed to model SMPP uncertainty, but have limited applicability for a variety of reasons. Nonparametric models such as that developed by Gebremichael et al. (2011) combine a positive continuous distribution and a point mass at 0 to create the conditional distribution of rainfall (Gebremichael et al., 2011). Because the positive continuous distribution is nonparametric, sampling from the conditional distribution is more difficult that selecting a random value on [0,1] and retrieving the associated quantile. Sarachi et al. (2015) developed a model of SMPP uncertainty using a generalized normal distribution (GND) model which provides a distribution of possible reference rainfall given an SMPP observation. However, the model is only defined for hit cases where both SMPP and a radar reference record rainfall above a set threshold (1 mm day^{-1}). Since false alarms and missed cases can comprise a significant portion of SMPP errors, this model is insufficient for characterizing SMPP uncertainty in its entirety.

CSGD–based Error Model

An error modeling framework based on the Censored Shifted Gamma Distribution (CSGD) was developed by Scheuerer and Hamill (2015) for post-processing ensemble numerical precipitation forecasts. It was adapted by Wright et al. (2017) to model errors in SMPPs, and is capable of

quantifying all three types of SMPP errors in a streamlined approach. The CSGD is an adaptation of the two-parameter gamma distribution with an additional "shift" parameter, which shifts the probability density function (PDF) leftward. The distribution is left-censored at zero, replacing all negative values with zero. In the resulting distribution, the area under the PDF left of zero represents the probability of zero precipitation while the probability density at any value greater than zero represents the relative likelihood of that amount of precipitation (Figure 2b). The CSGD is thus able to describe both precipitation occurrence and magnitude. The cumulative distribution function (CDF; Figure 2a) is defined by:

$$F_{k,\theta,\delta}(x) = \begin{cases} F_{k,\theta}(x-\delta)\ for\ x \geq 0 \\ 0 \qquad\qquad for\ x < 0 \end{cases} \quad \text{Eq. 1}$$

where μ, σ, and δ are the three climatological CSGD parameters and $F_{k,\theta}$ denotes the CDF of a two-parameter gamma distribution. CSGD parameter estimation is based on minimization of the continuous ranked probability score (Scheuerer & Hamill, 2015). The CSGD error model is trained using coincident historical records of SMPP and ground-truth reference precipitation observations. First, a "climatological CSGD" is fitted to the record of ground-truth precipitation, as described in Scheuerer & Hamill (2015) and Wright et al., (2017). Then, a nonlinear regression is trained using the climatological CSGD parameters and the historical timeseries of satellite and ground-truth observations. A linear regression can also be used in the CSGD framework, but for this work we always use the nonlinear regression approach. This regression model, defined with regression parameters $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, can generate unique "conditional" CSGD parameters μ(t), σ(t), and δ(t) conditioned on an SMPP observation, $R_s(t)$, at time *t* (Figure 2b).

$$\mu(t) = \frac{\mu}{\alpha_1}\log\left(1 + (e^{\alpha_1} - 1)\left(\alpha_2 + \alpha_3 \frac{R_s(t)}{\overline{R_s}}\right)\right) \quad \text{Eq. 2}$$

$$\sigma(t) = \alpha_4 \sigma \sqrt{\frac{\mu(t)}{\mu}} \quad \text{Eq. 3}$$

$$\delta(t) = \delta \quad \text{Eq. 4}$$

where $\overline{R_s}$ is the mean of the historical satellite observations. Regression model parameters $\alpha_1, \alpha_2, \alpha_3, and\ \alpha_4$ are estimated using minimization of the continuous ranked probability score. Conditional CSGDs describe the uncertainty surrounding any SMPP observation and represent the

estimated distribution of possible "true" precipitation at time t based on that observation (Figure 2b). The mean or median of the conditional CSGD can be used to deterministically correct systematic SMPP biases and the quantiles can generate confidence intervals of true precipitation. The CSGD framework implicitly involves statistical downscaling when the spatial resolution of the satellite and ground-reference

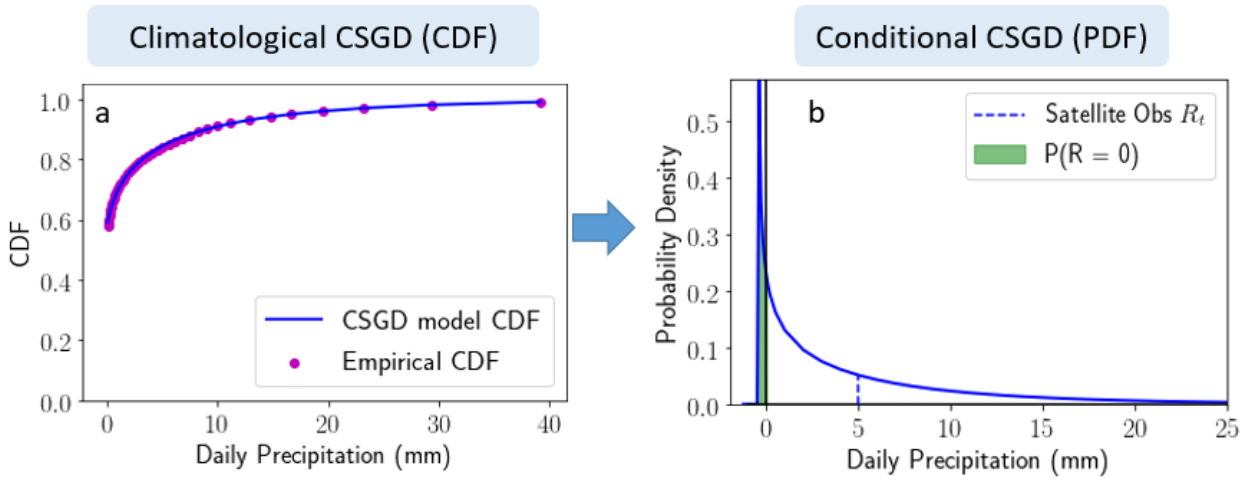

***Figure 2**: The CSGD Model fits ground reference precipitation to (a) a climatological CSGD for daily precipitation and optimizes a nonlinear regression to produce (b) conditional CSGDs based on satellite observations of daily precipitation.*

precipitation records differ, as shown in Scheuerer and Hamill (2015), who downscaled numerical precipitation forecasts from 0.5˚ to 0.125˚. The CSGD error modeling framework can be applied to daily and subdaily precipitation. Additionally, this framework allows for the incorporation of additional variables such as atmospheric data or geophysical characteristics (equations not shown). For a more detailed methodology of the CSGD Framework, see Wright et al. (2017) and Scheuerer and Hamill (2015).

2.3 Incorporating SMPP Uncertainty into Modeling Applications

While several SMPP error models are discussed in the previous section, there remains a gap between quantifying the range of SMPP error using these models and the incorporation of SMPP uncertainty in modeling applications. Several applications of SMPP data which don't account for the associated uncertainty are the Global Flood Monitoring System (GFMS; using TMPA; Wu et al., 2012), the Landslide Hazard Assessment for Situational Awareness (LHASA) model (using IMERG; Kirschbaum & Stanley, 2018), the Global Land Data Assimilation System (GLDAS; using TMPA; Rodell et al., 2004), and various disease tracking studies (Eisenberg et al., 2013;

Midekisa et al., 2012; among others). SMPP error models such as PUSH and the CSGD error model quantify uncertainty at a single time step, but don't necessarily provide a transferable method for environmental models which use SMPP input. Incorporating SMPP uncertainty may require simulating temporal autocorrelation and/or spatial autocorrelation in errors depending on the model purpose. For instance, flood modeling would require simulation of spatial autocorrelation in SMPP errors, while in land surface modeling, temporal autocorrelation of errors must be preserved, though spatial correlation is generally unnecessary. Additionally, all of the SMPP error models discussed are limited by their need for ground-reference data for model calibration. SMPP error model limitations related to autocorrelation, probabilistic form, and reliance on ground-reference data are discussed below.

Accounting for autocorrelation in SMPP error models

SMPP errors are autocorrelated in space and time, a difficult property to replicate during the generation of precipitation ensembles. SREM-2D addresses spatial and temporal autocorrelation by incorporating a lag-1 autocorrelation coefficient and spatially correlated random fields. Many other SMPP error models neglect autocorrelation of errors (Maggioni et al., 2014; Sarachi et al., 2015; Tian et al., 2013; Yan & Gebremichael, 2009). We have explored several approaches to adapt the CSGD-based error model framework for temporal autocorrelation. SMPP error can be quantified using the CSGD error model by finding the quantile of a conditional distribution generated at a given time step which corresponds to the recorded reference rainfall (Figure 3a,b,c,d). This is referred to as the 'true' quantile because picking that quantile of the conditional distribution, given satellite observed rainfall on a given day, would produce the correct rainfall. Even though the conditional distribution of possible precipitation values changes with the SMPP observation at each time step (Figure 3b,c,d), mapping these values into their quantiles can provide a more consistent measure of SMPP error. One issue in this approach so far is that the true quantile when both SMPP and ground-reference sources record no precipitation is always the same value (Figure 3a,b,e). When the ground-reference records no precipitation, the true quantile is equal to the probability of no precipitation (as depicted in Figures 2b, 3b). If the probability of zero precipitation is 0.8 for a conditional distribution based on an SMPP retrieval of 0 mm/day, the true quantile will only vary between 0.8 and 1 (Figure 3b). Since zero precipitation is the dominant case in a timeseries of daily precipitation, our timeseries of true quantiles contains many of these values (Figure 3e). This indicates that a timeseries of possible true quantiles would have to be

generated conditional on whether or not an SMPP retrieval recorded precipitation on each day. Additionally, the probability of zero precipitation changes based on the SMPP observation, such that different SMPP observations correspond to different ranges of true quantiles. To address this, we normalize true quantiles using:

$$q_{norm} = \frac{q - P(R=0)}{1 - P(R=0)} \quad \text{Eq. 5}$$

Timeseries of normalized quantiles in our study area generally have a lag-1 autocorrelation coefficient between 0.15 and 0.25 (Figure 3f). Preserving temporal autocorrelation in SMPP uncertainty estimates generated by the CSGD error model is an ongoing area of study.

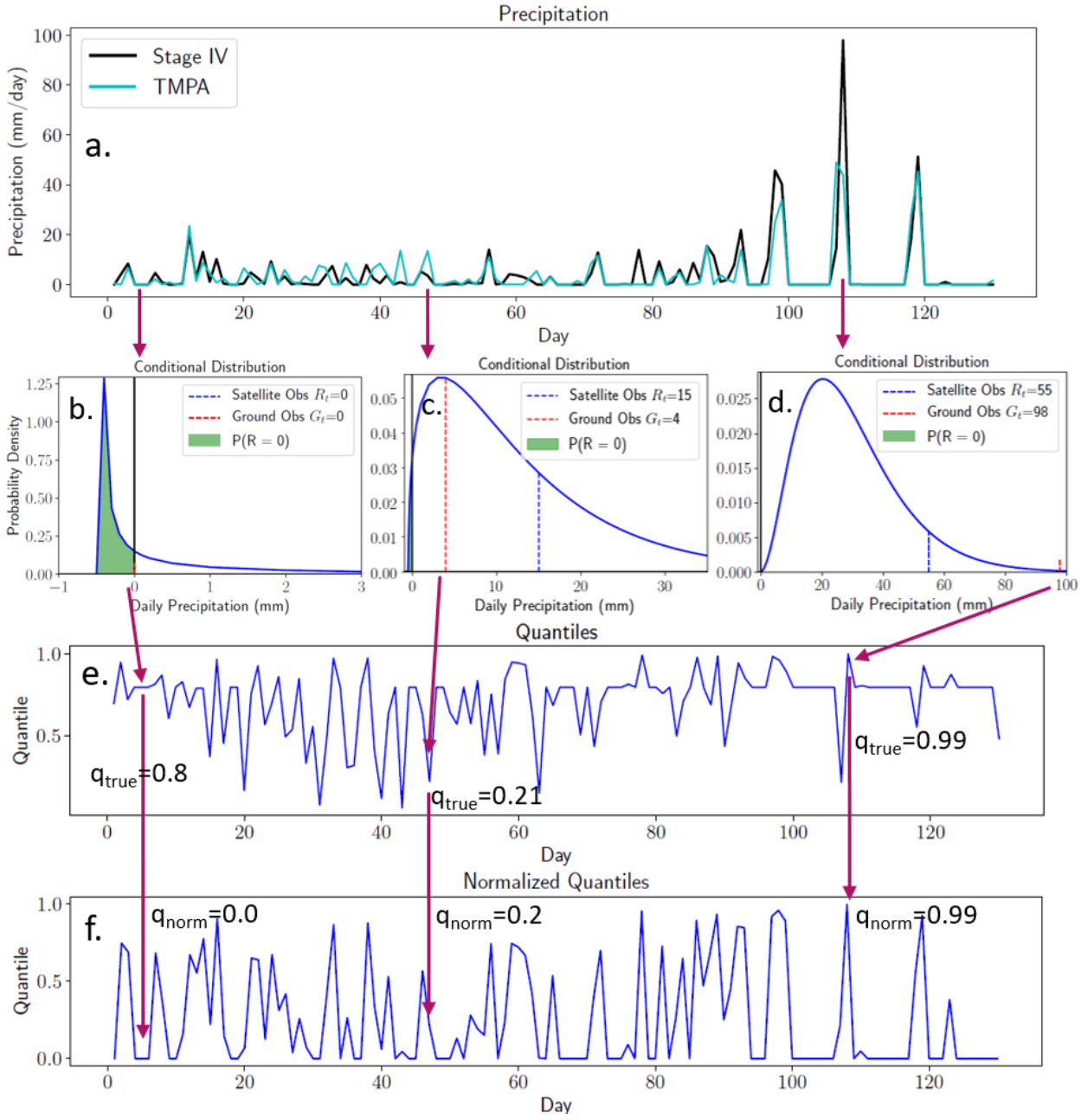

Figure 3: *(a) Timeseries of Stage IV radar and TMPA daily precipitation in Asheville, NC during 2004. Conditional distributions generated by CSGD error model for several select days are shown with SMPP observations of b) 0 mm/day, c) 15 mm/day, and (d) 55 mm/day. The actual ground observations on the corresponding days are shown in red to demonstrate how a true quantile is determined. e) Timeseries of true quantiles and (f) timeseries of normalized quantiles corresponding to Stage IV and SMPP precipitation in (a).*

Using probabilistic precipitation inputs

Another barrier to the application of current SMPP error models is the probabilistic form of their output. SREM-2D, PUSH, and the CSGD Framework all generate distributions to describe uncertainty in SMPP observations, reflecting the prevalence of random error in SMPP data. While it may seem logical to select the median or mean value of a distribution to use as the 'true' precipitation for an environmental model, this only addresses systematic bias and isn't a suitable way to account for uncertainty in SMPP input. Of course, most environmental models for land-surface processes, flooding, or landslide hazard are designed to intake a single value of precipitation at every time step and not a probabilistic estimate of precipitation. Further studies on the coupling of probabilistic precipitation input to hydrologic models would need to demonstrate the utility of this approach to potential users. Given the ensemble form of many current precipitation forecasts however, a transition towards probabilistic precipitation inputs is appropriate for certain hydrologic or environmental models (Hamill et al., 2013; Leutbecher et al., 2017). In Maggioni et al. (2013), SREM-2D ensemble members of simulated precipitation are individually input to a rainfall-runoff model, requiring 50 separate runs of the model to generate a 50-member ensemble of streamflow. While this is one 'brute force' method of incorporating probabilistic estimates of SMPP data in a model, it is fairly resource intensive. In Chapter 3, we describe an approach to adapting a decision tree model to probabilistic input in a manner that avoids the brute force of ensemble simulation. Though not straightforward, there are likely similar ways to adapt other types of environmental models such as LSMs to similarly "ingest" precipitation uncertainty information, rather than individual ensemble members.

Error–modeling in data–limited regions

The dependence of existing SMPP error models on ground-reference data for calibration may preclude the usage of SMPP error models in the data-limited regions that could most benefit from SMPP data and quantification of SMPP uncertainty. To address this limitation, Gebregiorgis & Hossain (2014) analyze the hydrologic predictability of several SMPPs as a function of geophysical features (topography, climate, seasons) in order to merge these SMPPs into a more accurate source of global precipitation data. The runoff error variance of three SMPPs, TMPA-RT (real time), CMORPH, and PERSIANN, when used in a land surface hydrologic model is assessed over several regions with available ground-reference data and used to create a global regression error model for each SMPP and geophysical feature combination (Gebregiorgis & Hossain, 2014). The strengths of each SMPP algorithm are combined in a performance based merged product

(PBMP), derived from a regression which linearly weighs the inverse of error variance. The merged product improves probability of detection and has comparable or smaller RMSE in validation regions across all seasons, but exhibits more variance in false alarm ratio. This model deterministically addresses SMPP uncertainty and it doesn't provide an estimate of uncertainty for the final merged product, which still exhibits some error. In a similar approach, Tang & Hossain (2012) investigate how SMPP uncertainty varies as a function of Koppen climate class in several regions of the world. They find that bias and RMSE are similar as a function of Koppen climate zones in the US and Australia, but run up against issues in contrasting ground validation data formats.

Using geophysical variables to predict error model parameters and SMPP uncertainty presents one approach to modeling SMPP uncertainty in data-limited regions. Another approach is to calibrate an error model using the most accurate SMPP retrieval data available, precipitation radar, though very limited sources of data are available for this (Precipitation Radar on the TRMM satellite and the Dual-frequency Precipitation Radar on the GPM Core Observatory). Training SMPP error models on minimal ground reference data is an option for regions that have at least a few ground-reference records. In Chapter 3, we demonstrate that a regionally-trained CSGD error model, calibrated using concatenated (i.e. "pooled") timeseries from across the study region, can obtain similar performance gains relative to a localized CSGD model, calibrated independently for each SMPP grid cell, and believe that several timeseries may be sufficient to calibrate the CSGD model in data-limited regions. Though this approach doesn't completely eliminate the model's dependence on ground reference data, it does reduce the data requirements for regional application of the CGSD framework in data-limited regions.

Incorporating SMPP uncertainty into a landslide hazard model

The global Landslide Hazard Assessment for Situational Awareness (LHASA) model issues hazard 'nowcasts' to regions around the world that are assessed to be at risk for landslide activity. LHASA uses SMPP input to calculate an antecedent rainfall index (ARI) and a global susceptibility map to assess landslide hazard at a 1km scale. In Chapter 3 we develop a Probabilistic LHASA model using the CSGD error model to general probabilistic estimates of antecedent rainfall as LHASA model input. Since landslides occur within relatively isolated spatial and temporal domains, we can neglect spatial autocorrelation in our model of SMPP uncertainty

and can calibrate the CSGD error model on antecedent rainfall timeseries to avoid having to simulate temporal autocorrelation. Our Probabilistic LHASA model demonstrates 1) the benefits of incorporating SMPP uncertainty into a landslide hazard model, 2) evidence that accounting for systematic bias alone actually reduces model performance, and 3) one approach to incorporating probabilistic precipitation data into a deterministic decision tree model.

3. Incorporating SMPP Uncertainty into a Landslide Hazard Model

This study incorporates SMPP uncertainty into the LHASA model framework using the CSGD error modeling framework of Wright et al. (2017). The CSGD framework characterizes SMPP error based on comparisons of historical SMPP and ground-based precipitation estimates, producing a statistical distribution of rainfall outcomes which can be used as input to LHASA. A new probabilistic version of LHASA is developed which takes in probabilistic rainfall input and produces probabilistic nowcasts (Section 3.3). Aside from incorporation of SMPP uncertainty, the probabilistic version of LHASA seeks to retain the original characteristics of the LHASA model as much as possible. The approach is implemented and evaluated in the mountainous Appalachian region of the southeastern United States (Section 3.4).

3.1 Background

Landslides result in thousands of fatalities, devastating property loss, and costly infrastructure damage around the world every year (Dilley et al., 2005; Froude & Petley, 2018; Petley, 2012). These natural hazards occur across a broad range of geographical, climatological, and land use settings and can range in scale from minor slope failures of several meters to kilometers-long debris flows. Factors that determine landslide hazard can be sorted into two categories: (1) static susceptibility factors which determine an area's pre-existing susceptibility to landslide events such as slope, aspect, forest loss, road cut activity, lithology, and distance to fault zones and (2) dynamic factors which trigger landslides such as rainfall, earthquakes, or anthropogenic activity (Dai et al., 2002; Sassa et al., 2014). Static factors can be conceptualized as determining *where* landslides are most likely to occur and dynamic factors as determining *when* landslides occur within susceptible areas. Though landslides can be initiated by seismic and human activity, rainfall is recognized as the most widespread and frequent trigger (Dai et al., 2002; Petley et al., 2005).

Most existing landslide hazard monitoring systems that cover city-to-nationwide scales use ground-based precipitation measurements. Japan and Norway, for example, operate nationwide early-warning systems using ground radar rainfall measurements (Krøgli et al., 2018; Devoli et al., 2015; Osanai et al., 2010), while Italy's early-warning system for rainfall-induced landslides, SANF, and Rio de Janiero's "Alerta Rio" system rely on rain gage networks (Calvello et al., 2015; Piciullo et al., 2017; Rossi et al., 2012). In many parts of the globe, however, including in steep

terrain and in developing countries, such measurements are usually lacking (Gebregiorgis & Hossain, 2014; Kidd et al., 2017); this limits the ability to monitor and warn of potential landslide hazards in real time.

Remote sensing products developed from satellite data have been used to estimate static susceptibility factors such as slope and land cover, and to statically characterize landslide hazard (Hong et al., 2007; Nadim et al., 2006). Light detection and ranging (LiDAR) products that repeatedly measure land surface elevations have also been used to monitor changes in land surface elevation to identify landslide events in remote regions. Using remotely-sensed satellite precipitation products to characterize dynamic landslide hazard on a global scale was first proposed by Hong et al. (2006) while introducing the first proof of concept for a near-realtime global satellite-based landslide hazard assessment system that combined TRMM rainfall data with a global susceptibility map (Hong et al., 2006). Farahmand & Aghakouchak (2013) developed a global landslide model based on Support Vector Machines, using PERSIANN precipitation data to globally classify landslide and non-landslide events. Kirschbaum et al. (2015b) developed the Landslide Hazard Assessment for Situational Awareness (LHASA) framework to assess landslide hazard in Central America based on TMPA. The LHASA model framework is now available at the global scale and uses precipitation data from IMERG to provide publicly-available near real-time nowcasts of landslide hazard around the world (D. Kirschbaum & Stanley, 2018). None of these models address the issue of SMPP uncertainty, however, potentially leading to mischaracterization of landslide hazard.

3.2 Study Region and Data

The study region extends from 81.0°-86.0°W and 37.0°-34.0°N, covering the Appalachian Mountains in western North Carolina and eastern Tennessee, extending north into Virginia and Kentucky, and south into Georgia and South Carolina (**Error! Reference source not found.**, 5). Extreme precipitation events (EPEs) have been reported as the primary natural hazard in this region, often causing floods and landslides that result in deaths and economic losses (Moore et al., 2015). It is characterized by complex terrain and a variety of EPE-producing weather modes, such as tropical cyclones, mesoscale convective systems, orographic lifts, and atmospheric rivers, which make regional rainstorm, flood and landslide monitoring over this region challenging (Barros et

al., 2014; Darby et al., 2016; Mahoney et al., 2014). In September 2004, Hurricane Frances was followed by

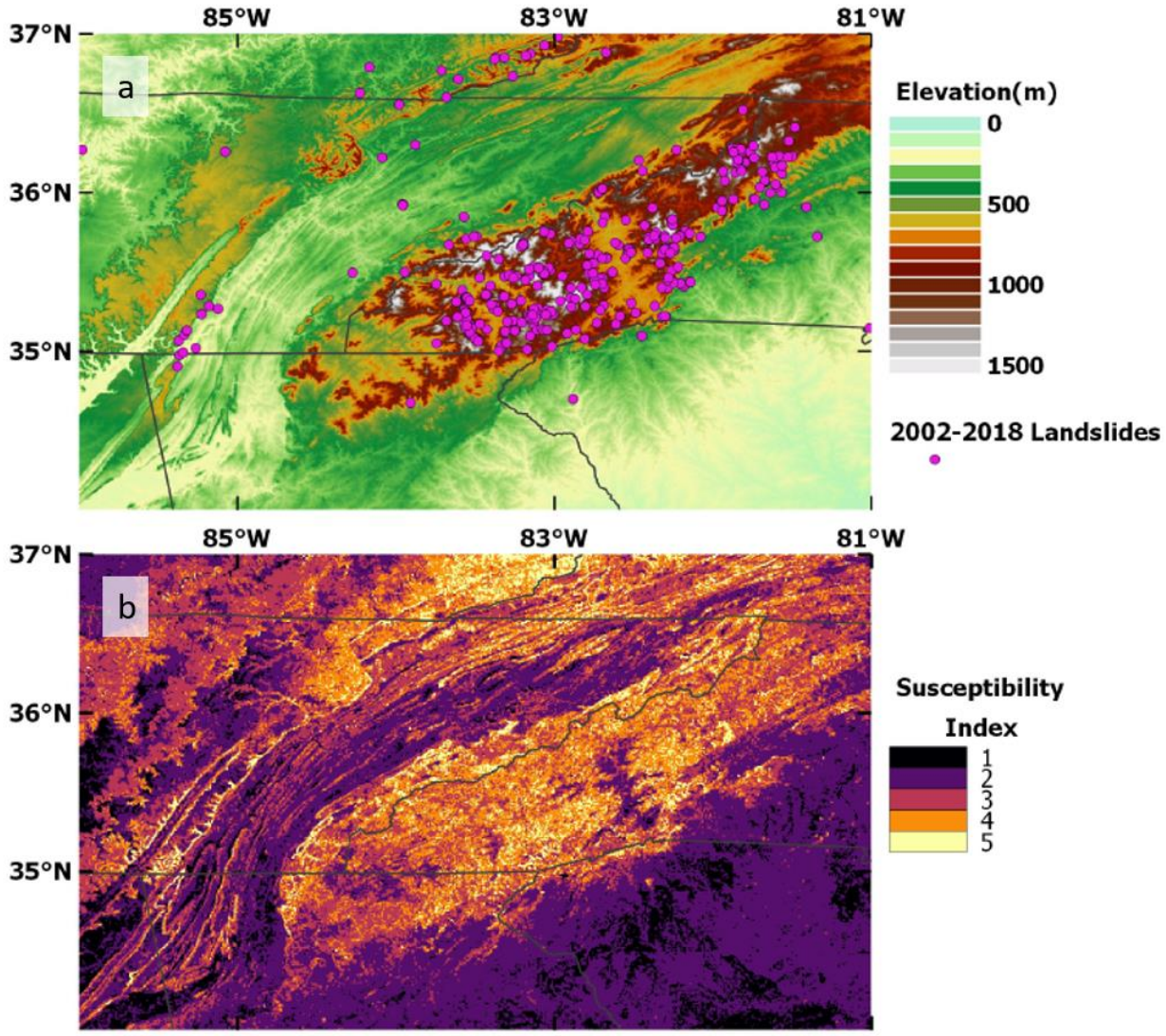

***Figure 4**: (a) Elevation and recorded landslides over 2002-2018 (b) Static susceptibility map at 1-km resolution*

Hurricane Ivan within a two-week period, causing approximately 400 landslides, 11 deaths, and widespread property damage (Boyle, 2014). Three fatalities occurred during May 2018, the wettest month on record in Asheville, NC, (Doom, 2018, Carter, 2018). SMPPs exhibit considerable errors relative to ground-based observations in the region (Figures 1, 5).

This study uses the TMPA-RT (real-time), IMERG-Early, and IMERG-Late products. TMPA is obtained for the period 2002-2018 (specifically, 2002-01-01 to 2018-06-30) and IMERG for the

period 2014-2018 (specifically, 2014-03-12 to 2018-06-30). The NEXRAD Stage IV radar-gage merged precipitation product, available over the continental U.S. from 2002-2018 at hourly,

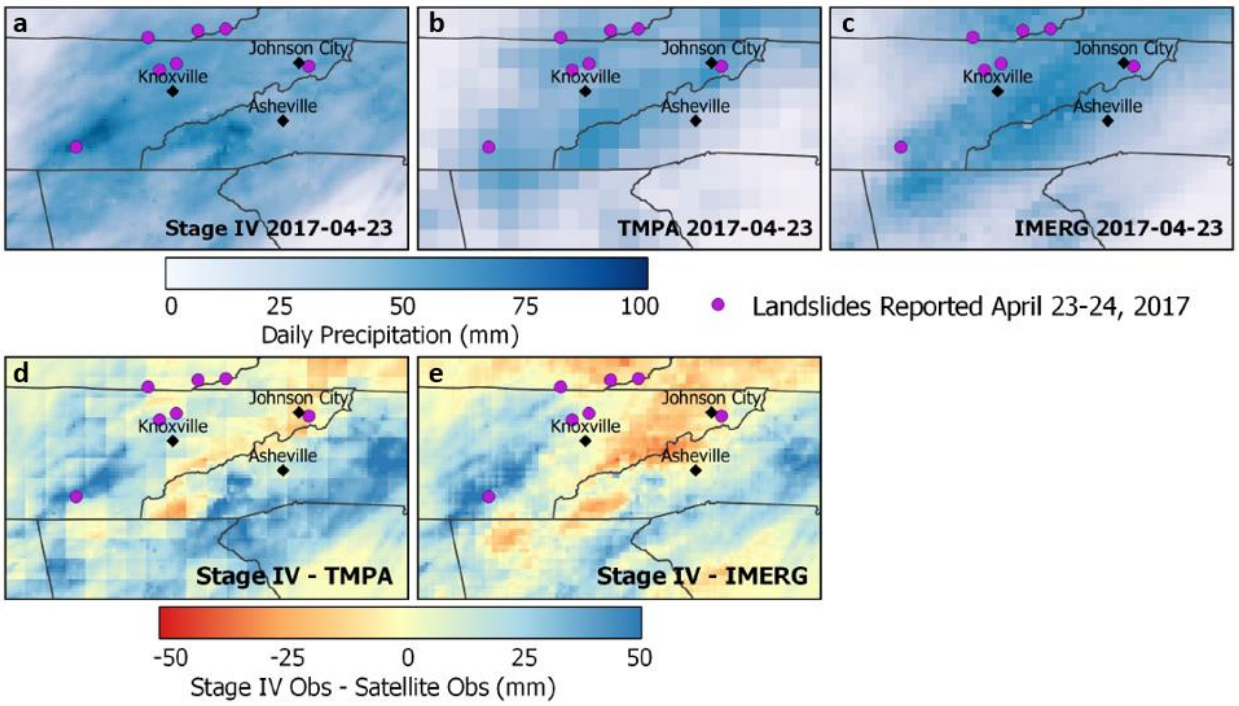

Figure 5*: Daily precipitation on April 23, 2017 observed by (a) Stage IV, (b) TMPA, and (c) IMERG. The seven landslides reported in the study region between 23-24 April 2017 are denoted by circles. Difference between daily precipitation observed on 23 April 2017 by (d) Stage IV and TMPA and (e) Stage IV and IMERG*

roughly 4 km resolution, is used as "ground-truth" reference precipitation. All precipitation data are aggregated to a daily timescale for this analysis.

Within the LHASA model (described further in Section 3.3), this study uses a 1-km global landslide susceptibility map (**Error! Reference source not found.**b) developed by Stanley and Kirschbaum (2017). The map depicts static Susceptibility Index (SI), determined using a fuzzy overlay model to combine global gridded data on five static factors: slope, geology, distance to fault zones, presence of roads, and forest loss. SI consists of integer values from one to five, corresponding to "very low," "low," "moderate," "high," and "very high" susceptibility. A more thorough review of the methodology used to create the susceptibility map can be found in Stanley and Kirschbaum (2017).

Landslide inventory data were obtained from the Global Landslide Catalog (GLC) and the Cooperative Open Online Landslide Repository (COOLR) (Kirschbaum et al., 2010; Kirschbaum et al., 2015a). Over 3,000 landslides have been recorded in the study region since 1916. Landslides without a reported location accuracy within 1 kilometer, as well as those with reported dates of occurrence spanning more than four days or which did not list rainfall as the event trigger were excluded to ensure that the LHASA validation (described in Section 3.3) is based upon the most reliable observations. This resulted in an inventory of 214 landslides over the period 2002-2018, approximately half of which are associated with Hurricane Frances and Hurricane Ivan in September 2004.

Table 1: Susceptibility zone coverage and landslide occurrence in study region using landslides reported from 1916-2018.

Susceptibility Zone	1 (very low)	2 (low)	3 (moderate)	4 (high)	5 (very high)
% of study area covered	8.0	42.0	29.0	16.0	5.0
Number (%) of reported landslides 1916-2018	0 (0)	8 (0.3)	830 (31.3)	1450 (54.7)	362 (13.7)
Number (%) of reported landslides 2002-2018	0 (0)	5(1.3)	68 (17.5)	225 (58.0)	90 (23.2)

3.3 Methods

The LHASA Model

The LHASA model characterizes rainfall-induced landslide potential worldwide in near real-time using a global susceptibility map and low-latency satellite precipitation data (Kirschbaum & Stanley, 2018) and is based on a regional model developed by Kirschbaum et al. (2015) that uses a heuristic decision tree model which combines static susceptibility with dynamic precipitation (Figure 6). LHASA outputs are “moderate-hazard” and “high-hazard” nowcasts issued only for areas that are currently assessed to have potential for landsliding. No nowcasts are issued for pixels which are assessed to be currently not at risk.

Many studies have used rainfall intensity-duration (I-D) thresholds to quantify landslide triggering behavior (Li et al. , 2011; Mathew et al., 2014). LHASA calculates a weighted measure of 7-day precipitation termed antecedent rainfall index (ARI):

$$ARI = \frac{\sum_{t=0}^{6} p_t w_t}{\sum_{t=0}^{6} w_t} \quad \text{Eq. 6}$$

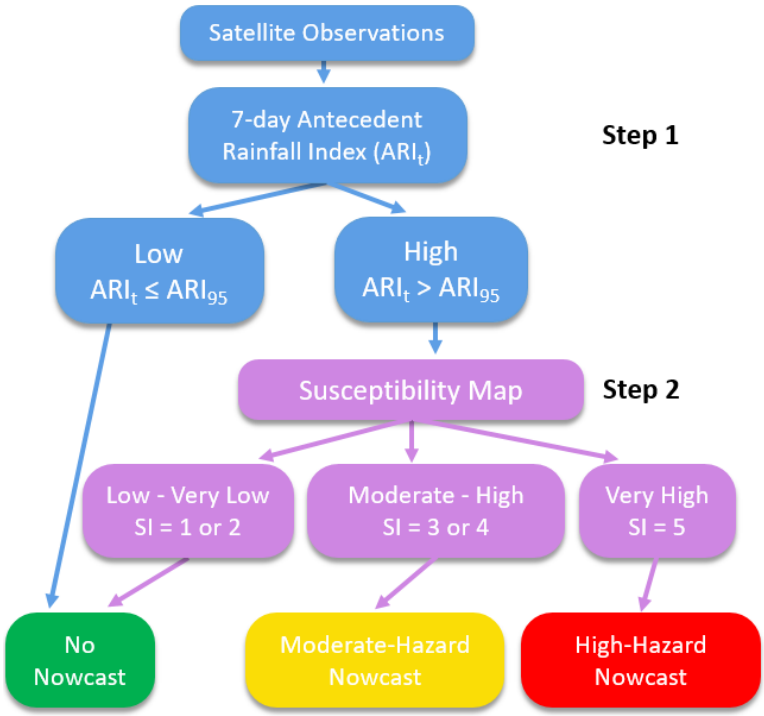

***Figure 6**: LHASA Model decision tree framework (Kirschbaum & Stanley, 2018)*

ARI is formulated to account for the combined impacts of current and recent precipitation on slope stability (Kirschbaum et al., 2015; Kirschbaum & Stanley, 2018), since soil moisture, pore water pressure, and other physical processes are not explicitly modeled in LHASA's empirical framework. The weighting scheme used in calculating ARI is derived from previous studies on landslide-triggering rainfall thresholds (Crozier, 1999; Glade et al., 2000). The operational global LHASA model uses IMERG, though due to its longer record, Kirschbaum & Stanley (2018) also used TMPA to estimate potential for landslide events and evaluate the model framework. Similarly, we evaluate our modified version of LHASA (Section 3.3) using both IMERG and TMPA.

The LHASA model decision tree assesses landslide hazard in two steps. In Step 1, daily ARI is calculated and compared against the 95th percentile of historical ARI for that location. The current

day is included in the ARI using IMERG-Early observations. IMERG-Late observations are used for the prior six days, due to their somewhat improved accuracy. TMPA-RT is used to calculate ARI for all 7 days. If the ARI value is below the 95th percentile threshold, a pixel is classified as "No Nowcast." If a pixel's ARI value exceeds the 95th percentile threshold, the process proceeds to Step 2, in which the susceptibility map (reviewed in Section 3.2) is consulted. If the ARI of a pixel exceeds the 95th percentile, the SI of that pixel is considered. A nowcast will never be issued for pixels with SI of one or two, regardless of the ARI value. ARIs above the 95th percentile, combined with SI of three or four will result in a Moderate-Hazard Nowcast. High-Hazard Nowcasts are issued only for very high susceptibility pixels (SI = 5) when the ARI is above the 95th percentile.

Kirschbaum & Stanley (2018) show that LHASA is able to broadly assess global landslide hazard and characterize landslide hazard by season. Their results also revealed areas with high nowcast rates, but few reported landslides, suggesting that landslides may be underreported in such areas. This existing version of LHASA will henceforth be referred to as "Deterministic LHASA" in recognition that only the SMPP observation, and not the range of uncertainty associated with it, is used. Deterministic LHASA is run on a daily scale using TMPA for the period 2002-2018 and IMERG for the period 2014-2018, as well as using Stage IV for the period 2002-2018. This latter simulation is conducted to show the optimal performance of LHASA, since Stage IV is more accurate than either SMPP and for the purposes of this study is assumed to be "ground truth." It can help contextualize how much of LHASA's predictive power is associated with precipitation uncertainty. Stage IV has been used to validate SMPPs (e.g. Aghakouchak et al., 2011).

CSGD–based Error Modeling Framework

The CSGD error modeling framework could be applied to daily SMPP data to generate possible precipitation values from which to calculate ARI values (Equation 6) and create a distribution of possible ARIs. This approach would require simulating the temporal autocorrelation in daily precipitation values in order to obtain realistic ARI values. Instead, we train the CSGD model on ARI timeseries derived from daily SMPP and ground reference timeseries. This removes the need to simulate temporal autocorrelation, since ARI is a time-integrated precipitation value (Equation 6). Climatological CSGDs fit to ARI timeseries in the CSGD model show good fit with empirical ARI CDFs, while probability of zero precipitation is much lower than that for daily precipitation (Figure 7). In this study, we use the CSGD to downscale from TMPA's (IMERG's) 0.25˚ (0.1˚) resolution to the 0.0417˚ of Stage IV. For a more detailed methodology of the CSGD Framework, see Wright et al. (2017) and Scheuerer and Hamill (2015).

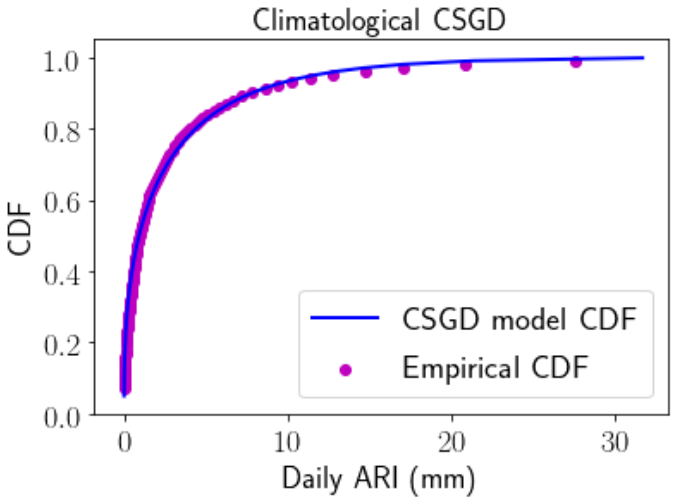

***Figure 7**: Climatological CSGD fit to Stage IV ARI in Asheville, NC during CSGD error model calibration*

Since we have Stage IV and SMPP observations of ARI over the entire study domain, we can obtain unique parameters $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ by training the error model on any two co-located SMPP and Stage IV ARI timeseries. This results in an optimal set of regression parameters, and thus a localized CSGD error model, for every individual Stage IV pixel. We indeed perform this localized parametrization, henceforth referred to as the "localized model". Previous SMPP error characterization and modeling efforts (Gebregiorgis & Hossain, 2014, 2013; Tang & Hossain, 2009, 2012), however, have noted that some form of "regionalization" of error estimates or error model parameters would be necessary in most locations, since ground-based precipitation data are typically limited (for example, to a limited number of rain gages within a region). We therefore adopt a regional parameterization scheme that is less optimal to better reflect the operational conditions. This consists of a regional set of regression parameters generated by concatenating the ARI timeseries of all pixels in the region and using these concatenated timeseries as input into the CSGD error model. This regional parameterization is henceforth referred to as the "regional

model." Regional and localized error models for daily ARI are trained on the six-year period 2002-2007 using TMPA and Stage IV and on the four-year period 2014-2017 using IMERG and Stage IV. This produces four CSGD error models: a regional and a localized model for application to TMPA ARI and a regional and a localized model for IMERG ARI. These models are compared in Section 3.4.

A Probabilistic LHASA Model

A new methodology based on the LHASA decision tree framework was developed to incorporate probabilistic precipitation estimates produced by the CSGD framework, henceforth referred to as "Probabilistic LHASA." This modified version is shown schematically in Figure 8.

In Step 1, a conditional CSGD is generated for each pixel based on the satellite-observed ARI and the trained CSGD model. The conditional CSGD is compared to the 95th percentile of historical SMPP ARI of the same pixel. We define the Rainfall Hazard Index (RHI) as $P(ARI > ARI_{95})$, or the area under the conditional CSGD PDF that lies above the 95th percentile theshold (highlighted as the light blue area in **Error! Reference source not found.**), quantifying the probability that the actual ARI is hazardous. In the CSGD PDF in Figure 8, the SMPP-observed ARI is depicted in red to illustrate how Deterministic LHASA evaluates rainfall hazard in comparison to probabilistic LHASA; while Deterministic LHASA model evaluates only whether the SMPP ARI is above or below the 95th percentile, Probabilistic LHASA considers the probability of the actual ARI being above this threshold.

After calculating the RHI, Probabilistic LHASA factors in susceptibility in Step 2 using the same global susceptibility map (Section 3.2). Pixels with very low or low susceptibility are assigned no nowcast. The rainfall hazard indices (RHIs) of pixels with moderate, high, and very high susceptibility are multiplied by 0.5, 0.75, and 1.0 respectively, to produce a landslide hazard index (LHI). Pixels with LHI less than 0.1 are reassigned as no nowcast. The resulting 0.1 to 1.0 LHI scale is continuous and reflects a range of moderate to very high landslide hazard, considering rainfall uncertainty. For convenience and potential communication purposes, this continuous scale

is divided into five equally-spaced categories: Category 1 LHI is (0.1,0.28], Category 2 LHI is (0.28,0.46], and so on up to Category 5 (Figure 8).

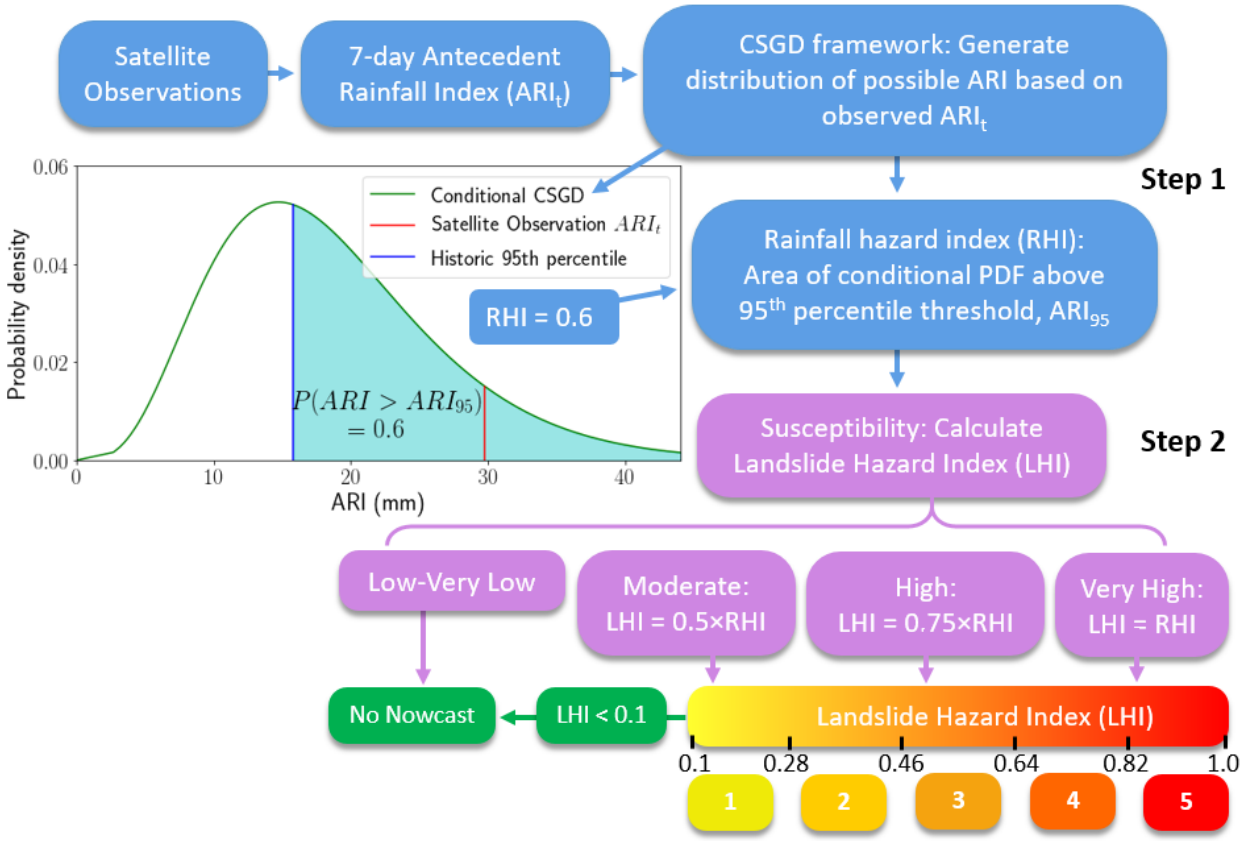

***Figure 8**: Probabilistic LHASA model framework*

No calibration of probabilistic LHASA version was performed to generate these categories. There are likely insufficient observed landslides in the study period to warrant such calibration. 86% of reported landslides in the study area occur in moderate and high susceptibility pixels (**Error! Reference source not found.**). This suggests that in Deterministic LHASA, it would be reasonable to allow not just moderate, but also high hazard nowcasts to be issued in moderate and high susceptibility zones under certain conditions. The 0.5 and 0.75 multipliers in Probabilistic LHASA were chosen to allow a Category 4 (3) hazard nowcast to be issued in high (moderate) susceptibility pixels under the condition that there is a near certain probability that ARI is above the 95th percentile. It could be beneficial to calibrate the relationship between RHI and LHI for moderate and high susceptibility zones, but this would require a more extensive landslide inventory. The

relationship between RHI and LHI in moderate to high susceptibility pixels may also vary based on the study region and could be tailored specifically to different regions.

Probabilistic LHASA is run at a daily scale using TMPA for the period 2002-2018 and IMERG for the period 2014-2018. Probabilistic LHASA is run using both the regional CSGD model and localized CSGD model.

CSGD–Adjusted Version of Deterministic LHASA Model

An additional version of Deterministic LHASA is implemented in which the median of the conditional CSGD for every satellite-observation of ARI is used as input. In Step 1, the regional CSGD model (for each respective SMPP) is used to generate "CSGD-adjusted" ARI values by taking the median of conditional CSGDs as the expected value of ARI. Taking the median of a conditional CSGD accounts for systematic bias because of how the CSGD framework has been trained; conditional CSGD parameters are generated such that the historical ground truth precipitation values most frequently associated with a specific SMPP retrieval (the 'expected values' for that retrieval) will be near the median. This derives from the CSGD error model calibration process which selects regression parameters which minimize the integrated quadratic distance between empirical and theoretical CSGD distribution functions (Wright et al., 2017). If the CSGD-adjusted ARI value is greater than the 95^{th} percentile threshold, then the model proceeds to Step 2 in the Deterministic LHASA decision tree (Figure 6). Using CSGD-adjusted SMPP data with Deterministic LHASA compares the utility of the median 'expected value' against using the entire conditional CSGD representing all potential ARI values. The median of the conditional CSGD corrects for systematic bias in SMPP data but does not address random error. Deterministic LHASA is run using CSGD-adjusted TMPA for the period 2002-2018 and CSGD-adjusted IMERG for the period 2014-2018.

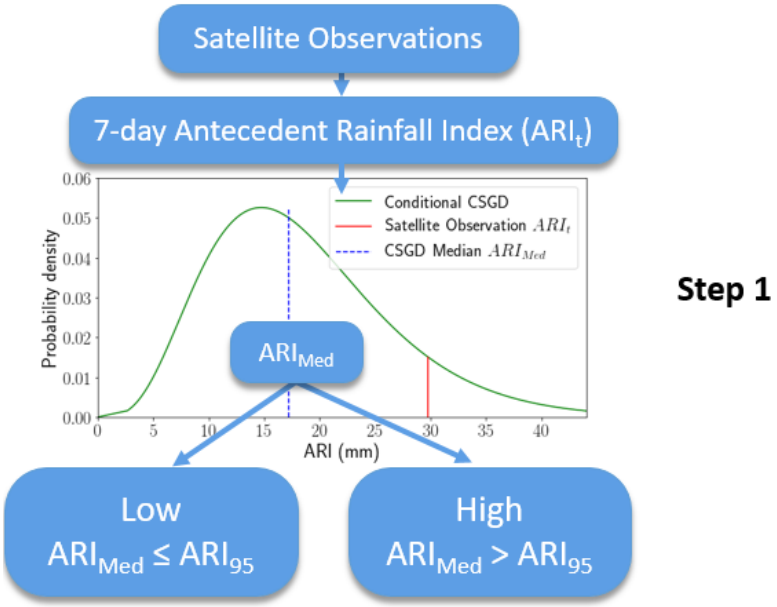

***Figure 9**: CSGD-Adjusted methodology for Step 1 of Deterministic LHASA. The median of the conditional CSGD for any given SMPP observation is used as input.*

Validation Approach

To assess the performance of Deterministic and Probabilistic LHASA, the true positive rate (TPR), false positive rate (FPR), and false alarm rates (FAR) are calculated. TPR is the percentage of reported landslides that are correctly detected by a landslide hazard nowcast. A nowcast is considered correct if it is issued the day of or the day before a recorded landslide in that pixel or any immediately adjacent pixel. A correct nowcast is said to "capture" a landslide. FPR is the percentage of instances when nowcast was issued but should not have been (i.e. no landslide was reported within 1 km). This is calculated as the number of pixel-days in which a nowcast was issued but no landslide occurred divided by the number of pixel-days in which no landslide occurred. FAR is the percentage of instances in which a nowcast is issued but no landslide occurs within 5 kilometers the day of or the day after the nowcast. FAR is calculated as the number of nowcasts issued that are not within 5 kilometers of a reported landslide the day of or the day after the report divided by the total number of nowcasts issued. The 5-kilometer spatial buffer is utilized in the FAR calculation to convey that nowcasts reported more than 1 kilometer, but less than 5 kilometers, away from a reported landslide are not necessarily false alarms.

3.4 Results

CSGD–based Error Model

CSGD localized and regionalized error models of TMPA-based ARI were trained for two time periods (2002-2007, 2002-2017) in Asheville, Knoxville, and Johnson City (Figure 10; locations are highlighted in Figure 5). The models are generally similar within each location except that the two time periods differ substantially in Knoxville and Johnson City. In Asheville and Johnson City, the regional model estimates higher ARI values than the local model; this difference becomes more pronounced at higher values of TMPA-based ARI. In Knoxville, the localized CSGD models estimate a higher range of ARI than the regionalized models. For ARI below 20 mm, the regional CI and median display similar fit to those generated from location-specific regression parameters. The lower boundary of the 90% CIs from all models differ by less by less than 10 mm for all

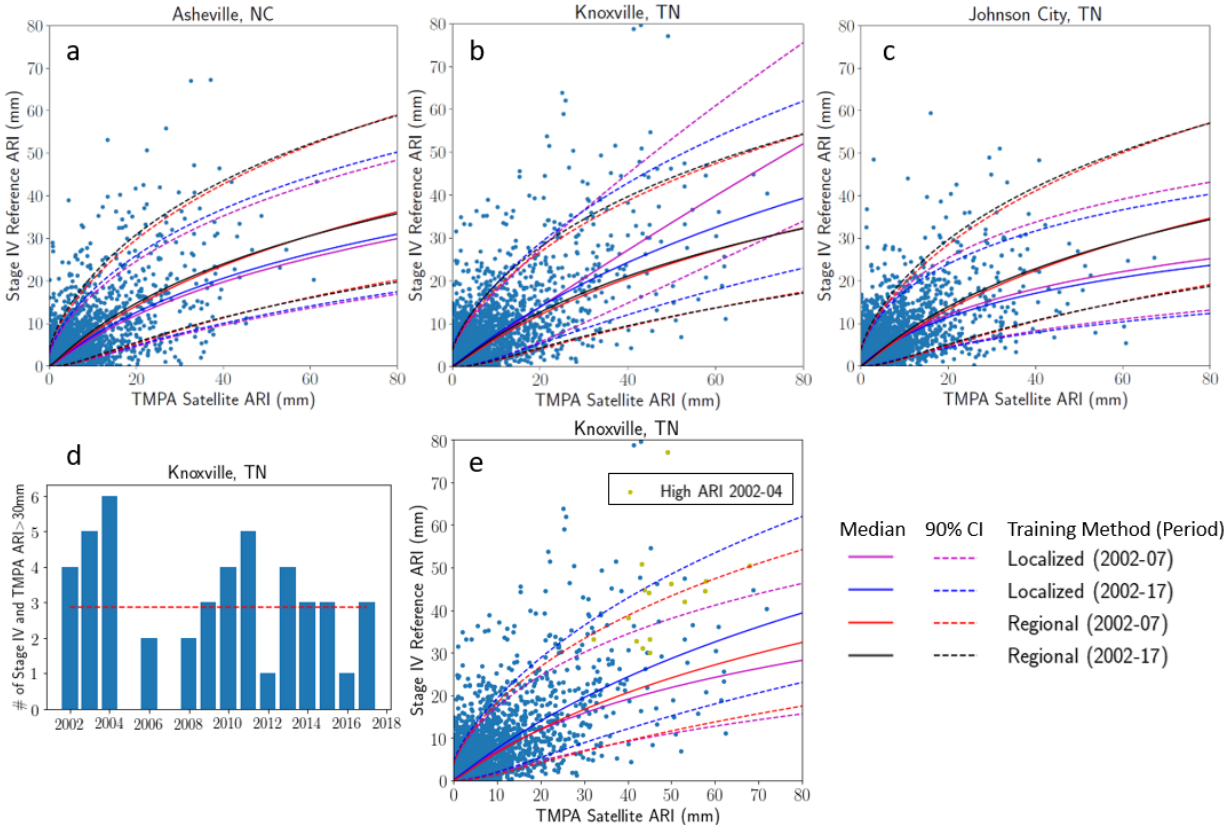

***Figure 10**: Median and 90% CI estimates of actual ARI given TMPA satellite observations using CSGD Models trained both regionally and locally during 2002-2007 and 2002-2017 in (a) Asheville, North Carolina, (b) Knoxville, Tennessee, and (c) Johnson City, Tennessee. (d) Number of days per year when Stage IV and TMPA concurrently record ARIs over 30 mm and (e) Knoxville, TN after all data points (in yellow) where Stage IV and TMPA concurrently record ARIs over 30 mm during 2002-2004 have been removed from the localized 2002-2007 model.*

locations except in Knoxville for the localized 2002–2007 model. When all data points from 2002-2004 where both Stage IV and TMPA concurrently record ARIs over 30 mm are removed from the training data for the localized 2002-2007 model, the 90% CI and median of the model are all within 10 mm of the regional 2002-2007 model (Figure 10e).

LHASA Model Performance

Fifty-five landslides were reported in the study region over September 16-17 when Hurricane Ivan brought high amounts of precipitation to the region. On September 17 hazard nowcasts are issued over much of the study area by both Deterministic and Probabilistic LHASA, encompassing all landslides (Figure 11a,b). Distributions of nowcast categories differ, however—Probabilistic LHASA issues high to very high hazard nowcasts mainly in the central part of the study area where most landslides were reported, while Deterministic LHASA issues high hazard nowcasts throughout the study area, including the northern portion where no landslides were reported. Conditional CSGDs for two TMPA pixels demonstrate how the deterministic and probabilistic models differ (Figure 11c,d; locations shown in Figure 11a,b). TMPA ARI for pixel at both locations 1 and 2 is above the 95th percentile, and so Deterministic LHASA issues moderate (high) hazard nowcasts for all moderate-to-high (very high) susceptibility areas within those pixels. Probabilistic LHASA, on the other hand, assesses the probability that the true ARI for each pixel is above the 95th percentile threshold by calculating the RHI. Pixel 1 in the north has an RHI of 0.37 while Pixel 2 in the center of the region has an RHI of 0.80 (Figure 11c,d). For this reason, Probabilistic LHASA issues higher hazard nowcasts in the center of the study area than the north.

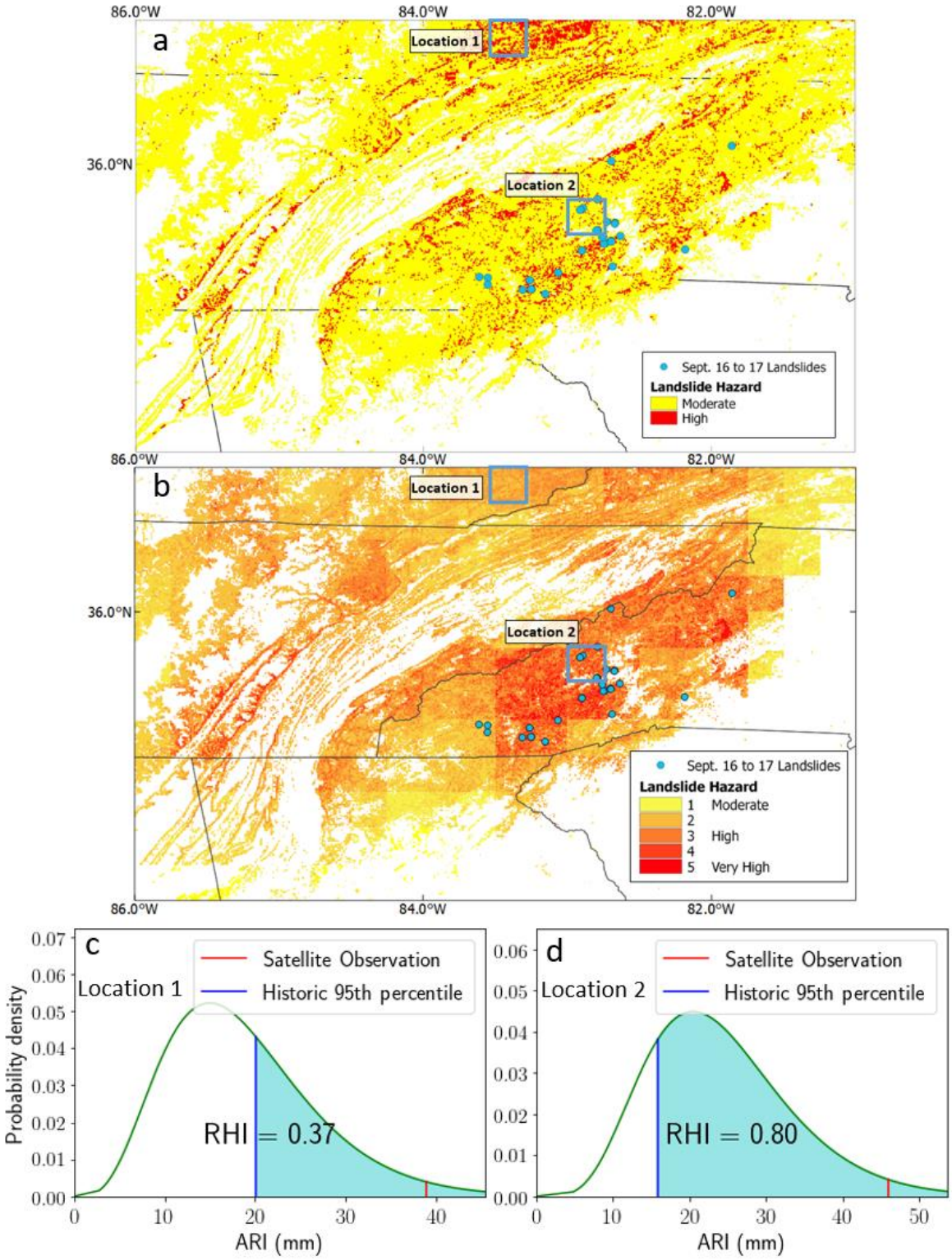

Figure 11*: (a) Deterministic LHASA model output using TMPA and (b) Probabilistic LHASA model output using regional CSGD model and TMPA on Sept. 17, 2004 and conditional CSGDs at (c) Location 1 and (d) Location 2 generated by probabilistic model to describe uncertainty surrounding SMPP-observed ARI.*

LHASA results for the 2002-2018 TMPA study period (Table 2) show that Deterministic (regional Probabilistic) LHASA using TMPA data captures 147 (155) of 214 reported landslides, while Deterministic LHASA using Stage IV data captures 192. TPR is fairly evenly dispersed for the probabilistic LHASA model between nowcast categories 2 through 4. The FPRs for categories 3-5 (high to very high hazard) are much lower than for the high hazard nowcasts in the deterministic model. The FPR for Category 5 nowcasts in Probabilistic LHASA is an entire order of magnitude lower than the FPR for high hazard nowcasts in Deterministic LHASA. Though Probabilistic LHASA using the localized instead of regional CSGD model captures five additional landslides, differences between FPR and FAR using the regionalized and localized models are minimal. Deterministic LHASA using CSGD-adjusted TMPA captures only 86 landslides but has lower FPR and FAR than Deterministic LHASA using TMPA (Table 2).

Table 2: Validation statistics for LHASA models over 2002-2018 study period. Probabilistic models use CSGD error model parameters trained for the 2002-2007 time period

LHASA Model and SMPP	Nowcast Category	TPR (%)	Landslides Captured	FPR (%)	FAR(%)
Deterministic TMPA	Moderate	23.4	49	2.06	99.95
	High	45.3	96	0.23	99.90
	Total	68.7	147 / 214		
Deterministic CSGD-adjusted TMPA	Moderate	15.4	33	0.58	99.94
	High	25.2	53	0.065	99.87
	Total	40.6	86 / 214		
Probabilistic (regional) TMPA	1	8.9	19	1.84	99.86
	2	21.0	45	0.59	99.74
	3	20.1	43	0.164	99.57
	4	16.8	36	0.045	99.12
	5	5.6	12	0.008	99.01
	Total	72.4	155 / 214		
Probabilistic (localized) TMPA	1	11.2	24	2.07	99.87
	2	19.6	42	0.64	99.74
	3	19.6	42	0.172	99.64
	4	15.9	34	0.045	99.40
	5	8.4	18	0.009	98.42
	Total	74.8	160 / 214		
Deterministic Stage IV	Moderate	28.0	60	2.29	99.94
	High	61.7	132	0.25	99.86
	Total	89.7	192 / 214		

Validation statistics calculated during the 2014-2018 study period demonstrate similar results (**Error! Reference source not found.**). Probabilistic LHASA using the localized CSGD error model is not included in Table 3 because validation statistics differed minimally from Probabilistic LHASA using the regional CSGD model (as during 2002-2018, Table 2). Probabilistic LHASA using the regional CSGD model with IMERG increases the number of captured landslides from 21 to 31 out of 46 (Table 3). With the exception of category 1, FAR for all Probabilistic LHASA nowcast categories is equal to or less than that of the Deterministic LHASA nowcast categories. Though Deterministic LHASA using TMPA produces a higher total TPR than with IMERG input, Probabilistic LHASA

Table 3: Validation statistics for LHASA models over 2014-2018 study period. Probabilistic models use CSGD error models trained over 2002-2007 for TMPA input and 2014-2017 for IMERG input.

LHASA Model and SMPP	Nowcast category	TPR (%)	Landslides Captured	FPR (%)	FAR (%)
Deterministic IMERG	Moderate	21.7	10	2.07	99.98
	High	23.9	11	0.22	99.97
	Total	45.6	21 / 46		
Deterministic TMPA	Moderate	21.7	10	2.17	99.98
	High	30.4	14	0.24	99.96
	Total	52.2	24 / 46		
Deterministic CSGD-adjusted IMERG	Moderate	13.0	6	0.90	99.97
	High	17.4	8	0.10	99.95
	Total	30.4	14 / 46		
Deterministic CSGD-adjusted TMPA	Moderate	13.0	6	0.66	99.97
	High	8.7	4	0.07	99.95
	Total	21.7	10 / 46		
Probabilistic (regional) IMERG	1	26.1	12	1.77	99.98
	2	10.9	5	0.74	99.96
	3	10.9	5	0.27	99.93
	4	13.0	6	0.10	99.91
	5	6.5	3	0.02	99.89
	Total	67.4	31 / 46		
Probabilistic (regional) TMPA	1	19.6	9	1.86	99.97
	2	23.9	11	0.62	99.96
	3	6.5	3	0.19	99.93
	4	13.0	6	0.06	99.91
	5	2.2	1	0.02	99.89
	Total	67.4	30 / 46		
Deterministic Stage IV	Moderate	26.1	12	2.52	99.97
	High	45.7	21	0.28	99.94

	Total	71.8	33 / 46		

with IMERG outperforms Probabilistic LHASA using TMPA. Probabilistic LHASA using IMERG increases landslides captured by 47.6% while Probabilistic LHASA using TMPA increased landslides captured by 25.0%.

The Deterministic LHASA model using CSGD-adjusted IMERG and TMPA performs poorly compared to both Deterministic LHASA and Probabilistic LHASA using original IMERG and TMPA. Using the CSGD median to adjust IMERG ARI in Deterministic LHASA produces a TPR that is less than half the TPR of Probabilistic LHASA using IMERG (Table 3). Deterministic LHASA using Stage IV highlights that even with "perfect" precipitation data, the model only captures 192 (33) out of 214 (46) reported landslides during the 2002-2018 (2014-2018) study period. Probabilistic LHASA model outputs are substantially closer to this "ground truth" performance than the SMPP-based Deterministic LHASA models.

Probabilistic LHASA has a higher total TPR using the localized CSGD model rather than the regional CSGD model for both the 2002-2018 and 2014-2018 study periods (Table 2, results not shown for 2014-2018), though this result is more varied when examined by nowcast category. The localized CSGD model increases the FPR in nowcast categories 1-3 (Figure 12) when using both TMPA and IMERG inputs.

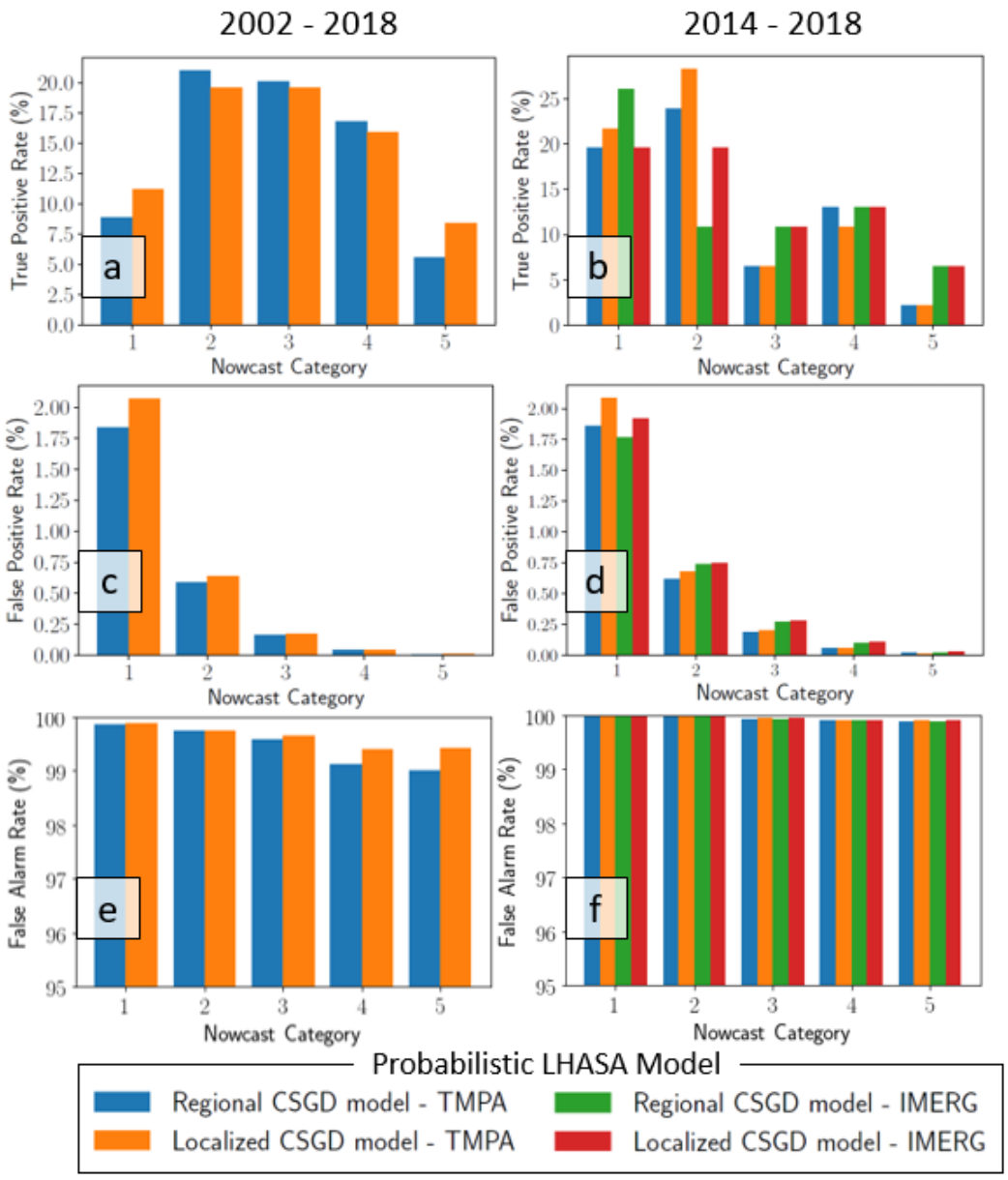

Figure 12: *(Top to Bottom) TPR, FPR, and FAR for both study periods using TMPA and IMERG.*

Seasonal and Spatial Variation

Nowcast rates over the study region, calculated as the percentage of days on which nowcasts are issued, are displayed in Figure 13 for 2002-2018 using TMPA. There is no differentiation between the category or hazard level of nowcasts. On average, Probabilistic LHASA with TMPA has a higher nowcast rate in the central mountains of the study area and a lower nowcast rate in the northwest (Figure 13a,b). The nowcast rate for both models is higher during winter months (Figure 13c,d), though Probabilistic LHASA still has a notably higher nowcast rate in the core of the

Appalachian Mountains. The nowcast rates from Probabilistic LHASA also exhibit significantly more spatial variability.

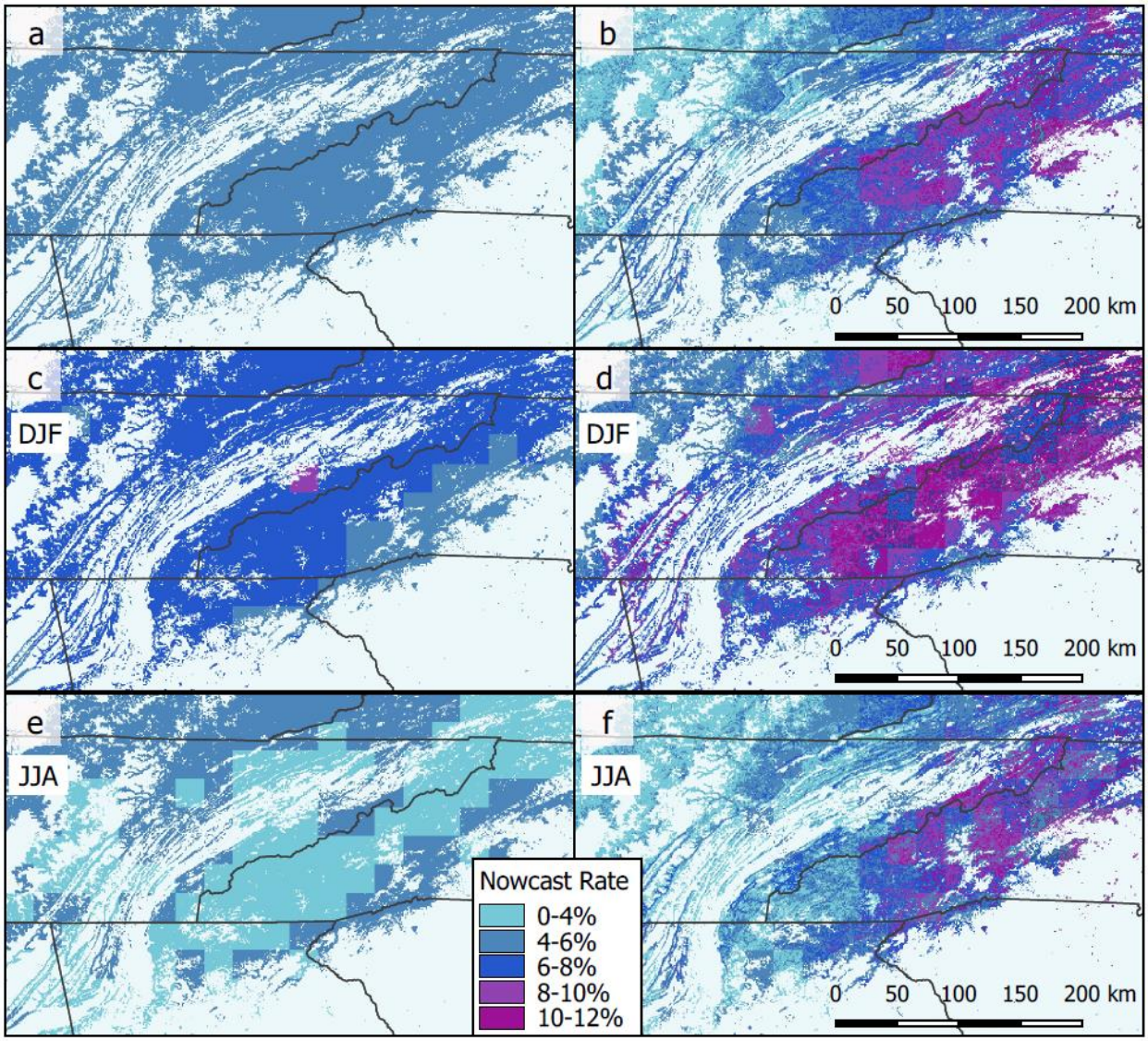

***Figure 13**: Nowcast rate (%) for (left) Deterministic LHASA and (right) Probabilistic LHASA using the regionalized CSGD error model and TMPA over 2002-2018 for (a,b) entire year (c,d) December, January, and February and (e,f) June, July, and August.*

3.5 Discussion

CSGD–based Error Model

Regional and localized CSGD error models trained using TMPA over 2002-2007 and 2002-2017 produce similar CIs, except in the case of the localized CSGD model for Knoxville, TN trained during 2002-2007 (Figure 10). The above-average number of high ARI observations recorded during 2002-2007 can explain the behavior of the localized CSGD model during that training period. 2002–2004, for instance, had an above average number instances of when Stage IV ARI and the observed TMPA ARI were both greater than 30 mm (Figure 10d). This explains why the 90% CI tends closer to the 1:1 line at high ARI values for the CSGD model trained locally over 2002-2007 than for other CSGD models (Figure 10b). When the 15 data points from 2002-2004 for which Stage IV and TMPA ARI are both over 30 mm are removed from the training data, the localized CSGD model trained over 2002-2007 bears a much stronger resemblance to the regional and localized CSGD models trained over 2002-2017 (Figure 10e). Though the CSGD error models trained locally and regionally over different periods may initially seem to differ based on Figure 10, their performance in Probabilistic LHASA, reviewed below, is the best way to evaluate how significant the differences between the models really are.

LHASA Model Performance

Probabilistic LHASA improves both total TPR and landslides captured relative to Deterministic LHASA (Tables 2 and 3), but comparing the distribution of TPR across nowcast categories is less straight forward since the two model frameworks use different nowcast categories. For Deterministic LHASA, the TPR for high hazard nowcasts is higher than for moderate nowcasts (Tables 2 and 3), and thus can be said to have higher skill. On the other hand, TPR for Probabilistic LHASA is, with the exception of category 1, fairly evenly dispersed across nowcast categories. Though this suggests equal skill, the FPR shows that there are fewer instances of higher category nowcasts being issued in error. The FPR of category 5 nowcasts is an order of magnitude smaller than the FPR of high hazard nowcasts from Deterministic LHASA. The lower TPR in categories 4 and 5 of Probabilistic LHASA can be attributed to the conditions required for such nowcasts to be issued; category 5 hazard nowcasts are only issued under the condition that there is 0.8 to 1.0 probability of the ARI being above the 95^{th} percentile threshold in a very high susceptibility area. Though the TPRs of categories 4 and 5 are not as high as the TPR of high nowcasts from

Deterministic LHASA, they offer the advantage of much lower FPRs and FARs. Low FPR and FAR are important for stakeholders who need to establish how confident they can be in a nowcast. FAR for both models is high, though, given the uncertainty in regional landslide hazard modeling, this is not unexpected.

When Probabilistic LHASA captures a reported landslide that Deterministic LHASA fails to catch, the SMPP ARI is below the 95th percentile. Though the probability of actual ARI being hazardous may not be high, Probabilistic LHASA can still generate a category 1 or 2 nowcast. This demonstrates how the Probabilistic LHASA could be especially beneficial in regions where satellite observations frequently underestimate heavy precipitation events.

The decision tree of Deterministic LHASA does not permit high hazard nowcasts to be issued to moderate or high susceptibility areas (**Error! Reference source not found.**). However, given that over 50% of landslides in the study region occur in pixels categorized as high susceptibility (SI = 4) by the static susceptibility map (Table 1), it seems appropriate for such areas to merit high hazard nowcasts under certain circumstances. The deterministic LHASA framework is limited because the model cannot issue a high hazard nowcast to all high susceptibility areas without greatly increasing the FPR and FAR of high hazard nowcasts. If the framework were to allow high susceptibility pixels to differ in nowcast category, some moderating factor would need to be introduced to decide whether a high susceptibility pixel is issued a moderate or high hazard nowcast. The probabilistic LHASA model framework (Figure 8) is advantageous in that it allows the level of hazard nowcasts to vary based on the probability of hazardous ARI. A category 3 or 4 nowcast can be issued to high susceptibility pixels under the circumstance that there is a very high probability the ARI is above the 95th percentile threshold. Thus, Probabilistic LHASA is able to differentiate between two pixels of the same static susceptibility, for which SMPP-observed ARI is above threshold, but which in actuality have varying probabilities of experiencing hazardous precipitation (Figure 11).

The poor performance of Deterministic LHASA using CSGD-adjusted TMPA and IMERG (Tables 2 and 3) demonstrates the importance of considering the entire range of possible errors provided by conditional CSGDs and not using expected value (median) to address SMPP uncertainty. Using the CSGD median adjusts ARI values for systematic bias but does not incorporate random error into the LHASA framework. Additionally, adjusting for systematic bias

minimizes SMPP observations of extreme ARIs, resulting in fewer nowcasts issued and fewer landslides captured. Though the FPR and FAR are lower for Deterministic LHASA using CSGD-adjusted values than using the original SMPP values of ARI (Tables 2 and 3), this is outweighed by the model's inability to correctly issue nowcasts for even half of the reported landslides during the study period.

The nowcast categories in Probabilistic LHASA are presented here as equally subdivided portions of the continuous LHI scale. However, the number and LHI range of these categories are flexible, since the continuous LHI scale (Figure 8) can be divided in any number of ways. This flexibility could be used to optimize specific performance metrics such as TPR or FPR. A stakeholder could, for example, use four categories instead of five, and specify that the highest category have a TPR of at least 0.25 or a maximum FPR of 0.001. This supports the original goal of the LHASA framework—to provide a flexible model for providing situational awareness of landslide hazard that can be fine-tuned as required (Kirschbaum et al., 2015; Kirschbaum & Stanley, 2018).

Regional and Localized CSGD Models

The performance of the probabilistic LHASA model using regionally-trained and locally-trained CSGD models is compared in Table 2 and Figure 12. The localized CSGD model increases total TPR during both study periods, capturing five additional landslides over 2002-2018 using TMPA and one and two additional landslides using IMERG and TMPA, respectively, over 2014-2018. The localized CSGD model also increases the FPR of nowcast categories 1 and 2 (Figure 12c,d). The regional and localized CSGD models produce near identical FPR and FAR for nowcast categories 3-5, with the exception of category 4 and 5 FAR for 2002-2018, when the localized model increases FAR (Figure 12e,f). Though validation statistics show that the localized CSGD model performs incrementally better than the regional model within Probabilistic LHASA, both substantially outperform Deterministic LHASA (Tables 2 and 3), indicating that inclusion of SMPP uncertainty is more critical than exactly how the incorporation is done. Even though regional and localized CSGD models produce varying estimates of uncertainty surrounding ARI observations (Figure 10), their incorporation into the Probabilistic LHASA model produces very similar results. In data-limited regions, localized uncertainty estimation is impossible since ground observations of precipitation are lacking, and thus is it probable that error modeling would require

a regionalization approach. Our results suggest that such an approach could in principle be successful, though nontrivial to implement.

Seasonal and Spatial Variation

The largest differences in nowcast rates between Deterministic and Probabilistic LHASA occur in the steepest terrain (Figure 13). This may indicate higher SMPP uncertainty in such terrain, and the impact that accounting for SMPP uncertainty may have in such mountainous regions where landslides typically occur. It is also noteworthy that the probabilistic LHASA model does not increase the nowcast rate across the entire study region, but actually decreases the nowcast rate in the lower-elevation northwest. It seems appropriate that the central mountainous terrain merit more frequent hazard nowcasts since it has experienced the majority of the landslides in the study region (Figure 4a).

The increased spatial variability in nowcast rate for Probabilistic LHASA is a result of how the probabilistic nowcast scale is generated; while every moderate to very high susceptibility area with a satellite-observed ARI above the 95th percentile will be issued a nowcast from Deterministic LHASA, Probabilistic LHASA allows for nowcasts to be issued to only very high or high and very high susceptibility pixels under certain conditions. For instance, if a TMPA-observed ARI is barely above the 95th percentile, Deterministic LHASA will issue nowcasts for all areas with moderate to very high static susceptibility. In contrast, Probabilistic LHASA will estimate the probability (i.e. RHI) that the true ARI is above the threshold and use this to determine which nowcasts to issue. An RHI of 0.2 would result in nowcasts being issued only for high and very high static susceptibility pixels. So while all moderate to very high susceptibility areas within an SMPP grid cell will have the same nowcast rate in Deterministic LHASA, the nowcast rate in Probabilistic LHASA will vary according to static susceptibility (**Error! Reference source not found.**).

Limitations

Like other SMPP error models, the CSGD Framework requires a ground-based reference record of precipitation, which is available in the study area of this particular analysis. In order to apply the probabilistic LHASA model elsewhere, ground-based data would also need to be available or some predictor of CSGD framework parameters would need to be established based on available data such as elevation, land use, slope, or climate. In data-scarce regions with limited ground-based precipitation records, a regional set of regression parameters could be estimated, as is done

in this analysis, by training the regression in the CSGD Framework on any available satellite and ground-based observations at different locations concatenated together into continuous timeseries.

4. Summary & Future Work

In Chapter 2, current barriers to incorporating SMPP uncertainty into environmental models are discussed and addressed in relation to the CSGD error model. We demonstrate how data-limitations and probabilistic form of SMPP uncertainty can be addressed within a landslide hazard model. Modeling temporal autocorrelation in SMPP errors will be addressed in future work with land surface models.

In Chapter 3, our study demonstrates how accounting for SMPP uncertainty can improve hazard assessments from a landslide hazard model over a medium-sized study region in the mountainous southeastern United States. Historical records of satellite and ground-based precipitation are used to train a CSGD-based error model, which characterizes the uncertainty in current and antecedent rainfall. A probabilistic version of the LHASA model framework is developed that can translate this rainfall uncertainty into landslide hazard nowcasts. The hazard nowcast scale adopts a probabilistic form based on a five category landslide hazard index. Probabilistic LHASA performs well in the study region, capturing more landslides than the deterministic version of LHASA and reducing false positive and false alarm rates in high-hazard nowcast categories. Using a CSGD-based error model to generate expected antecedent rainfall values (i.e. median of conditional distribution) as input for the deterministic LHASA model leads to poor model performance, confirming that correcting systematic bias alone is not sufficient for addressing SMPP uncertainty in models. For all models that utilize SMPPs, this demonstrates the need to account for the range of SMPP error in a probabilistic fashion. While our analysis confirms that slightly better results can be obtained using localized, as opposed to regional, parameterization of the CSGD error model within the probabilistic LHASA model, our objective is not to produce optimal nowcast results but rather demonstrate the value of precipitation uncertainty information in landslide prediction. We are successful in this regard, in that large performance gains are realized compared with the deterministic version of LHASA, even using the regional parameterization of the error model. In future work, a similar regional parameterization of the error model may be performed for regions with limited reference data. This study shows that moving from SMPPs as unknown sources of uncertainty to SMPPs as known sources of uncertainty can significantly improve model

performance, even without being able to deterministically derive true precipitation from SMPP observations.

While the performance of Probabilistic LHASA is primarily assessed using TMPA in this paper because of the longer study period (and therefore greater number of reported landslides for validation), evaluation using IMERG from 2014-2018 also shows positive performance. Once IMERG is post-processed to 1998, Probabilistic LHASA will be evaluated with the longer IMERG record. Though the probabilistic LHASA model framework was developed over a limited region in the U.S., its performance is promising for application to other regions, including data-limited areas. The probabilistic nowcast scale provides a way to adapt landslide hazard nowcast categories to stakeholder requirements.

Accounting for SMPP uncertainty allows landslide hazard models to better utilize SMPPs globally and in data-limited regions. Other environmental prediction models, such as the Global Flood Monitoring System, could likely benefit from deeper consideration of SMPP uncertainty. Since precipitation forecasts from the Global Forecasting System (Saha et al., 2006) are generally produced as ensembles, this methodology could be used to incorporate probabilistic precipitation forecasts as input to the probabilistic LHASA model to generate nowcasts at decreased latencies, as well as short-term landslide hazard forecasts. In future work, we will apply Probabilistic LHASA to data-limited regions and develop methods to account for SMPP uncertainty in land surface modeling applications.

References

Aghakouchak, A., Behrangi, A., Sorooshian, S., Hsu, K., & Amitai, E. (2011). Evaluation of satellite-retrieved extreme precipitation rates across the central United States. *Journal of Geophysical Research Atmospheres*, *116*(2), 1–11. https://doi.org/10.1029/2010JD014741

AghaKouchak, A., Nasrollahi, N., & Habib, E. (2009). Accounting for uncertainties of the TRMM satellite estimates. *Remote Sensing*, *1*(3), 606–619. https://doi.org/10.3390/rs1030606

Alemohammad, S. H., McLaughlin, D. B., & Entekhabi, D. (2015). Quantifying Precipitation Uncertainty for Land Data Assimilation Applications. *Monthly Weather Review*, *143*(8), 3276–3299. https://doi.org/10.1175/mwr-d-14-00337.1

Artan, G., Gadain, H., Smith, J. L., Asante, K., Bandaragoda, C. J., & Verdin, J. P. (2007). Adequacy of satellite derived rainfall data for stream flow modeling. *Natural Hazards*, *43*(2), 167–185. https://doi.org/10.1007/s11069-007-9121-6

Barros, A. P., Petersen, W., Schwaller, M., Cifelli, R., Mahoney, K., Peters-liddard, C., … Starr, D. (2014). *NASA GPM - Ground Validation: Integrated Precipitation and Hydrology Experiment 2014 Science Plan* (Vol. 12). https://doi.org/http://dx.doi.org/10.7924/G8CC0XMR

Bitew, M. M., & Gebremichael, M. (2011). Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands. *Hydrology and Earth System Sciences*, *15*(4), 1147–1155. https://doi.org/10.5194/hess-15-1147-2011

Calvello, M., D'Orsi, R. N., Piciullo, L., Paes, N. M., Magalhaes, M. A., Coelho, R., & Lacerda, W. A. (2015). The Community-Based Alert and Alarm System for Rainfall Induced Landslides in Rio de Janeiro, Brazil. In *Engineering Geology for Society and Territory - Volume 2* (pp. 653–657). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-09057-3_109

Crozier, M. (1999). Prediction of rainfall-triggered landslides: a test of the antecedent water status model. *Earth Surface Processes and Landforms*, *24*, 825–833. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291096-9837%28199908%2924%3A9%3C825%3A%3AAID-ESP14%3E3.0.CO%3B2-M

Dai, F. C., Lee, C. F., & Ngai, Y. Y. (2002). Landslide risk assessment and management: An overview. *Engineering Geology*, *64*(1), 65–87. https://doi.org/10.1016/S0013-7952(01)00093-X

Darby, L., Hughes, M., Cifelli, R., Sukovich, E., Jackson, D. L., Mahoney, K., … Neiman, P. (2016). Understanding the Role of Atmospheric Rivers in Heavy Precipitation in the Southeast United States. *Monthly Weather Review*, *144*(4), 1617–1632. https://doi.org/10.1175/mwr-d-15-0279.1

Dilley, M., Chen, R. S., Deichmann, U., Lerner-Lam, A., Arnold, M., Agwe, J., … Yetman, G. (2005). Natural disaster hotspots: A global risk analysis. *World Bank Disaster Risk Management Series*. https://doi.org/10.1080/01944360902967228

Dinku, T., Connor, S. J., Ceccato, P., & Ropelewski, C. F. (2008). Comparison of global gridded precipitation products over a mountainous region of Africa. *INTERNATIONAL JOURNAL OF CLIMATOLOGY Int. J. Climatol*, *28*, 1627–1638. https://doi.org/10.1002/joc.1669

Ebert, E. E., Janowiak, J. E., & Kidd, C. (2007). Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bulletin of the American Meteorological Society*, *88*(1), 47–64. https://doi.org/10.1175/BAMS-88-I-47

Eisenberg, M. C., Kujbida, G., Tuite, A. R., Fisman, D. N., & Tien, J. H. (2013). Examining rainfall and cholera dynamics in Haiti using statistical and dynamic modeling approaches. *Epidemics*, *5*(4), 197–207. https://doi.org/10.1016/j.epidem.2013.09.004

Farahmand, A., & Aghakouchak, A. (2013). A satellite-based global landslide model. *Natural Hazards and Earth System Science*, *13*(5), 1259–1267. https://doi.org/10.5194/nhess-13-1259-2013

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*. https://doi.org/10.5194/nhess-18-2161-2018

Gebregiorgis, A., & Hossain, F. (2014). Making Satellite Precipitation Data Work for the Developing World. *IEEE Geoscience and Remote Sensing Magazine*, *2*(2), 24–36. https://doi.org/10.1109/mgrs.2014.2317561

Gebregiorgis, A. S., & Hossain, F. (2013). Understanding the dependence of satellite rainfall uncertainty on topography and climate for hydrologic model simulation. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(1), 704–718. https://doi.org/10.1109/TGRS.2012.2196282

Gebregiorgis, A. S., Kirstetter, P.-E., Hong, Y. E., Carr, N. J., Gourley, J. J., Petersen, W., & Zheng, Y. (2017). Understanding Overland Multisensor Satellite Precipitation Error in TMPA-RT Products. *Journal of Hydrometeorology*, *18*(2), 285–306. https://doi.org/10.1175/jhm-d-15-0207.1

Gebregiorgis, A. S., Kirstetter, P. E., Hong, Y. E., Gourley, J. J., Huffman, G. J., Petersen, W. A., … Schwaller, M. R. (2018). To What Extent is the Day 1 GPM IMERG Satellite Precipitation Estimate Improved as Compared to TRMM TMPA-RT? *Journal of Geophysical Research: Atmospheres*, *123*(3), 1694–1707. https://doi.org/10.1002/2017JD027606

Gebremichael, M., Liao, G. Y., & Yan, J. (2011). Nonparametric error model for a high resolution satellite rainfall product. *Water Resources Research*, *47*(7), 7504. https://doi.org/10.1029/2010WR009667

Glade, T., Crozier, M., & Smith, P. (2000). Applying Probability Determination to Refine Landslide-triggering Rainfall Thresholds Using an Empirical "'Antecedent Daily Rainfall Model.'" *Pure Appl. Geophys*, *157*. Retrieved from https://link.springer.com/content/pdf/10.1007/s000240050017.pdf

Graziella, D., Ingeborg, K., Monica, S., Nils-Kristian, O., Ragnar, E., Erik, J., & Hervé, C. (2015). Landslide early warning system and web tools for real-time scenarios and for distribution of warning messages in Norway. *Engineering Geology for Society and Territory - Volume 2: Landslide Processes*, 625–629. https://doi.org/10.1007/978-3-319-09057-3_104

Hamill, T. M., Bates, G. T., Whitaker, J. S., Murray, D. R., Fiorino, M., Galarneau, T. J., … Lapenta, W. (2013). NOAA's Second-Generation Global Medium-Range Ensemble Reforecast Dataset. *Bulletin of the American Meteorological Society*, *94*(10), 1553–1565. https://doi.org/10.1175/bams-d-12-00014.1

Hong, Y., Adler, R., & Huffman, G. (2006). Evaluation of the potential of NASA multi-satellite precipitation analysis in global landslide hazard assessment. *Geophysical Research Letters*, *33*(22), L22402. https://doi.org/10.1029/2006GL028010

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

Hossain, F., & Anagnostou, E. N. (2004). Assessment of current passive-microwave- and infrared-based satellite rainfall remote sensing for flood prediction. *Journal of Geophysical Research: Atmospheres*, *109*(D07102), n/a-n/a. https://doi.org/10.1029/2005jd005831

Hossain, F., & Anagnostou, E. N. (2006). A two-dimensional satellite rainfall error model. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(6), 1511–1522. https://doi.org/10.1109/TGRS.2005.863866

Hossain, F., & Huffman, G. J. (2008). Investigating Error Metrics for Satellite Rainfall Data at Hydrologically Relevant Scales. *Journal of Hydrometeorology*, *9*(3), 563–575. https://doi.org/10.1175/2007jhm925.1

Huffman, G. J., Adler, R. F., Bolvin, D. T., & Nelkin, E. J. (2010). The TRMM Multi-Satellite Precipitation Analysis (TMPA). In *Satellite Rainfall Applications for Surface Hydrology* (pp. 3–22). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-90-481-2915-7_1

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., … Xie, P. (2015). NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). *Algorithm Theoretical Basis Doc., Version 4.5*, (November), 26. Retrieved from https://pps.gsfc.nasa.gov/Documents/IMERG_ATBD_V4.pdf

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., … Stocker, E. F. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, *8*(1), 38–55. https://doi.org/10.1175/JHM560.1

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology*, *5*(3), 487–503. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., & Kirschbaum, D. B. (2017). So, how much of the Earth's surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, *98*(1), 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1

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

Kirschbaum, D. B., Stanley, T., & Simmons, J. (2015). A dynamic landslide hazard assessment system for Central America and Hispaniola. *Hazards Earth Syst. Sci*, *15*, 2257–2272. https://doi.org/10.5194/nhess-15-2257-2015

Kirschbaum, D., & Stanley, T. (2018). Satellite-Based Assessment of Rainfall-Triggered Landslide Hazard for Situational Awareness. *Earth's Future*. https://doi.org/10.1002/2017EF000715

Kirschbaum, D., Stanley, T., & Zhou, Y. (2015). Spatial and temporal analysis of a global landslide catalog. *Geomorphology*, *249*, 4–15. https://doi.org/10.1016/j.geomorph.2015.03.016

Krøgli, I. K., Devoli, G., Colleuille, H., Boje, S., Sund, M., & Engen, I. K. (2018). The Norwegian forecasting and warning service for rainfall-and snowmelt-induced landslides. *Hazards Earth Syst. Sci*, *18*, 1427–1450. https://doi.org/10.5194/nhess-18-1427-2018

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., & Simpson, J. (1998). The Tropical Rainfall Measuring Mission (TRMM) Sensor Package. *Journal of Atmospheric and Oceanic Technology*, *15*. Retrieved from https://journals.ametsoc.org/doi/pdf/10.1175/1520-0426(1998)015%3C0809%3ATTRMMT%3E2.0.CO%3B2

Leutbecher, M., Lock, S.-J., Ollinaho, P., Lang, S. T. K., Balsamo, G., Bechtold, P., … Weisheimer, A. (2017). Stochastic representations of model uncertainties at ECMWF: state of the art and future vision. *Quarterly Journal of the Royal Meteorological Society Q. J. R. Meteorol. Soc*, *143*, 2315–2339. https://doi.org/10.1002/qj.3094

Levizzani, V., Amorati, R., & Meneguzzo, F. (2002). *A review of satellite-based rainfall estimation methods*. *Commission Project MUSIC Report*. Retrieved from www.isao.bo.cnr.it/~meteosat

Li, C., Ma, T., Zhu, X., & Li, W. (2011). The power-law relationship between landslide occurrence and rainfall level. *Geomorphology*. https://doi.org/10.1016/j.geomorph.2011.03.018

Maggioni, V., Sapiano, M. R. P., Adler, R. F., Tian, Y., & Huffman, G. J. (2014). An Error Model for Uncertainty Quantification in High-Time-Resolution Precipitation Products. *Journal of Hydrometeorology*, *15*(3), 1274–1292. https://doi.org/10.1175/jhm-d-13-0112.1

Maggioni, V., Vergara, H. J., Anagnostou, E. N., Gourley, J. J., Hong, Y., & Stampoulis, D. (2013). Investigating the Applicability of Error Correction Ensembles of Satellite Rainfall Products in River Flow Simulations. *Journal of Hydrometeorology*, *14*(4), 1194–1211. https://doi.org/10.1175/jhm-d-12-074.1

Mahoney, K. M., Moore, B. J., Hamill, T. M., Sukovich, E. M., & Cifelli, R. (2014). Climatology and Environmental Characteristics of Extreme Precipitation Events in the Southeastern United States. *Monthly Weather Review*, *143*(3), 718–741. https://doi.org/10.1175/mwr-d-14-00065.1

Mathew, J., Babu, D. G., Kundu, S., Kumar, K. V., & Pant, C. C. (2014). 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).

https://doi.org/10.1007/s10346-013-0408-2

Midekisa, A., Senay, G., Henebry, G. M., Semuniguse, P., & Wimberly, M. C. (2012). Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia. *Malaria Journal*, *11*. https://doi.org/10.1186/1475-2875-11-165

Nadim, F., Kjekstad, O., Peduzzi, P., Herold, C., & Jaedicke, C. (2006). Global landslide and avalanche hotspots. *Landslides*, *3*(2), 159–173. https://doi.org/10.1007/s10346-006-0036-1

Osanai, N., Takeshi, I., Kazumasa, S. I., Shinichi Kojima, K. I., & Noro, T. (2010). Japanese early-warning for debris flows and slope failures using rainfall indices with Radial Basis Function Network, *7*, 325–338. https://doi.org/10.1007/s10346-010-0229-5

Petley, D. (2012). Global patterns of loss of life from landslides. *Geology*, *40*(10), 927–930. https://doi.org/10.1130/G33217.1

Petley, D. N., Dunning, S. A., & Rosser, N. J. (2005). The analysis of global landslide risk through the creation of a database of worldwide landslide fatalities. In *Landslide Risk Management*.

Piciullo, L., Gariano, S. L., Melillo, M., Brunetti, M. T., Peruccacci, S., Guzzetti, F., & Calvello, M. (2017). Definition and performance of a threshold-based regional early warning model for rainfall-induced landslides. *Landslides*, *14*(3). https://doi.org/10.1007/s10346-016-0750-2

Rodell, B. Y. M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., … Toll, D. (2004). THE GLOBAL LAND DATA ASSIMILATION SYSTEM. *Bull. Amer. Meteor. Soc.*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Rossi, M., Peruccacci, S., Brunetti, M. T., Marchesini, I., Luciani, S., Ardizzone, F., … M. Rossi S. Peruccacci, M. T. B. I. M. S. L. F. A. V. B. C. B. M. C. F. F. A. C. M. P. R. P. S. M. S. D. B. S. L. G. M. P. G. V. A. V. L. A. L. B. A. M. D. G. I. F. L. M. P. (2012). SANF : National warning system for rainfall-induced landslides in Italy. *Landslides and Engineered Slopes: Protecting Society through Improved Understanding – Eberhardt et Al. (Eds)*, (June 2014), 1895–1899. https://doi.org/10.13140/2.1.4857.9527

Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., … Xie, P. (2006). The NCEP Climate Forecast System. *Journal of Climate*. https://doi.org/10.1175/JCLI3812.1

Sapiano, M. R. P., & Arkin, P. A. (2008). An Intercomparison and Validation of High-Resolution Satellite Precipitation Estimates with 3-Hourly Gauge Data. *Journal of Hydrometeorology*, *10*(1), 149–166. https://doi.org/10.1175/2008jhm1052.1

Sarachi, S., Hsu, K., & Sorooshian, S. (2015). A Statistical Model for the Uncertainty Analysis of Satellite Precipitation Products. *Journal of Hydrometeorology*, *16*(5), 2101–2117. https://doi.org/10.1175/jhm-d-15-0028.1

Sassa, K., Canuti, P., & Yin, Y. (2014). *Landslide science for a Safer Geoenvironment: Volume 2: Methods of landslide studies*. *Landslide Science for a Safer Geoenvironment: Volume 2: Methods of Landslide Studies*. https://doi.org/10.1007/978-3-319-05050-8

Scheuerer, M., & Hamill, T. M. (2015). Statistical Postprocessing of Ensemble Precipitation Forecasts by Fitting Censored, Shifted Gamma Distributions*. *Monthly Weather Review*. https://doi.org/10.1175/MWR-D-15-0061.1

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B., & Braithwaite, D. (2000). Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, *81*(9), 2035–2046. https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2

Stanley, T., & Kirschbaum, D. B. (2017). A heuristic approach to global landslide susceptibility mapping. *Natural Hazards*. https://doi.org/10.1007/s11069-017-2757-y

Tang, L., & Hossain, F. (2009). Transfer of satellite rainfall error from gaged to ungaged locations: How realistic will it be for the Global Precipitation Mission? https://doi.org/10.1029/2009GL037965

Tang, L., & Hossain, F. (2012). Investigating the similarity of satellite rainfall error metrics as a function of Köppen climate classification. *Atmospheric Research*. https://doi.org/10.1016/j.atmosres.2011.10.006

Tian, Y., Huffman, G. J., Adler, R. F., Tang, L., Sapiano, M., Maggioni, V., & Wu, H. (2013). Modeling errors in daily precipitation measurements: Additive or multiplicative? *Geophysical Research Letters*, *40*(10), 2060–2065. https://doi.org/10.1002/grl.50320

Tian, Y., & Peters-Lidard, C. D. (2007). Systematic anomalies over inland water bodies in satellite-based precipitation estimates. *Geophysical Research Letters*, *34*(14), 1–5. https://doi.org/10.1029/2007GL030787

Tian, Y., & Peters-Lidard, C. D. (2010). A global map of uncertainties in satellite-based precipitation measurements. *Geophysical Research Letters*, *37*(24), 1–6. https://doi.org/10.1029/2010GL046008

Tian, Y., Peters-Lidard, C. D., Choudhury, B. J., & Garcia, M. (2007). Multitemporal Analysis of TRMM-Based Satellite Precipitation Products for Land Data Assimilation Applications. *Journal of Hydrometeorology*, *8*(6), 1165–1183. https://doi.org/10.1175/2007jhm859.1

Tian, Y., Peters-Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., … Zeng, J. (2009). Component analysis of errors in Satellite-based precipitation estimates. *Journal of Geophysical Research Atmospheres*, *114*(24), 1–15. https://doi.org/10.1029/2009JD011949

Villarini, G., Krajewski, W. F., & Smith, J. A. (2009). New paradigm for statistical validation of satellite precipitation estimates: Application to a large sample of the TMPA 0.25° 3-hourly estimates over Oklahoma. *J. Geophys. Res*, *114*, 12106. https://doi.org/10.1029/2008JD011475

Wright, D. B., Kirschbaum, D. B., & Yatheendradas, S. (2017). Satellite Precipitation Characterization, Error Modeling, and Error Correction Using Censored Shifted Gamma Distributions. *Journal of Hydrometeorology*, *18*(10), 2801–2815. https://doi.org/10.1175/JHM-D-17-0060.1

Wu, H., Adler, R. F., Hong, Y., Tian, Y., & Policelli, F. (2012). Evaluation of Global Flood Detection Using Satellite-Based Rainfall and a Hydrologic Model. *Journal of Hydrometeorology*, *13*(4), 1268–1284. https://doi.org/10.1175/jhm-d-11-087.1

Yan, J., & Gebremichael, M. (2009). Estimating actual rainfall from satellite rainfall products. *Atmospheric Research*, *92*(4), 481–488. https://doi.org/10.1016/j.atmosres.2009.02.004