

¹ Effect of digital elevation model on Mohr-Coulomb
² geophysical flow model output

³ E. R. Stefanescu¹, M. Bursik², and A.K. Patra¹

⁴ ¹Department of Mechanical and Aerospace Engineering, University at
⁵ Buffalo, Buffalo, NY 14260

⁶ ²Department of Geology, University at Buffalo, Buffalo, NY 14260

⁷ June 25, 2014

⁸ **Abstract**

⁹ Digital Elevation Models(DEMs) used in geospatial analysis like the simulation of
¹⁰ geophysical flows such as floods, landslides and block and ash flows, differ in resolution,
¹¹ acquisition time and generation methodology, which results in varied representation
¹² of topographic features. This study investigates the effects of DEMs on the output
¹³ of a granular flow model, TITAN2D by comparing the output using different DEMs
¹⁴ to that obtained with a "true" representation of the terrain, which is considered to
¹⁵ be that obtained by using TOPSAR5m data. Seven DEMs at four resolutions from
¹⁶ four sources were used for Mammoth Mountain, California, a cumulodome volcano.
¹⁷ TITAN2D was run for seven different locations of an eruption of a potential dome
¹⁸ and two different collapse volumes. The resulting outputs were subsequently compared

19 with TOPSAR5m output, and qualitative and statistical inferences were drawn. DEMs
20 with different resolutions and sources generated different outputs which led to different
21 flow maps. For moderate and smaller scale flows ($\mathcal{O}(10^4) \text{ m}^3 - \mathcal{O}(10^5) \text{ m}^3$), different
22 representations can affect the computation of accurate footprint of the flow and fine
23 DEM resolution is critical for correct characterization of these flows.

24 1 Introduction

25 1.1 Overview

26 The use of computational models to calculate inundation areas by floods, landslides and
27 volcanic flows is increasing [Bates and Roo, 2000]. These models depend on proper rep-
28 resentation of topography. Digital elevation models (DEMs) are digital representations of
29 a terrain surface and are used to define topographic morphology. A DEM is an array of
30 squared cells (pixels) with an elevation associated to each pixel. There are different methods
31 to obtain digital topographic data with different levels of accuracy. The most commonly
32 used methods/instruments are: Shuttle Radar Topography Mission (SRTM) interferome-
33 try, Advance Spaceborne Thermal Emission and Reflection Radiometer (ASTER) digital
34 stereophotography, National Elevation Dataset (NED) and LiLight Detection And Ranging
35 (LiDAR). The traditional algorithms to obtain terrain surface use neighborhood operations
36 to calculate slope, aspect and points of inflection [Jenson and Dominique, 1988]. Additional
37 data corrections may include: edge matching, minimizing artefacts or filling of sliver areas
38 of missing data [Zhang et al., 2008]. These processes confirm the fact that the topographic
39 reality can be represented digitally only to a certain level of accuracy. However, the implicit
40 spatial relationships between elevation values can be determined, since the dimensions of
41 grid and the number of observations in each row is known [Fisher and Tate, 2006].

42 A common method to obtain a continuous DEM surface from contour lines or scattered
43 point elevations is by interpolation. The interpolation technique plays an important role in
44 achieving a high accuracy of DEM and the choice of interpolation technique depends on the
45 type of topography and the distribution of the sample points [Xiao et al., 2010, Binh and
46 Thuy, 2008]

47 Although DEMs are extensively used in geologic mapping, civil engineering, landscape
48 planning, visibility analysis, hydrology and geophysical modeling, and are critical to many
49 natural hazard risk applications, little work has been done on the effect of topographic
50 representation (point spacing, method of construction,etc.) on model output. Past analyses
51 generally included a comparison of numerical model output for digital elevation models
52 (DEMs) created with the same technology but at different resolution, due to the limited
53 number of available technologies to create DEMs. Zhang and Montgomery [1994] focused
54 on the effect of DEM grid size (2m, 4m, 10m, 30m and 90m) on landscape representation
55 and hydrological modelling with TOPMODEL. Their analysis based on variation of the
56 saturation area within a catchment suggested that 10-m resolution data presents a rational
57 compromise between increasing resolution and data volume for simulating geomorphic and
58 hydrological processes. Wolock and Price [1994] evaluated the effect of DEM data resolution
59 and map scale on a topographic index, for DEMs in eastern Pennsylvania, southern New
60 York, and northern New Jersey. The topographic index is a parameter used to characterize
61 the topography of a basin. It is defined as $\ln(\alpha / \tan \beta)$, where α is the upslope area per unit
62 contour length and $\tan \beta$ is the local slope angle. They compared TOPMODEL inputs and
63 outputs calculated with DEMs digitized from different sources at different resolution (30m
64 and 90m). They concluded that DEMs digitized from higher resolution maps yield a more
65 detailed representation of topography, but are not necessarily a better source of topographic
66 information for TOPMODEL inputs. They also found out that the mean of the topographic

67 index distribution increased as data resolution became more coarser. Another hydrological
68 model [Chanbey et al., 2005] for which the effect of input data resolution was studied was
69 SWAT (Soil and Water Assessment Tool). Seven scenarios were run at increasing DEM
70 grid sizes (30x30m, 100x100m, 150x150m, 200x200m, 300x300m, 500x500m, 1000x1000m)
71 on Moores Creek watershed (near Lincoln, AR, USA). Results of this study indicated that
72 DEM data from 100 to 200 m resolution achieved less than 10% error in the model output.
73 In another study, six DEMs at three resolutions from three sources were used to calculate
74 the inputs to a soil-erosion model (WEPP) [Zhang et al., 2008]. The study found the 10-m
75 LIDAR DEM to be a satisfactory input source for the model compared to 30m LIDAR,
76 4m LIDAR, 30m NED, 10m NED and 30m SRTM DEM. Stevens et al. [2006]investigated
77 the effects of grid-spacing on lahar inundation zones predicted by LAHARZ by comparing a
78 10-m gridded DEM with resampled derivatives of 30 m and 90 m resolution. No significant
79 effect could be detected except for some lahars falling short in travel distance by 1–2 km
80 when using the coarser-spaced grids.

81 Some work has also been done to study the link between ways of obtaining high-resolution
82 DEMs and the eventual use of the DEMs. Cochrane [1999] showed that by decreasing the
83 interval between sampled elevation points, the cost of creating a finer resolution DEM in-
84 creases. Finer resolution DEMs are also commonly generated from a coarser resolution DEM
85 by interpolation. Mitasova et al. [1996] used the method of regularized spline interpolation
86 with tension to create a finer DEM from a 30-m DEM. However, this method has been crit-
87 icized by Zhang and Montgomery [1994] and Desmet and Grovers [1997] perhaps because it
88 performs poorly when the appropriate parameters are not selected. Wise [2007] created high
89 resolution DEMs using different methods to run the hydrological model TOPMODEL. He
90 found that although the total flow, which mostly consisted of groundwater flow, was little
91 affected by creation method, the characteristics of the surface flow component were very

92 sensitive to creation method.

93 There is interest in developing computational models that simulate the propagation of
94 granular flows such as pyroclastic flows and rockslide-avalanches as well as the propagation
95 of water. For such hazards related models, it is furthermore often important to calculate
96 inundation areas of the flows under consideration. The TITAN2D model [Patra et al., 2005]
97 performs dry granular flow simulations on a DEM of a user-define region. DEMs used as
98 TITAN2D input consist of files containing large numbers of measurements representing the
99 height of the surface of the earth. These DEMs are obtained using spaceborne, airborne and
100 photogrammetric elevation data products, such as SRTM, ASTER, TOPSAR and NED.
101 In a recent paper, [Capra et al., 2011] explore the impact of topographic resolution alone
102 on TITAN2D simulations of debris flow/ pyroclastic density currents. Different DEMs of
103 Colima volcano (Mexico) at different cell sizes (5, 10, 30, 50, and 90 m) were explored.
104 They performed a qualitative analysis of the velocity and planview trajectories of the flow
105 centerlines, and concluded that several high resolution DEMs give similar results, while
106 coarser DEM results are largely different and inaccurate. In this paper we perform a thorough
107 quantitative analysis of TITAN2D output to investigate the effects of different DEMs on flow
108 height and flow inundation area calculations.

109 1.2 DEM sources and methods of interpolation

110 The SRTM (Shuttle Radar Topography Mission) was the first space-borne, fixed-baseline in-
111 terferometric synthetic aperture radar (InSAR). In 2000, the SRTM collected interferometric
112 radar data over 80 % of the Earth's land surface [Zhang et al., 2008], using a dual-radar an-
113 tenna. The goal of a radar interferometer is to measure the difference in range between
114 two observations of a given ground point with sufficient accuracy to allow topographic re-
115 construction. The SRTM 1-arc second DEM ($\sim 30 \times 30\text{m}$) represents the raw form of the

116 SRTM DEM and has a higher resolution than its derivative, the SRTM 3-arc second DEM
117 ($\sim 90 \times 90$ m). For the SRTM 1-arc second dataset the mission planners anticipated linear
118 vertical absolute height error to be less than 16m, linear vertical relative height error less
119 than 10m, circular absolute geolocation error less than 20 m and circular relative geolocation
120 error less than 15 m. These accuracies were achieved at 90 % confidence level, consistent
121 with the National Map Accuracy Standards (NMAS) [Farr et al., 2007].

122 ASTER (Advance Spaceborne Thermal Emission and Reflection Radiometer) on board
123 the Terra spacecraft is a multispectral optical sensor that provides imagery with visible
124 and near-infrared (VNIR) bands in 15-m spatial resolution, shortwave infrared (SWIR) and
125 thermal infrared (TIR) bands with 30-m and 90-m resolution, respectively [Abrams, 2000].
126 The ASTER instrument produces two types of Level-1 data — Level-1 A and Level-1 B
127 [Fujisada et al., 2005]. Level-1 A data are defined as reconstructed, unprocessed instrument
128 data at full resolution. Level-1 A data are used as source data to generate DEM products,
129 because useful instrument geometric parameters and spacecraft information are included.
130 Level-1 B data are created by applying the radiometric calibration and geometric correction
131 coefficients to the Level-1 A data. Several study results show that the vertical accuracy of
132 an ASTER DEM approaches 25m [Fujisada et al., 2005], but in areas with less vegetation
133 cover, the accuracy can rise to 9–11m [Goncalves and Oliveira, 2004].

134 The NED (National Elevation Dataset) is a seamless elevation dataset assembled for the
135 continental US from various data sources. The NED is a combination of many datasets
136 that consist of digitized USGS quadrangles, gravimetric surveys, ortho-rectified airphotos,
137 laser altimeter surveys and SAR surveys [Shaffer, 2006]. Available horizontal resolutions
138 are 10m and 30m. The USGS creates the original DEMs by autocorrelation or manual
139 profiling from aerial photographs. Data corrections are made in the NED assembly process
140 to minimize artefacts, perform edge matching and fill sliver areas of missing data. The NED

¹⁴¹ dataset vertical accuracy is 2.44m RMSE at 90% confidence interval according to the NMAS
¹⁴² (USGS, 2009).

¹⁴³ TOPSAR data are a polarimetric Synthetic Aperture Radar(SAR) dataset generated by
¹⁴⁴ an airborne SAR system [Mouginis-Mark and Garbeil, 2005]. The TOPSAR system was
¹⁴⁵ implemented via two antennas mounted nearly vertically on the side of a DC-8 aircraft. The
¹⁴⁶ TOPSAR DEM for our target area has a resolution of 5×5 m, with a vertical accuracy of
¹⁴⁷ 1.3–2.0 m RMSE [Hooper et al., 2003] at 90% confidence interval in accordance with the
¹⁴⁸ NMAS.

¹⁴⁹ The reason the analysis was performed on these particular DEMs was to study model
¹⁵⁰ characteristics on standard, widely available topographic products. NED is the standard
¹⁵¹ topographic model for the US, while SRTM and ASTER are widely, almost globally available.
¹⁵² It is known that ASTER DEMs contain higher vertical errors than SRTM DEMs, while the
¹⁵³ TOPSAR DEM was constructed from high-resolution airborne SAR data with dense areal
¹⁵⁴ coverage. The TOPSAR DEM is thus little affected by interpolation errors, although the
¹⁵⁵ data lose accuracy on slopes facing away from the sensor, or in areas obscured from the
¹⁵⁶ line-of-sight of the radar. In many of the above methods of raw data acquisition can result
¹⁵⁷ large amounts of missing data. This can be caused by the type of relief in the area of interest
¹⁵⁸ (e.g SRTM has large amounts of missing data in high relief areas compared to ASTER),
¹⁵⁹ around bodies of water and data may be completely missing around areas of persistent cloud
¹⁶⁰ cover. We have no information regarding the amount of data missing in the raw data or the
¹⁶¹ dates when the data points were collected.

¹⁶² Since elevation data usually consist of discrete, irregularly spaced sample points, surface
¹⁶³ interpolation techniques are used to transform discrete elevation data to a DEM surface with
¹⁶⁴ regularly spaced data points. Algorithms to perform the interpolation include techniques
¹⁶⁵ such as inverse distance weighted (IDW) or bi-cubic spline interpolation, Triangular Irregular

₁₆₆ Network construction, kriging, finite elements, etc. [Florinsky, 1998, Cebecauer et al., 2002].
₁₆₇ Some recommendations on choosing interpolation technique were made: for mountainous
₁₆₈ areas the spline regularized is the most suitable, for gentle hilly and flat areas, IDW or
₁₆₉ kriging are recommended [Mitasova and Hofierka, 1993, Binh and Thuy, 2008]. To produce
₁₇₀ a DEM dataset to which we had ourselves applied a standard interpolating methodology, we
₁₇₁ calculated a 30 m DEM obtained from the 5 m TOPSAR DEM using spline interpolation.
₁₇₂ *The spline interpolation* estimates the elevation of a specific point using a mathematical
₁₇₃ function that minimizes the overall surface curvature, resulting in a smooth surface that
₁₇₄ passes exactly through the input points [Kamada and Enkhabat, 2009, Xiao et al., 2010].
₁₇₅ There are two spline methods: regularized and tension. The regularized method creates
₁₇₆ a smooth, gradually changing surface with values that may lie outside the sample data
₁₇₇ range. The tension method is based on the assumption that the interpolation function
₁₇₈ should pass through (or close to) sample points and should be at the same time as smooth
₁₇₉ as possible [Mitasova and Hofierka, 1993, Kamada and Enkhabat, 2009, Kvasov, 2011]. The
₁₈₀ main parameters of the spline interpolation are the number of sampled points used for
₁₈₁ interpolation, and the weight. For the regularized spline, the higher the weight, the smoother
₁₈₂ the output surface. For the tension spline, the higher the weight, the coarser the output
₁₈₃ surface. The parameter values used in a particularly instance of splining are quantitative
₁₈₄ measures that describe only a statistical aspect of interpolation quality, so applicability of
₁₈₅ the interpolated DEM for a specific purpose is not known a priori (e.g. correct representation
₁₈₆ of undersampled features important to a flow model).

₁₈₇ Hypothesis tests are usually conducted in all fields in which theory can be tested against
₁₈₈ observations. The null hypothesis is a statement that we want to test. In general, the null
₁₈₉ hypothesis is that things are the same, or the same as a theoretical expectation. In this
₁₉₀ paper, we test the simple hypothesis that different DEMs result in different model outputs.

191 To test the hypothesis, an analysis is carried out by comparing the results against the null
192 hypothesis that different DEMs yield the same model output. We furthermore analyze the
193 results to determine if the test outcome is dependent on flow volume or on the particular
194 DEM used. Because it is not possible to have a DEM that represents the “true” topography,
195 the hypothesis is tested using the best DEM available for the area. In our case, the TOPSAR
196 5-m DEM is considered the “truth”.

197 This paper focuses on the datasets available for Mammoth Mountain, performing a qual-
198 itative and statistical analysis to address how different the datasets are in terms of their
199 effect on the computation of granular flow propagation.

200 **2 Technical Approach and Methodology**

201 The TITAN2D code used in this effort conducts numerical simulations of flows on digital
202 representations of the natural terrain. It is based on a depth-averaged model for an in-
203 compressible granular material, governed by Coulomb-type friction interactions [Savage and
204 Hutter, 1989]. The governing equations are obtained by applying conservation laws to the
205 incompressible continuum, providing appropriate constitutive modeling assumptions, and
206 then taking advantage of the shallowness of the flows (flows are much longer and wider than
207 they are deep) to obtain simpler depth-averaged representations [Patra et al., 2005]. The
208 motion of the material is considered to be gravitationally driven and resisted by both in-
209 ternal and bed friction forces. The stress boundary conditions are: no stress at the upper
210 free-surface and a Coulomb-like friction law imposed at the interface between the material
211 and the basal surface. Note that the primary factor driving the flow is the component of
212 gravity along the surface which depends on a local slope computed from the elevation data.
213 The resulting hyperbolic system of equations is solved using a finite-volume scheme with a

214 second-order Godunov solver. We have thus used a fixed-rheology model, which we agree
215 cannot give accurate interpretations or predictions of debris flow motion because of their
216 complex behavior, but the Coulomb rheology has proven to be a reliable option for coarse
217 solid constituents and flow models based on the mechanics of granular media. In a Coulomb
218 model the shear stress is a function of a bed friction and/or internal friction angle and the
219 effective normal stress, and hence is, contrary to rheologic approaches for viscoplastic fluids,
220 independent of the shear rate. Mohr-Coulomb models are used to describe 'hard suspension'
221 models, when the frictional interactions between particles comprise the dominant process.
222 'Soft suspensions' correspond to viscoplastic flow behavior and phenomenological laws like
223 the Bingham or the generalized Herschel-Bulkley for rheological behavior.

224 The program runs in parallel, using the message passing interface standard (MPI) to al-
225 low communication between multiple processors, increasing computational power, decreasing
226 computational time and allowing use of large datasets. The algorithm uses local adaptive
227 mesh refinement for shock capturing, and dynamic load balancing for the efficient use of
228 computational resources. Topographic data are included in the simulation by using a pre-
229 processing routine in which digital elevation data are imported. The DEM file containing
230 X,Y,Z data (typically UTM easting, UTM northing and elevation in meters) must be prop-
231 erly configured to operate with TITAN2D through the use of a header. The other input
232 parameters for the model which describe the initial geometry of the dome and control the
233 flow movement are presented in Table 1.

234 TITAN2D was run for a few combinations of the parameters that characterized the initial
235 failure (Table 2): two different-sized volumes, and seven different starting locations. The
236 initial pile shape is a paraboloid for which the volume is calculated as $V = \frac{\pi}{2} \cdot r_{min} \cdot r_{max} \cdot h_{max}$.

237 The study area is Mammoth Mountain — a large, geologically young, composite dome
238 volcano located on the southwestern rim of Long Valley Caldera, California [Bailey, 1989].

The summit ridge is composed of 18 domes and/or lava flow units [Burkett, 2007]. The summit itself is 3,371 m above sea level, and more than 1,200 m above the general level of the adjacent caldera floor [Farrar, 1995]. Phreatic eruptions from vents on the lower portion of the mountain in 667–737 calBP directly preceded the magmatic 605–625 calBP North Mono eruption and the 1350 AD (= 600 calBP) Inyo eruption. Recently, an anomalously high level of CO_2 has been discovered in the local soil and groundwater, suggesting the existence of a large mass of free magmatic gases and/or a water reservoir beneath Mammoth Mountain saturated with magmatic gases. These data taken together suggest that magmatism underneath the mountain has been rejuvenated, and that this rejuvenation could lead to explosive events, including the generation of both pumiceous and dome-collapse pyroclastic flows.

Common hazardous geophysical mass flows on Mammoth Mountain might contain $O(10^5 - 10^8)$ m³ material [Patra et al., 2005, Burkett, 2007]. We consider two flows — a low-volume flow with an initial volume of 1.54×10^5 m³, with a maximum initial thickness of $h_{max} = 20m$, major axis $r_{max} = 70m$ and minor axis $r_{min} = 70m$, and a high-volume flow of 2.82×10^7 m³ corresponding to $h_{max} = 200m$ and $r_{max} = r_{min} = 300m$. By using two endmember volumes, we can observe extrema in the effect of volume on the output for different DEMs.

By choosing seven potential starting locations, we are trying to cover an area near the summit where we observed that potential vents could be situated, based on positions of ancient vents [Bailey, 1989, Burkett, 2007] (Figure 1). The flow maps were constructed by combining the maximum pile height of the individual flows. In case of an flow overlap the maximum of the piles height was plotted. The properties of the flowing material are taken into account by choosing appropriate values for the bed (ϕ_{bed}) and internal friction angles (ϕ_{int}): 20 degrees and 25 degrees, respectively.

The DEM's available for experiment are: TOPSAR (5 × 5m and 30 × 30m resolution), NED (10 × 10m and 30 × 30m resolution), SRTM (30 × 30m and 90 × 90m resolution) and

264 ASTER 30×30 m resolution. The DEMs cover an area of 93.75 km^2 centered on the volcano.
265 Two DEM's used in the present analysis were obtained from different re-sampling methods:
266 an interpolation method and decimation method. The interpolated 30m DEM was obtained
267 using regularized spline with tension (RST) interpolation (Fig. 2 a,b,c; Fig. 3 a,b,c) using
268 GRASS GIS *r.resamp.rst* on the TOPSAR DEM. It was used with the default parameters
269 since the tension works as a rescaling parameter. This means that high tension increases the
270 distances between the points and reduces the range of impact of each point, while low tension
271 decreases the distance and the points influence each other over longer range [Cebecauer et al.,
272 2002]. This interpolation method is usually used when there is a large difference in resolution
273 between original and target (1:3 or more) or when there are large gaps in the raster data.
274 The resulting surface is similar to that produced by bicubic interpolation. The decimation
275 method is a resampling method which involves retaining every fifth grid cell resulting in a
276 30m DEM constructed from TOPSAR data. We refer to the DEM obtained by interpolation
277 as *spline interpolation / interp* while the DEM obtained by decimation as *decimated*. The
278 height difference between available DEMs over the area of interest varies between 20m and
279 90m (for the case where both DEMs in the difference have available elevation information
280 at a particular grid cell). The slope value varies between 0 degrees and 65 degrees.

281 One of the output files produced by TITAN2D contains the maximum flow depth over
282 the entire simulated time at each grid cell. We use this file in creating flow maps as this
283 map represents the maximum extent of inundation for a given set of input parameters and
284 boundary conditions. Its outline is in fact the “inundation area” commonly referred to
285 in analyses of volcanic flow and landslide hazards [Procter et al., 2010]. We consider the
286 TOPSAR 5-m DEM to best represent the terrain features and therefore to be “ground truth”;
287 and the resultant flow map is considered the reference output dataset. Full statistical error
288 analysis is done on outputs using NED 10m and SRTM 90m DEMs, and a generalization

289 to the other DEMs is performed in the end. The two DEMs were chosen on the basis of
290 their expected outputs: the output of NED 10m is expected to be most similar to that of
291 TOPSAR 5m based on their close resolution, while the SRTM 90m output is expected to be
292 most different.

293 Two types of analysis are performed: a qualitative analysis and a statistical analysis.
294 The qualitative analysis is specific for geophysical mass flows, for which it is common to
295 simply compare the shapes of model and field-data generated inundation areas [Takahashi
296 and Tsujimoto, 2000]. We therefore extend the analysis to a grid cells based classification
297 in which we consider the TOPSAR 5m flow map to be the reference data and the other flow
298 maps to be the classified data.

299 The validity of a simulation model is determined by the accuracy of reproducing natural
300 events. However, comparing the model's output with reality is not a true validity test. One
301 way of evaluating the goodness-of-fit of a model output is through an accuracy assessment.
302 Accuracy assessment can be performed by comparing two maps: one generated by the model
303 when using the Topsar 5m (the reference data), and the other obtained when using the rest
304 of the DEMs (data to be assessed).

305 The grid cells are classified into inundated and noninundated classes based on whether a
306 grid cell is within the inundation area. We count the number of grid cells in each class, and
307 construct a matrix (error matrix) telling us what percentage of those grid cells ended up being
308 placed into each of the two classes in both reference and classified data for each combination
309 of reference with classified data. The error matrix gives us information about how accurate
310 our classification is as well as a general accuracy for the map as a whole, by calculating
311 different statistics such as overall accuracy, producer's accuracies (omission error), user's
312 accuracies (commission error) and model efficiency [G.Legorreta-Paulin and Bursik, 2009].
313 In an ideal situation, all the nonmajor diagonal elements of the error matrix would be zero,

314 indicating that no area has been misclassified [Congalton, 2005]. An omission error occurs
315 when inundated grid cells are classified as noninundated grid cells and vice-versa. It is also
316 referred as the “producer’s accuracy” because from this measurement the producer of the
317 classification will know how well a certain inundated area category was classified [Congalton,
318 2005]. The commission error is when grid cells were included in a given class (inundated or
319 noninundated) but should have been in the other [Lillesand and Keifer, 2000]. This is also
320 called “user’s accuracy”, indicating for the user of the map the probability that a grid cell
321 classified on the map actually represents that category on the reference map. The producer’s
322 accuracy and user’s accuracy thus characterize the probability of a cell in the classified
323 map being correct relative to the number in the reference map and in the classified map,
324 respectively. The overall accuracy summarizes the total agreement/disagreement between
325 the maps. It only incorporates the major diagonal of the error matrix where both classified
326 and reference data matched, and excludes the omission and commission errors.

327 For the statistical analysis we performed the Kolmogorov-Smirnov(K-S) [Barr and David-
328 son, 1973, Lopes et al., 2007] test, which is a widely used non-parametric statistical test.
329 Although it is mainly used as a one-sample test to compare the frequency distribution of
330 a sample to some known distribution, such as a normal distribution, it can also be used as
331 a two-sample test. As a two-sample test, the K-S test checks if two data sets X_1 and X_2
332 of length n_1 and n_2 differ significantly. The null hypothesis for this test is that X_1 and X_2
333 come from the same distribution. The alternative hypothesis is that the two samples do not
334 come from the same distribution. Mathematically the test statistic can be written as

$$k = \max(|F_1(x) - F_2(x)|) \quad (1)$$

335 where $F_1(x)$ is the proportion of X_1 values less than or equal to x and $F_2(x)$ is the proportion
336 of X_2 values less than or equal to x . When the p -value ($p = P(k > k_{critic})$) is greater than
337 or equal to the level of significance α , it can be concluded that X_1 and X_2 have the same

338 distributions, otherwise the null hypothesis has to be rejected [Biswas et al., 2008].

339 3 Results

340 3.1 Qualitative analysis

341 For low and high volume flows using different DEMs, maps were constructed displaying
342 the maximum depths of the seven flows at each grid cell (Fig. 2, Fig. 3). The qualitative
343 analysis performed here is a visual comparison of the flow shapes to the TOPSAR 5 m
344 shape. First of all, we remark that the terrain features are captured differently for each
345 DEM, following a smoothing trend compared to the reference data as the DEM becomes
346 coarser. Thus, the deposit outline for the TOPSAR flow has a complex finger-like spread
347 digitation and numerous small-scale granular outliers. The flow depth is highly variable
348 within the outline on the granular scale. Generally with decreasing granularity/ details of
349 the DEM, the complexity of the outline and the flow depth variations decrease. An exception
350 is seen in the ASTER flow map, which appears to have a smoother variation in flow depth
351 than other maps; as well as fewer flow depth variations resulting from linear artefacts in the
352 DEM, except perhaps for TOPSAR 5m, and 30m decimated flow maps. For high-volume
353 flows, for all DEMs, the runout and of the flow outline is longer than it is for low-volume
354 flows. Valleys and channels exert less control over the inundation area for high-volume flows.
355 In fact, the largest difference among the high-volume flows occurs in the extreme western
356 extension of Flow 1, where the flows interact differently with a low sill, resulting in different
357 volumes of material flowing into the drainage of the Middle Fork of the San Joaquin River.
358 The granularity of the DEM is still reflected in the digitation on the flow margin, but there
359 are relatively fewer granular outliers and internal variations in flow depth on the scale of the
360 resolution of the DEM.

361 The “match” of inundation areas can be defined as the intersection of two flow areas
362 from different DEMs divided by their union [Galas et al., 2007]. The results from Table 3
363 indicate that NED 10m and TOPSAR 30m constructed by decimation of 5m data (the first
364 of which has a resolution “close” to TOPSAR 5m) give a “good” match while ASTER and
365 SRTM 90m give the poorest results. We also can conclude that the “match” improves and
366 converges with increasing flow volume.

367 The differences among datasets can be attributed to different reasons. For example,
368 we can compare the TITAN2D output of two DEM’s obtained from different re-sampling
369 methods. Many details of the map were lost in the interpolation method, which resulted in
370 a completely different flow shape than either TOPSAR5m or TOPSAR30m decimated.

371 The error matrix gives us information regarding how much the inundated area and the
372 non-inundated area in each map were classified similarly to the reference DEM (Table 4). For
373 a low volume flow, it can be seen that the reference map (TOPSAR 5m) has 11836 inundated
374 grid cells and 13416 grid cells noninundated, while the NED 10m classified map has 11904
375 noninundated grid cells and 13348 grid cells with no flow. If the classified map is the SRTM
376 90m, the reference map has 11836 inundated grid cells and 13416 grid cells noninundated,
377 while the classified map has 9524 inundated grid cells and 15728 noninundated grid cells.

378 There are two types of error in the classified map relative to the reference map: omission
379 error and commission error. The omission error corresponds to non-diagonal elements relative
380 to the column dataset. 1362 grid cells that should have been classified as inundated were
381 omitted from that category by NED 10m, and 1430 cell girds that should have been classified
382 as noninundated were omitted from that category. SRTM 90m failed to classify 3699 as
383 inundated and 1387 as noninundated. The commission error corresponds to non-diagonal
384 elements relative to the row dataset. NED 10m improperly includes 1430 grid cells in the
385 inundated category, and 1362 grid cells in noninundated category, while SRTM 90m includes

³⁸⁶ 1387 grid cells in the inundated category and 3699 in the noninundated category.

³⁸⁷ The producer's and user's accuracies characterize the omission and the commission error,
³⁸⁸ respectively. The producer's accuracy is calculated as the ratio of the number of correctly
³⁸⁹ classified grid cells in each category to the total number of true (reference) grid cells for each
³⁹⁰ category. The value of the producer's accuracy shows how well the data are classified into
³⁹¹ inundated and noninundated areas in the classified map. The producer's accuracy for TOP-
³⁹² SAR 5m as reference data and NED 10m as classified data is 88.49% ((10474/11836)*100)
³⁹³ (Table 5). This means NED 10m failed to detect 11.51 % of the "true" inundation area (omis-
³⁹⁴ sion error). For the noninundated area NED 10m predicted 89.34 % ((11986/13416)*100)
³⁹⁵ correctly. When the classified data are SRTM 90m, the producer's accuracy for the inundated
³⁹⁶ area is 68.75 % (8137/11836) while for the noninundated area it is 89.66 % (12029/13416).

³⁹⁷ User's accuracies are calculated as the ratio of the number of correctly classified grid cells
³⁹⁸ in each category to the total number of grid cells that are classified in that category. The
³⁹⁹ value of the user's accuracy shows the percentage that grid cells classified in a given category
⁴⁰⁰ actually represent that category in the reference data. If we consider NED 10m data, the
⁴⁰¹ user's accuracy is 87.99% ((10474/11904)*100) for the inundated area and 89.80% for the
⁴⁰² noninundated area ((11986/13348)*100), due to the fact that 1362 and 1430 grid cells are
⁴⁰³ incorrectly included in the inundated and noninundated categories, respectively (Table 5).
⁴⁰⁴ SRTM 90m data failed to include 3699 in the inundated category and 1387 in the noninun-
⁴⁰⁵ dated category, resulting in a 85.44 % ((8137/9524)*100) user accuracy for the inundated
⁴⁰⁶ area and a 76.48 % ((12029/15728)*100) user accuracy for the noninundated area.

⁴⁰⁷ Overall accuracy indicates how reliable the model results are. It is calculated as the total
⁴⁰⁸ number of correctly classified grid cells of inundated and noninundated (the main diagonal
⁴⁰⁹ data) divided by the total number of grid cells in the study area. NED 10m and SRTM 90m
⁴¹⁰ have an overall accuracy of 88.94% and 79.86 %, respectively (Table 5). This is consistent

411 with our hypothesis that use of the NED 10m data results in inundated or noninundated
412 areas closer to those of TOPSAR than does the SRTM 90m data.

413 An important output variable that we explore graphically is the height of the flowing
414 mass at the flow centerline, as a function of the elevation at the same location. In this way
415 we can gain an intuitive understanding of the dependence of the output (flow height) on the
416 terrain. In this analysis, it is critical to emphasize the presence of two regimes in the motion
417 of masses of granular materials [Lube et al., 2005, Lajeunesse et al., 2004] that are almost
418 exclusively characterized by the ratio a , $a = h_{max}/r_{max}$. For our high-volume flows, this ratio
419 is 2/3, and for the low-volume it is 2/7. High a flows (our high-volume flows) are dominated
420 by the collapse of the initial mass of material. The descent of the upper parts of the collapsing
421 mass pushes the toe outward radially. These can be deep flows dominated by inertia of the
422 entire collapsing mass. Low a flows are dominated by avalanching, which occurs mostly on
423 the face of the initial pile. These are therefore shallow flows dominated by grain interaction.
424 We can for convenience name these two regimes the collapsing and avalanching regimes,
425 respectively. In Figure 4 and Figure 5 the normalized elevation and normalize pile height
426 versus the distance for low volume flow and high volume flow, respectively, are presented.
427 The normalization is done using the following formula: $\frac{x - \mu_x}{\max_x - \min_x}$, where x represents the
428 variable to be normalized (elevation or pile height), μ_x is the mean value of x and \max_x , \min_x
429 represents maximum of x and minimum of x , respectively. Figure 4 shows the results of this
430 analysis applied to low-volume (low a) flows. Tracking of the pile height starts approximately
431 100 m from the source, as closer than this distance the flow is essentially a collapsing mass
432 governed by internal rather than bed interactions. For flows on all DEMs the difference with
433 the TOPSAR 5m output is greater at the beginning than it is at the end of the flow. This
434 is because all flows feather to a vanishingly small thickness that is not well illustrated on a
435 non-logarithmic scale. Figure 4a) shows that NED10m etc datasets produce flow depths that

are similar to that for TOPSAR 5m, while SRTM 90m, TOPSAR 30m (interp) and ASTER produce the most divergent depth calculations. The output from the TOPSAR 30m (interp) is so different from the TOPSAR 5m because the underlying topography used by the flow is in fact quite different (Fig. 4b)). Note especially that at distances below 300 m, the slope and curvature of the topography are particularly divergent. The TOPSAR 30m(dec) dataset produces flow depths quite similar to those for TOPSAR 5m (Fig. 4c)). This is because not only is the along-profile topography necessarily quite similar, but also because the cross-flow topography would be similar. (The topography along the TOPSAR 30m(dec) flowline is not a simple subset of the TOPSAR5m topography because the centerlines are in fact in different positions in mapview.) A comparable analysis holds for flows on all other DEMs that diverge sharply from TOPSAR5m, and those that do not (Fig. 4d)-h)). These flows are dominated by grain and bed interaction, which is characterized by the friction angles. Over much of the runout, these are essentially in an avalanching regime. The analysis for high-volume (high *a*) flows (Fig. 5) shows that in this case, the topography has a considerably muted effect on flow behavior as characterized by pile height. This is because high-volume flows inundate the topography to great depth and therefore the effect of variations in surface topography is muted in the inertia-dominated flow body and hence on the upper surface of the flow. In essence, the flows are in the collapsing regime throughout their runout.

3.2 Statistical analysis

Descriptive statistics were developed for a better interpretation of the datasets (Table 6). Looking at the flow depth data that resulted from running TITAN2D with different DEMs, but the same initial volume and same starting location, it is hard to interpret any significant trends in the computed flows. First, maximum inundation depth data are not normally distributed. Since the distribution of the data is unknown, and potentially complex, a

460 non-parametric Kolmogorov-Smirnov test was performed to determine if output data us-
461 ing different DEMs is different from those obtained using TOPSAR 5m output data. The
462 Kolmogorov-Smirnov test makes no assumption about the distribution of the data.

463 In Table 7, results of the Kolmogorov-Smirnov test are shown for outputs using all DEMs
464 with respect to output using the TOPSAR 5m DEM when the sample size of the (randomly
465 sampled) dataset of flow depth is equal to 50, 500, 1200 and the total size of the population,
466 respectively. It was observed that for low volume flow, 1200 sample points are sufficient
467 to differentiate the outputs from all DEMs from TOPSAR 5m. However, 50 points are
468 too few and the null hypothesis fails to be rejected. If 500 points are randomly chosen, the
469 rejection of the null hypothesis depends on the choice of DEM and also on the location of the
470 initial pile. For NED 30m, low-volume at Location 1, the test statistic (maximum absolute
471 difference) for 50 points is 0.1400 and *p*-value is 0.6779, while for the entire population of
472 flow depths the test statistic is 0.0692 and *p*-value is 7.9688×10^{-26} . So, at the 5% level of
473 significance, we fail to reject the null hypothesis in the first case, which is consistent with
474 the two datasets coming from the same distribution; in the second case we can reject the
475 null hypothesis, the datasets come from different probability distributions.

476 For the high-volume flow the results are in accordance with our expectations, due to the
477 fact that the inundation area is less affected by minor topographic features than that for the
478 low-volume flow and there is more consistency in the volume per area. 1200 sample points are
479 not enough to conclude that the datasets are not from the same distribution for all DEMs.
480 As is shown in Table 7 for high-volume flow, the null hypothesis fails to be rejected when
481 the sample size equals 1200 for NED10m. It was found that the minimum sample size such
482 that the null hypothesis rejected is around 5200 points and is significantly different for the
483 rest of the DEMs. The minimum number of sample points increases considerably for high-
484 volume compared to low-volume flows. It is different for different DEMs and is dependent

485 on the starting location of the initial volume. In this case 50 sample points always yields a
486 failure to reject the null hypothesis. Use of the entire population results in rejection under
487 all conditions.

488 4 Conclusions

489 Earlier analyses of computational models using output of different DEMs suggested that since
490 creating finer-resolution DEMs is an expensive process, one often needs to find a compromise
491 between increasing resolution and computational cost. In the present work, we have tried to
492 characterize the DEM that can be a satisfactory topographic input source on which to run
493 the geophysical flow model TITAN2D, by testing TITAN2D outputs using different DEMs.

494 The differences between the DEMs do affect the characteristics of the output flow maps,
495 as tested by classification of the grid cells into inundated or noninundated classes, and by
496 characterizing the minimum number of measurements of flow depth that can be used to
497 differentiate the DEMs. Over much of the earth's surface, usage of DEMs as input for
498 geophysical models has been limited by the relatively coarse spatial resolution of available
499 datasets [Zhang et al., 2008]. High-resolution DEMs are created by decreasing the interval
500 between sampled elevation points with more sophisticated instruments, or computationally
501 regularized spline with tension to interpolate to a finer DEM. The analysis of the output from
502 two computationally resampled 30×30 m resolution TOPSAR DEMs from a 5×5 m TOPSAR
503 DEM suggests that for geophysical flow modeling, spline interpolation has a measurable
504 negative effect on model output. In fact the quality of the splined TOPSAR 30m DEM is
505 similar to or worse than the SRTM90m DEM, using a variety of measures. This suggests
506 that the common methods of changing DEM resolution or filling of DEM holes by spline
507 interpolation do not in general produce a DEM suitable for geophysical flow modeling. Spline

508 interpolation should be used with caution.

509 The analysis suggests that flow volume is an important factor in accuracy of outputs for
510 a given DEM. This means that fine DEM resolution is critical for correct characterization of
511 low-volume flows ($\mathcal{O}(10^5)$ m^3 and lower). For high-volume flows ($\mathcal{O}(10^7)$ m^3 and greater),
512 a coarse DEM can be suitable for cost-effective model runs. Thus, the globally available
513 SRTM90m DEM should only be used with great caution for hazard and risk assessment if
514 the concern is the low-volume flow hazard.

515 An important result for researchers is the minimum number of sample points necessary
516 to determine if DEMs yield different inundation maps. Using the Kolmogorov-Smirnoff test,
517 we found that for low-volume flows, because of the fact that there are more significant
518 differences in flow depth, area covered and flow shape than for high-volume flows, DEMs are
519 differentiable given a small number of samples of inundation depth. The conclusion for both
520 low-volume and high-volume flows is that the lower the number of samples, the greater the
521 relative difference between datasets. The implication of this result for risk assessment is the
522 following. Assume we have a city for which we would like to calculate some risk factor. For
523 any DEM, the larger the city, given a sampling frequency the more likely it is that we have
524 calculated a statistic that properly characterizes the relative risk for the city as a whole.

525 It was found that SRTM90m and TOPSAR30m (interp) require a small number of sample
526 points to reject the null hypothesis, which means that there are significant topographic
527 differences between these two DEMs and TOPSAR5m. We also found that for SRTM30m,
528 ASTER and NED30m DEMs, the minimum number of sample points required to reject
529 the null hypothesis is larger than it is for SRTM90m and TOPSAR30m (interp) DEMs,
530 but is less than the number for NED10m and TOPSAR30m (decimated). Therefore, we
531 can say that we have three categories of DEMs in which the differences with respect to
532 TOPSAR5m are of the same order. In the first category we include the DEMs that are most

533 different: SRTM90m and TOPSAR30m (interp); in the second category are SRTM30m,
534 ASTER and NED30m, while in the last category we include the ones that are least different:
535 TOPSAR30m(decimated) and NED10m. The implication of this result for risk assessment is
536 as follows. Assume a specific city that occupies a certain number of grid cells. It is most likely
537 that the relative risk calculated using the TOPSAR30m (decimated) and NED10m DEMs
538 is correct, whereas it is most likely that the relative risk calculated using the SRTM90m
539 and TOPSAR30m (interp) DEMs is flawed. In general, the paper shows DEMs are not free
540 of errors and limitation, and imperfections in DEM will negatively affect the flow path and
541 direction of the flow. Those DEMs derived from the most precise instrumentation or most
542 thoroughly vetted for elevation errors (NED10m, for example) are also those that produce
543 the most reliable flow paths. Those produced with less precise instrumentation and less
544 thoroughly vetted (SRTM90m and ASTER without GCPs as used herein) will produce the
545 least reliable computational output. It is of utmost importance therefore to deal with DEMs
546 with as little noise and as few errors as possible. Since DEM quality mostly depends on
547 the source of data and method of interpolation, detailed, careful DEM analysis and pre-
548 processing is required. Different output of the model can be a result of choice of DEM
549 grid size also. SRTM90m, for example, is measurably worse at producing accurate flow
550 paths than is SRTM30m. A particular grid size will provide a specified accuracy of surface
551 representation using a minimum number of points. Again, using SRTM90m and SRTM30m
552 as an example, SRTM90m will accurately represent coarser features than will SRTM30m,
553 and will therefore adequately represent fewer channels, which are so critical to reasonable
554 calculation of flow path. Finally, lower resolution DEMs under-estimate slope and curvature,
555 which results locally in different spreading characteristics even on slopes without channels.

556 **5 Acknowledgment**

557 This research was supported by NASA grant NNX08AF75G. We would like to thank Keith
558 Dalbey for his invaluable comments and feedback.

559 **References**

- 560 M. Abrams. The Advanced Spaceborne Thermal Emission and Reflection Radiometer
561 (ASTER): Data Products for the High Spatial Resolution Imager on NASAs Terra Plat-
562 form. *International Journal of Remote Sensing*, 21(847-859), 2000.
- 563 R.A. Bailey. *Geologic Map of Long Valley Caldera, Mono-Inyo Craters Volcanic Chain and*
564 *Vicinity, Eastern California*. Department of the Interior, Reston, VA (US), 1989.
- 565 D. Barr and T. Davidson. A Kolmogorov-Smirnov Test for Censored Samples. *Technometrics*,
566 15, 1973.
- 567 P.D. Bates and A.P.J. De Roo. A simple raster-based model for flood inundation simulation.
568 *Journal of Hydrology*, 236:54–77, 2000.
- 569 T.Q. Binh and N.T. Thuy. Assessment of the influence of interpolation techniques on the
570 accuracy of digital elevation model. *VNU Journal of Science, Earth Sciences*, 24:176–183,
571 2008.
- 572 S. Biswas, S. Ahmad, M.K.I. Molla, K. Hirose, and M. Nasser. Kolomogrov-Smirnov Test
573 in Text-Dependent Automatic Speaker Identification. *Engineering Letter*, 2008.
- 574 S.M. Burkett. Geomorphic Mapping and Petrography of Mammoth Mountain, California.
575 Master’s thesis, State University of New York at Buffalo, 2007.

- 576 L. Capra, V.C. Manea, M. Manea, and G. Norini. The importance of digital elevation model
577 resolution on granular flow simulations: a test case for Colima volcano using TITAN2D
578 computational routine. *Nat Hazards*, 59:665–680, 2011.
- 579 T. Cebecauer, J. Hofierka, and Marcel Suri. Processing digital terrain models by regularized
580 spline with tension: tuning interpolation parameters for different input datasets. *Proceed-
581 ing of Open source GIS - GRASS users conference*, 2002.
- 582 I. Chanbey, A.S. Cotter, T.A. Costello, and T.S. Soerens. Effect of DEM data resolution on
583 SWAT output uncertainty. *Hydrological Processes*, 19(3):621–628, 2005.
- 584 T.A. Cochrane. *Methodologies for watershed modeling with GIS and DEMs for the
585 parametrization of the WEPP model*. PhD thesis, Purdue University, 1999.
- 586 R.G. Congalton. Thematic and Positional Accuracy Assrssment of Digital Remotely Sensed
587 Data. *Proceedings of the Seventh Annual Forest Inventory and Analysis Symposim*, 2005.
- 588 P.J.J. Desmet and G. Grover. Comment on modeling topographic potential for erosion
589 and deposition using GIS. *International Journal of Geographical Information Systems*, 11:
590 603–610, 1997.
- 591 T.G. Farr, P.A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, and M. Kobrick. The
592 Shuttle Radar Topography Mission. Technical report, Jet Propulsion Laboratory, 2007.
- 593 C.D. Farrar. Forest-killing diffuse co_2 emission at Mammoth Mountain as a sign of magmatic
594 unrest. *Nature*, 376:675, 1995.
- 595 P. Fisher and N.J. Tate. Causes and consequences of error in digital elevation models.
596 *Progress in Physical Geography*, 30:467–489, 2006.

- 597 I.V. Florinsky. Combined analysis of digital terrain models and remotely sensed data in
598 landscape investigations. *Physical Geography*, 22, 1998.
- 599 H. Fujisada, G.B. Bailey, G.G. Kelly, S. Hara, and M.J. Abrams. ASTER DEM performance.
600 *IEEE Transactions on Geoscience and Remote Sensing*, 43(12), 2005.
- 601 S. Galas, K. Dalbey, D. Kumar, A.K. Patra, and M. Sheridan. Benchmarking TITAN2D
602 mass flow model against a sand flow experiment and the 1903 Frank slide. *The 2007*
603 *International Forum on Landslide Disaster Management*, Ho & Li editors, 2007.
- 604 G.Legorreta-Paulin and M.I. Bursik. *Assessment of Landslides Susceptibility*. VDM and Co.
605 KG, Saarbrucken, Germany, 2009.
- 606 J.A. Goncalves and A.M. Oliveira. Accuracy Analysis of DEMs derived from ASTER im-
607 agery. *ISPRS XX*, 2004. Istanbul, Turkey.
- 608 D.M. Hooper, M.I. Bursik, and F.H. Webb. Application of high-resolution, interferometric
609 DEMs to geomorphic studies of fault scarps, Fish Lake Valley, Nevada-California, USA.
610 *Remote Sensing of Environment*, 84(2):255–267, 2003.
- 611 S.K. Jenson and J.O. Dominque. Extracting Topographic Structure from Digital Elevation
612 Data for Geographic Information System Analysis. *Photogrammetric Engineering and*
613 *Remote Sensing*, 54:1593–1600, 1988.
- 614 M. Kamada and R. Enkhbat. Spline Interpolation in Piecewise Constant Tension. *SAMPTA*,
615 2009.
- 616 B. Kvasov. Parallel mesh methods for tension splines. *Journal of Computational and Applied*
617 *Mathematics*, 2011.

- 618 E. Lajeunesse, A. Mangeney-Castelnau, and J.P. Vilotte. Spreading of a granular mass on
619 a horizontal plane. *Physics of Fluids*, 16:2371–2381, 2004.
- 620 T.M. Lillesand and R.W. Keifer. *Remote sensing and image interpretation*. New York, NY:
621 Wiley, 4 edition, 2000.
- 622 R.H.C Lopes, I. Reid, and P.R. Hobson. The two-dimensional Kolmogrov - Smirnov test.
623 *XI International Workshop on Advanced Computing and Analysis Techniques in Physics*
624 *Research*, 2007.
- 625 G. Lube, H.E. Huppert, R. Stephen, J. Sparks, and A. Freundt. Collapses of two-dimensional
626 granular columns. *Physical Review*, 72, 2005.
- 627 H. Mitasova and J. Hofierka. Interpolation by Regularized Spline with Tension: II. Appli-
628 cation to Terrain Modeling and Surface Geometry Analysis. *International Association for*
629 *Mathematical Geology*, 25, 1993.
- 630 H. Mitasova, J.Hofierka, M. Zlocha, and L.R. Iverson. Modeling topographic potential for
631 erosion and deposition using GIS. *International Journal of Geographical Information Sys-*
632 *tems*, 10:629–641, 1996.
- 633 P.J. Mouginis-Mark and H. Garbeil. Quality of TOPSAR topographic data for volcanology
634 studies at Kilauea Volcano, Hawaii: An assessment using airborne LiDAR data. *Remote*
635 *Sensing of Environment*, 96:149–164, 2005.
- 636 A.K. Patra, A.C. Bauer, C.C. Nichita, E.B. Pitman, M.F. Sheridan, M. Bursik, B. Rupp,
637 A. Webber, A.J. Stinton, L.M. Namikawa, and C.S. Renschler. Parallel adaptive numerical
638 simulation of dry avalanches over natural terrain. *Journal of Volcanology and Geothermal*
639 *Research*, 139:1–21, 2005.

- 640 J.N. Procter, S.J. Cronin, I.C. Fuller, M.F. M. Sheridan, V.E. Neall, and H. Keys. Lahar hazard
641 assessment using TITAN2D for an alluvial fan with rapidly changing geomorphology:
642 Whangaehu River, Mt. Ruapehu. *Geomorphology*, 116(1-2):162–174, 2010.
- 643 S.B. Savage and K. Hutter. The motion of a finite mass of granular material down a rough
644 incline. *Journal of Fluid Mechanics*, 199:177–215, 1989.
- 645 W. J. Shaffer. Evaluation of the Shuttle Radar Topography Mission digital elevation model
646 and application to the modeling of permanent deformation. Master’s thesis, State University
647 of New York at Buffalo, 2006.
- 648 N.F. Stevens, V. Manville, and D.W. Heron. The sensitivity of a volcanic flow model to digital
649 elevation model accuracy: experiments with digitised map contours and interferometric sar
650 at Ruapehu and Taranaki volvanoes, New Zealand. *Journal of Volcanology and Geothermal*
651 *Research*, 119:89–105, 2006.
- 652 T. Takahashi and H. Tsujimoto. A mechanical model for Merapi-type pyroclastic flow.
653 *Journal of Volcanology and Geothermal Research*, 98:91–115, 2000.
- 654 S.M. Wise. Effect of differing DEM creation methods on the results from a hydrological
655 model. *Computers & Geosciences*, 33(1351-1365), 2007.
- 656 D.M. Wolock and C.V. Price. Effects of digital elevation model map scale and data resolution
657 on a topography-based watershed model. *Water Resources Research*, 30(11):3041–3052,
658 1994.
- 659 L.L. Xiao, H.B. Liu, and X.G. Zhao. Impact of digital elevation model resolution on stream
660 network parameters. *2nd Conference of Environmental Science and Information Applica-*
661 *tion Technology*, 2010.

- 662 J.X. Zhang, K.T. Chang, and J.Q. Wu. Effects of DEM resolution and source on soil erosion
663 modelling: a case study using the WEPP model. *International Journal of Geographical
664 Information Science*, 22(8):925–942, 2008.
- 665 W.H. Zhang and D.R. Montgomery. Digital elevation model grid size, landscape represen-
666 tation and hydrologic simulations. *Water Resources Research*, 30(4):1019–1028, 1994.

Table 1: Description of the flow parameters

Parameter	Description
Number of Piles	This feature is used to simulate events that start at different locations
Maximum Initial Thickness	The initial pile shape is defined
Major and Minor Axes	Used to calculate the volume of the pile
Center of Initial Volume	The location of the initial pile center
Orientation	The orientation of the initial pile center, defined in degrees from X-direction to the major axis
Use of GIS Material Map	For the spatially distributed basal friction
Internal Friction Angle	The internal friction angle of the pile in degrees
Bed Friction Angle	The basal friction angle in degrees

Table 2: Initial pile parameters

	r_{min} (m)	r_{max} (m)	h_{max} (m)	<i>east</i> (m)	<i>north</i> (m)	ϕ_{bed} (deg)	ϕ_{int} (deg)
<i>Low vol</i>	70	70	20	319694.259259	4167433.7037	20	25
				321181.111111	4166645.74074		
				322058.148148	4167488.51852		
<i>High vol</i>	300	300	200	321174.259259	4166871.85185	20	25
				320639.814815	4166810.18519		
				320242.407407	4167043.14815		
				321071.481481	4166405.92593		

Table 3: Percent match for low and high volume at location 1

	TOPSAR	TOPSAR	ASTER	NED	SRTM	SRTM
	30m	30m	10m	30m	30m	90m
	(decimated)	(interpolated)				
TOPSAR						
5m	73.82%	42.10%	57.50%	79.02%	68.43%	70.75%
<i>(Low flow)</i>						60.50%
TOPSAR						
5m	92.98%	75.28%	86.98%	91.60%	90.06%	93.21%
<i>(High flow)</i>						89.36%

Table 4: Error matrix between TOPSAR 5m and SRTM 90m low and high flow

TOPSAR 5m					
		Flow	Nonflow	Total	
<i>Low flow</i>	NED 10m	Flow	10474	1430	11904
		Nonflow	1362	11986	13348
		Total	11836	13416	25252
<i>High flow</i>	SRTM 90m	Flow	8137	1387	9524
		Nonflow	3699	12029	15728
		Total	11836	13416	25252
<i>Low flow</i>	NED 10m	Flow	24107	2029	26136
		Nonflow	666	26058	26724
		Total	24773	28087	52860
<i>High flow</i>	SRTM 90m	Flow	23884	1977	25861
		Nonflow	889	26110	26999
		Total	24773	28087	52860

Table 5: Producer's accuracy and user's accuracy for low and high flow at location 1

		Prod acc1 ^a	Prod acc2 ^b	User acc1 ^c	User acc2 ^d	Overall acc ^e
<i>Low flow</i>	TOPSAR5m					
	vs	88.49 %	89.34 %	87.99 %	89.80 %	88.94 %
	NED10m					
	TOPSAR5m					
<i>High flow</i>	vs	68.75 %	89.66 %	85.44 %	76.48 %	79.86 %
	SRTM90m					
	TOPSAR5m					
	vs	97.31 %	92.78 %	92.24 %	97.51 %	94.90 %
<i>High flow</i>	NED10m					
	TOPSAR5m					
	vs	96.41 %	92.96 %	92.36 %	96.71 %	94.58 %
	SRTM90m					

^aProducer accuracy for the flow area

^bProducer accuracy for the nonflow area

^cUser accuracy for the flow area

^dUser accuracy for the nonflow area

^eOverall accuracy

Table 6: Descriptive statistics for high-volume flow from Location 1

		Mean	St.dev.	Mean	St.dev.
		200	200	all	all
		points	points	population	population
<i>Low flow</i>	Topsar 5m	0.5822	1.4637	0.6490	1.7586
	NED 10m	0.520	0.7461	0.7667	1.8301
	SRTM 90m	0.7367	1.7791	0.8228	1.9509
<i>High flow</i>	Topsar 5m	14.2688	29.6266	13.2949	25.7137
	NED 10m	11.8160	21.930700	12.8503	25.3877
	SRTM 90m	16.5245	33.0382	12.8347	25.0674

Table 7: Kolmogorov-Smirnov parameters for Topsar 5m, NED 10m and SRTM 90m datasets, low and high flow, location1

		K-S parameter	50 points	500 points	1200 points	All points
TOPSAR	30m	h	1	1	1	1
	(interp)	p	0.0021	1.5032x10 ⁻²⁰	1.9315x10 ⁻⁴²	0
		k	0.3600	0.3020	0.2825	0.2772
TOPSAR	30m	h	0	0	0	1
	(decimated)	p	0.6779	0.9762	0.9870	1.1424x10 ⁻⁹
		k	0.1400	0.0300	0.0183	0.0416
<i>Low flow</i>	NED 10m	h	0	0	1	1
		p	0.6779	0.0774	0.0408	7.9688x10 ⁻²⁶
		k	0.1400	0.0800	0.0567	0.0692
NED 30m		h	0	0	1	1
		p	0.5077	0.1685	1.9785x10 ⁻⁵	4.774x10 ⁻⁷¹
		k	0.1600	0.0700	0.0975	0.1207
SRTM 30m		h	0	0	1	1
		p	0.5077	0.1685	1.9785x10 ⁻⁵	4.774x10 ⁻⁷¹
		k	0.1600	0.0700	0.0975	0.1207
SRTM 90m		h	0	1	1	1
		p	0.0951	2.9860x10 ⁻⁹	1.0572x10 ⁻¹⁶	7.1187x10 ⁻⁸⁷
		k	0.2400	0.200	0.1758	0.1355
Aster		h	0	0	1	1
		p	0.6779	0.0774	0.0408	7.9688x10 ⁻²⁶

Continued on next page

Table 7 – continued from previous page

		K-S	50	500	1200	All
		parameter	points	points	points	points
		k	0.1400	0.0800	0.0567	0.0692
	TOPSAR	h	0	0	0	1
	30m	p	0.3584	0.3599	0.1878	1.0711×10^{-11}
	(interp)	k	0.1800	0.0580	0.0442	0.0328
	TOPSAR	h	0	0	0	1
	30m	p	0.8409	0.4031	0.0708	1.3836×10^{-9}
	(decimated)	k	0.1200	0.0560	0.0525	0.0287
		h	0	0	0	1
<i>High</i>	NED 10m	p	0.8409	0.1907	0.2236	4.3098×10^{-15}
<i>flow</i>		k	0.1200	0.0680	0.0425	0.0364
		h	0	0	0	1
	NED 30m	p	0.3584	0.9572	0.8962	2.5004×10^{-22}
		k	0.1800	0.0320	0.0233	0.0442
		h	0	0	0	1
	SRTM 30m	p	0.9541	0.4493	0.3615	6.3138×10^{-15}
		k	0.1000	0.0540	0.0375	0.0363
		h	0	1	1	1
	SRTM 90m	p	0.8409	0.0027	1.0881×10^{-4}	2.7947×10^{-31}
		k	0.1200	0.1140	0.0900	0.0528
		h	0	0	0	1
	Aster	p	0.0951	0.2491	0.9792	5.0967×10^{-24}

Continued on next page

Table 7 – continued from previous page

K-S parameter	50 points	500 points	1200 points	All points
k	0.2400	0.0640	0.0192	0.0457

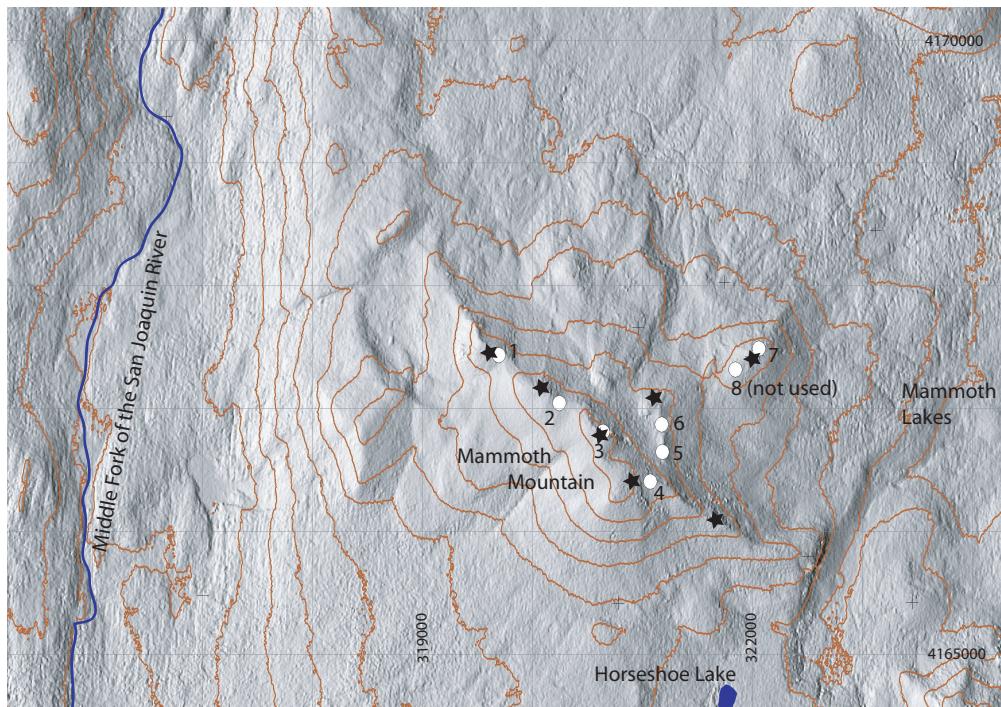
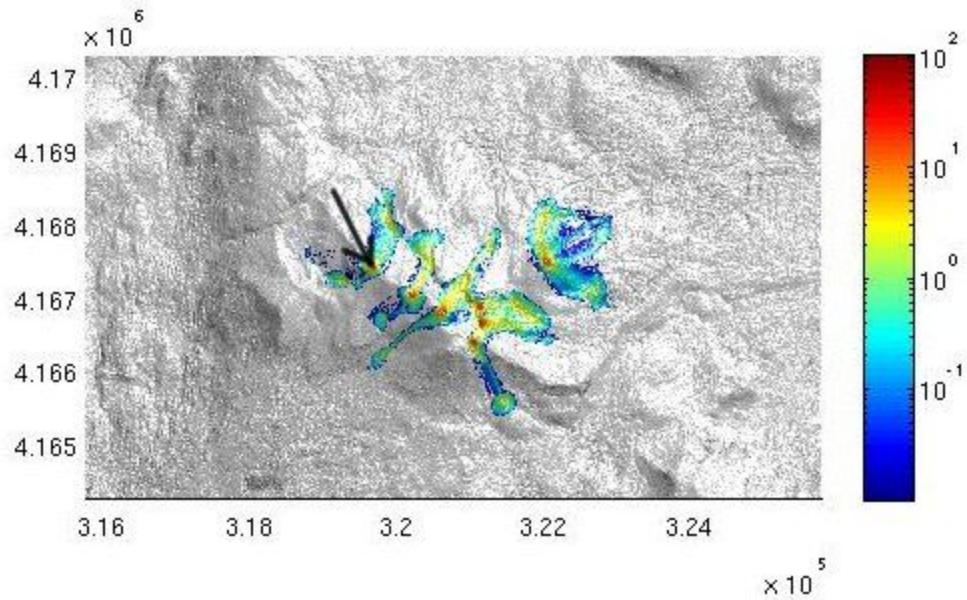
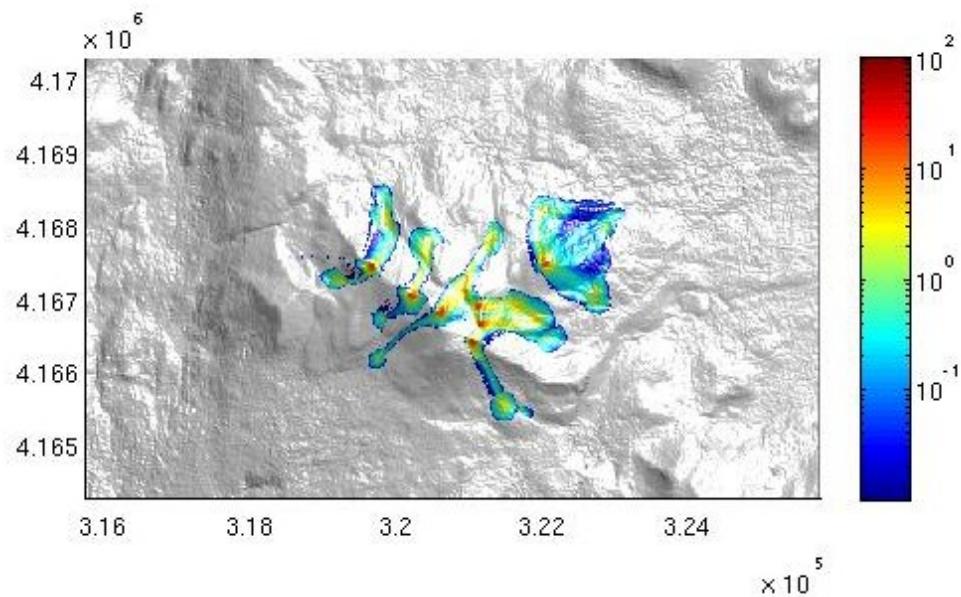


Figure 1: UTM grid for WGS84 reference system, zone 11. Contour interval is 100 m. Stars are locations of vent according to Bailey (1989); circles are locations of our starting piles.

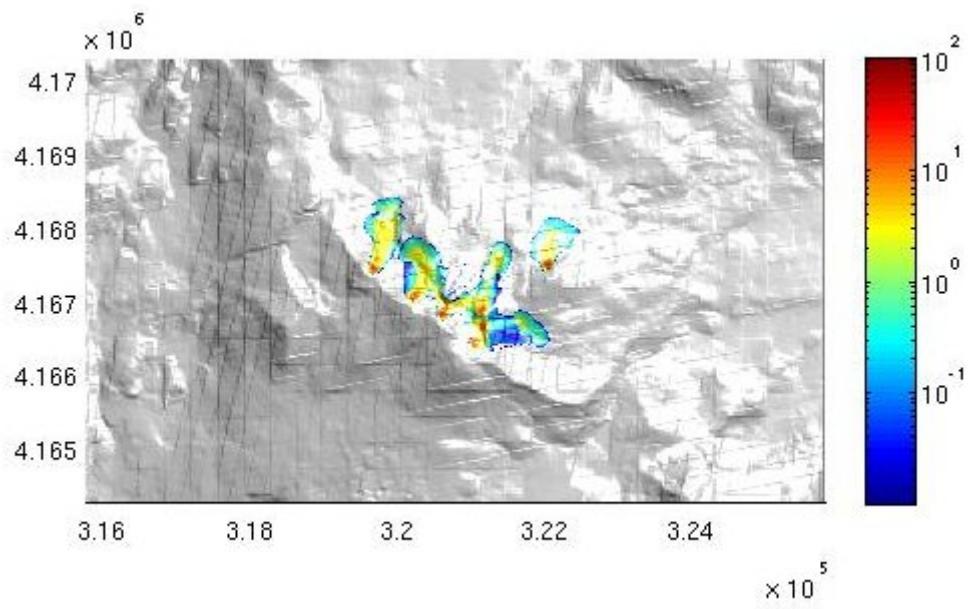


(a)

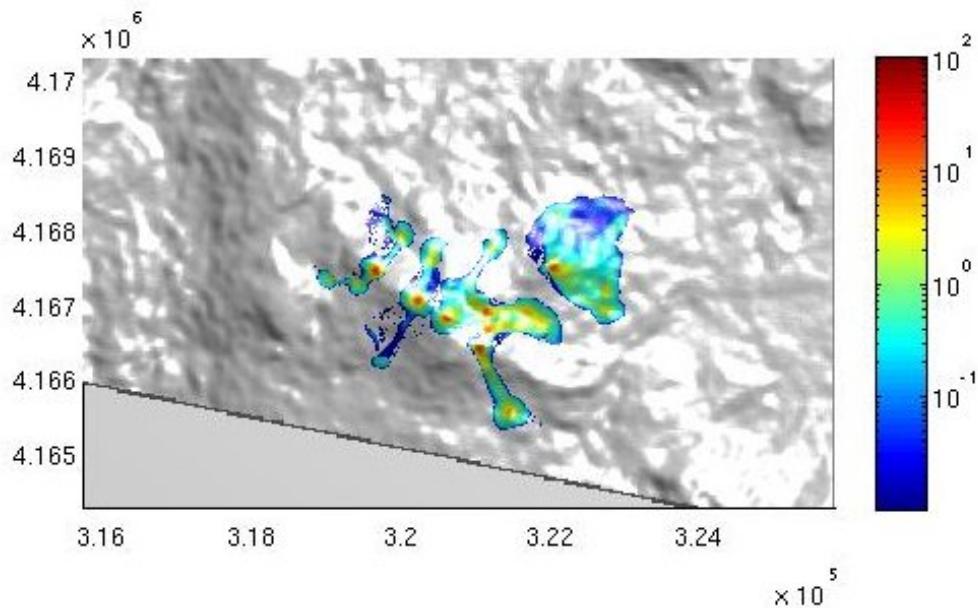


(b)

Figure 2: Flow depth for low-volume in UTM easting, northing coordinates (a) TOPSAR5m, the arrow indicates Location 1 (b) TOPSAR30m (decimated)

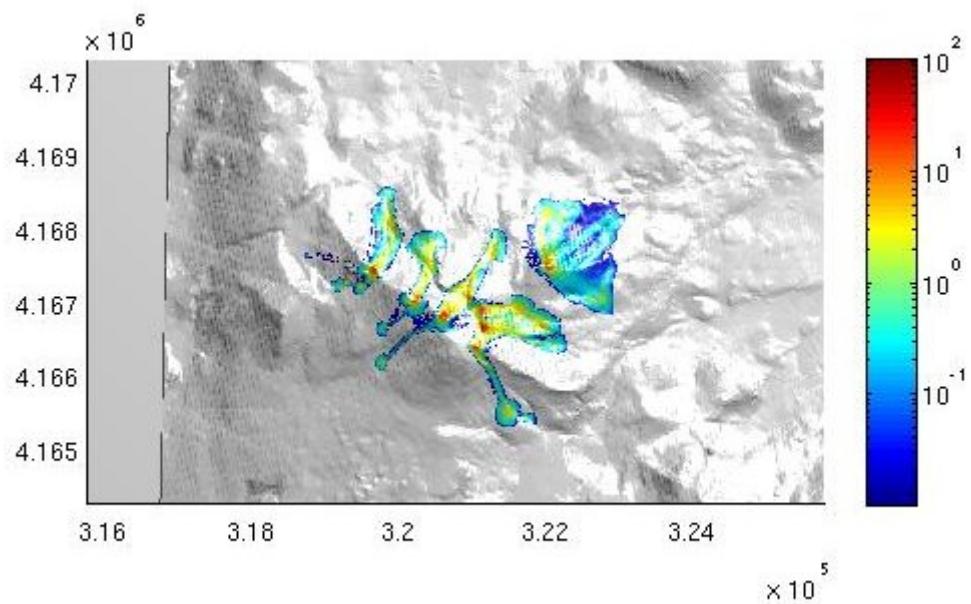


(c)

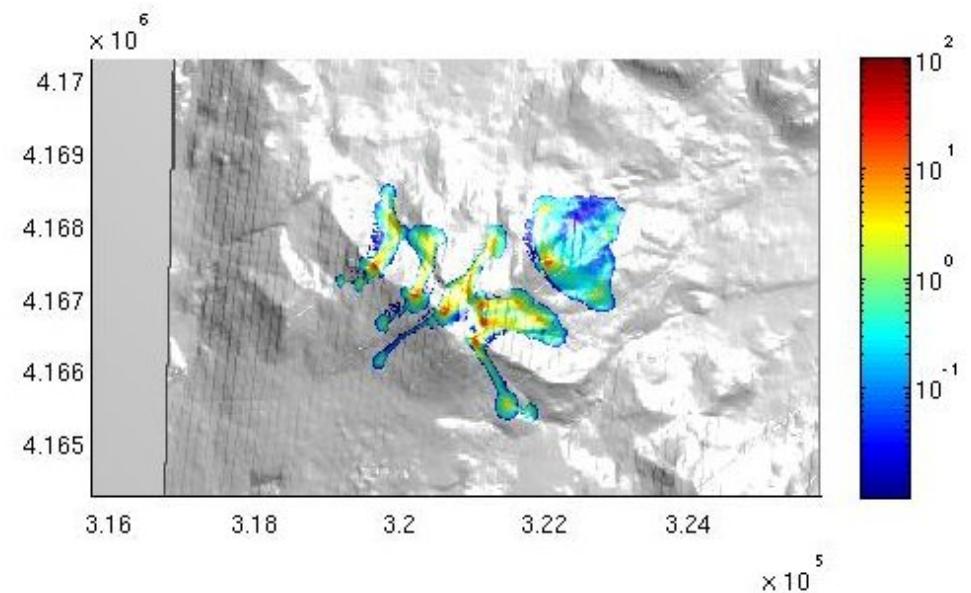


(d)

Figure 2: Flow depth for low-volume in UTM easting, northing coordinates (c) TOPSAR30m (spline interpolation) (d) ASTER

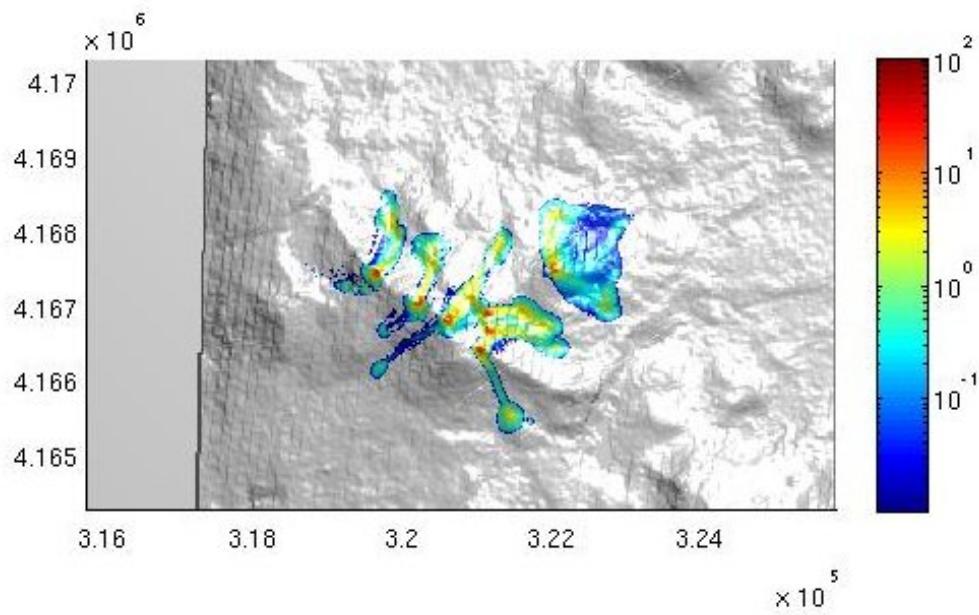


(e)

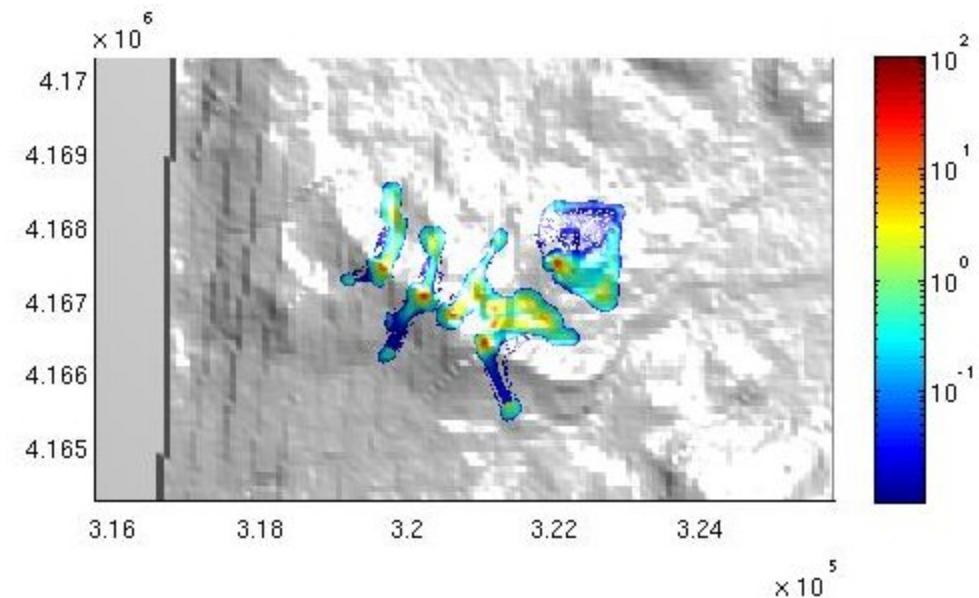


(f)

Figure 2: Flow depth for low-volume in UTM easting, northing coordinates (e) NED10m (f) NED30m

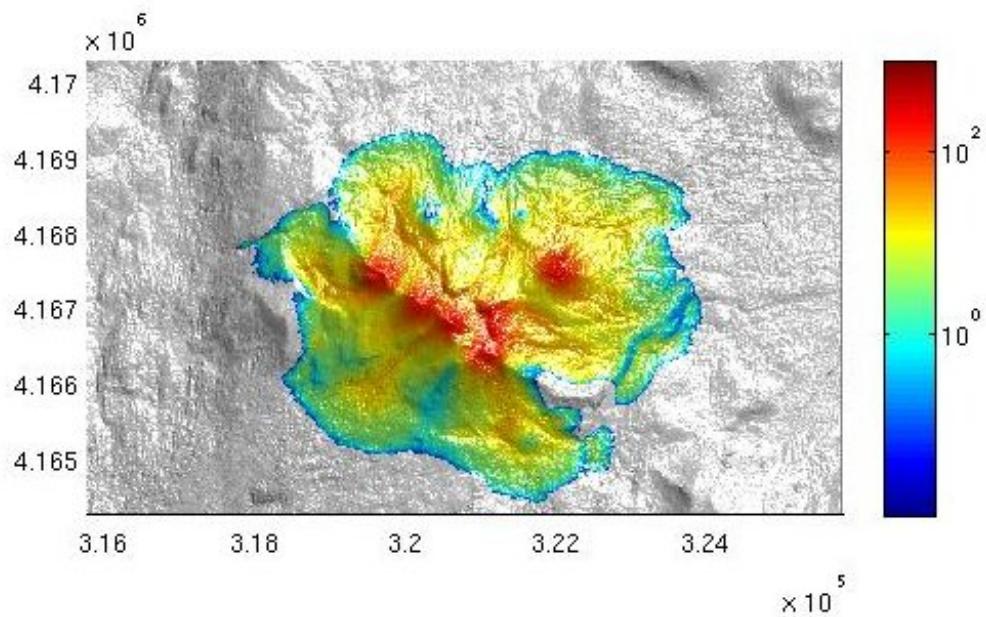


(g)

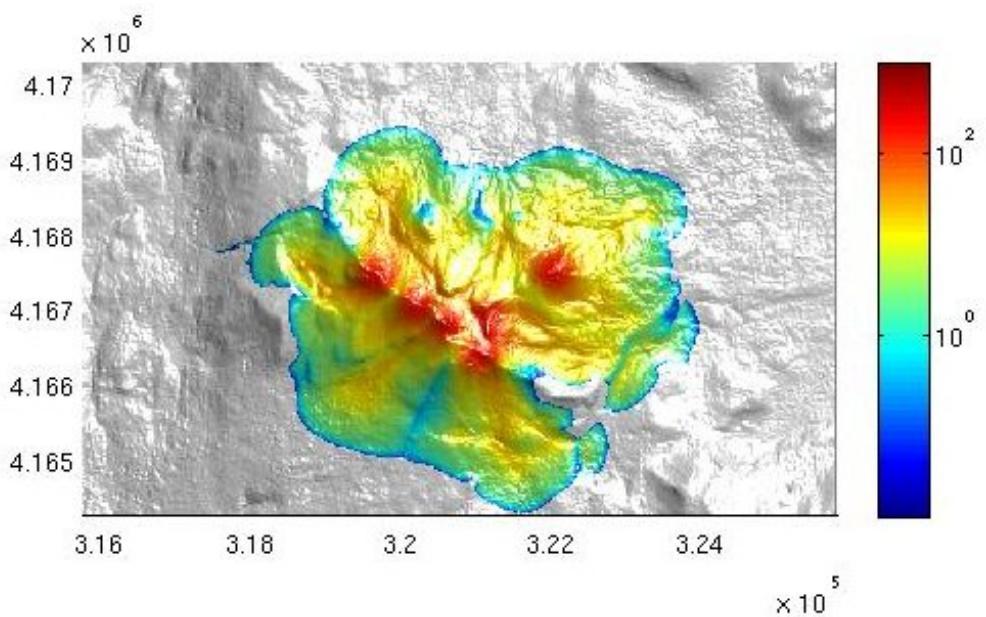


(h)

Figure 2: Flow depth for low-volume in UTM easting, northing coordinates (g) SRTM30m
(h) SRTM90m

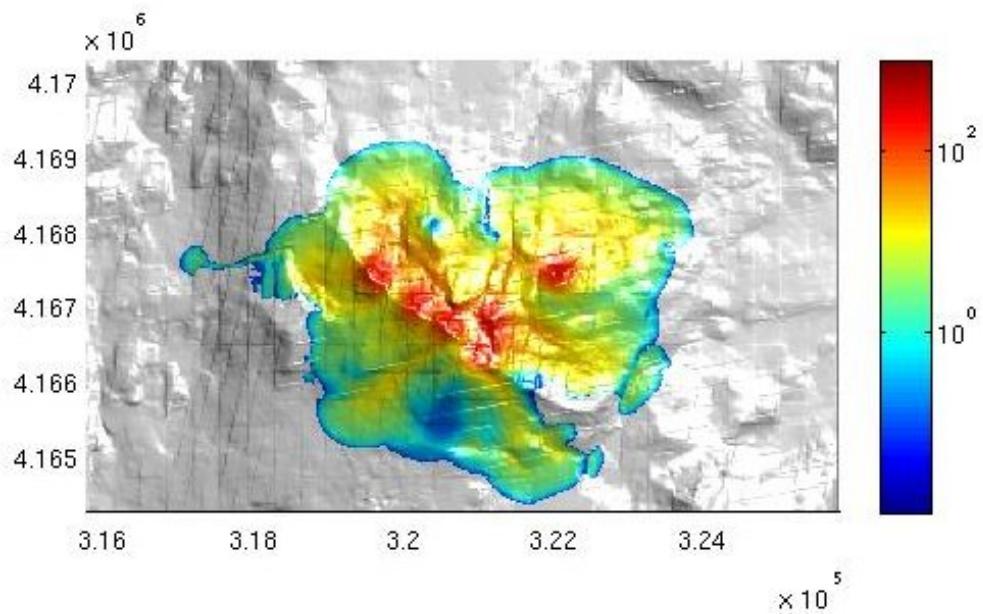


(a)

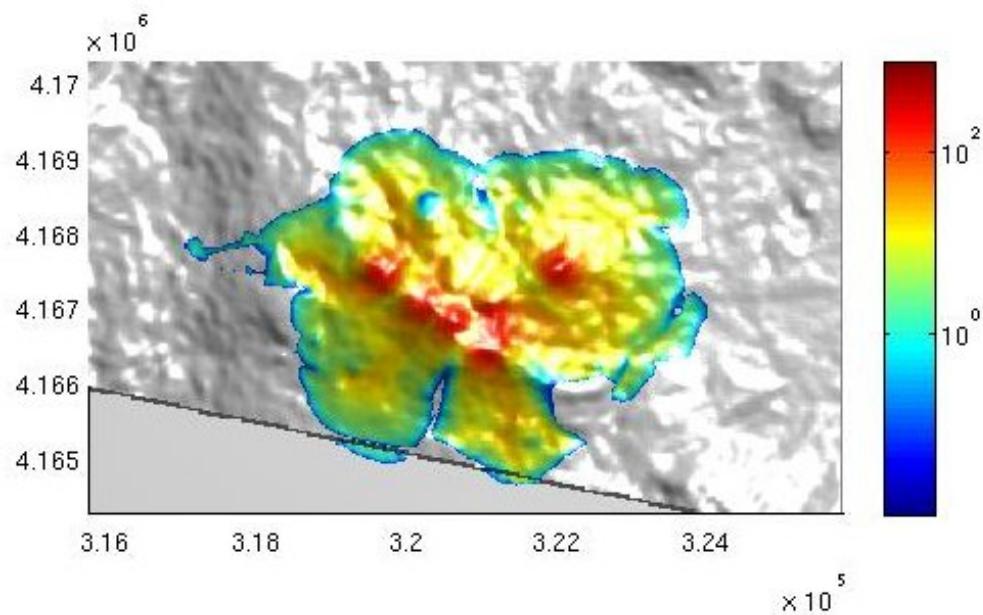


(b)

Figure 3: Flow depth for high-volume in UTM easting, northing coordinates (a) TOPSAR5m
(b) TOPSAR30m (decimated)

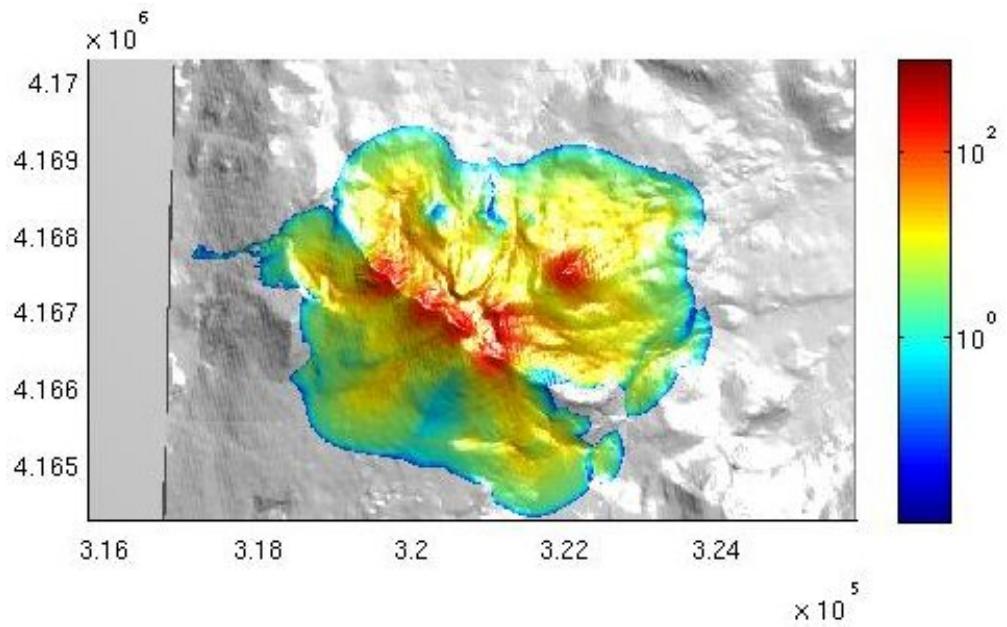


(c)

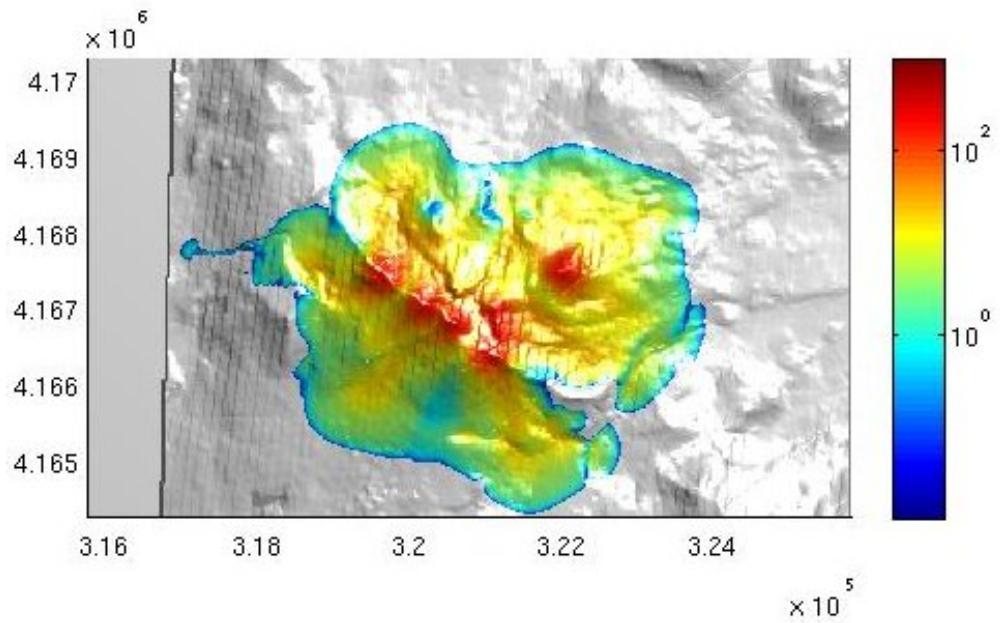


(d)

Figure 3: Flow depth for high-volume in UTM easting, northing coordinates (c) TOP-SAR30m (spline interpolation) (d) ASTER

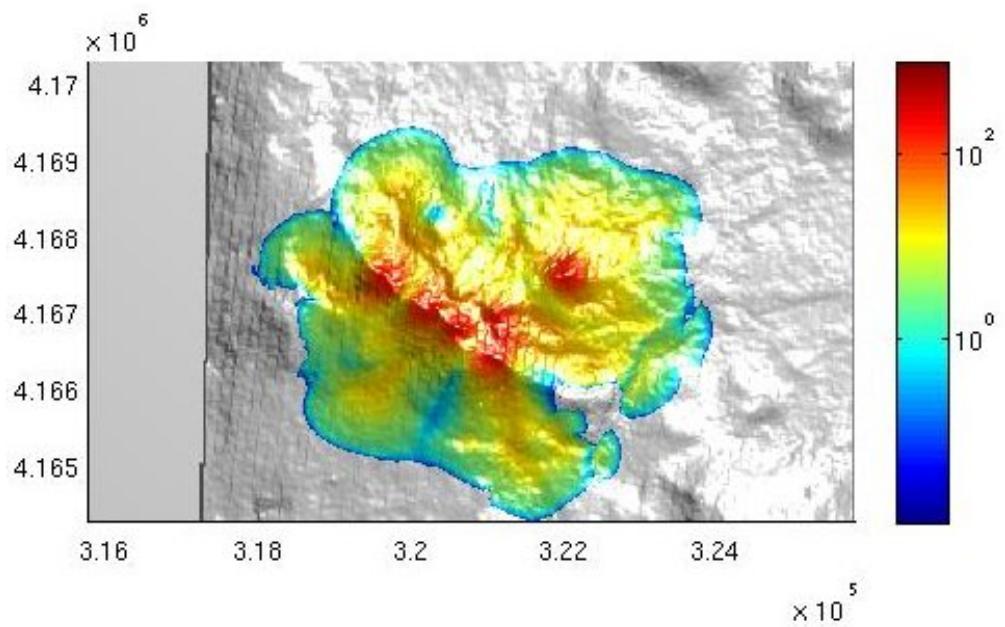


(e)

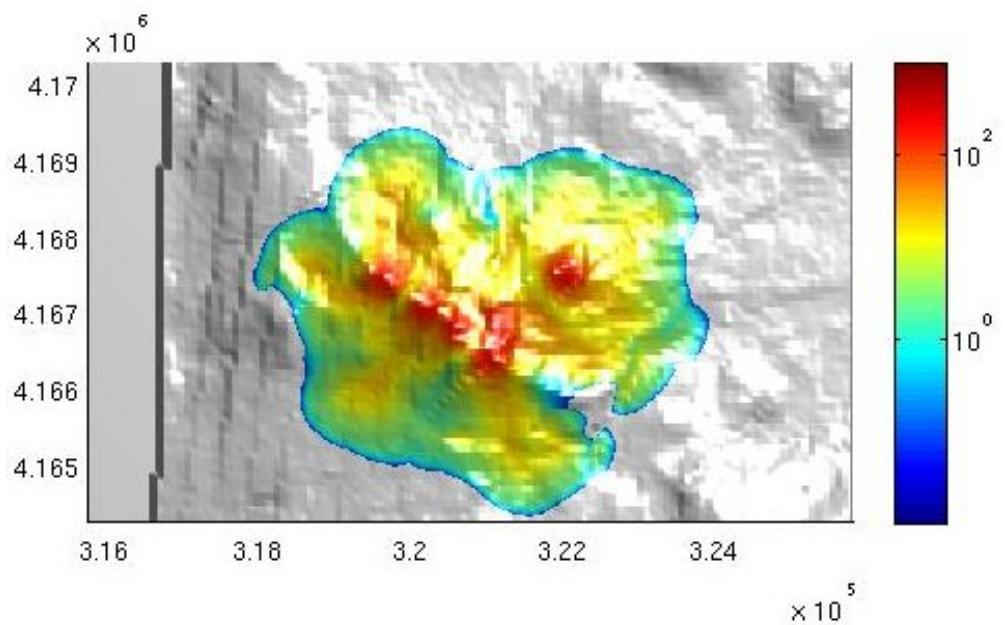


(f)

Figure 3: Flow depth for high-volume in UTM easting, northing coordinates (e) NED10m
(f) NED30m



(g)



(h)

Figure 3: Flow depth for high-volume in UTM easting, northing coordinates (g) SRTM30m
(h) SRTM90m

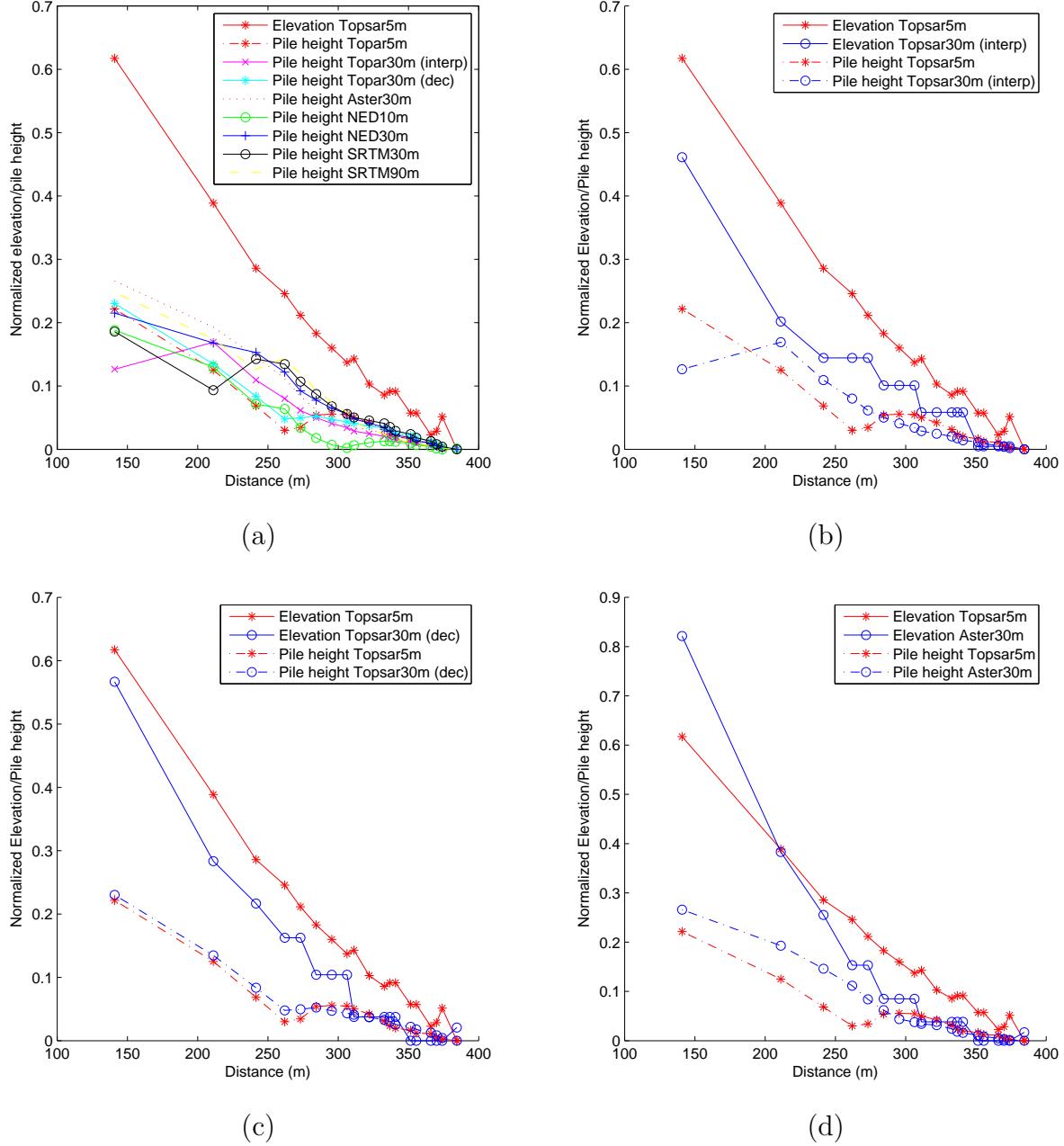


Figure 4: Normalized elevation and maximum pile height over time at the center line of the flow versus the distance – low flow volume a) Topsar 5m normalized elevation vs normalized pile height of all output flow b) Topsar 5m vs Topsar 30m(interp) c) Topsar 5m vs Topsar 30m(dec) d) Topsar 5m vs Aster 30m

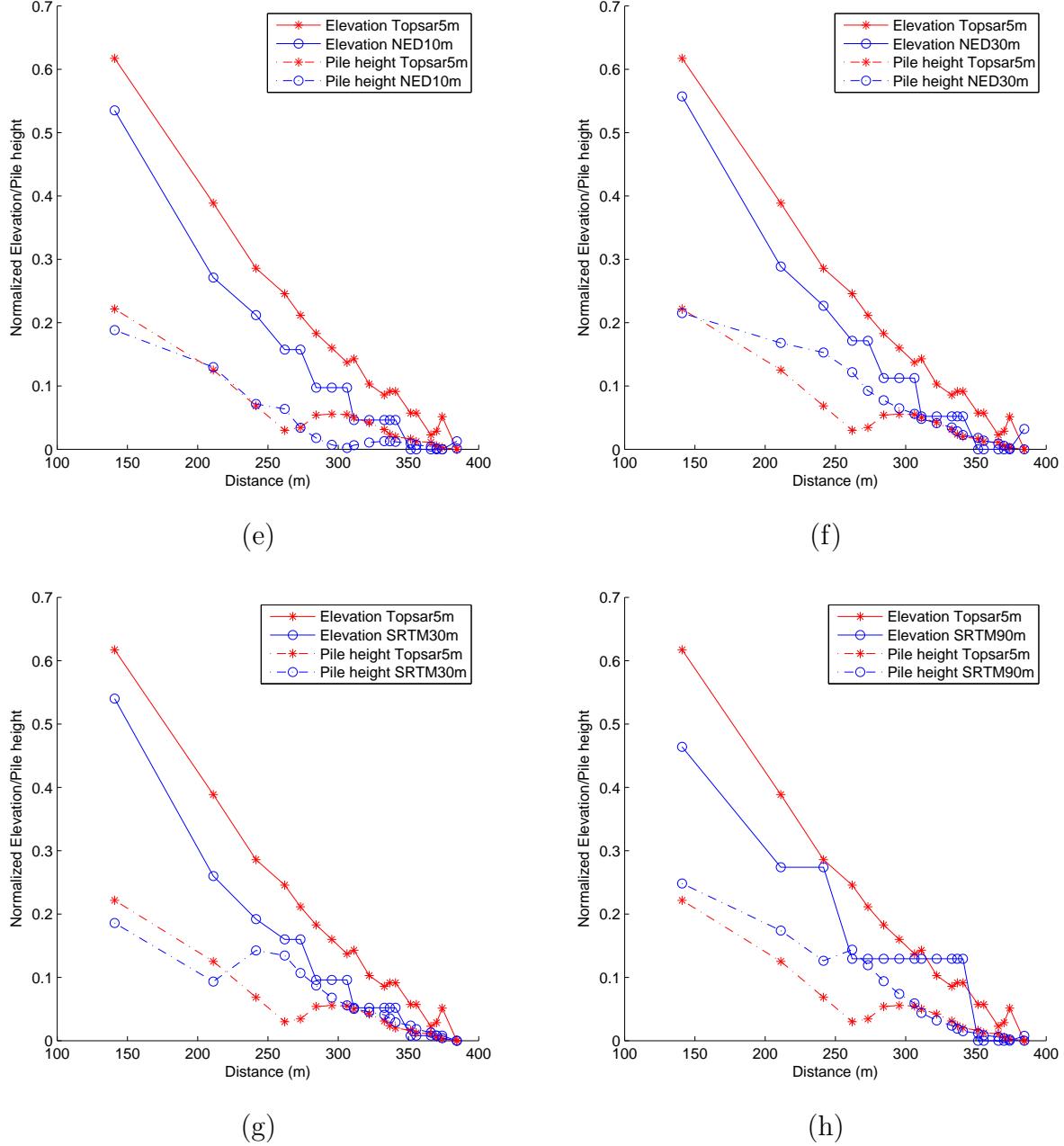


Figure 4: Normalized elevation and maximum pile height over time at the center line of the flow versus the distance – low flow volume a) Topsar 5m vs NED 10m b) Topsar 5m vs NED 30m c) Topsar 5m vs SRTM 30m d) Topsar 5m vs SRTM 90m

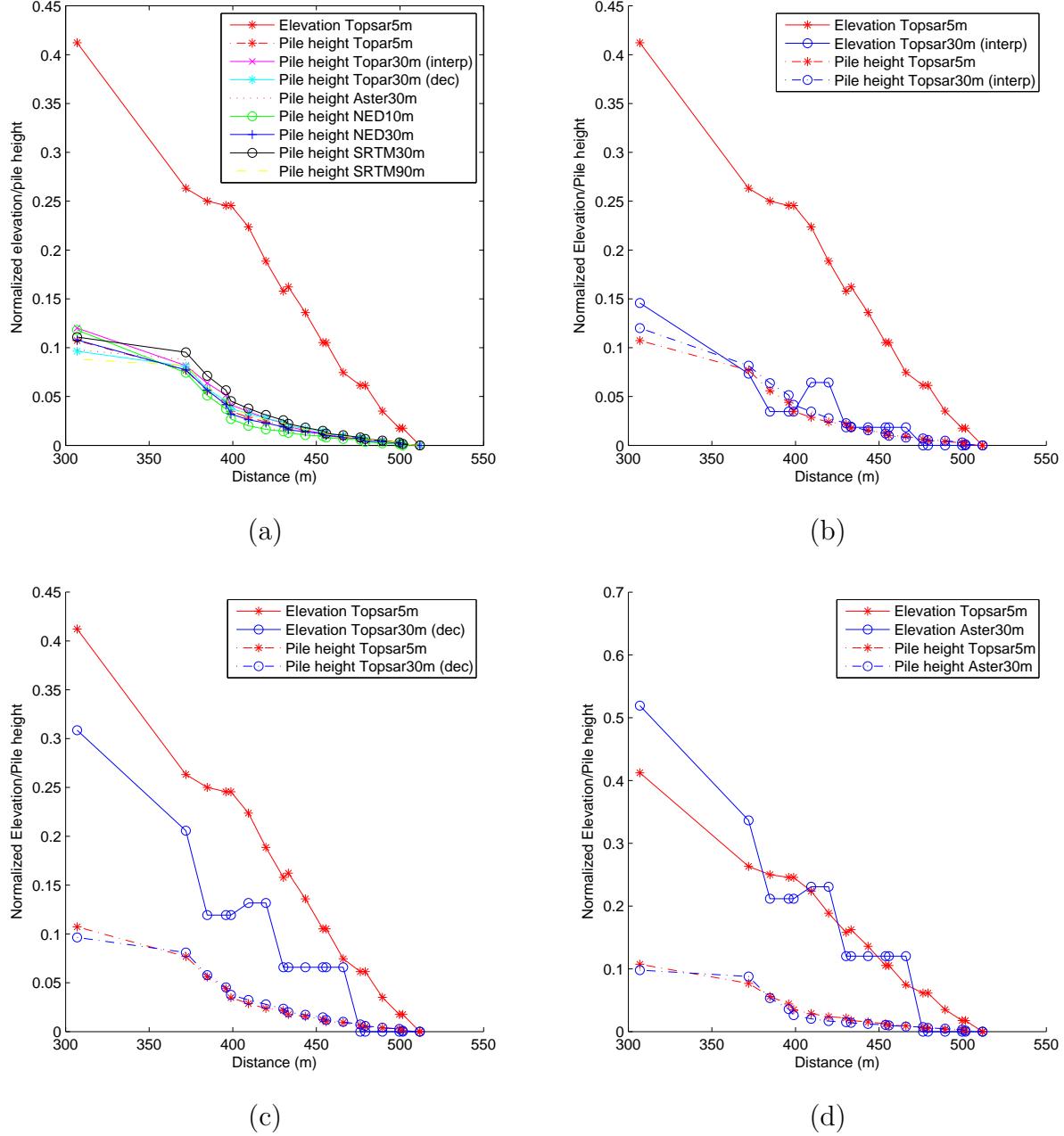


Figure 5: Normalized elevation and maximum pile height over time at the center line of the flow versus the distance – high flow volume a) Topsar 5m normalized elevation vs normalized pile height of all output flow b) Topsar 5m vs Topsar 30m(interp) c) Topsar 5m vs Topsar 30m(dec) d) Topsar 5m vs Aster 30m

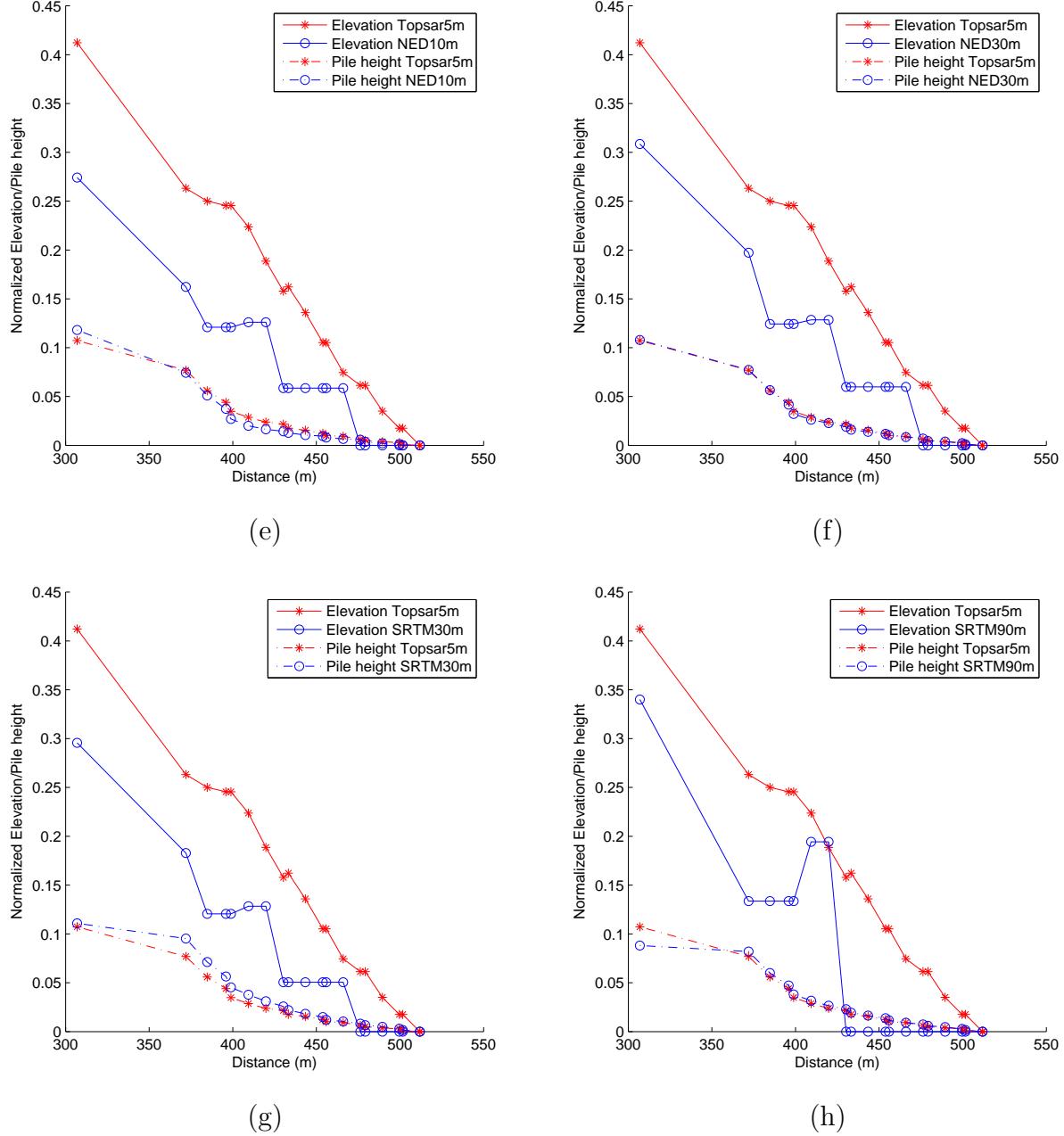


Figure 5: Normalized elevation and maximum pile height over time at the center line of the flow versus the distance – high flow volume a) Topsar 5m vs NED 10m b) Topsar 5m vs NED 30m c) Topsar 5m vs SRTM 30m d) Topsar 5m vs SRTM 90m