Comparison of a Global Landslide Event Inventory to a Satellite-based Landslide Algorithm

Final Paper for GSSP 2007

Dalia Bach

Mentors: Bob Adler, Yang Hong

Abstract

A global, satellite-based landslide algorithm has been developed using surface information and multi-satellite rainfall data. The technique integrates surface parameters such as slope, land cover, soils, and elevation with satellite precipitation data to obtain an estimate of areas susceptible to landslides in near-real time. This research compares the predictions from the global landslide algorithm run retrospectively for individual years with global landslide inventories to assess both the relative skill of the technique and the value of currently available landslide information on a global scale.

Results indicate that the general pattern of landslide activity (number of total events, geographic distribution, etc.) can be reproduced, but finer-scale distributions and individual events are difficult to match between the forecast and the event inventory. Preliminary results indicate that three-fourths of the landslide events correspond to locations with high susceptibility values based on the satellite-based Landslide Susceptibility Index map. Probability of Detection and False Alarm Rate statistics are presented for the global database, with results varying based on the size of area used for event validation. Results are also shown to be a function of population density with more densely populated areas having higher scores, as expected.

This global algorithm represents the first phase in identifying landslide hazards at this scale. With adjustment, the algorithm shows great promise in approaching landslide hazard assessment globally and providing information for the research community to address landslide issues in a broader context. The evaluation also provides insight into the necessary considerations and potential adaptations to the algorithm for improved landslide hazard forecasting on a global scale and the need for international efforts for developing accurate landslide inventories.

I. Introduction/Background

Landslide hazards pose a significant threat to populations throughout the world due to their rapid onset, multiple triggers, and impact on many different land cover types. These events are also one of the most challenging hazards to evaluate based on their small spatial and temporal scales. Identification and mapping of these hazards varies by region and represents heterogeneous coverage in even the most well-observed areas. The processes governing landslide initiation can be viewed as a function of several components, namely a background forcing including regional climatology and surface characteristics (slope and elevation), and a triggering mechanism (Dai et al. 2002). Other contributing variables such as anthropogenic surface modification can also significantly influence landslide susceptibility. Rainfall, including antecedent conditions and distribution, is a key climatological feature that affects landslide susceptibility. Intense rainfall over short durations can then serve as a triggering mechanism to initiate an event (Sidle and Ochiai 2006).

Studies generally characterize the potential for a rainfall-triggered landslide event according to a rainfall intensity-duration threshold from past landslide and rainfall events. These thresholds have been produced on the global (Caine 1980), regional (Larsen and Simon 1993; Chien-Yuan et al. 2005), and local (Wieczorek 1987; Calcaterra et al. 2000) scales and rely on landslide mapping and rainfall gauge information (Figure 1). The development of a threshold relationship requires information on the intensity (mm/hour) and duration (hours) of the rainstorm event that corresponds to the location of the landslide and time the landslide(s) occurred. This information is often challenging if not impossible to retrieve due to the relatively sparse network of rain gauges in most parts of the world and the inability to determine exactly when landslides were initiated. Therefore, establishing a relationship between rainfall events and landslide hazards is limited by the spatial and temporal scale of the events and the availability of in situ rainfall data.

Past landslide hazard research has relied on spatially and temporally heterogeneous rainfall gauge data at small spatial scales. Similarly, the majority of landslide studies focus on smaller scale evaluation at the slope or watershed level. The

high resolution nature of these studies limits their applicability to larger areas. Remote sensing products both for rainfall and surface features can fill the void to offer continuous data up to the global scale. An algorithm developed by Hong et al. (2006; Hong et al. 2007) utilizes global remote sensing data to develop a landslide susceptibility and forecasting algorithm which draws on surface parameters and a merged satellite rainfall product to identify potential landslide events on the global scale. This research compares the predictions from the global landslide algorithm run retrospectively for a 2003 global landslide inventory to assess both the relative skill of the technique and the value of currently available landslide information on a global scale. The goal of this work is to better understand the interaction between different surface parameters and rainfall triggers at large spatial scales.

This paper first provides a short discussion of the research issues associated with using satellite products for landslide hazard research, citing specific case studies. Section 3 presents a more detailed explanation of the model being evaluated as well as the methodology used to evaluate the accuracy of the model outputs. Sections 4 and 5 outline the preliminary results of this evaluation study and provide a discussion of the uncertainties associated with the input parameters. Lastly, section 6 describes the conclusions of this work as well as provides insight into the next steps for this research.

II. Remotely Sensed Precipitation Data: Opportunities and Challenges

For studies at the slope or watershed level, gauge or radar data remain the only reliable sources available. However, for large scale studies and for evaluating areas without an in situ network, satellite data can provide crucial information for general landslide hazard analysis and susceptibility mapping.

Rainfall gauges provide accurate measurements of precipitation in a small area, but they can be highly affected by wind and gauge placement, causing significant underestimations of rainfall. Radar measurements such as operational surface radar provide near-complete coverage over some developed countries at high spatial and temporal resolution, although results in mountainous areas (where landslides occur) are sometimes questionable due to beam shielding errors.

Rainfall estimates from remote sensors on orbiting satellites can provide frequent and consistent coverage over large areas. Low-orbiting satellites such as NASA's Tropical Rainfall Measuring Mission (TRMM) image the rainfall structure using active and passive microwave sensors in the tropics and into mid-latitudes. Other passive microwave instruments on polar-orbiting satellites also provide precipitation information over land. While TRMM and other individual orbiting platforms can provide relatively accurate instantaneous rainfall rates, the return period of individual satellites limits the ability to estimate storm intensity and duration on small temporal scales.

Infrared cloud-top brightness temperature sensing from geostationary infrared data provide the highest spatial and temporal coverage over the equatorial and mid-latitudes at 4 km, 15 minute resolution, estimating rain-rates from cloud top temperature and other information. However, the generally poor correlation between the coldest cloud areas and rainfall intensity suggests that these rain rates may be difficult to use independently but can enhance other precipitation products.

To improve upon single-instrument precipitation estimates, many studies have used more sophisticated methodologies or merged satellite information to produce enhanced products. The TRMM Multi-Satellite Precipitation Analysis (TMPA) (Huffman et al. 2007) integrates data from multiple microwave satellites to provide a 3-hourly 0.25º product. More than nine years of the TMPA product is now available for research with a real-time version available for use in hydrological applications. PERSIANN-CCS algorithm (Hong et al. 2004) is an example of a product that uses geosynchronous IR data, which provides quasi-global rainfall estimates at 4 km hourly resolution by incorporating information on cloud temperature, texture, and geometry.

Table 1 summarizes the range of rainfall products and describes their coverage, resolution, and limitations. Figure 2 illustrates the resolution differences between point and gridded satellite coverage over the island of Puerto Rico. These satellite products are generally applied on longer spatial and temporal scales to get a sense of the climatology or anomalies of a region. However, for landslide research purposes it is critical to have a representation of the rainfall structure at high resolution. Two case studies provide insight into the ability of different rainfall products to resolve individual rainfall events that were instrumental in triggering landslides.

In 2004, the eye of Tropical Storm Jeanne traveled over Puerto Rico, leaving extensive flooding and landslides in its wake. A gauging station in the south eastern portion of the island, NEXRAD radar, and multiple satellite instruments were able to capture rainfall accumulation over this period. Figure 3 illustrates the 3-hourly cumulative rainfall signature of this event as recorded by many rainfall products. The majority of the estimates accurately track the rapid rainfall accumulation as the storm made landfall on September 16th.

In a second example, a major storm over Southern California's La Conchita area caused a massive mudslide that killed 10 people and injured 14 on January 10th, 2005. Figure 4 plots the 3-hourly cumulative rainfall over this period. These examples are not meant to serve as an intercomparison among products but rather to illustrate the variability of rainfall products at high temporal and spatial scales. The statistics indicate that the remote sensing products correlate well with each other but have a relatively poor correlation to gauge measurements. The RMSE and Bias also supports this finding. This is likely due to the resolution differences between point and gridded data as well as the natural variability existing in rainfall structure over different areas.

Another common issue that limits the accuracy and availability of the majority of rainfall products is orography, which has sizeable consequences for landslide hazard assessment. Precipitation generally increases with elevation as moist air is forced upwards, causing significant rainfall on the windward side of a mountain. For in situ and remote sensing instruments mountains can bias the areas in which gauges are installed, cause large signal interference of radar beams, and produce different atmospheric patterns which distort the typical cloud-top temperature-rain rate relationship. Studies have attempted to correct for orographic precipitation effects by introducing additional surface data such as terrain, elevation, and surface wind direction. While these studies assist rainfall estimation in mountainous settings, they do not provide the spatial and temporal resolution required for landslide hazards assessment.

Landslide hazard assessment is performed on an array of scales. For studies at the slope or watershed level, gauge or radar data remain the only reliable sources available. However, for large scale studies and for evaluating areas without an in situ network,

satellite data can provide crucial information for general landslide hazard analysis and susceptibility mapping with its continuous spatial and temporal coverage.

III. Landslide Susceptibility and Prediction Model Evaluation

A global, satellite-based landslide algorithm has been developed using surface information and multi-satellite rainfall data Hong et al. (2006, 2007). The technique integrates surface parameters such as slope, land cover, soils, and elevation with satellite precipitation data to obtain an estimate of areas susceptible to landslides in near-real time. The global landslide hazard assessment model is composed of three parts: a susceptibility map, rainfall data, and combined product (Figure 5). The surface data including slope, land cover data, soil information, and geology were weighted in order of their relative importance in affecting landslide susceptibility and their products were combined into a global susceptibility map with values ranging from 0 – least susceptible to 5 – most susceptible. Rainfall data from the TMPA product was then evaluated cumulatively at every pixel globally and flagged when the average rain-rate exceeded the algorithm's intensity-duration threshold. Lastly, these two products were combined so that when the rainfall rate at a pixel exceeds the threshold value and the pixel has a susceptibility of class of 4 or 5, the pixel is highlighted as having a predicted landslide event. The result of algorithm is a global indication of where landslide events are predicted at a 0.25º scale in near real-time.

This model was then evaluated for its skill in predicting landslides by using a landslide inventory and comparing the model predications with actual events. In order to perform the model validation, a landslide inventory was compiled for 2003 landslide events globally from three data sources: the International Landslide Centre and the University of Durham, International Consortium on Landslides, and EM-DAT International Disaster Database (CRED) (EM-DAT 2006; ICL 2007; ILC 2007). The events reported in this database come from news or other media reports and only account for landslide events which have caused fatalities. This inventory is not representative of the total population of landslide events for the 2003 year but rather represents a cross-section of events that occurred near large population centers, in areas with efficient and

widespread reporting measures, or in areas which landslide researchers have focused their attention. The limited availability and unequal representation of landslide events is a major source of error in the validation process and research is ongoing.

Once the landslide inventory was compiled and locations were assigned to each event, the 2003 inventory was used to evaluate the susceptibility model. As introduced above, the susceptibility map represents a combination of weighted surface parameters to denote a landslide susceptibility level. Figure 6 illustrates the susceptibility map developed by Hong et al. (2006) along with the 228 events in the 2003 landslide inventory. When the landslide locations were matched to a corresponding susceptibility pixel, the distribution showed that 73% of the landslides fell within susceptibility Category 4 or 5 (Figure 7). The events falling within Category 0 are primarily due to landslides that occurred on small islands where there was no available surface data. The distribution of the landslide susceptibility suggests that on a global scale, the susceptibility map can largely resolve areas which have the potential to be susceptible to landslide events.

In order to test the second component of the algorithm involving the precipitation trigger, the global rainfall intensity-duration threshold curve was used to determine when the cumulative rainfall rate for each landslide location exceeded the curve value. Each day's cumulative rainfall was considered for the 10 days prior to the recorded event at the exact pixel as well as the maximum and mean of a 3 x 3 pixel window. The number of times the threshold was exceeded was recorded for each of the three scenarios. In order to determine if the algorithm would have detected the landslide event, the events that had rainfall values exceeding the threshold curve were compared to the susceptibility categories for the same location. If the susceptibility category was a value of 4 or 5, the event was considered a successfully detected event. The statistics were summarized as a Probability of Detection (POD) for the 2003 landslide inventory.

The algorithm was then evaluated from the opposite standpoint by using the 2003 TMPA rainfall database as the input into the algorithm to test how many and where landslides were predicted compared to the actual 2003 inventory, termed the False Alarm Rate (FAR). Due to the large volume of detections, we considered only events that had rainfall accumulations which were flagged on the seventh day before the event. The

resulting predictions were then merged temporally and spatially to cluster events which had been listed on consecutive days and represented adjoining pixels. 338 separate events were identified as algorithm predictions for landslide events in 2003. Due to uncertainties in both the landslide inventory and algorithm predictions, the predicted events database was considered at range of radii to determine how many landslide events fell within specific distances of the predicted events on time scale of ± 1 day and ± 15 days from the reported event. The results and discussion are presented below.

IV. Results

The POD and FAR statistics for this evaluation represent preliminary and likely incomplete results. For the exact pixels, the POD was 14.5%, the 3x3 mean had a POD of 11.9%, and the 3x3 maximum had a POD of 46%. The large POD value for the maximum pixel is likely misleading because it represents the maximum rainfall value of all the pixels for each day, meaning that the pixel recorded could change over the course of 10 days but would still be recorded as a cumulative measurement.

For the FAR, the areas considered at different buffers indicate that predictably, as the buffer size increases, more landslide events fall within the predicted range. Results indicate that 1 landslide event fell within the 10 km at a time resolution of ± 15 days from the reported event. When the radius increased to 100 km, 9.2% of the events matched those of the algorithm predictions by ± 1 day. Lastly, at a buffer of 200 km 69 events were detected for both the 1 and 15-day windows, totaling 20.5% of the algorithm prediction dataset. Because this false alarm calculation only considers the 7 day accumulation before a landslide event, it limits the number of landslides reported and thus likely underestimates the ability of the algorithm to accurately predict events. The same model runs were calculated for the 1 and 3-day accumulations but due to the size of the databases we were unable to process these results.

V. Discussion

The statistical results from the algorithm evaluation tests represent preliminary calculations and are expected to improve given more reliable validation data and adjustments to the algorithm inputs. The initial POD and FAR calculations likely underestimate the algorithm's potential for landslide detection at the global scale. While rough, this evaluation brings about some interesting points and new research questions. Figure 8 shows a global comparison of the 2003 landslide events and the predictions made by the algorithm for the same year. While it is clear that these points rarely overlap, the distribution of predictions indicate that there seems to be an over-prediction of events along the western coast of India while the algorithm under-predicts events in the mountainous regions of northern India and Tibet and central and eastern China. This could be due to the difficulty of the TMPA product to accurately record rainfall over orographically complex regions.

A second reason for the discrepancy between reported and predicted events may stem from population density and reporting biases. Figure 9 compares mean population density within a 50 km radius for the prediction and event databases. The histogram indicates that the 2003 landslide events generally occur in areas with larger population densities compared to that of the predicted events. This suggests that either the algorithm is under predicting events in more populated areas, or the 2003 landslide events are not capturing landslides which occurred in areas of low population density. It is more likely that both issues are impacting the evaluation results.

At present, there are a range of uncertainties associated with each input parameter which are difficult to quantify and when introduced into the system result in multiple sources of error. The 2003 landslide inventory was compiled from news reports, government papers, articles, and other sources and therefore only represents events that occurred in proximity to a population center. Due to biases in reporting by country and region as well as the ability to obtain such reports, it is extremely difficult if not impossible to quantitatively assess the completeness of the inventory. We are presently working to compile a landslide database for 2007, beginning in late June. Increasing the

amount of global inventory data available for analysis will help to provide insight on the global distribution of reported landslides as well as how these inventories may be biased.

VI. Conclusions and Future Work

The global landslide hazard assessment and prediction algorithm is in its first phase of implementation and evaluation. As a result, the algorithm is associated with an array of uncertainty stemming from the accuracy of the TMPA rainfall data at high spatial and temporal resolutions, the weighting of the surface parameters that make up the global susceptibility map, and the assumption of a universal rainfall intensity-duration threshold relationship. These input parameters can most likely be improved when considered and evaluated at the regional scale.

There has been a large push in precipitation research to compare and evaluate different rainfall products on a variety of scales. Landslide hazard research requires rainfall information on the highest spatial and temporal resolutions possible, which also represent the most uncertain measurements at the point source and gridded satellite products. While research must be continued to more closely understand the relationship between products at varying length scales, the merged satellite products can provide relatively accurate rainfall information for landslide hazard assessment when being considered on larger spatial scales.

With a better understanding of the accuracy in satellite rainfall data, the rainfall intensity-duration threshold curves may be able to be improved at the regional and local scales. Past calculations primarily used gauge information and nearby landslides to derive threshold curves; however, more continuous rainfall information and more complete landslide inventories can improve the threshold curves at the regional scale. When considered on smaller spatial scales, these curves may also be adjusted for regional climatology and other forcings which could influence the accuracy of the results.

Results for the global susceptibility map indicated that the current weighting scheme of the surface parameters was effectively able to categorize 73% of the 2003 landslide events. When considered on the regional scale, these weights can be further attuned to the specific surface processes influencing landslide susceptibility. This will

likely increase the POD of the susceptibility values by better categorizing surface cover with respect to landslide hazard as well as including some of the low lying or island regions which are currently absent in the susceptibility calculations.

Lastly, because the 2003 landslide database is compiled from reported events and any future inventory we may assemble will be similarly biased, it may be useful to incorporate population density into the algorithm's susceptibility weighting to better predict landslides that have the potential to occur in populated areas. Figure 10 illustrates population density in 2005 with events in the 2007 inventory. By focusing more narrowly on populated regions, it may provide more accurate statistics for this algorithm. Due to the large influence of road cuts on landslide potential, this parameter may also be valuable to incorporate into the algorithm.

The concept of using satellite rainfall data to assess and predict the occurrence of landslides on the global scale is pioneering in the field of landslide hazard assessment. As such, this global landslide hazard assessment and prediction algorithm represents the first phase in tackling these issues. Given the uncertainties associated with each of the input parameters, it is difficult to determine as this point the absolute skill of this algorithm in detecting events on the global scale. However, with fine tuning and regional evaluation, this algorithm show great promise in approaching landslide hazard assessment globally and opening up the research community to addressing these issues in a broader context. The evaluation also provides insight into the necessary considerations and potential adaptations to the algorithm for improved landslide hazard forecasting on a global scale and the need for international efforts for developing accurate landslide inventories.

References:

Caine, N. (1980). "The rainfall intensity–duration control of shallow landslides and debris flows." Geografiska Annaler **62**(1/2): 23 - 27.

Calcaterra, D., M. Parise, B. Palma and L. Pelella (2000). The influence of meteoric events in triggering shallow landslides in pyroclastic deposits of Campania, Italy. Proceedings of the 8th International Symposium on Landslides. E. Bromhead, N. Dixon and M. L. Ibsen, eds. Cardiff, Wales, A.A. Balkema. **1:** 209-214.

Chien-Yuan, C., C. Tien-Chien, Y. Fan-Chieh, Y. Wen-Hui and T. Chun-Chieh (2005). "Rainfall duration and debris-flow initiated studies for real-time monitoring." Environmental Geology **47**(715 - 724).

Dai, E. C., C. F. Lee and Y. Y. Nagi (2002). "Landslide risk assessment and management: an overview." Engineering Geology **64**: 65 - 87.

EM-DAT (2006). The OFDA/CRED International Disaster Database Brussels – Belgium, Université Catholique de Louvain, www.em-dat.net

Hong, Y., R. Adler and G. Huffman (2006). "Evaluation of the Potential of NASA Multi-satellite Precipitation Analysis in Global Landslide Hazard Assessment." Geophysical Research Letters **33**(L22402): doi:10.1029/2006GL028010.

Hong, Y., R. Adler and G. Huffman (2007). "An Experimental Global Monitoring System for Rainfall-triggered Landslides using Satellite Remote Sensing Information." IEEE Transactions on Geosciences and Remote Sensing **45**(6): 1671-1680.

Hong, Y., K.-L. Hsu, S. Sorooshian and X. Gao (2004). "Precipitation estimation from remotely sensed imagery using Artificial Neural Network-Cloud Classification System." Journal of Applied Meteorology **43**(12): 1834–1853.

Huffman, G., R. Adler, D. Bolvin, G. Gu, E. Nelkin, K. Bowman, Y. Hong, E. Stocker and D. Wolff (2007). "The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales." Journal of Hydrometeorology **8**(1): 38-55.

ICL (2007). "International Consortium on Landslides: Recent Landslides." http://icl.dpri.kyoto-u.ac.jp/.

ILC. (2007). "The International Landslides Centre landslide fatality database." http://www.landslidecentre.org/database.htm Retrieved June 20th, 2007.

Larsen, M. C. and A. Simon (1993). "A Rainfall Intensity-Duration Threshold for Landslides in a Humid-Tropical Environment, Puerto Rico." Geografiska Annaler **75**(1-2): 13 - 23.

Sidle, R. C. and H. Ochiai (2006). Landslide Processes, Prediction, and Land use. Washington, DC, American Geophysical Union.

Wieczorek, G. F. (1987). Effect of rainfall intensity and duration on debris flows in central Santa Cruz Mountains. Debris flow/avalanches: process, recognition, and mitigation. J. E. Costa and G. F. Wieczorek, eds. Geological Society of America, Reviews in Engineering Geology. **7:** 93–104.

Figures and Tables for Final GSSP Paper

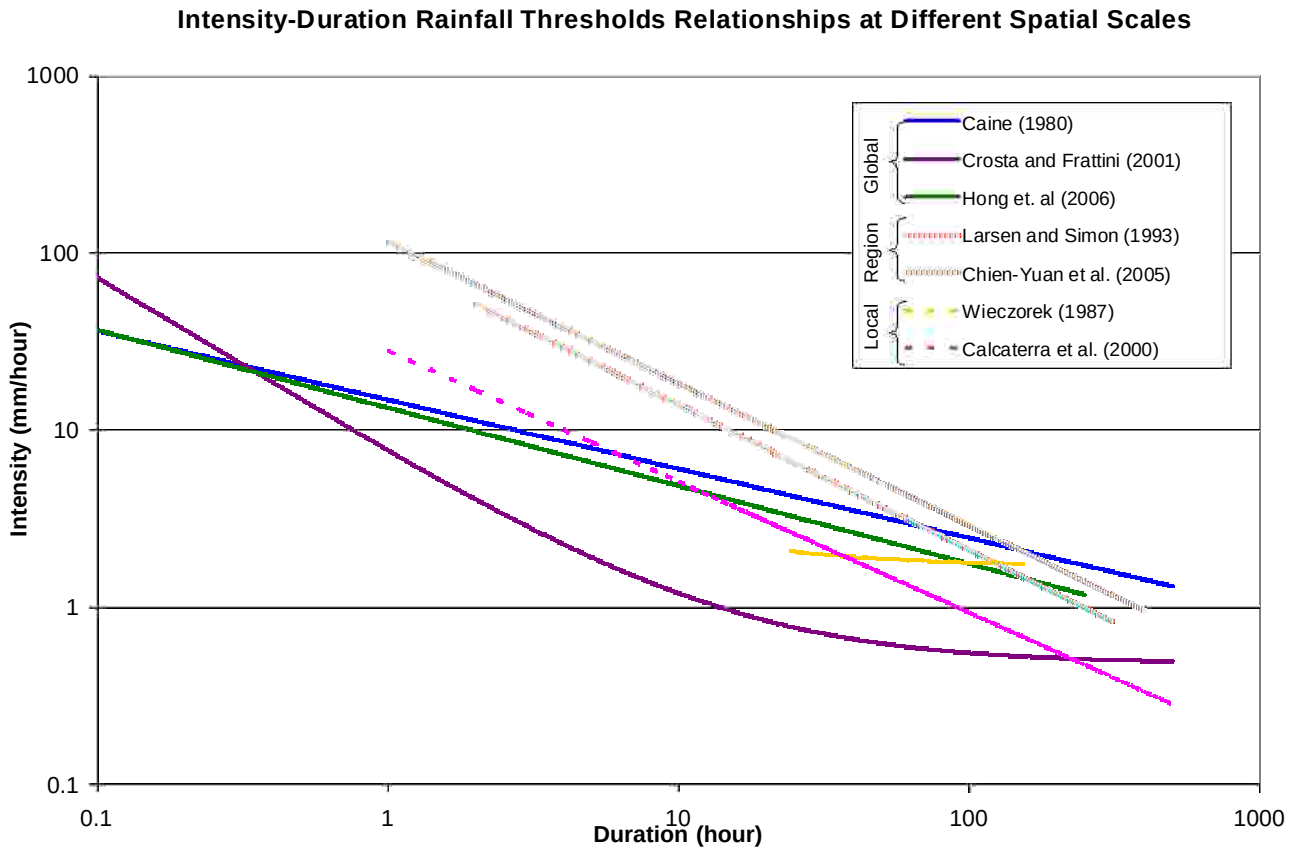

Figure 1: Rainfall Intensity-Duration threshold at the global regional and local scales as derived from seven different studies.

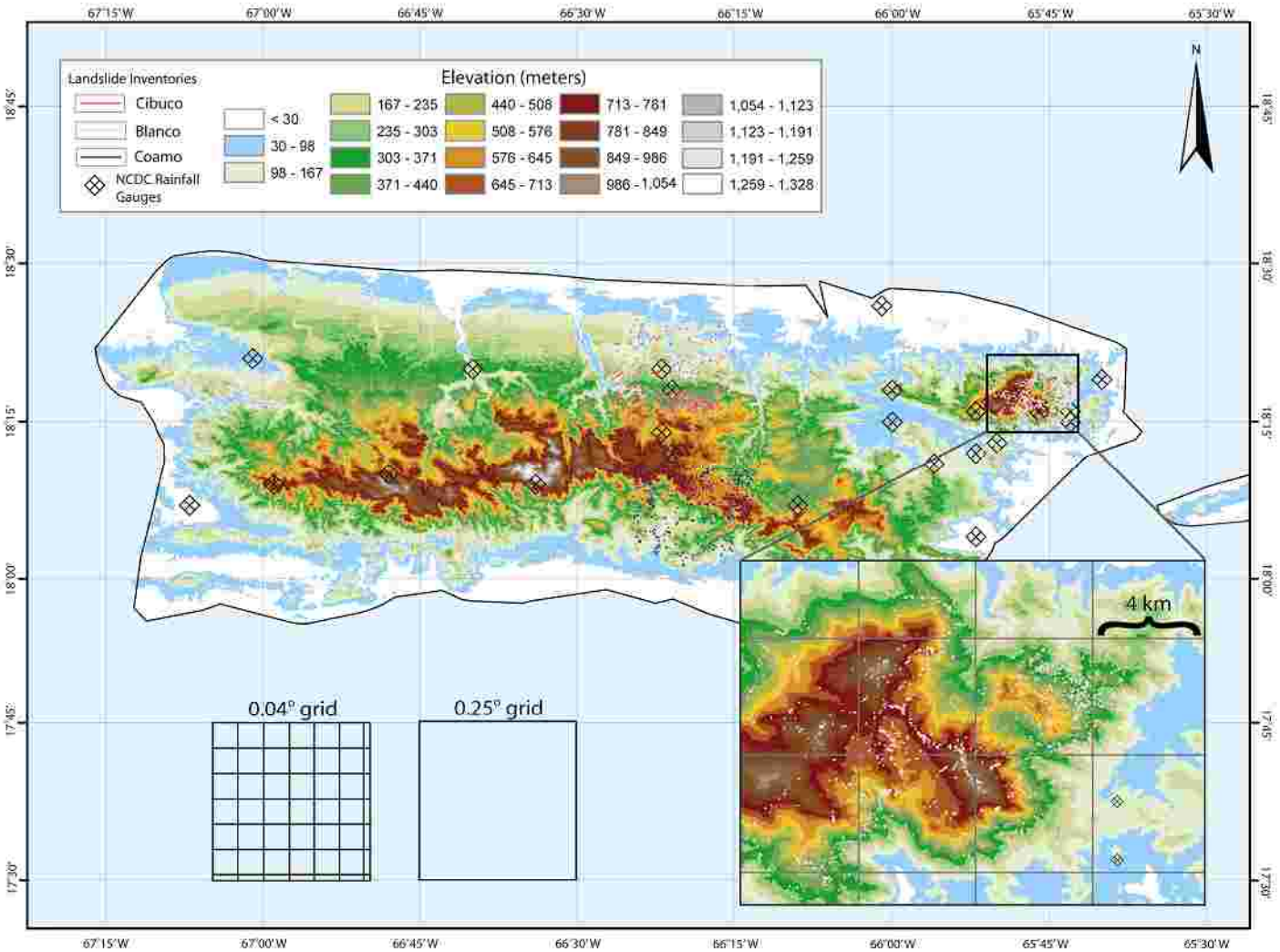

Figure 2: Map of Puerto Rico showing a 30 m DEM and location of rainfall gauging stations. The grid boxes indicate the resolution of the current precipitation products. Three landslide inventories done by Larsen and Torres-Sanchez (1998) are included for comparison

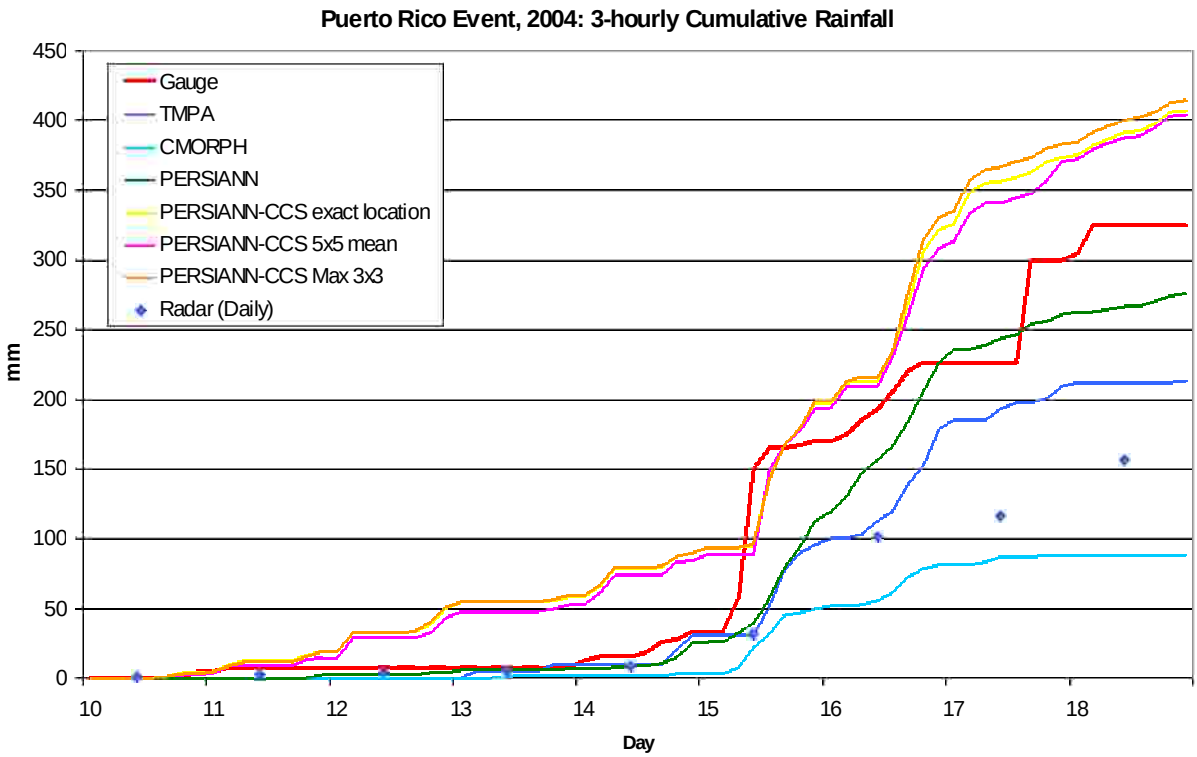

Figure 3: Comparison of gauge, radar, and satellite products for September 10th – 18th, 2004 in south-eastern Puerto Rico during Tropical Cyclone Jeanne.

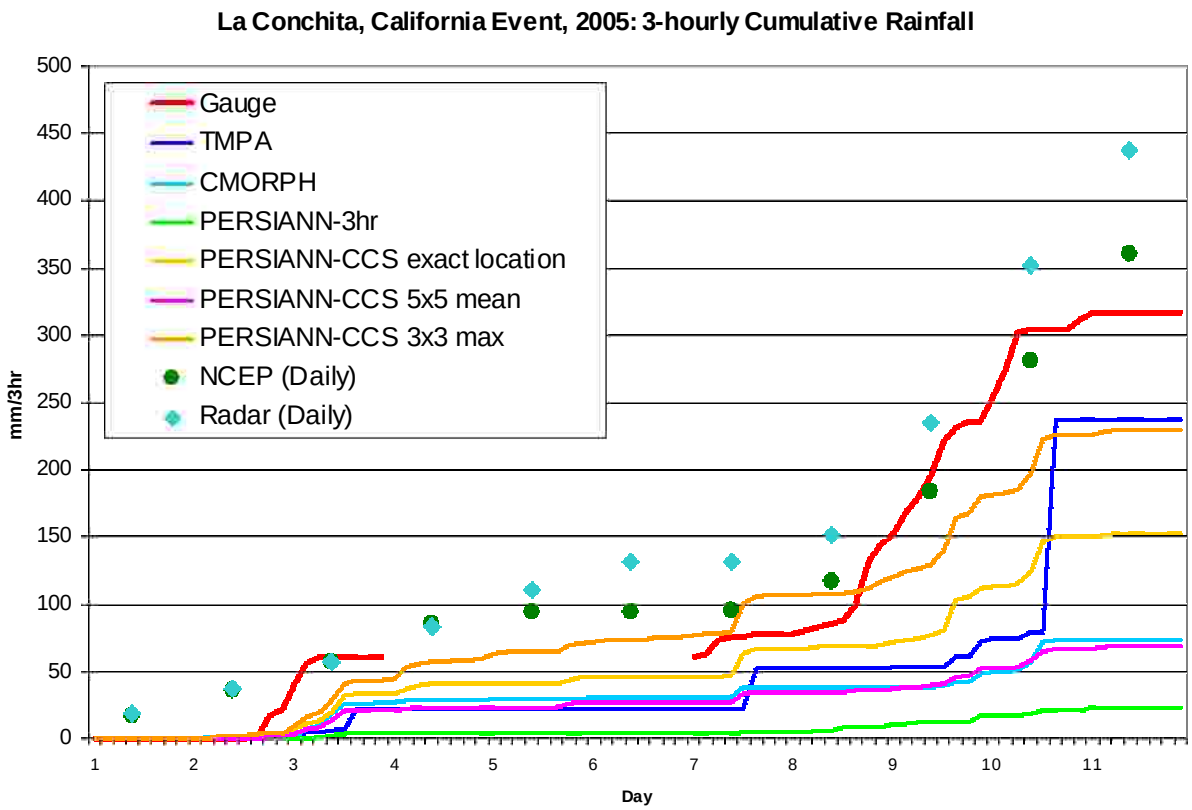

Figure 4: Comparison of gauge, radar, and satellite products for the January 1st-11th, 2005 event in La Conchita, California.

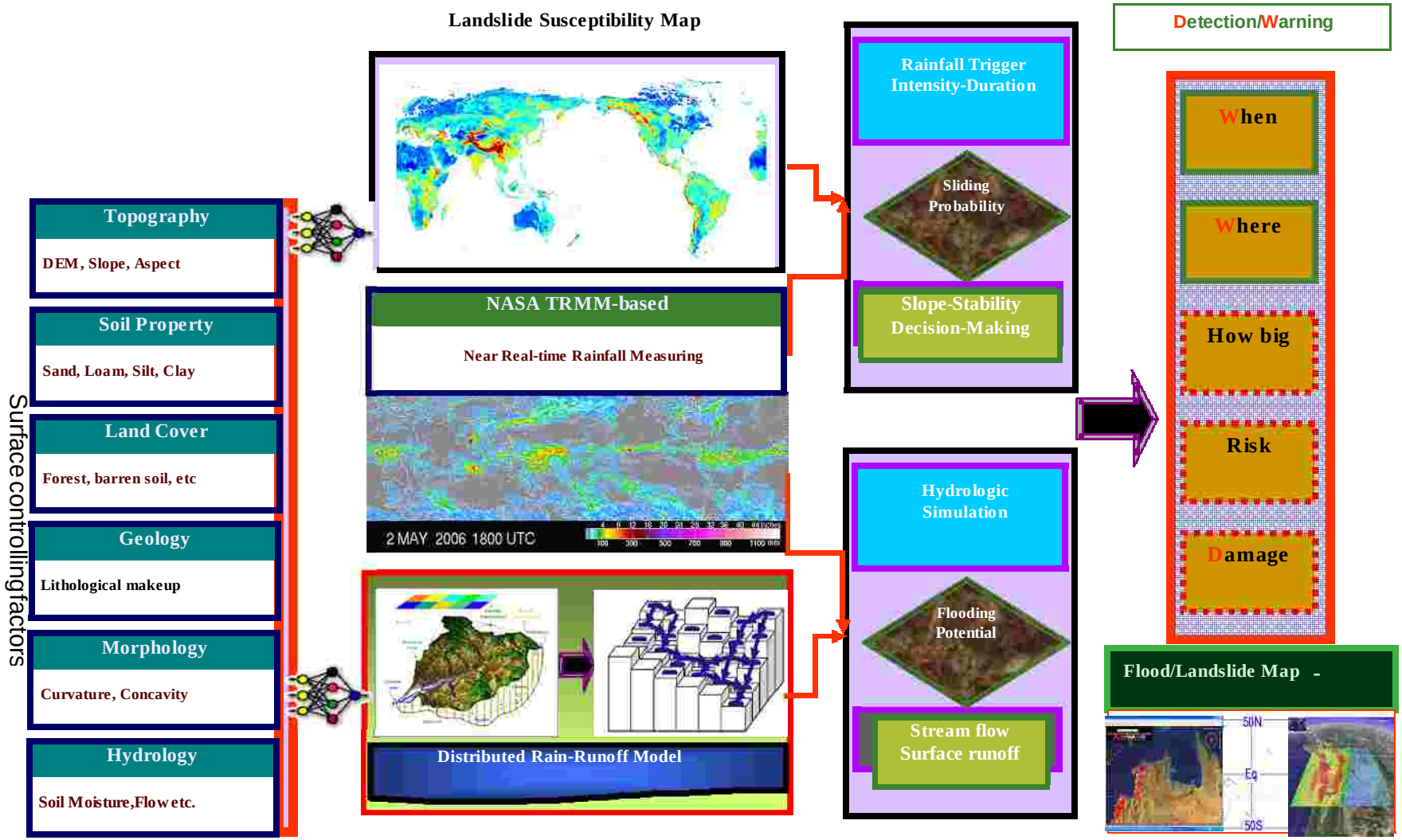

Figure 5: Outline of Global Landslide Hazard Assessment Model methodology. Hong et al. (2007).

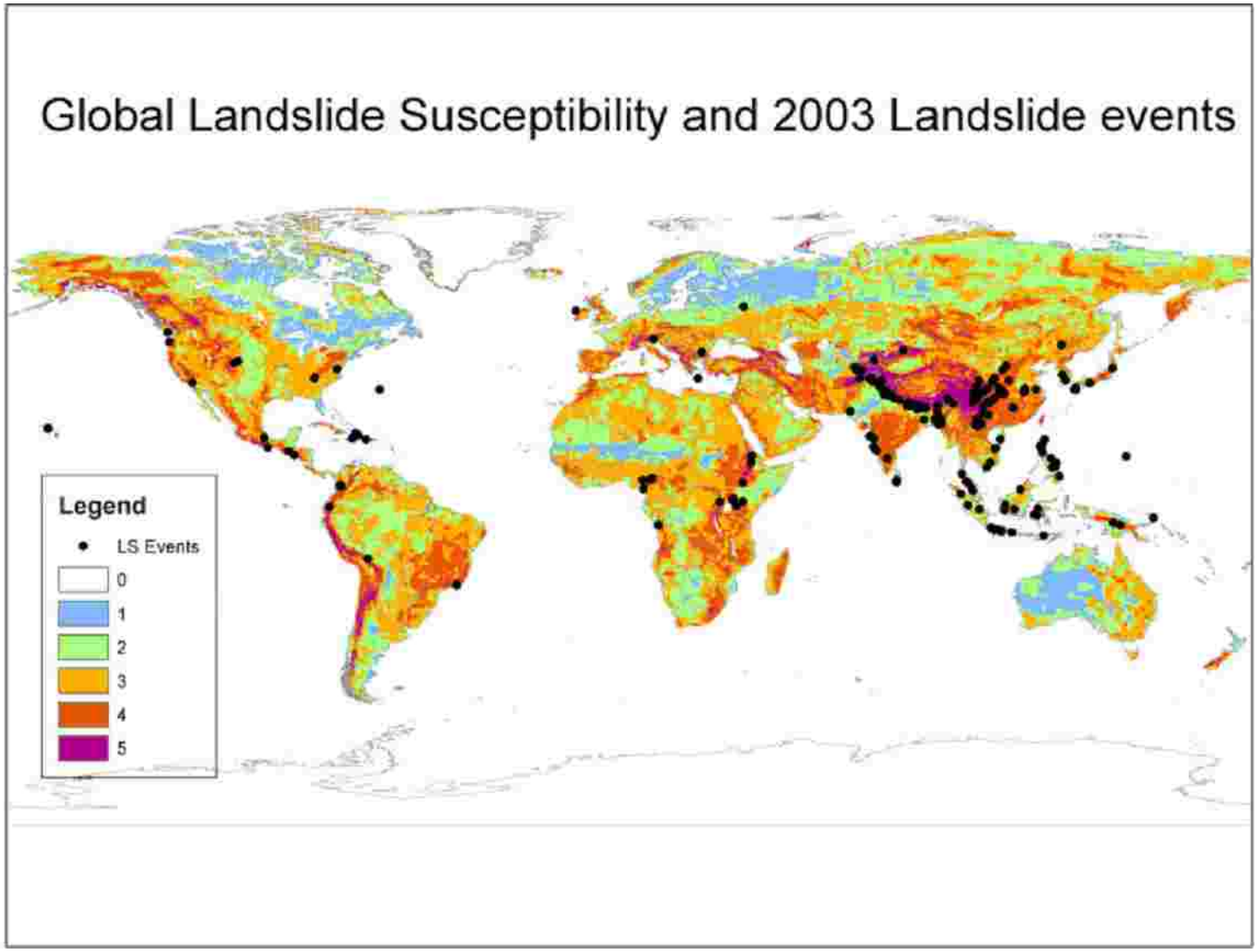

Figure 6: Global landslide susceptibility map from Hong et al. (2007) and 2003 landslide inventory compiled from multiple sources.

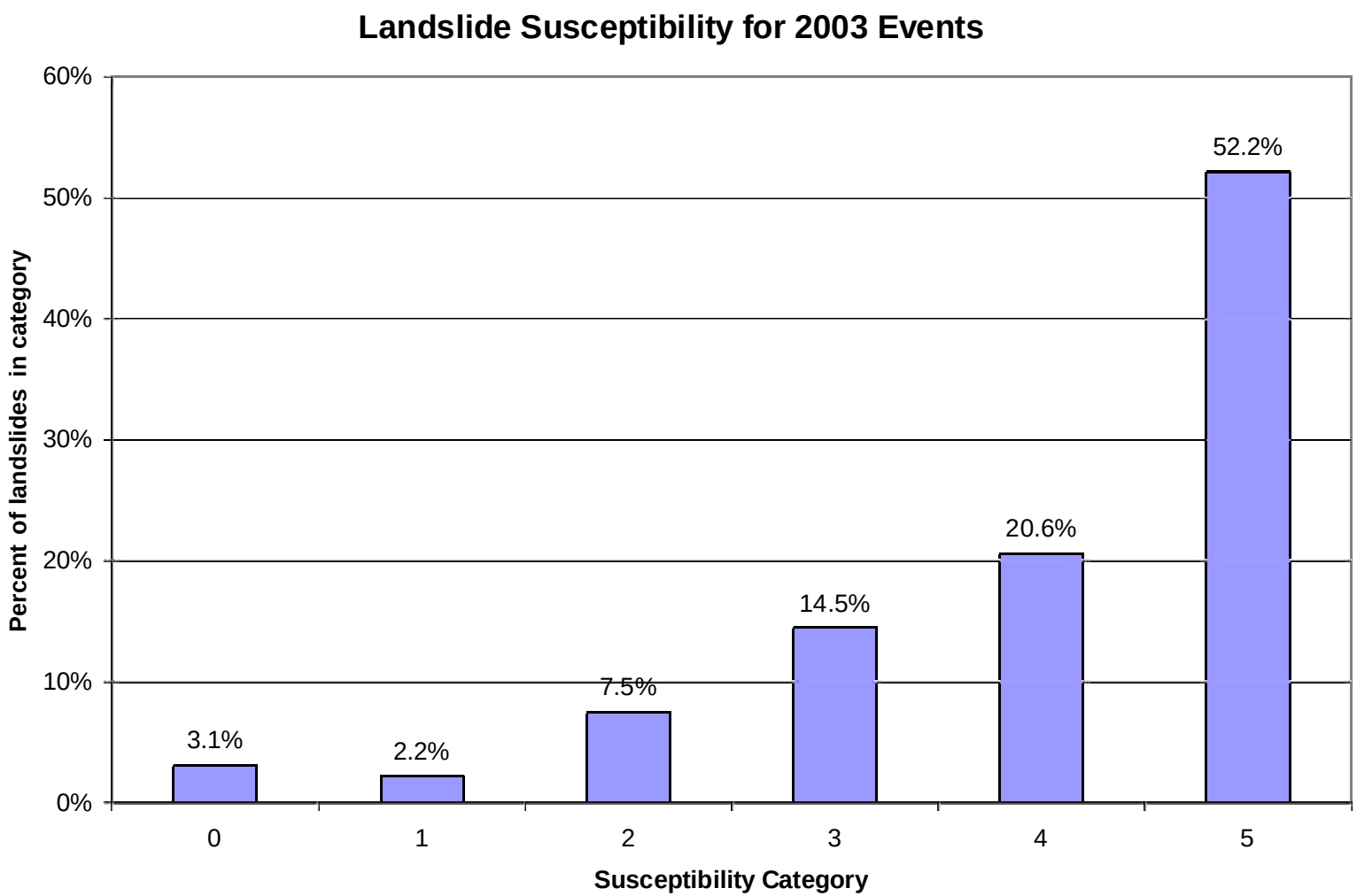

Figure 7: Distribution of 2003 landslide events for corresponding to the susceptibility category.

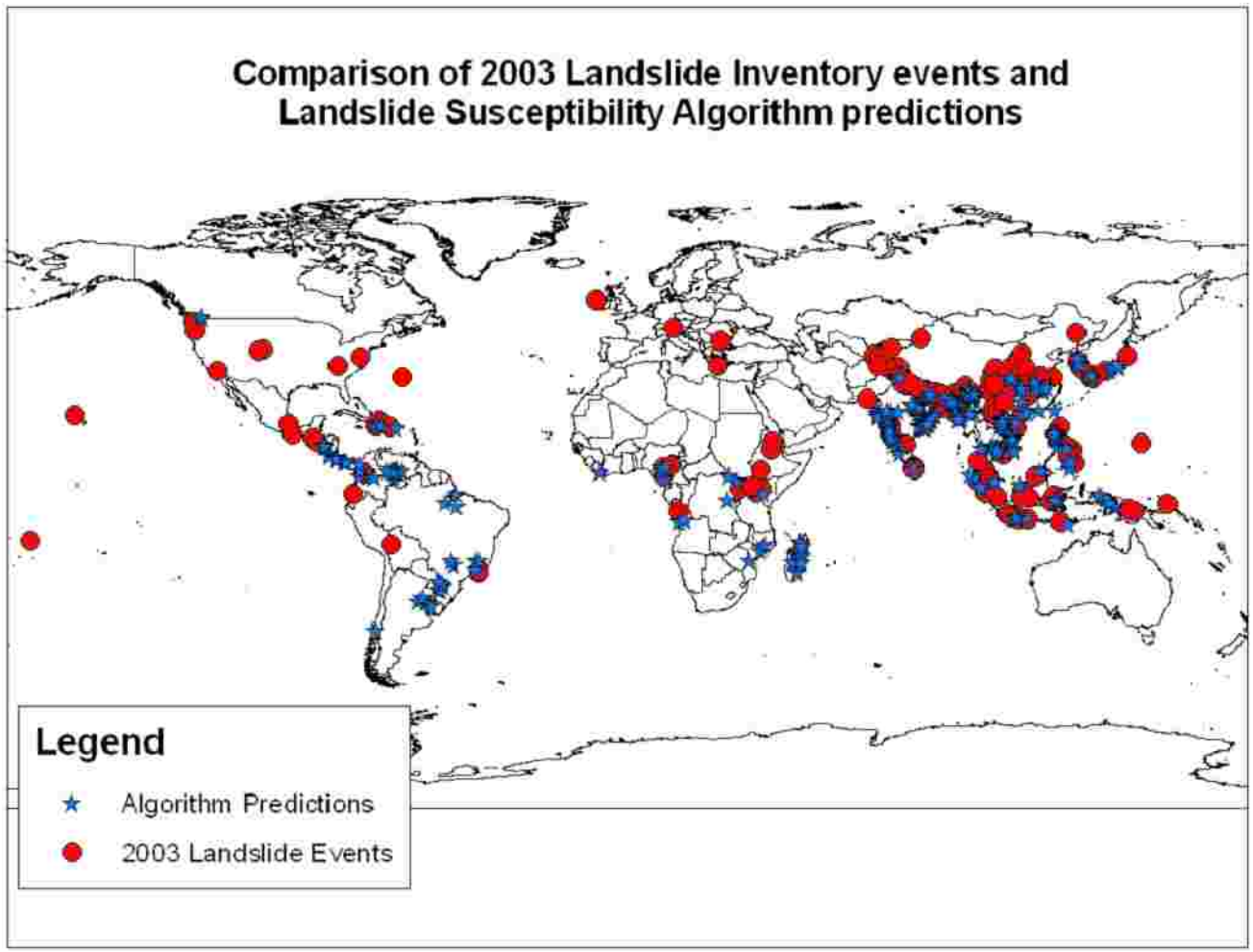

Figure 8: Comparison of 2003 landside inventory events and the algorithm predictions

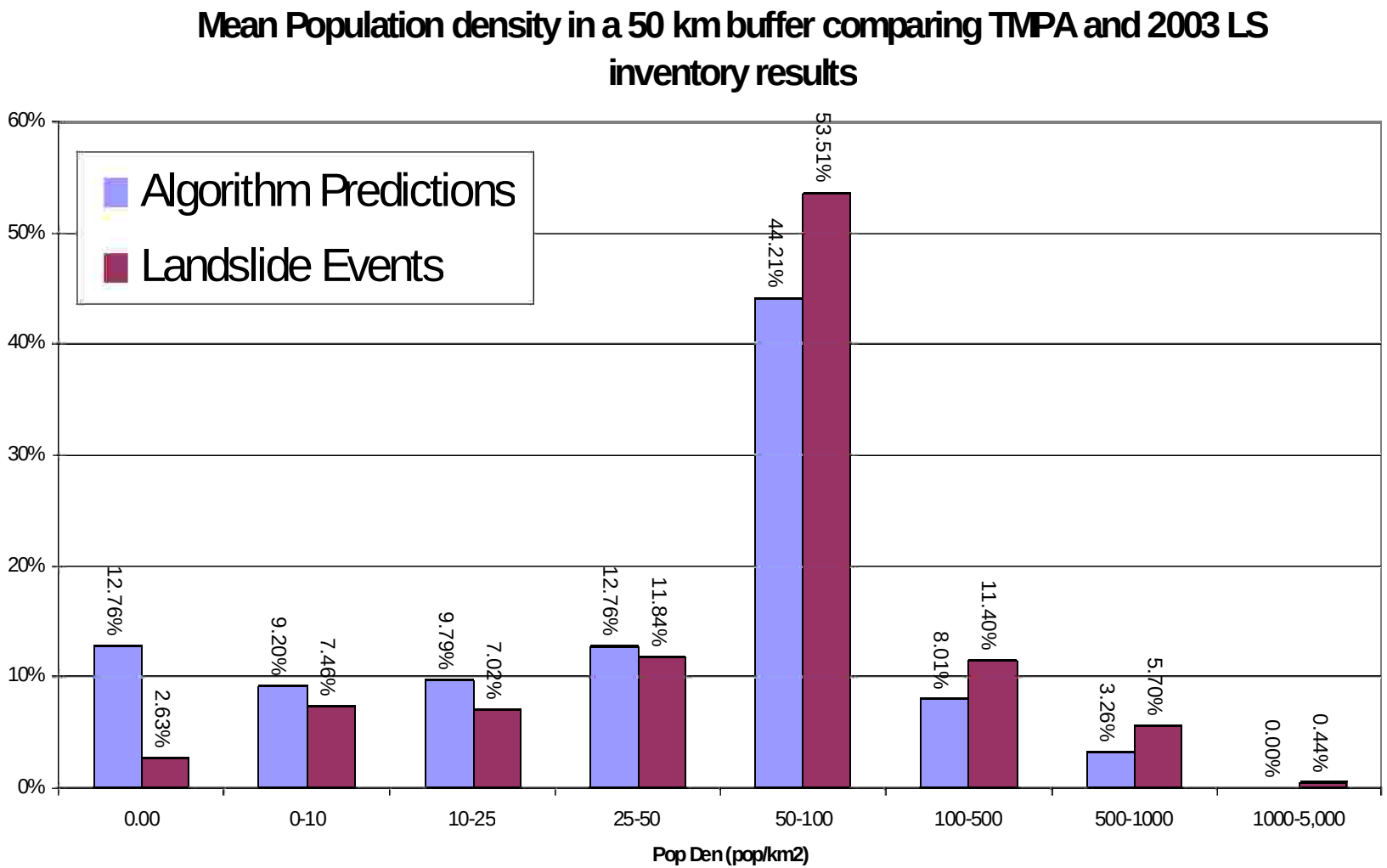

Figure 9: Comparison of mean population density distribution within a 50 km radius of the 2003 landslide events and algorithm predictions with corresponding percentages of the total.

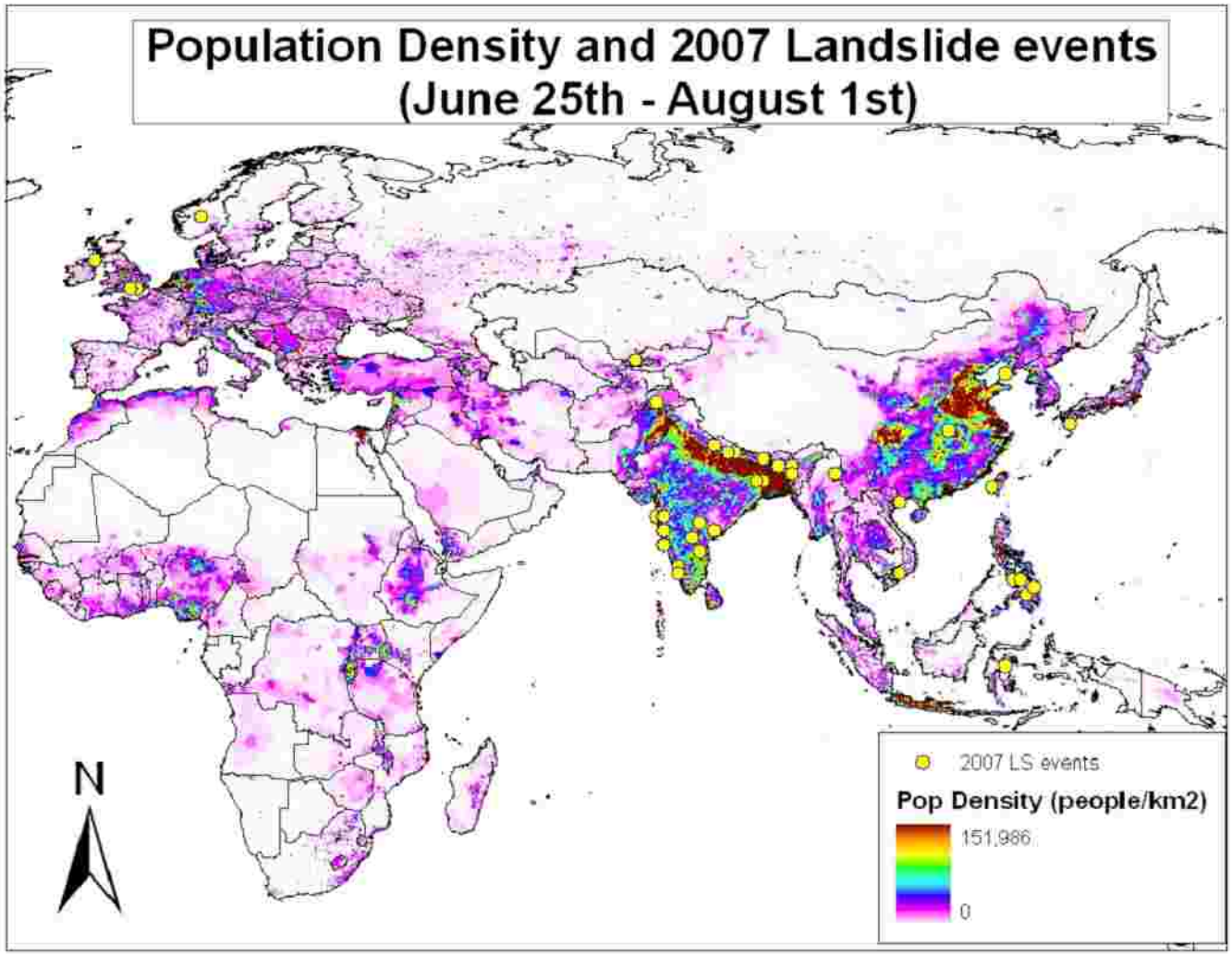

Figure 10: Population density and 2007 landslide events in the Asian region (June 25th – August 1st)

Rainfall Product	Coverage (geographic)	Spatial Resolution	Temporal Resolution	Limitations
Rainfall Gauges (NCDC,USGS)	Variable, 10,000 stations globally monitored by NCDC	Point Source (~30 cm radius)	hourly, 1970s to present	Wind errors, reporting issues, gauge placement
Surface Radar (NEXRAD Doppler systems)	158 stations in US, 400 km maximum radius from station	4 km gridding	4-10 minutes, hourly products, 1988 – present	Beam shielding and errors in mountain areas
IR-based estimates PERSIANN-CCS	Global	~4 km	Hourly, 2001 to present	Poor correlation to actual rain rate, indirect proxy
Low-Orbit Satellite TRMM (TMI, PR)	50 ° N-S	0.25º	3 h, 1997 to present	Limited temporal sampling, algorithm inconsistencies
TMPA CMORPH	50º N-S 59º N-S	0.25º	3 h 1998 – Present 2002 - Present	Not available at small spatial scales (<25km)

Table 1: Overview of different precipitation data sets, their coverage, resolution, time of operation and general limitations.