Hidden Markov Models for Detecting Aseismic Events in Southern California

Robert Granat
Jet Propulsion Laboratory
California Institute of Technology

Objectives

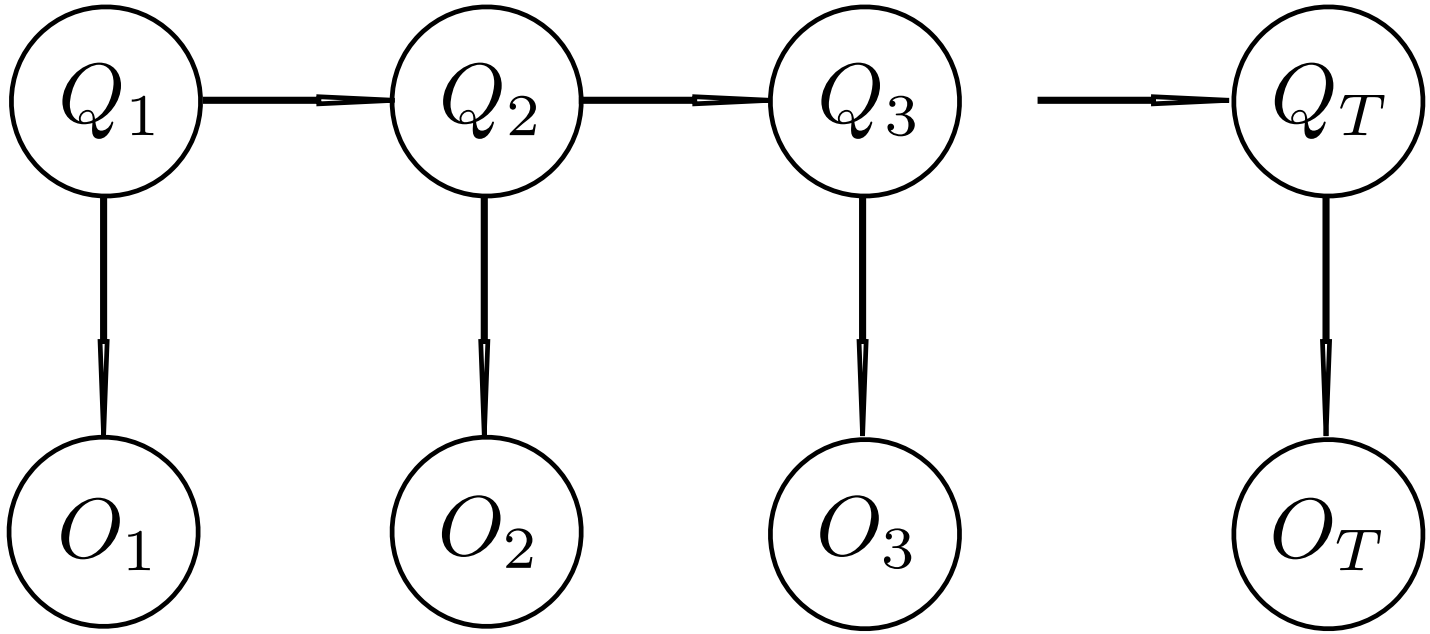
- Develop pattern recognition methods for location and classification of unusual signals in GPS data.
- Use these methods for identification of regional phenomenon indicative of aseismic activity.
- Make these methods available to the scientific community through the QuakeSim web portal.
- Integration of the method with the SCIGN data collection network.

Hidden Markov Model

- Statistical model with discrete states.
- Can be trained in supervised and unsupervised modes.
- Successful in many areas, notably speech processing and protein sequence analysis.
- These areas offer a great deal of information about the underlying system.
- In scientific exploration, however, little information is available.

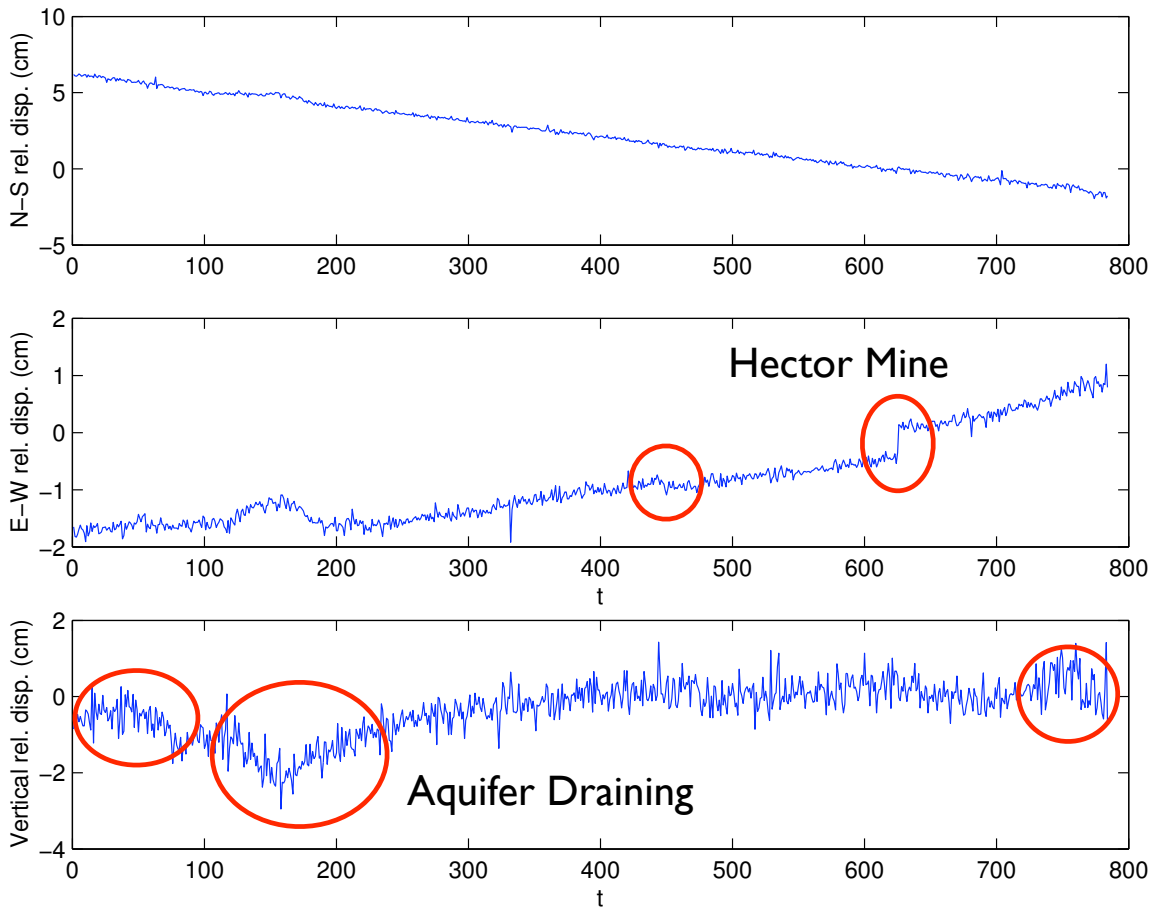
Hidden Markov Model

Discrete Hidden States



Observations

GPS Data

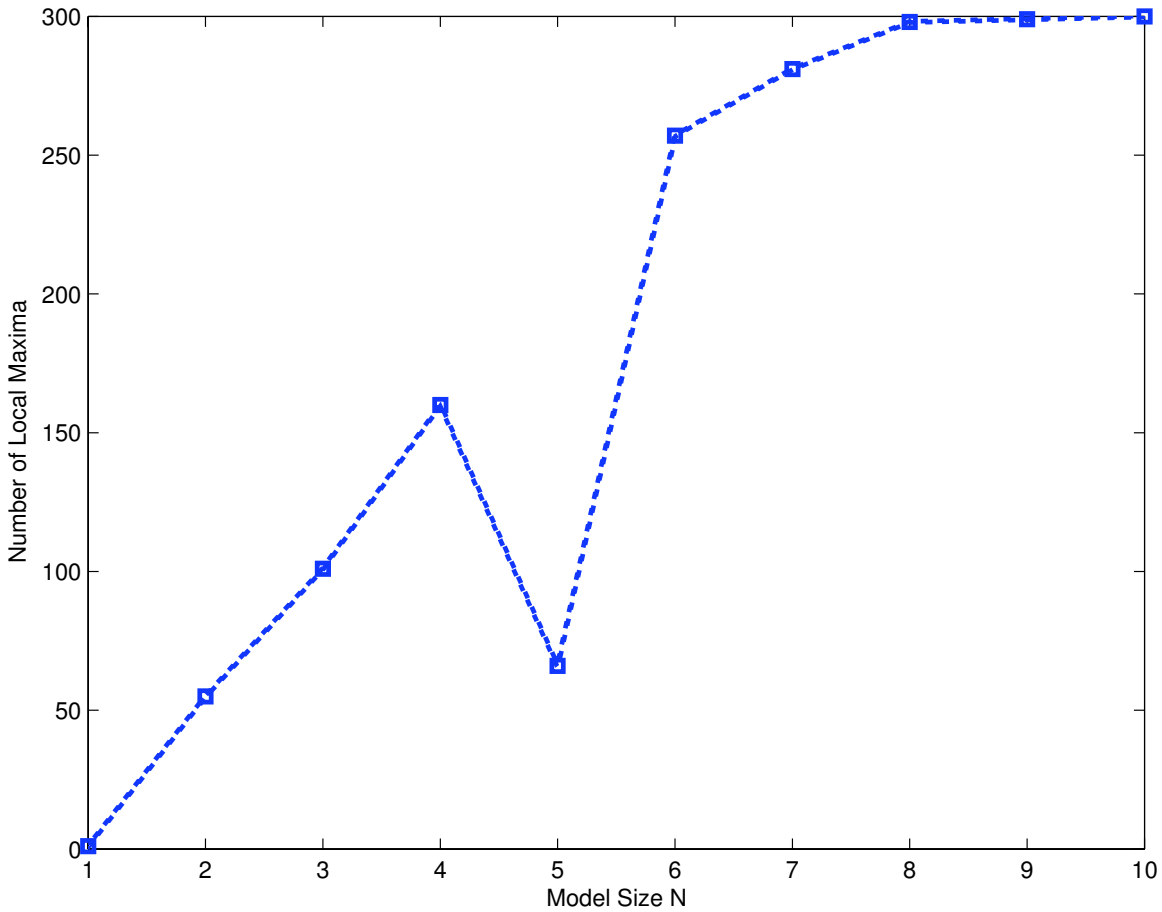


Daily displacement data from Claremont, CA

Geophysics HMMs

- Little or no labeled training data - so we are almost always using unsupervised training.
- Classification/segmentation results using the HMM state assignment are the goal.
- Data is complicated enough to induce large numbers of local maxima.

Empirical Local Maxima



1000 tests on Claremont GPS data (EM).

Dealing With Local Maxima

- EM alone is not sufficient.
- Use Deterministic Annealing EM (DAEM) to avoid local maxima [Ueda 98].
- DAEM modifies the objective function according to a computational temperature.
- At hot temperatures ($1/T \sim 0$), the objective function is nearly flat, but at the coolest temp. ($1/T = 1$) it is identical to the original.
- Major features emerge first and are tracked.

DA Problem

- DAEM tends to get stuck in local maxima where there are redundant states ([Whiley 02] demonstrates this for Bayes networks).
- Intuitively, consider a perfectly flat objective function. All probability distributions are identical - but this is a fixed point of EM for all objective functions!
- If this occurs at any temperature, the solution can't escape.

DA Problem: Solution

- Our approach is to devise regularization terms that guide the solution away from these redundant state local maxima.
- Results are not biased against any specific property of the observation data.
- This preserves the usefulness of the method for exploratory research.

Regularization

For the output distribution, there is no set prior.

In the case of Gaussian outputs, we can choose a prior based on the Euclidean distance:

$$P_3(B) = \prod_{i=1}^{N} \prod_{j=1}^{N} \exp\left(\frac{\omega_{Q_3}}{2} (\mu_i - \mu_j)^T (\mu_i - \mu_j) \right)$$

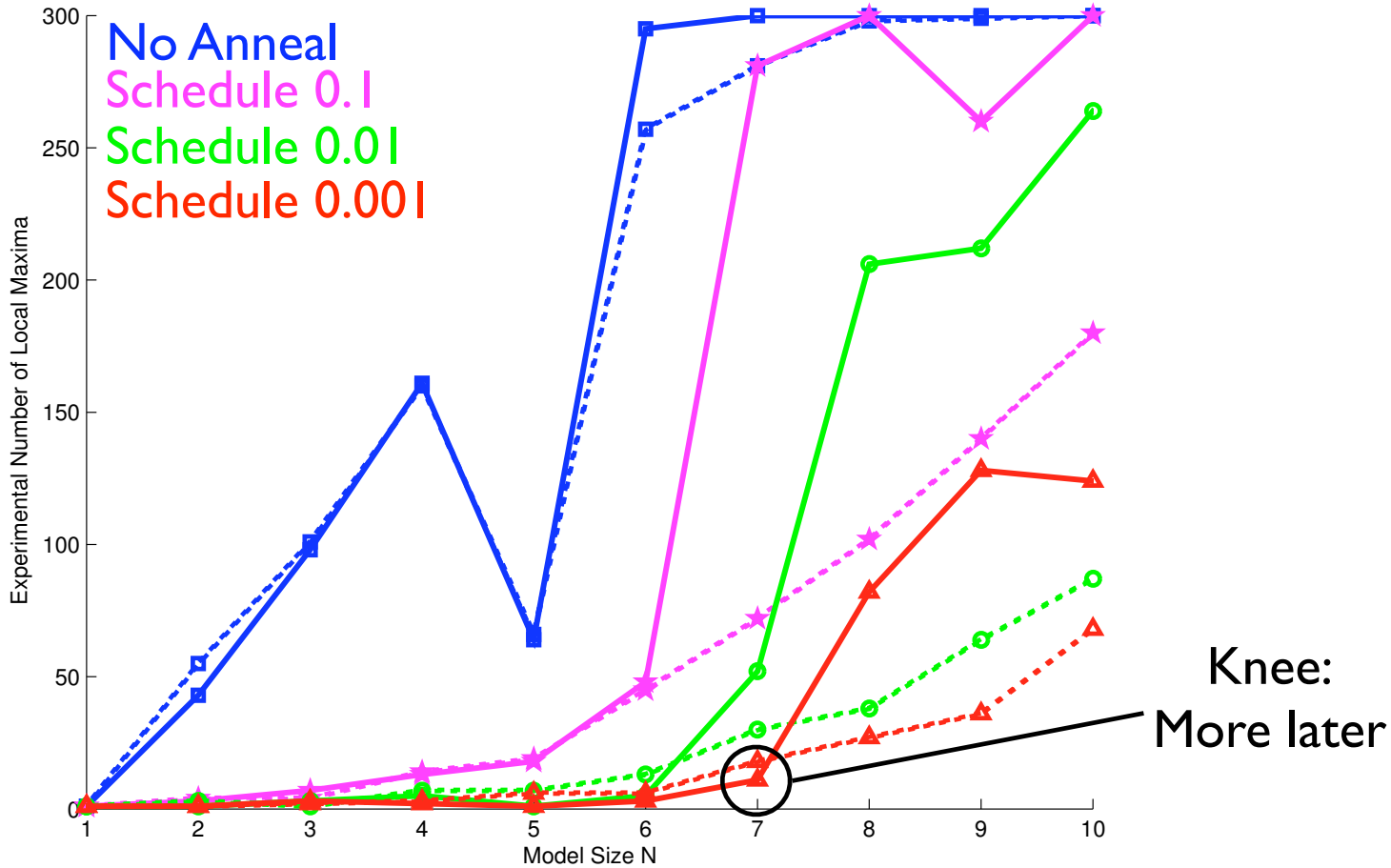
Which produces the regularization term

$$\sum_{i=1}^{N} \sum_{t=1}^{T} \tau_{it}^{(k)} \frac{\omega_{Q_3}}{2} \sum_{j=1}^{N} (\mu_i - \mu_j)^T (\mu_i - \mu_j)$$

To maintain concavity of the Q-function, we require

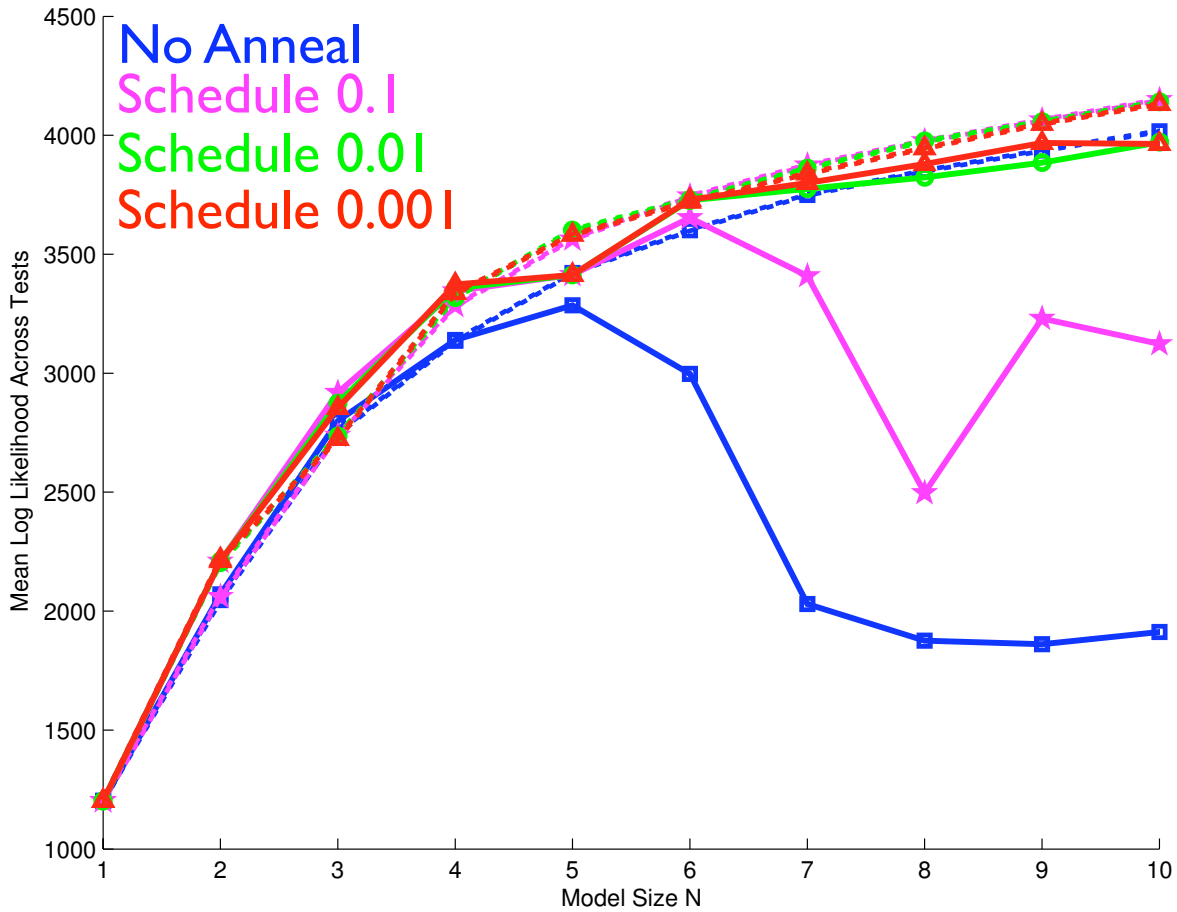
$$\omega_{Q_3} \leq ||\Sigma_i^{-1}||/2N$$

Regularization Results: Claremont (1)

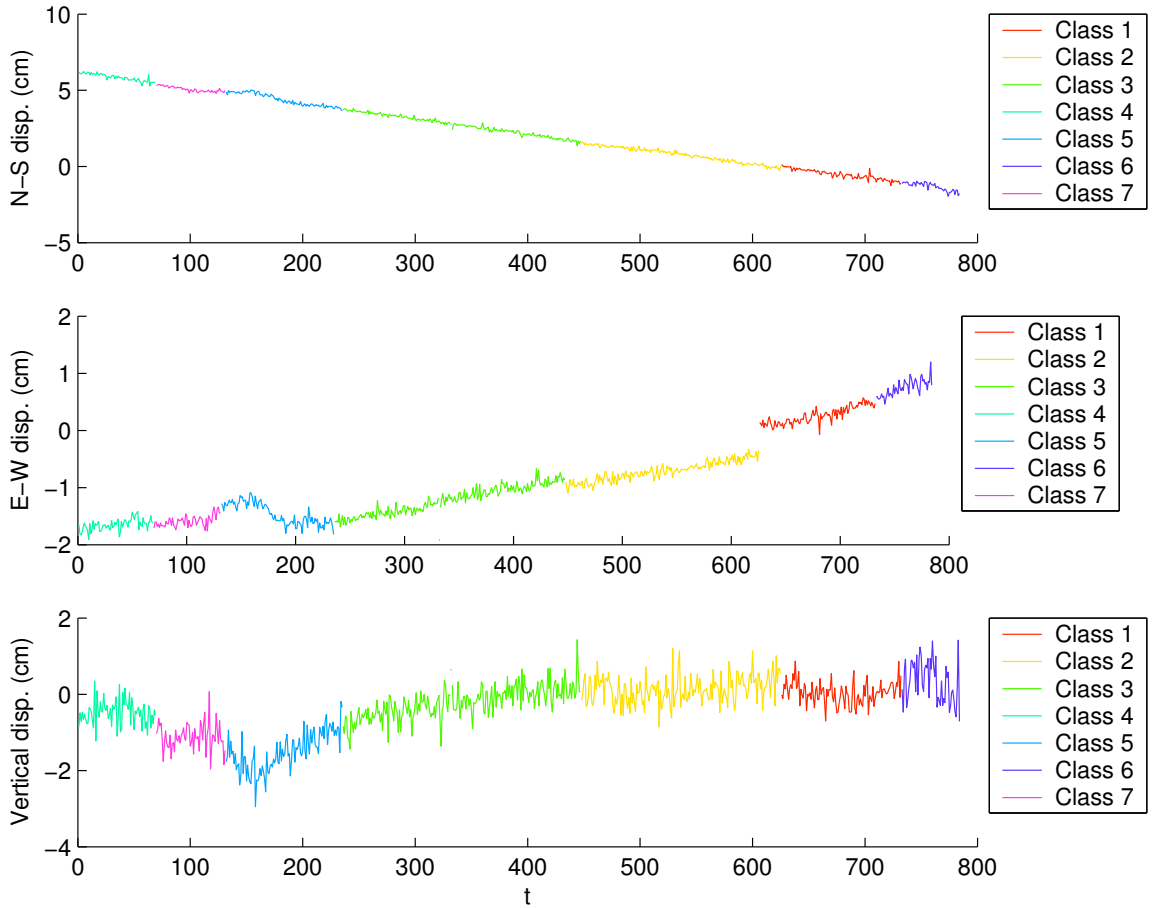


Less than 3 maxima at slowest annealing rate until N =6;
11 maxima for N = 7 (3 comprise 92% of results)

Regularization Results: Claremont (2)



Regularized Classification: Claremont

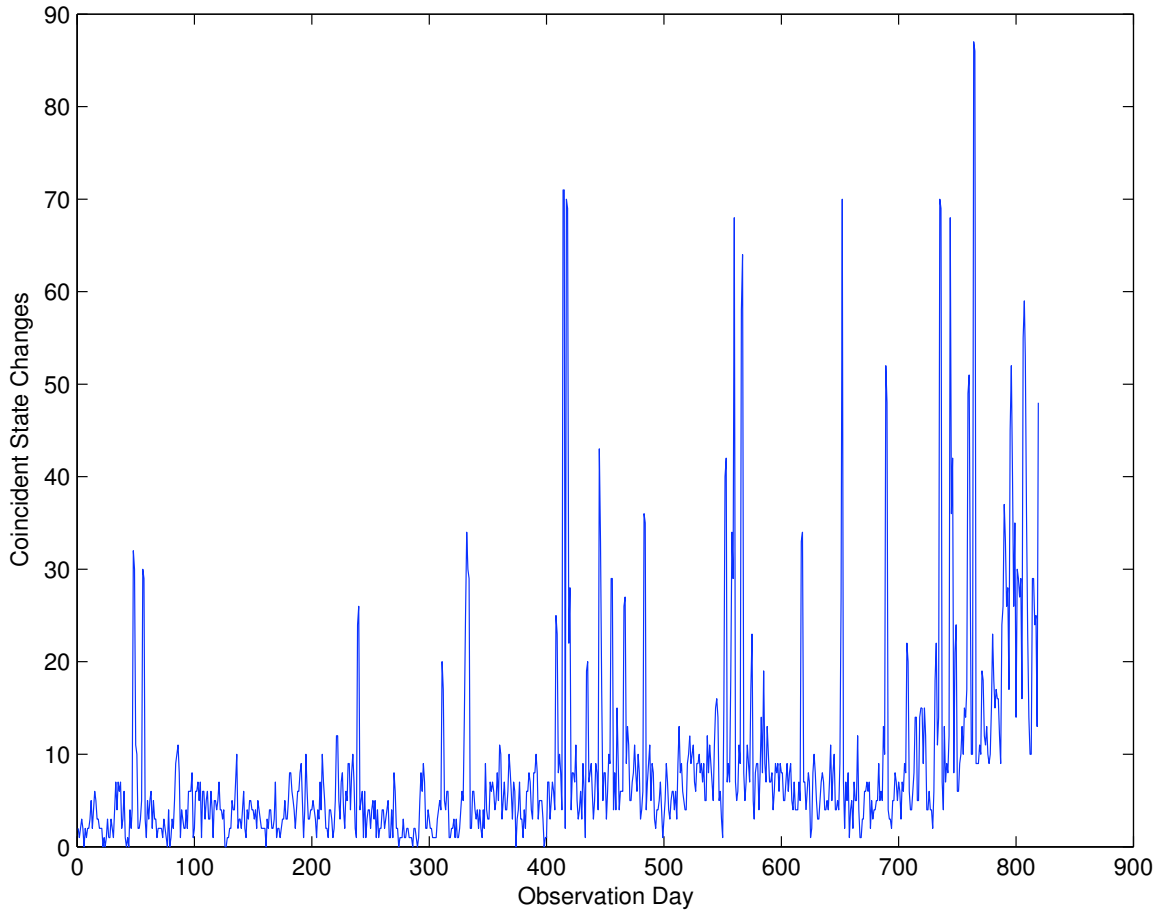


Majority Solution

SCIGN Data

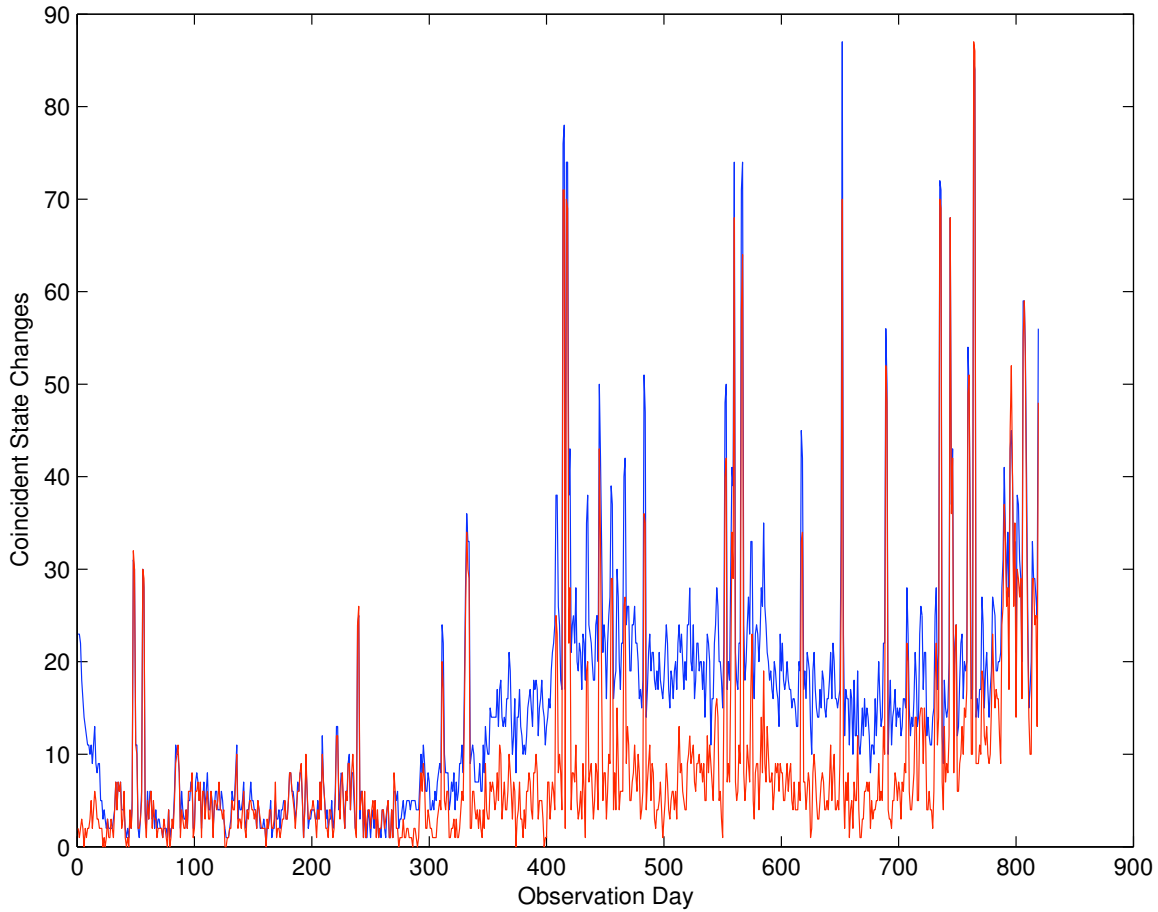
- GPS-integrated measurements of 3-D surface displacement calculated daily.
- Data collected over 800 days starting Jan 1, 1998 from 127 stations.
- Trained 6-state HMMs separately on each data sequence. Classified observations based on trained HMMs.
- Identified correlated state changes across different stations.

SCIGN State Change Correlations



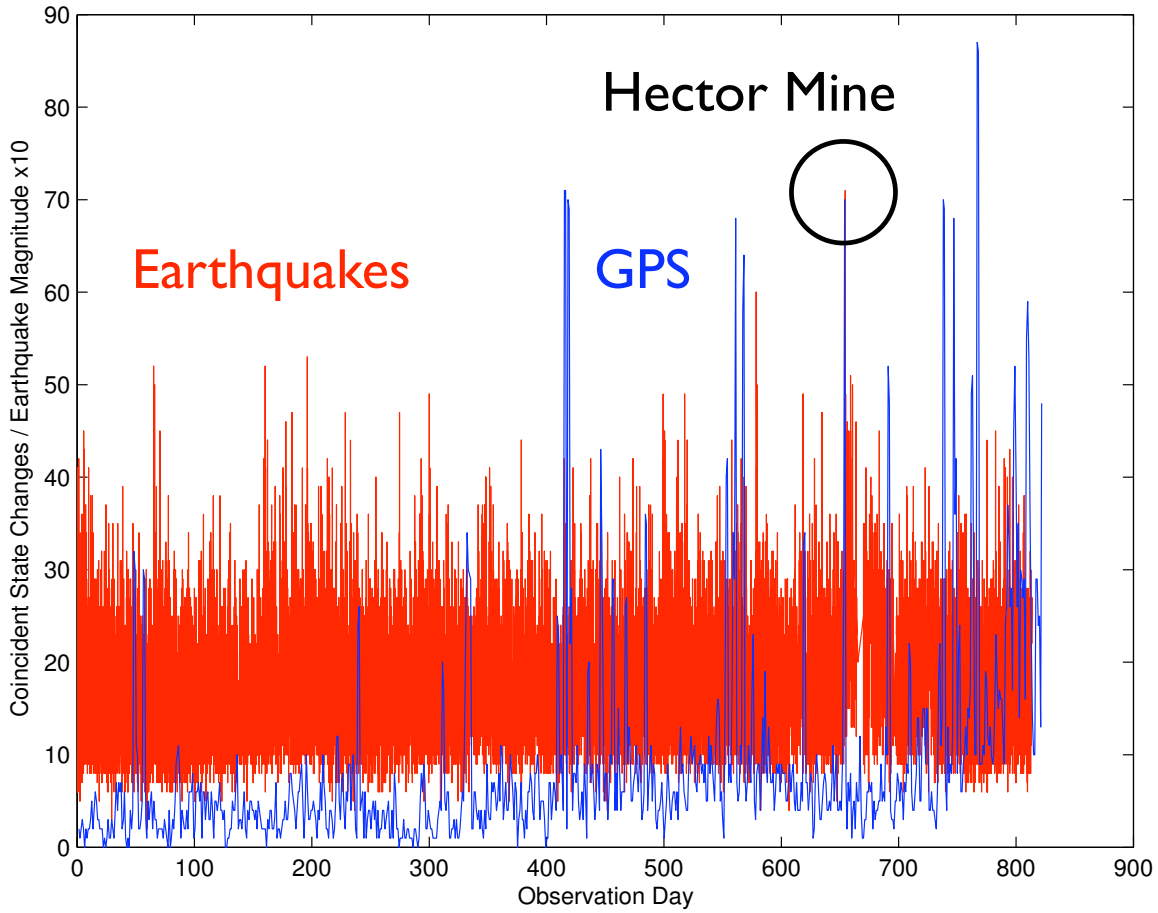
Spikes indicate the number of stations that change state on a given day (e.g. 72 stations change on day ~420)

SCIGN: EM and RDAEM Comparison



RDAEM (red) has greatly reduced noise.

SCIGN: Seismic Record Comparison



Only the Hector Mine is correlated with a SCIGN regional activity spike - aseismic activity?

Conclusions

- Regularization, along with deterministic annealing, can be used to find optimal hidden Markov model parameters for unconstrained science data.
- HMMs can be used to segment/classify individual GPS time series.
- By comparing segmented individual time series in a GPS network, correlations can be found that indicate behavioral changes across the region.

References

Granat, R. A., Regularized Deterministic Annealing EM for Hidden Markov Models, Doctoral Dissertation, University of California, Los Angeles, 2004.

Heki, K., S.-i. Miyazaki, and H. Tsuji, Silent fault slip following an interplate thrust earthquake at the Japan Trench, Nature, 386 (6625): 595-598, 1997.

Hirose, H., K. Hirahara, F. Kimata, N. Fujii, and S.i. Miyazaki, A slow thrust slip event following the two 1996 Hyuganada earthquakes beneath the Bungo Channel, Southwest Japan, Geophysical Research Letters, 26 (21): 3237-3240, 1999.

Melbourne, T. I., and F. H. Webb, Slow But Not Quite Silent, Science, 300: 1886-1887, 2003.

Miller, M.M., T.I. Melbourne, D.J. Johnson, and W.Q. Sumner, Periodic Slow Earthquakes from the Cascadia Subduction Zone, Science, 295: 2423, 2002.

Rogers, G. and H. Dragert, Episodic Tremor and Slip on the Cascadia Subduction Zone: The Chatter of Silent Slip, Science, 300:1942-1943, 2003.

Ueda, N. and R. Nakano, Deterministic Annealing EM Algorithm, Neural Networks, 11(2): 271-282, 1998.

Whiley, M. and D.M. Titterington, Applying the Deterministic Annealing Expectation Maximization Algorithm to Naive Bayesian Networks, Univ. Glascow Tech. Report, 2002.