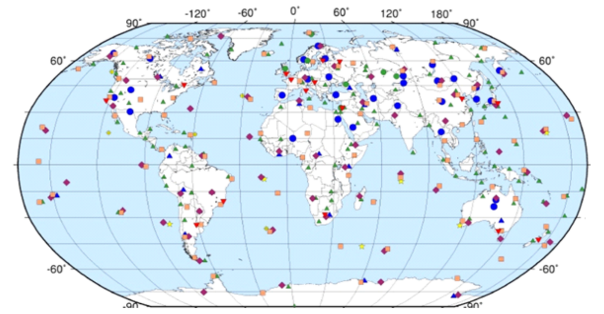
Progress in Signal-Based Bayesian Monitoring

David A. Moore[1], Kevin Mayeda[1], Stephen C. Myers[2], Min Joon Seo[1], Stuart J. Russell[1]
University of California, Berkeley[1] and Lawrence Livermore National Laboratory[2]

Introduction

• Global seismic monitoring for the **Comprehensive Nuclear-Test-Ban Treaty (CTBT)** aims to recover the time, location, depth, and magnitude for all seismic events in the magnitude range of interest.
• Data from the International Monitoring System (IMS) are processed in real time at the International Data Centre (IDC) in Vienna. Our goal is to improve the sensitivity and accuracy of automated processing at IDC.

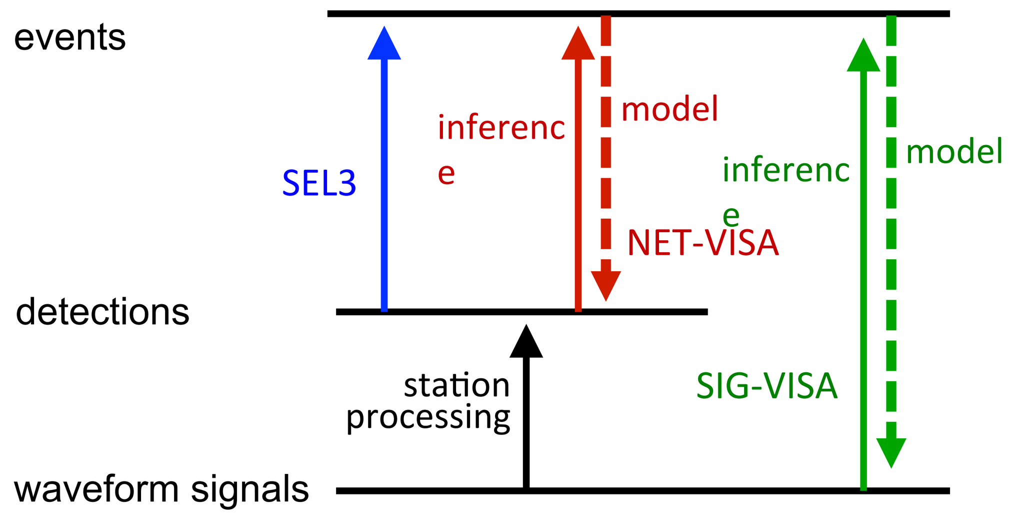


Blue dots and triangles are primary seismic stations.

• The current automated system (**SEL3**) detects 69% of real events and creates twice as many spurious events.
• 16 human analysts find more events, correct existing ones, throw out spurious events, and generate **LEB** ("ground truth").
• Unreliable below magnitude 4 (about 1 kiloton).

• NET-VISA is a *detection-based* Bayesian monitoring system whose performance is limited by the classical, bottom-up, threshold-based detections algorithms used in station processing. *It misses about 2-3 times fewer events than SEL3.*
• SIG-VISA, a *signal-based* system, uses generative models that span the range from events to waveform traces. This approach has several qualitative advantages over NET-VISA, with the potential for significantly improved sensitivity and localization performance.

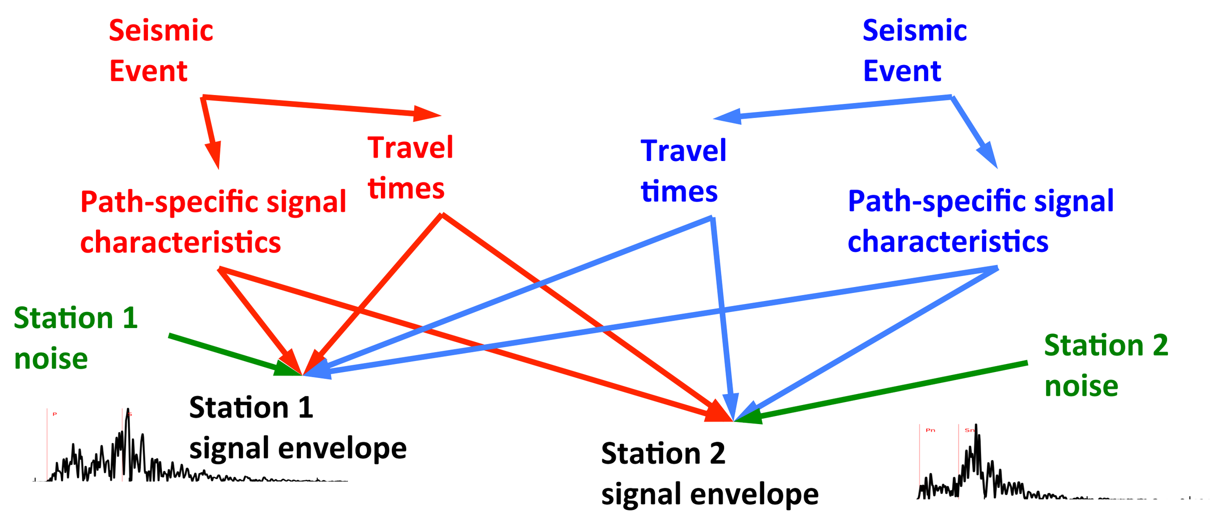
Signal-Based vs. Detection-Based Monitoring



Bayesian monitoring with a generative approach
P_θ(world) describes prior probability for what *is* (*events*)
P_θ(signal | world) describes forward model (propagation, measurement, etc.)
Detection-based Bayesian monitoring:
P(world | *f* (signal)) ∝ P_θ(*f* (signal) | world) P_θ(world)
where *f* (signal) = set of all detections
Signal-based Bayesian monitoring:
P(world | signal) ∝ P_ϕ(signal | world) P_θ(world)

Signal Envelope Model

SIG-VISA is a probabilistic generative model of seismic event origins, propagation, and observed waveform envelopes, including event signals along with station background noise:

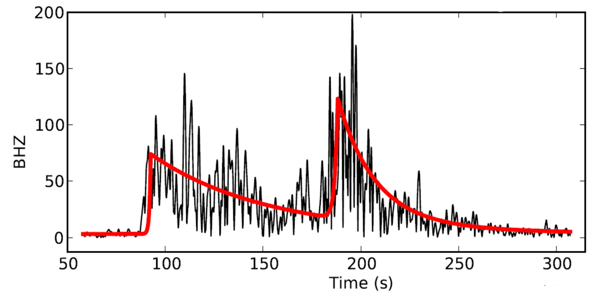


• The signal model encodes a distribution over waveform envelopes at each station given parameters for all hypothesized events.

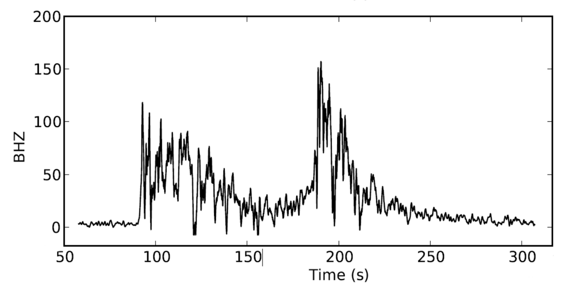
• Each observed envelope is modeled with a **shaped template** combined with a stochastic **modulation signal** for each arriving phase, and a **background noise** process at each station:

observedEnvelope = template x exp(modulation) + stationNoise

• SIG-VISA uses a paired-exponential envelope template whose arrival time, amplitude, and decay rate depend on event magnitude and the event->station path. Template fits to historical signals are used to train predictive models for new observations.



Left: An observed envelope showing the P and S arrivals and subsequent coda decays. The red line indicates the template fit.



Right: A synthetic envelope generated from the template with an autoregressive modulation process.

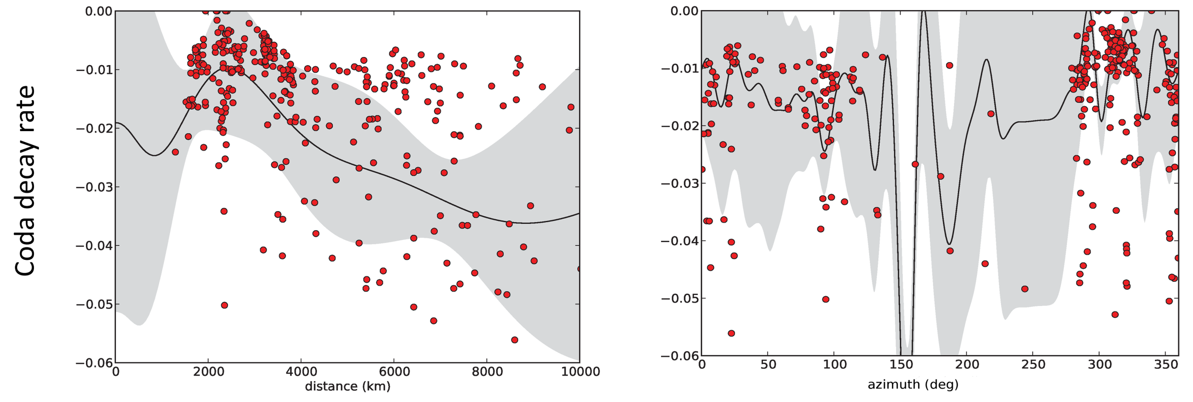
• Station noise is modeled by an autoregressive process, retrained hourly at each station to capture regional and temporal variation in noise properties.

• Observed signals are modeled independently across a range of narrow frequency bands, automatically learning spectral characteristics and frequency-specific decay rates.

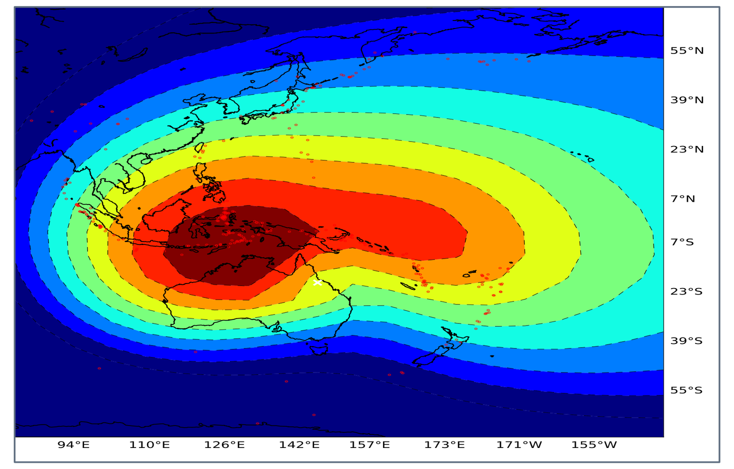
• Amplitude for each phase within each frequency band is predicted from event magnitude using a physics-based source model (Brune, 1971; Mueller and Murphy, 1971) and a path-specific learned transfer function (next column).

Path-Specific Characteristics

Observed envelope amplitudes and decay rates vary with the event source location, with effects from distance and depth as well as local variation in seismic structure.

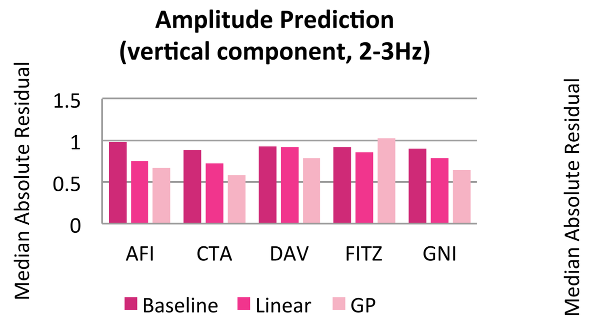


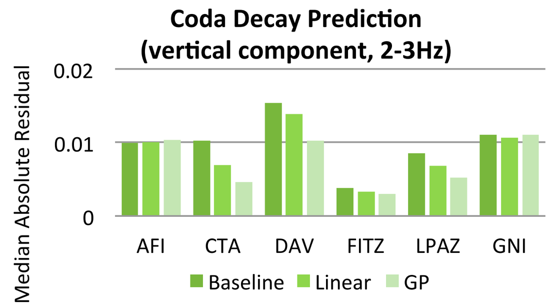
We model these effects empirically using a nonparametric spatial **Gaussian Process regression** (kriging) model, which learns for each seismic station a probability distribution over template parameters for events at every point on the Earth:



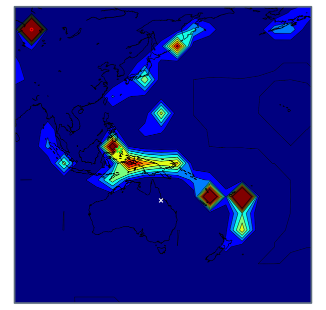
Spatial model of coda decay rates observed by Charter Towers station. (red = sharper decay).

The nonparametric model captures local variation, yielding improved predictions relative to a distance-dependent linear model:

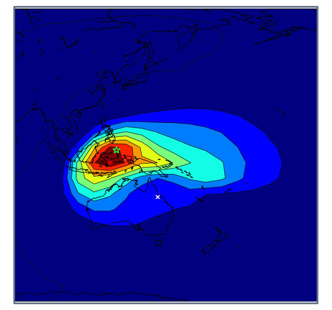




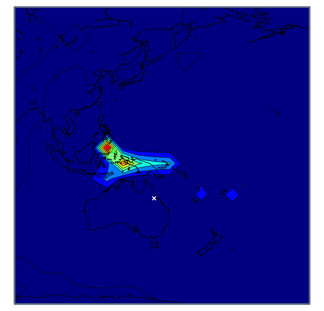
The spatial model enables inference from signal characteristics to possible event locations:



Prior location density



Coda decay rate likelihood

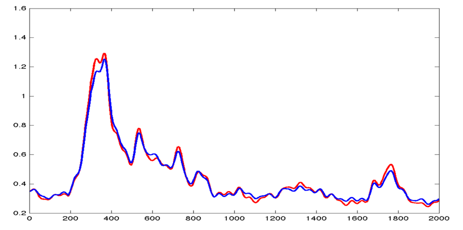


Posterior location density given coda decay rate

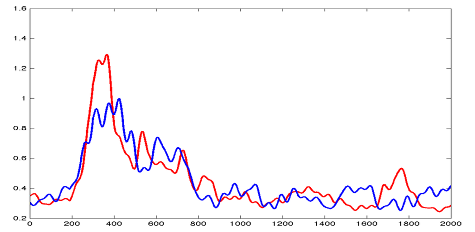
Waveform Matching

• Waveform shape is known to be **highly repeatable** across events with the same location and source mechanism (Thorbjarnardottir and Pechmann, 1987; Harris, 1991).

• SIG-VISA captures this effect by replacing the independently sampled modulation signal with a signal conditioned on the event location, causing nearby events to generate correlated waveforms:



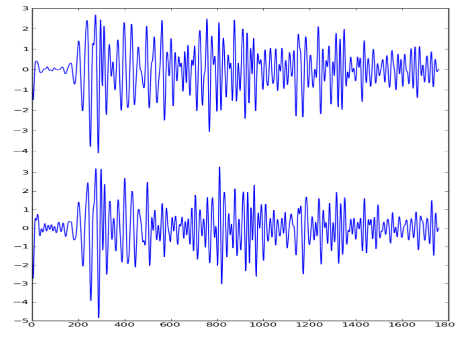
Synthetic reference envelope (red) and a sampled envelope (blue) from a Gaussian-process generative model for a nearby event location.



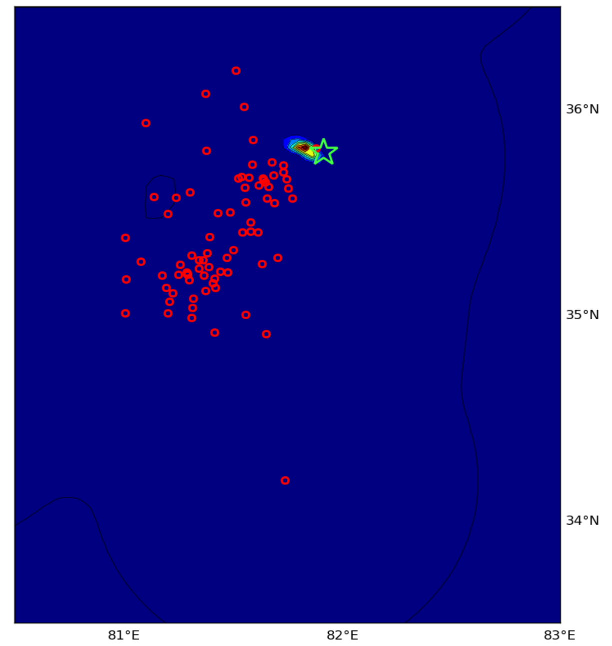
The same reference envelope (red) and a sampled envelope (blue) from a distant event location.

• This causes a statistical "waveform matching" effect to emerge from inference in the probabilistic model.

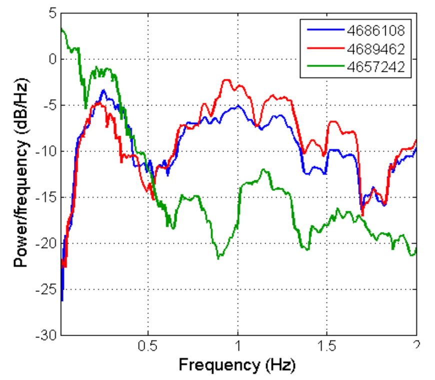
• Signals are represented parametrically as a **sum of basis functions** (e.g. Fourier basis), with coefficients modeled by a spatial Gaussian process.



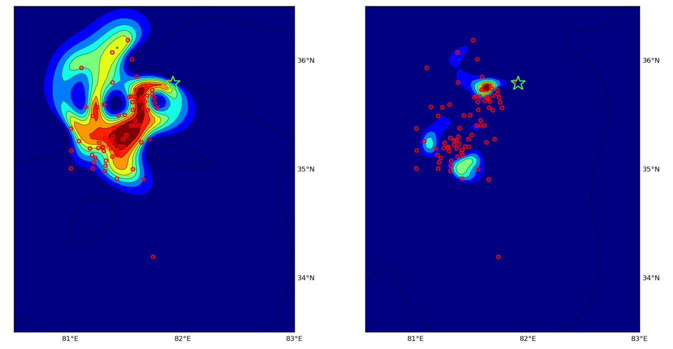
Correlated event waveforms (IMS events 4686108, 4689462, correlation .60).



Posterior location density of IMS event 4689462 from a single station observation (MKAR) using a waveform matching model trained on data from 99 other events, shown in red. The posterior peak is 8.4km from the true event location, marked by a green star.



Fourier features for two nearby and correlated events (red, blue) vs. a distant event (green).



Partial posterior densities conditioned on one (left) and five (right) Fourier coefficients.

CONCLUSIONS

• Bayesian methods are a powerful and promising approach to monitoring problems. In particular, signal-based monitoring improves on traditional methods by making full use of the available data, with potential for significant performance improvements.
• The SIG-VISA prototype system includes an empirically-calibrated, physics-based model of observed waveform envelopes.
• Gaussian Process regression enables learning of flexible, data-driven probabilistic models of spatial parameters.
• Waveform matching / correlation effects can be recovered from Bayesian inference in a spatially conditioned model, enabling localization from a single detection in some cases.

FUTURE WORK

• Computationally efficient algorithms for training and probabilistic inference at global scale.
• Model refinement based on geophysical expertise.
• Dependent modeling of array station elements.
• Putting it all together: a practical, effective global monitoring system.

REFERENCES

Arora, Nimar S., Stuart J. Russell, Paul Kidwell, and Erik Sudderth, "Global seismic monitoring as probabilistic inference," In *NIPS 23*, MIT Press, 2011.
Harris, D.B, "A wave-form correlation method for identifying quarry explosions," *Bull. Seism. Soc. Am.*, 81, 2395-2418, 1991.
Harris, D. B. and D. A. Dodge, "An autonomous system for grouping events in a developing aftershock sequence," *Bull. Seismol. Soc. Amer* ., vol. 101, no. 2, pp. 763-774, 2011.
Mayeda, K., Hofstetter, A., O'Boyle, J.L., and Walter, W.R. (2003). Stable and transportable regional magnitudes based on coda-derived moment-rate spectra. *Bulletin of the Seismological Society of America*, 93(1), 224-229.
Myers, S.C., G. Johannesson, and W. Hanley (2007). A Bayesian hierarchical method for multiple-event seismic location. *Geophys. J. Int.*, 171, 1049-1063.
Schaff, D. P., Bokelmann, G. H. R., Ellsworth, W. L., Zanzerkia, E., Waldhauser, F., & Beroza, G. C. (2004). Optimizing Correlation Techniques for Improved Earthquake Location. *Bull. Seism. Soc. Amer*, 94(2), 705-721.
Slinkard, Megan E., Dorthe B. Carr, Stephen L. Heck, and Christopher J. Young (2011). Towards an automated waveform correlation detector system. In *Monitoring Research Review*, Tucson, Arizona.
Thorbjarnardottir, B. S. and J. C. Pechmann (1987). Constraints on relative earthquake locations from cross-correlation of waveforms. *Bull. Seismol. Soc. Am.*, 77, 1626-1634.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of DTRA for this work; also the support of CTBTO for the work on NET-VISA and the technical cooperation of CTBTO personnel including Ronan Le Bras and Jeff Given. We thank Sheila Vaidya of LLNL and Ola Dahlman, former Chair of CTBTO Working Group B, for their moral support and encouragement in this project. We also thank Megan Slinkard and Regina Eckert of Sandia National Labs for sharing with us their data on correlated arrival waveforms.