ACI1339765 SI2-SSI: Collaborative Research: Building Sustainable Tools and Collaboration for Volcanic and Related Hazards

A. K. Patra, T. Kosar, M. D. Jones, S.M. Gallo, G. Valentine, M. I. Bursik
University at Buffalo, SUNY, Buffalo, NY 14260
C. B. Connor, L. J. Connor
University of South Florida, Tampa, FL
P. Webley,
Univ of Alaska-Fairbanks, Fairbanks, AK

GOALS OF PROJECT

I. Provide "**operational infrastructure**" for a large class of researchers in volcanology and mass flows (not necessarily computationally sophisticated).

II. Enable routine use of proper **validation and verification** processes under the VHub platform.

III. Provide a standard well managed hardware/software platform and approaches to **standardize the documentation associated with input data, source code, and output data** ensuring that model calculations are **reproducible** across machine and software environment.

IV. Collaborative Model Development and Testing.

V. Broaden Accessibility of Sophisticated Modeling

Architecting & Reengineering

- Simple Practices for Sustainability
  - Github, issue tracking, unit testing, shared communications
- Reengineering to Promote Reuse
  - Modularize
  - Separate Concerns
  - Code Reuse by independent Vhub release of Components
- Validation, Inversion and Hazard Analysis Workflows
  - Support validation and UQ Workflows – data and compute intensive work
- Parallelism, Adaptivity and GPGPUs
  - Hierarchical parallelism, multithreading, standard APIs for GPU

Validation and UQ Framework

- " Is the model correct?" – wrong question
- "With probability γ, the prediction is within (tolerance) τ of the true y^R"
  - If model is too expensive use a surrogate y^M
  - Training of surrogate with well chosen ensemble of expensive simulator
  - Surrogate – Polynomial Chaos(PC), Gaussian Processes, Bayes Linear Models (BLM)

Data

- Observation data
- Experimental Data
- Metadata – "paper" and doi?
- Distributed data access – in situ – during simulations

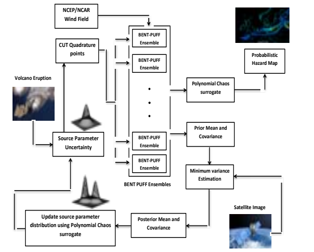


Figure 1: Schematic view of hazard analysis process using source uncertainty propagation and satellite imagery based source estimation.

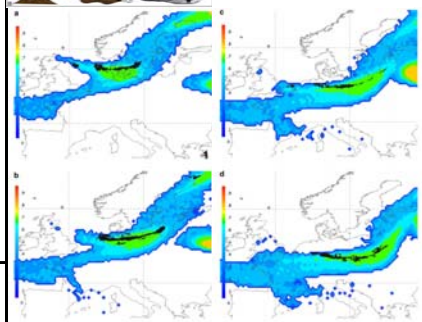
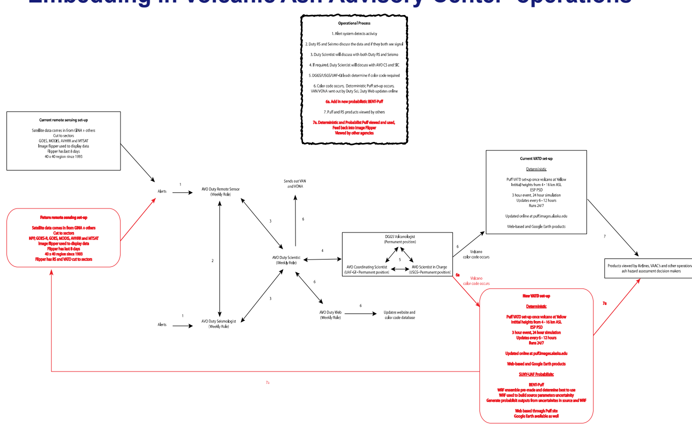


Fig. 8 Meteosat-9 SEVIRI cloud outline (*filled black region*) compared with model generated probabilities of ash presence based on source parameter uncertainty propagation. Color scale bar in fractional probability. Outer edge of blue area is at 20 % probability. **a** 16 April 2010 00Z, **b** 16 April 2010 06Z, **c** 16 April 2010 12Z, **d** 16 April 2010 18Z

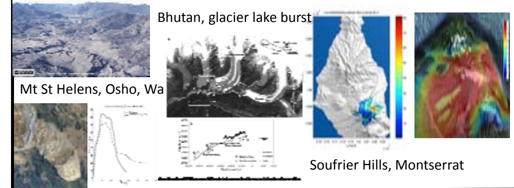
PUFFIN: Puffin is a theoretical model of a volcanic plume, based on applying the equations of motion in a plume-centered coordinate system. Puffin suggests that the interaction between a volcanic plume and wind causes enhanced entrainment of air and horizontal momentum, plume bending, and a decrease in plume rise height at constant eruption rate. Because of rapid dilution in the high windspeeds of the polar jet, plumes that vary over more than one order of magnitude in mass eruption rate (10 to 10 kg/s), if injected into the polar jet, may all attain rise heights only slightly different from that of the core of the jet, ~10 km, as opposed to 17 – 33 km in a still atmosphere. The model outputs plume trajectories and rise heights, as well as pyroclast loadings as a function of height, and can therefore be used to produce input for advection-diffusion and volcanic ash transport models such as tephra2 or puff. Puff simulates the transport, dispersion and sedimentation of volcanic ash. It requires horizontal wind field data as a function of height on a regular grid covering the area of interest. Puff output includes the location (in 3 dimensions), size, and age-since-eruption of representative ash particles. puff can also produce gridded data of relative and absolute ash concentration in the air and on the ground. Puff is a fast and efficient research and operational tool for predicting the trajectories of ash particles, which is essential for hazard assessment

Embedding in Volcanic Ash Advisory Center operations



TEPHRA2: Numerical models used for tephra hazard assessment (Hazard Models) typically result from the combination and integration of different theories and modeling approaches depending on the specific eruptive scenario and mitigation program required. They can be grouped within two main categories: particle-tracking models and advection-diffusion models. Particle-tracking models are Eulerian or Lagrangian models that can forecast volcanic-cloud position at specific times and space.

Advection-diffusion models are Eulerian models that describe the solution of the equations of particle diffusion, transport, and sedimentation and can forecast tephra accumulation on the ground relative to a particle-release sourceThe code uses a closed-form solution of the advection-diffusion equation, particle fall velocities that depend on local Reynold's number, and stratified wind field to forecast tephra accumulation, expressed as kilogram per cubic meter, particle size distribution at specific locations from the vent, and maximum clast size expected as a function of distance from the vent.



Simulation ensembles used to create hazard analysis and probabilty of catastrophe map – see right.

TITAN2D: The Titan2D toolkit is used for modeling of geophysical mass flows over natural terrain (volcanic flows, avalanches, landslides etc.). Titan2D uses an adaptive finite volume scheme, and assumes that a large class of granular flows of geophysical origin can be modeled as an incompressible continuum satisfying a Mohr-Coulomb law. The solution is parallelized using the message-passing interface (MPI), and simulations for difficult terrains can be relatively time consuming even on large computational clusters. An essential input is the digital elevation model (DEM), since the flow simulation requires elevation, slope, curvature, and material information at every cell to be computed.

Workflow Needs

Complex inference from simulation ensembles used in dealing with the Uncertainty Quantification

Twin computational challenges of managing large amount of data and performing cpu intensive processing.

Sample workflow used for hazard map generation.

2048 sets of input parameters were generated using Latin Hypercube sampling

Titan2D simulations of these inputs were performed.

2048 flow depth profiles were merged using ad hoc python scripts and loaded on a Netezza data warehouse server

Data was used to create a statistical surrogate (a Bayes linear model using a mix of SQL, python and C code) which was in turn sampled to create a probabilistic hazard map.