Blue Waters Exploratory Proposal

# Building a data assimilation framework for forecasting volcanic activity during periods of unrest

**PI:** *Patricia M. Gregg*,
Department of Geology
University of Illinois at Urbana-Champaign

## Executive Summary

A primary goal of the UIUC Volcano Lab is to develop innovative strategies for combining volcano-monitoring data sets with thermomechanical finite element models in order to assess hazards at volcanoes experiencing unrest. The purpose of our Exploratory Allocation on Blue Waters is to develop and test a framework for multi-data stream data assimilation by conducting a series of eruption “hind casts” for recent hazardous eruptions at Sinabung Volcano in Indonesia. Blue Waters is uniquely capable of handling the computational expense for our ensemble-based approach, which has compute times and storage requires considerably outside the capabilities of traditional HPC resources. Ultimately, this work will provide a critical foundation for future interdisciplinary efforts to model volcano evolution and mitigate volcano disasters worldwide.

## Description of research activities and results

### *Key Challenges and Why Blue Waters*

The Ensemble Kalman Filter (EnKF) is an ensemble based sequential data assimilation method that requires calculating hundreds to thousands of finite element models at each time step. While the EnKF analysis step has been optimized to run very swiftly, the computational expense of running and storing hundred to thousands of finite element models for each time step in the EnKF analysis is cost prohibitive for even the large EXEDE clusters. Blue Waters is uniquely positioned to handle our computational needs and has allowed our group to make rapid progress and think ambitiously without being hampered by computational limitations. Our exploratory allocation is paving the way for developing large-scale data assimilation strategies to aid volcano-monitoring agencies worldwide.

The key challenge for this exploratory allocation was determining how to adapt our workflow for High Performance Computing (HPC). Our ensemble-based approach creates bottlenecks during the distribution of our multiphysics models and in exporting model outputs from hundreds to thousands of models.

Our goal was to utilize the power of Blue Waters to overcome two major obstacles:

(1) The computational cost of our multiphysics FEMs. Each 3D FEM requires several minutes of run time on a single core. In our Monte Carlo ensemble approach, thousand of FEMs can be run simultaneously. This level of computation is cost-prohibitive on local clusters, and as our FEMs routinely can take up to 1 TB of online storage, we were unable to implement our approach on EXEDE.

(2) Optimizing handling of big geodata. Satellites collect observations of volcanic unrest at the meter scale resulting in data matrixes on the order of 1Mx1M. Our current workflow requires us to vastly under sample these observations, which may result in “throwing out” critical information.

*Why it Matters*

A primary motivation for investigating volcanic systems is developing the ability to predict eruptions and mitigate disaster for vulnerable populations. In the past decades, the evaluation of volcanic activity has been greatly enhanced by coupling remote (e.g., satellite geodetic and global seismic arrays) and local observations (e.g., GPS, tiltmeters, gas emissions, campaign gravity, and seismometers) to provide early warning of immanent eruption or the evolution of a magma system during a volcano crisis. Concurrently, thermomechanical models of magma reservoirs have significantly advanced our understanding of eruption-triggering mechanisms beyond the temporal and spatial limitations of our observations (Grosfils, 2007; Gerbault et al., 2012; Gregg et al., 2012; Gregg et al., 2013; Grosfils et al., 2015). This new generation of models has pushed the field beyond the classic analytical approaches by providing new insights into the mechanics of magma chamber growth and failure.

Volcano monitoring datasets are commonly analyzed using analytical inversions techniques or by optimizing finite element models (Battaglia and Hill, 2009; Parks et al., 2012; Hickey et al., 2015). While these approaches work well for combining models with one or potentially two data streams, they are static assessments of the system state and do not provide updates or forecasts. Alternatively, statistical data assimilation methods systematically link data with models to provide model updates. Significant advancements have been made in hydrologic, physical oceanography, and climate modeling to incorporate disparate datasets into dynamic, nonlinear models and provide model forecast (Evensen, 2009; Williams et al., 2009; Wilson et al., 2014; Wilson and Hay, 2015). Sequential data assimilation methods provide a framework for integrating large disparate data sets into time forward, forecast models. Data are used to nudge the model trajectory and provide updates of the system’s evolution and the models inform future data targets and strategies.

Our current efforts on Blue Waters are focused on developing strategies for rapid assimilation of monitoring data sets into evolving geodynamic models to provide near real-time forecasts and assessment of volcanic unrest. To that end, we are adapting data assimilation strategies developed in other fields to combine observations of volcanoes experiencing unrest with thermomechanical finite element models to calculate volcano evolution (Gregg and Pettijohn, 2016; Zhan and Gregg, 2017).

*Accomplishments*

The most ciritical accomplishment in this year-long investigation was determining that the Kray architecture is not well suited for our FEM approach. The software we are utilizing: COMSOL Multiphysics, does not parallelize appropriately when implemented on a Kray system. We worked closely with the COMSOL tech support and Blue Waters to find strategies to work around this issue, and were able to create a brute-force approach to run the FEMs, but our efforts on Blue Waters illustrated several bottlenecks in our workflow that must be addressed to take advantage of HPC resources.

The initial tests of implementation on Blue Waters are key for providing critical information to devise strategies for adapting our approach for HPC. As such, the primary outcome was providing necessary testing in a targeted and efficient manner. The benchmarking and test results allowed us to move our approach forward and develop new strategies. Our outcomes were folded into a NSF OCE proposal, which was selected for funding, as well as an NCSA Fellowship proposal that was selected for funding. Over the next year will be working with Prof. Shoawen Wang and NCSA CyberGIS to take advantage of the unique architecture of the new ROGER (Resourcing Open Geospatial Education and Research) cluster to implement our approach and develop new workflow strategies directly resulting from our efforts on Blue Waters. Without the very helpful and efficient testing on Blue Waters, we would have not been able to make such swift progress.

*Next Generation Work*

Over the next 2 years we will continue to work closely with NCSA to develop strategies for utilizing HPC in volcano hazards. The UIUC Volcano Lab currently has 4 PhD students and 1 Postdoc working to implement data assimilation strategies to better understand volcanic systems. We have several papers in submission and preparation phases as well as proposals submitted to NSF. While, ultimately the Kray architecture of Blue Waters was the ultimate roadblock for implementing our approach, we are looking for new methods including coding up basic open-source multiphysics implementations in Python to avoid commercial software. Our longterm goals are to establish a partnership with the NCSA that will be further enhanced through immediate proposals to NSF and future proposals to NASA. We hope to position the University of Illinois as leaders in volcano data assimilation.

*References:*


Battaglia, M. and Hill, D.P., 2009. Analytical modeling of gravity changes and crustal deformation at volcanoes: The Long Valley caldera, California, case study. Tectonophysics, 471(1-2), 45-57, 10.1016/j.tecto.2008.09.040.

Evensen, G., 2009. The Ensemble Kalman Filter for combined state and parameter estimation - Monte Carlo techniques for data assimilation in large systems. IEEE Control Syst. Mag., 29(3), 83-104, 10.1109/mcs.2009.932223.

Gerbault, M., Cappa, F. and Hassani, R., 2012. Elasto-plastic and hydromechanical models of failure around an infinitely long magma chamber. Geochemistry Geophysics Geosystems, 13, 10.1029/2011gc003917.

Gregg, P.M., de Silva, S.L. and Grosfils, E.B., 2013. Thermomechanics of shallow magma chamber pressurization: Implications for the assessment of ground deformation data at active volcanoes. Earth and Planetary Science Letters, 384, 100-108, 10.1016/j.epsl.2013.09.040.

Gregg, P.M., de Silva, S.L., Grosfils, E.B. and Parmigiani, J.P., 2012. Catastrophic caldera-forming eruptions: Thermomechanics and implications for eruption triggering and maximum caldera dimensions on Earth. Journal of Volcanology and Geothermal Research, 241-242, 1-12, doi: 10.1016/j.jvolgeores.2012.06.009.

Gregg, P.M. and Pettijohn, J.C., 2016. A multi-data stream assimilation framework for the assessment of volcanic unrest. Journal of Volcanology and Geothermal Research, 309, 63-77, 10.1016/j.jvolgeores.2015.11.008.

Grosfils, E.B., 2007. Magma reservoir failure on the terrestrial planets: Assessing the importance of gravitational loading in simple elastic models. Journal of Volcanology and Geothermal Research, 166(2), 47-75, 10.1016/j.jvolgeores.2007.06.007.

Grosfils, E.B., McGovern, P.J., Gregg, P.M., Galgana, G.A., Hurwitz, D.M., Long, S.M. and Chestler, S., 2015. Elastic models of magma reservoir mechanics: A key tool for understanding planetary volcanism. In: T. Platz, M. Massironi, P. Byrne and H. Hiesinger (Editors), Volcanism and Tectonism Across the Inner Solar System. Geological Society, London, Special Publications, pp. 139-158.

Hickey, J., Gottsmann, J. and Mothes, P., 2015. Estimating volcanic deformation source parameters with a finite element inversion: The 2001–2002 unrest at Cotopaxi volcano, Ecuador. J. Geophys. Res, 120, 1473-1486, 10.1002/2014JB011731.

Parks, M.M., Biggs, J., England, P., Mather, T.A., Nomikou, P., Palamartchouk, K., Papanikolaou, X., Paradissis, D., Parsons, B., Pyle, D.M., Raptakis, C. and Zacharis, V., 2012. Evolution of Santorini Volcano dominated by episodic and

rapid fluxes of melt from depth. Nat. Geosci., 5(10), 749-754, 10.1038/ngeo1562.
Williams, M., Richardson, A.D., Reichstein, M., Stoy, P.C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D.Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. and Wang, Y.-P., 2009. Improving land surface models with FLUXNET data. Biogeosciences Discuss, 6, 2785-2835, 10.5194/bgd-6-2785-2009.
Wilson, G.W. and Hay, A.E., 2015. Acoustic backscatter inversion for suspended sediment concentration and size: A new approach using statistical inverse theory. Continental Shelf Research, 106, 130-139, 10.1016/j.csr.2015.07.005.
Wilson, G.W., Ozkan-Haller, H.T., Holman, R.A., Haller, M.C., Honegger, D.A. and Chickadel, C.C., 2014. Surf zone bathymetry and circulation predictions via data assimilation of remote sensing observations. J. Geophys. Res.-Oceans, 119(3), 1993-2016, 10.1002/2013jc009213.
Zhan, Y. and Gregg, P.M., 2017. Data assimilation strategies for volcano geodesy. Journal of Volcanology and Geothermal Research(special issue on Volcano Geodesy), 10.1016/j.jvolgeores.2017.02.015.