Characterizing the effect of Parametric Uncertainty on Flow Models Using Surrogates, Localized Covariances, and Other Related "Crimes" When Constructing Hazard Maps

Geophysical Mass Flow Group,
University at Buffalo, SUNY

K. Dalbey,
Sandia National Laboratories

MOSES-III, Petropolis, Brazil, August 8-12, 2011
One day, while we were at work in the crater, a huge slice of the precipitous wall of rock that had been bared by the explosion fell suddenly and crashed with a tremendous uproar down the steep incline beneath. This slab fell from a place about 300 metres high. The great masses of earth and rocks were shattered as they fell, and broken up into pieces, ever growing smaller as they descended. The behaviour of this pulverized mass resembled the rush of a headlong torrent. Although boulders measuring 10 metres or more in diameter were mixed up with finer matter, as a whole the movement approximated to that of a fluid. No words can describe the fierceness and force of that impetuous downpour—its mad surgings this way and that, and the bold leaps with which it would now and then bound over low ridges that hindered its progress, and shoot onward down the neighbouring depression.

Sekiya and Kikuchi (1889)
Over the last 10 years we have created a calibrated, validated parallel adaptive tool TITAN2D to simulate such flows on complex natural terrains -- many computational challenges addressed! http://vhub.org

Pitman et al 2004, Patra et al 2005a,b, 2006, ... 2009
Thorthormi and Raphstreng glacial lakes have significant altitudinal differences from the lakes themselves, with a moderate gradient of about 10% (Table 1, cases 2 and 3). On the other hand, a significant altitudinal difference with a quite large gradient is found between the adjacent two lakes, Thorthormi and Raphstreng (Table 1, case 4). Although the possibility of a GLOF through the moraine between the two lakes has been pointed out (Richardson and Reynolds, 2000), the altitudinal difference has not been measured precisely. Because the altitudinal differences and gradients before the GLOF of 1994 are, respectively, 30–34 m and 5–6%, all cases other than Lugge glacial lake show some potential for GLOF, which should not be ignored.

Volume estimation in case of GLOFs provides significant information for planning risk mitigation. Bathymetric profiles were measured for Raphstreng glacial lake (Bhargava, 1995) and Lugge glacial lake (Yamada and others, 2004). However, it is practically impossible to measure the bathymetric profiles of ponds on Thorthormi glacier, because many ponds are now expanding and aggregating on the glacier. Table 1 shows that the glacial lakes of Lunana have high potential for a GLOF with a volume similar to that of 1994.

CONCLUSIONS

The potential for GLOFs in the Lunana region was assessed using DEMs generated from ground survey carrier-phase differential GPS, ASTER and SRTM. We evaluated the relative accuracy of elevations indicated by the ASTER DEM and SRTM DEM, by comparing with the ground-survey data. Topographical classification allows us to partition the total error in both DEMs into those in the terrain class. This classification enables better evaluation of altitudes of lakefronts, glacier surfaces and riverbeds, although it is less useful for moraine ridges and hill slopes because the terrain slope is significantly correlated with the topography. Using satellite images and the DEMs, we have re-evaluated the volumes and examined causes of the 1994 GLOF. In addition, we point out the sections where future GLOFs could occur, showing altitudinal differences and gradients around the glacial lakes.

One of the GLOF-triggering events is considered to be the melting of ice inside the moraine, which is damming water. Photogrammetric analysis of aerial photographs, which showed high accuracy, helped monitor the possible outburst trigger in the Swiss Alps (Haeberli and others, 2001). However, our study reveals that the remote-sensing DEMs from space are not applicable for monitoring altitudinal changes in a moraine ridge because of its inferior altitudinal relative accuracy. Monitoring at the site, therefore, is still crucial in addressing GLOF problems, even in this age of remote-sensing technology. GLOFs have been a serious problem in Himalayan countries. It is our hope that this study might not only present some sort of scientific advance, but also contribute to the daily lives of the local people who even now face the ongoing risk of GLOFs.
What do we need to know?

✦ Q1: Given a location $x$ and time $T$ -- what is the hazard of a catastrophic event? e.g. $P(\text{flow} > 1 \text{m in } T) \sim 0.000001$?

✦ Q2: Given an area what is the hazard of an event in the next $T$ time period of all locations?

✦ Q3: Given locations $A$ and $B$ should we evacuate from $A$ to $B$ at cost $\$\$\$\$ \$ ?

The Objective: Hazard Maps

a) Map of the extent of the phase III pyroclastic flows from the 1913 eruption of Colima (Saucedo et al 2005), b) Traditional hazard map based on field study only (Navarro & Cortez 2003), c) Hazard map based on deterministic calculations using the FLOW3D model (Saucedo et al 2005), d) Probability of flow exceeding 1 m given an event that generates a $10^7$ to $10^8$ [m$^3$] of flow volume. (Dalbey et. al 2008)
What do we have?

- Models of the Physics of individual flows (PDE based)
- Data on past events -- detailed and precise for some aspects, sparse and poor for most aspects; Use data to fit distributions to input data. e.g. Flow volume -- generalized Pareto ...(see Bayarri et. al. 2009 for details); Topography -- two parameter Gaussian error model of topography map sampled to generate ensemble of possible topography maps ...
- Expert belief and intuition
- Methodology for quantifying uncertainty
- High end Computing and Data services
Hazard Map Construction

Historical flow data and expert belief converted into recurrence probability of largest events e.g. Bayarri et al. 2009

Probability of flow exceeding 1m for initial volume ranging from 5000 to $10^8$ m$^3$ and basal friction from 28 to 35 deg at Colima and Pico de Orizaba
a few difficulties!

- Complex unpredictable physics
  - Johnson ‘70, Savage -Hutter ‘89, Iverson ‘97, Pitman-Le ‘05 … complex physics is still not perfectly represented

\[
\begin{align*}
U_i + F(U)_x + G(U)_y &= S(U) \\
U &= (h, hv_x, hv_y), F(U) = (h v_x, h v_x^2 + 0.5 kg h^2, hv_y v_y)
\end{align*}
\]

\[
S_x = g_x h - hk \text{sgn}(\frac{\partial v_x}{\partial y}) \frac{\partial}{\partial y} (g_x h) \sin \phi_{\text{int}} - \frac{v}{|v|} \left[ g_x h (1 + \frac{v}{rg_z}) \right] \tan \phi_{\text{bed}}
\]

h: flow depth; hv: depth averaged momentum; g: gravity; \(\phi\): friction

- Uncertain Inputs
  - \(\phi_{\text{bed}}\)
  - \(\phi_{\text{int}}\)
  - Initial location
  - Initial volume
  - Initial velocity
  - Terrain elevation

- Ensemble computations needed for hazard map constructions are EXPENSIVE!
- Single calculation -- 20 minute on 64 proc => Monte Carlo type computation needs \(10^6\) calculations in days!!

We need the answer by tomorrow? “The manana map.” -- Sheridan

Tuesday, August 9, 2011
Approach

Approach 0: Using standard Monte Carlo type sampling probability of hazard computation at all points on hazard area requires $O(10^6)$ simulations, each of which requires a range of 5 mts to $O(1\text{hour})$ and analytics using subsets from $O(1\text{PB})$ distributed data!

Approach 1: Given a simulator with “well defined input data uncertainties” -- use well chosen ensemble (Latin Hypercube, Quadrature driven …) to propagate uncertainty and use simple expectation computations to make hazard map. [Dalbey et. al. 2008, J. Geophys. Res.]

Approach 2: Given a location and sparse data create estimates of predictions and associated uncertainty using Bayesian methodology by using simulator to create emulator and use emulator in appropriate statistical methodology.[Bayarri et. al. 2009 Technometrics, Dalbey et. al. in rev.]
Approach 2

Stage 1: Evaluate an ensemble of several hundred to several thousand multiprocessor landslide simulations, dynamically assigning simulations to processors as they become available to continually use the entire pool of processors efficiently.

Stage 2: Create a multi-level hierarchical emulator (a statistical model) from the output of the ensemble of simulations. Its hierarchical nature allows the emulator’s components to be constructed (and evaluated) concurrently. Emulator acts as a fast surrogate of the simulator.

Stage 3: Use the emulator through importance sampled Monte Carlo to compute a map of the probability that a hazard criterion will be met at hundreds of thousands (or more) of locations.
Statistical Surrogate Model

Gaussian Process Model, Bayes Linear Models (BLM) (Goldstein, ...’05,’07)
• an approximate deterministic mean function, (a least squares fit) plus a Gaussian error model

BLM updates are a projection of the prior beliefs onto the span of the data
under the standard variance inner product!

\[ s(x) = \mu(x) + \epsilon(x) \]
\[ \text{Cov}(\hat{\epsilon}(x), \hat{\epsilon}(y_i)) = \sigma^2 r_i(x) \]
\[ r_i(x) = r(x - y_i) = \exp \left( -\sum_{i}^{N_{in}} \theta_{i_{in}} (x_{i_{in}} - y_{i_{in}})^{p} \right) \]
\[ R_{i,j} = r_i(y_j) = r_j(y_i) \]

\[ E(s_{BLM}(x) | s_y) = g(x)^T \beta + r(x)^T R^{-1} \epsilon \]
\[ \text{Var}(s_{BLM}(x) | s_y) = \sigma^2 \left( 1 - r(x)^T R^{-1} r(x) \right) \]

“s” represents the simulator output, “x” is an arbitrary input, “g” are the least squares basis function, \( \beta \) are their coefficients, a \( \epsilon(x) \) is “Gaussian” model of the error, \( p=1 \) (absolute val.), 2. Good values for \( \sigma \) and \( \theta \) must be inferred -- lots of heuristics!
Statistical Model

Difficulties

• “Hazard Map Emulator” must represent both high dimensional input and output -- need full flow field emulation

• Conti & O’Hagan ’07 suggest treating the spatial indexing of the field variable output as additional input variables -- total input $O(10^6)$ field locations plus up to 8 input parameters, friction, initial volume, starting location

• Inverting the covariance $R^{-1}$ repeatedly is the key computation whose costs need to be managed

• $R$ may be almost singular and hard to invert

• Equations show inherent sequential nature of emulator.
Strategies for $R^{-1}$

- Large scale parallel linear algebra -- see for e.g. Bekas, Curioni Fedulova ’10
- Use a properly localized covariance function (Wood’95, Gneiting’97,’99,...)
  - Truncated power law function -- “virtually identical to the Gaussian” -- we use “this” and also examine $p=1$ which meets the Wood criteria for truncation
  - Replace single global emulator by ensemble of $m$ emulators of size $n \Rightarrow$ cost reduces from $O(N^3)$ to $O(mn^3)$ and introduce concurrency!
  - Wood’95 -- sufficient conditions to allow truncation of valid covariance function defined on $(-\infty, \infty)$ to $[-K,K]$, $K>0$; proved Gaussian form ($p=2$) cannot be truncated!
- Need to maintain positivity of Fourier coefficients of chosen function
Strategies for $R^{-1}$

- Similar ideas in Localized Ensemble Kalman Filter Gasparri & Kohn'99
- Can use Moore-Penrose inverse for local emulator construction.
- Methodology is closely related to classical domain decomposition methods -- additive Schwartz with no coarse grid -- current work
Ensemble Emulator

• Hierarchical emulator is an ensemble of smaller emulators each covering a portion of the uncertain input space -- introducing concurrency and combined using weighted averaging.

• Method is similar to treed Gaussian process of Gramacy, Lee, Macready ’04, windowed or local neighborhood Kriging methods ...

• Tessellate data to be used for surrogate construction

• Use “1 or 2-hop” neighborhood to isolate moving windows on which to construct a local emulator

• Assemble local emulators into global emulator using a barycentric weighting (partition of unity)
Localized Model

Upper left subplot: Delaunay tessellation of “simulation input space.” Remaining subplots: East-North Delaunay triangulation of simulations A, B, and C output data. Gray rectangles indicate regions with flow, colored triangles cover flow. Black asterisk indicates a resample point.
Ensemble Emulator

Upper left subplot: color indicates partition of unity weight given to mini-emulators A, B, and C during assembly. Remaining subplots: shading indicates weight given to mini-emulators A, B, and C, colored lines indicate simulations used to construct mini-emulators.
Statistical Model

A sample two-level piecewise hierarchical emulator approximating the response of the simulator for different inputs. Starting center of mass is normally distributed about summit of Colima volcano with standard deviation of 150m in East and North directions. Output is North coordinate of centroid, 600 seconds after initiation.
Schwarz Domain Decomposition Theory?

$R \approx \sum_{i=1}^{m} R_i$

Ri are overlapping additive decompositions?

Need coarse grid? Multilevel methods?

$R \approx R_0 + \sum_{i=1}^{m} R_i$
Hazard Map Emulator

Construction

• 3 level construction

• Micro Emulator -- in physical space -- heuristically downsamped -- create BLM of 2-hop neighborhood of specified node \(O(2000)\) nodes

• Mini Emulator -- Tessellate input parameter space -- for each vertex use appropriate 2 hop neighborhood and weighted averaging of micro-emulator output to construct this emulator

• Macro Emulator -- Weighted averaging of mini emulators in input space
Hazard Maps at Mammoth Mt. CA

Uncertain Inputs are Initial Volume, location and frictional resistance to flow -- “Gaussian” covariance, left -- p=1 right
Hazard maps for Mammoth Mountain computed using 64 multi-processor TITAN simulations and $10^{15}$ resamples of hierarchical emulator. A) Input random variables are volume, basal and internal friction angles and DEM (using one RV) B) Input random variables are volume, initial pile aspect ratio, starting location, basal and internal friction angles and DEM (2 random variables.)
list of simulator input draws

Generate Simulator Data
- TITAN simulation (itself multiprocessor for each sample point)
- Thousands of sample points on available clusters (thousands of processing cores)
- Each TITAN run typically on order of 10 cpu-hours

down sampled data

Build Macro Emulator
- Delaunay Tesselation of simulations/sample points
- Determine Neighborhood of each simulation as those sample points within N hops along simplex edges

list of resample input draws

resample inputs and assembly info

Build Mini Emulators
- Built in parallel from sample points within N hops
- An East–North Delaunay triangulation of data points defines micro emulator neighborhoods
- Micro emulators are built by traditional means

initial hazard map

Evaluate Hierarchical Emulator
- In parallel, mini–emulators are evaluated at resample inputs and evaluations combined in weighted by barycentric coordinates sum to produce a set of component hazard maps

resample hazard maps

Merge Hazard Maps
- Can be done in tree fashion

final hazard map
Workflow Parallelization Strategy

• Each stage has a similar parallelization strategy – master/worker daemon to allocate tasks to available CPUs

• I/O contention can be a serious issue (hundred of files per processor, tens of GB), so data is managed locally first, then critical inter-stage files are put on fastest available shared filesystem

• TITAN simulations (Stage 1) scale well, but still require more than 6h on 1024 processors (for only 2048 initial simulations)

• Emulator evaluation (Stage 3) is very fast, near real-time responsiveness for 512 available processors

• Principal objective achieved: simulator + emulator strategy provides very fast surrogate for pure direct simulation
Proposed workflow optimizing computing and data device use
Performance speedup of three stages of the hazard map workflow: Stage 1 is generation of direct simulation inputs, Stage 2 is emulator construction, and Stage 3 is emulator evaluation (only Stage 3 needs to be redone to produce a new hazard map based on the range covered by the initial direct simulations)
Fig 6. This figure displays an East-North map of Montserrat of the probability that the flow depth will exceed $h$ [m] in the next ten years. It was calculated from an ensemble of $zx48$ Titan2D simulations using our Hierarchical Emulator approach, the Fréchet volume distribution of Bayarri et al, and assuming that the preferred initial direction of slope failure was uniformly distributed between $0°$ and $360°$. The start to finish time needed to compute this hazard map was under 9 hours.

Fig. 7. This figure displays an East-North map at Montserrat of the spatially varying non-confidence in the probability-of-hazard map depicted in Figure 6. The non-confidence is defined as the computed standard deviation of probability-of-hazard $\sigma_P$ divided by the probability-of-hazard $P$. When calculated by standard means, as was done here, the ratio $\sigma_P/P$ only measures the non-confidence in the statistics due to insufficient re-sampling. It makes no statement regarding the quality of the simulations or the quality of the emulator, although it is possible to obtain a measure of the latter with some minor modification and the re-evaluation of the re-sample inputs.

2048 Titan2D simulations using 1024 processors, 4 inputs, initial volume (Frechet distribution), flow direction (uniform), basal and internal friction.

Hazard Map completed in 9hrs on 1024 processors.
Conclusions

1: Efficient hazard analysis procedure using large scale simulation and modest supercomputing.

2: Introduced ensemble hierarchic construction to accelerate emulator construction with properly localized covariance.

3: Full “workflow” scaling requires an integrated strategy for computations and analytics.

4: New challenges in load balancing; data access and migration need to be addressed.

5: Life at petascale is interesting but at exascale will be even more so!
A Vision

- Use premonitory information like "weather", field data, models in "real time" to produce "hazard forecast".

- Integrate a priori analysis of terrain, weather forecast into models that change this into a probability of failure (WEPP and TransTAB) and then do suitable ensembles of flow models TITAN2D etc.
