Time Series Pattern Recognition for Classifying Seismicity Records and Geodetic Sensor Trajectories

Michael Turmon and Robert Granat and
Kenneth Hurst and Andrea Donnellan
Jet Propulsion Laboratory
California Institute of Technology

February 2002

Executive Summary:

This proposal is aimed at the topography and surface change challenge in the SENH NRA. We address items B.2.b, "Applications of Southern California Integrated GPS Network and similar geodetic networks," and B.1.a, "develop techniques which contribute to a better understanding of crustal dynamics for the mitigation of earthquake risk." We propose to construct statistical models of geophysical data, including GPS sensor trajectories and seismic history. The information learned from these models will enable breakthroughs including earthquake classification, automatic anomaly detection, sensor motion clustering, advanced signal modeling, and data cleaning. The modeling tools are general-purpose and will be released to the geophysics community along with derived data products.

Contents

1 Abstract

We address SENH research announcement items B.2.b, "Applications of Southern California Integrated GPS Network and similar geodetic networks," and B.1.a, "develop techniques which contribute to a better understanding of crustal dynamics for the mitigation of earthquake risk," by using novel statistical methods to learn models for, and to classify, seismicity and geodetic data. These methods include Hidden Markov Models (HMMs) which track the evolving state of a fault system or GPS signal, and Kalman filters, which will be used to classify GPS sensor responses and track microblock motion. These statistical approaches offer inherent confidence estimates and error bars so their predictions can be evaluated, or fed into other modeling systems. They allow investigators to extract meaning from Earth Science datasets and to focus on interesting and unusual behaviors.

The innovative science that will result from these methods will include:

Earthquake classification: Automatic and objective classification of events into categories such as aftershocks, swarm events, and foreshocks. We can then link events to underlying physical processes, and learn how processes interact over time. This information will improve hazard assessment.

Anomaly detection: Detection of mode changes in GPS signals to better understand when new behaviors arise. With this tool we can rapidly find and determine the extent of anomalous behavior.

Sensor clustering: Detection of related patterns of crustal motion to group GPS sensor sites into coherently-moving blocks, allowing the objective discovery of the extent and velocity of crustal microblocks.

Signal modeling: Modeling and understanding spatially linked distortions to GPS signals, such as tropospheric delays. Accounting for and removing these distortions improves signal to noise ratio in the calculated position estimates.

Data cleaning: Intelligent filling in of missing data segments to allow application of methods requiring continuous data.

The end product of our work will be innovative scientific analysis of at least three geophysical data sets along these lines, as well as the software tools we develop that make the analysis possible. This software will be general enough to be applied to many other related sources of geophysical data. To facilitate their eventual use in a wide variety of geophysical analysis efforts, our software tools will be distributed through the NASA/HPCC GEM web portal to the larger geophysical research community.

2 Technical Plan

This proposal addresses the *topography and surface change* challenge in the Solid Earth and Natural Hazards research announcement. It will contribute to understanding how the Earth's surface is being transformed and use this understanding to help predict future changes and assess natural hazard. Our project will further NASA's mission by enhancing the application of space-based technologies. To do this, we propose to construct *models of geophysical data* including GPS sensor trajectories and seismic history. The information learned from these models will make possible numerous breakthroughs:

Earthquake classification: Automatically and objectively classify events into categories such as aftershocks, swarm events, and foreshocks. We can then link events to underlying physical processes, and learn how processes interact over time. This information will improve hazard assessment.

Anomaly detection: Detect mode changes in GPS signals to better understand when new behaviors arise. With this tool we can rapidly find and determine the extent of anomalous behavior.

Sensor clustering: Detect related patterns of crustal motion to group GPS sensor sites into coherently-moving blocks, allowing the objective discovery of the extent and velocity of crustal microblocks.

Signal modeling: Model and understand spatially linked distortions to GPS signals, such as tropospheric delays. Accounting for and removing these distortions improves signal to noise ratio in the calculated position estimates.

Data cleaning: Intelligent filling in of missing data segments to allow application of methods requiring continuous data.

2.1 Objectives and Significance of the Proposed Work

The Historical Moment The earthquake research community is standing at the edge of a watershed in the quantity of available data. Consider the situation in Southern California: With the completion of the SCIGN network in June 2001, Southern California now has 250 continuously recording GPS stations. Trinet and the Southern California Seismic Network now provide seismic data which is unprecidented in both quality and quantity. InSAR measurements are starting to become available which provide detailed deformation information with pixel spacings of tens to hundreds of meters. Active seismic experiments contribute a wealth of information about the subsurface structure. Other areas of the world such as Japan and Taiwan are also experiencing a buildup of measurement capability.

Automated techniques are necessary to assist in coping with the deluge of information. They have been effectively deployed in other domains: discovering high-redshift quasars in astronomy [9], locating volcanoes in planetary science [5], reliably identifying sunspots in solar physics [29], and uncovering new dynamics in atmospheric modeling [27]. These techniques help assist scientific understanding and hazard assessment in a number of ways: they can *analyze large quantities of data* that would overwhelm human analysts, they can *find subtle changes* in the data that might evade an unassisted human expert, they allow investigators to *focus on interesting phenomena* by modeling typical behavior and noting deviations from it, and they *assist in objective decision making* in cases where even experts disagree (for example, identifying aftershock sequences, or modes in GPS signals). If systems of significant generality are developed, other geophysics investigators can *re-use the same model family and fitting algorithms* on their data. These techniques are not expected to replace human analysis, but to be tools for human experts to use as part of the research cycle.

Hidden-Variable Models The organizing principle of the proposed work is a general state-space framework for modeling, classifying, and predicting the dynamics of an underlying geophysical process. (See figure 1.) Earthquakes and other seismic events are the observable product of a complex, nonlinear evolution within a high-dimensional phase space, shown schemati-

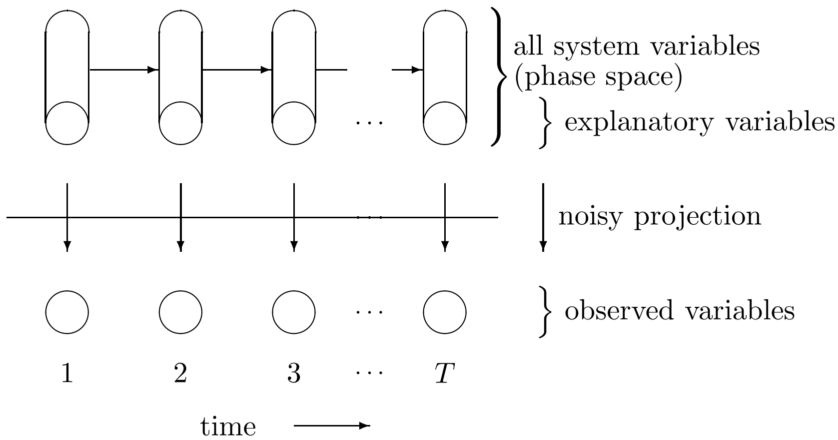


Figure 1: System phase-space evolution, partially observed through a noisy projection

cally in the top of the figure. This underlying system is not fully observable — indeed, most of the state information is relatively uninformative about natural hazards. At each time step, the system evolves and a new snapshot of its behavior is taken. These snapshots are a noisy projection of the full system state into a vector of just a few components, as depicted in the lower part of the figure.

Observable patterns are emergent processes that reflect the structures, dynamics, and properties of the underlying high-dimensional system. The job of the analyst is to use the observations to look back into the physical system and gather information about it — not to infer the full state, but just the explanatory variable that is most related to the processes being studied.

For example, during an earthquake the system jumps to a radically different part of phase space, and this transition can be explained by a change in the value of a discrete explanatory variable. In the example of the SCIGN network, with 250 stations it is necessary to develop automated tools for classification of the signals. The signals of interest must be coherent between stations in the light of a geophysical model. We know of several sources of coherent and incoherent signals in the SCIGN data. The signal of primary interest is of course generated by slip on a fault or collection of faults. We have also observed coherent signals which we attribute to ground water pumping and recharge. Incoherent or coherent signals could come from tropospheric effects, and possibly thermal effects. It is not possible using unaided human inspection to identify and understand all the signatures in the SCIGN data. We need automated generic tools to aid in identifying and classifying the signals we are observing.

Proposed Work We propose to develop and use state-of-the-art, high-performance pattern recognition techniques to analyze geophysical data in this framework — using various data sets to infer both the values of explanatory variables and the models that link them to observations. The proposal targets seismic record and GPS data, but the methods are general and will provide a starting point for investigation of other time-dependent geophysical data such as gravity or interferometric synthetic aperture radar (InSAR). The resulting tools will be provided for the use of the geophysics research community. These techniques can provide not only understanding of the physical processes that generate earthquakes, but also the potential for new classes of practical earthquake forecast algorithms.

For definiteness, consider a sample application: the classification of catalogs of seismic events using seismic record data. Earthquake activity in a geographic area can be regarded as a time series, with each seismic event representing a point in the series. Each event has observable quantities such as location and magnitude associated with it, but each is also tagged with one of a number of discrete hidden states representing the *current underlying pattern of activity.* These hidden states group earthquakes that occur as a result of some particular physical process: activity along a particular fault, stress buildup in a region, or aquifer activity. The grouping of seismic events according to the value of the hidden state allows scientists to investigate the underlying physical processes indicated by those states. Furthermore, analysis of the transitions between states uncovers relationships between physical processes. For example, activity along one fault may be statistically related to activity along another fault.

In summary, we propose to develop general-purpose tools for handling salient objects as they evolve in time within massive spatiotemporal datasets. We identify *four thrusts* in this effort:

- Development of discrete-valued hidden Markov models for identifying hidden classes governing seismicity and GPS data.
- Development of continuous-valued Kalman filter models for modeling, and then classifying, GPS sensor trajectories.
- Release of software to allow other practitioners the ability to perform these analyses on their own data.
- And finally, in the second year, to integrate the discrete and continuous variable modeling to better model and classify seismic activity.

2.2 Impact of Proposed Work

This investigation will use hidden-variable techniques to accomplish several geophysical analysis tasks:

- Earthquake classification: Automatically and objectively identify sets of events as, for example, aftershocks, swarm events, and possibly foreshocks. See figure 3.

- Anomaly detection: Detect mode changes in GPS signals to better understand when new behaviors arise. See figure 4.

- Sensor clustering: Detect related patterns of crustal motion to group GPS sensor sites. See section 2.3.3.

- Signal modeling: Model and understand distortions to GPS signals, such as tropospheric delays. See section 2.3.3.

- Data cleaning: Intelligent filling in of missing data segments to allow application of methods requiring continuous data. See figure 6.

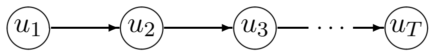
In performing this analysis, we will not do one-off experiments but develop the technology and software for learning models relating explanatory variables to geophysics time series. This technology will allow researchers to better understand data from new observation systems and large-scale simulations that are now becoming available. At the conclusion of this scientific analysis, we will distribute this software and integrate it into GEM (General Earthquake Models), a new NASA/HPCC-supported modeling and analysis facility. The tools we develop are also applicable to other domains within the purview of natural hazards assessment, such as studies of volcanism. See also section 2.3.4.

2.3 Technical Approach and Methods

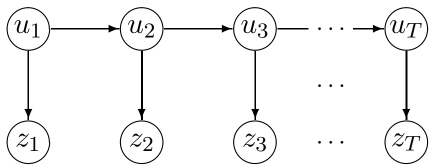
To establish the technical basis for this work, we first introduce its fundamental basis, hidden-variable models for geophysical time series. We describe the applications of these models to geophysical data analysis and hazard assessment, and our plan for embodying the components via open software tools.

2.3.1 Knowledge Discovery with Hidden-State Models

As outlined in the introduction, our fundamental outlook is to model the *unobserved state* of evolving phenomena, and then to use the resulting state estimates to classify phenomena, to learn about their dynamics, and to group their dynamics into similarly-behaving classes. The key recent advance allowing use of these explanatory models is the development of computational



The basic Markov chain.



Partially observed Markov chain.

Figure 2: Fundamental models for dynamic data, with hidden variables u representing underlying states of the physical system, and observable variables z.

methods for acyclic directed graph (ADG) statistical models, also known as Bayesian networks [21, 15]. We can use these networks to specify any of a large family of related hidden-variable structures, which can account for many patterns of temporal evolution and incorporation of related data. And then, as we outline below, the hidden states of the network can be learned from the observed data using standard methods.

General-purpose models The ADG is a graph-based formalism for representing a multivariate probability distribution; it is specified by selecting a model structure G and distributional parameters. The ADG notation was developed to concisely encapsulate the pattern of dependence encountered in real applications, such as medical diagnosis [28], and has since been extended to many domains. The first step in setting out a model structure is to identify the collection of variables, observed and not, to be modeled — each variable corresponds to one in the graph G. Then, the chain of dependence is set out by drawing directed links between pairs of variables that are related, running in direction of causation. For geophysical data sequences, this typically means chains of dependence ordered by time as in figure 2.

The formalism incorporates numerous well-known models as special cases. For example, a Markov model (figure 2, top) is an ADG across time where $u(t)$ directly depends only on $u(t-1)$. In Kalman filters and hidden Markov models (figure 2, bottom), the evolution of the state $u(t)$ is seen noisily through the observations $z(t)$; typically one wants to recover the hidden state. This graph is the mathematical expression of the schematic diagram in figure 1. The hidden state variable can take on various physical meanings. For example, a discrete state variable might represent a type of earthquake (aftershock, mainshock, swarm event, and so forth) while a continuous state variable can represent various physical attributes, such as position and velocity. The continuous model can be generalized by augmenting the state space with other evolving but measurable covariates such as tidal forces, ground water levels, and so on.

In either case, the important quantities are the probabilities

$$Pr(u(t) \,|\, u(t-1)) \quad \text{(state evolution)} \qquad (1)$$

$$Pr(z(t) \,|\, u(t)) \quad \text{(output distribution)} \quad . \qquad (2)$$

If there are, say, N discrete states, the state evolution (1) is quantified by the $N \times N$ matrix A of state-transition probabilities, while the observations (2) are described by the N output distributions for $z(t)$. When the state $u(t)$ is continuous, the state evolution (1) is Gaussian with a mean and covariance depending on $u(t-1)$, and the output distribution is again Gaussian. We explain below how these system parameters may be learned from data, but inferring the hidden variables is the more basic problem.

Estimating variables The first fundamental problem arising from the use of ADG models in scientific problems is inference, or the problem of calculating estimates for variables of interest in the presence of observed data. Here, this means finding the maximum a posteriori (MAP) estimate of the most likely state of the unobserved variables of the partially observed Markov chain, given the observation:

$$\arg \max_{u(1),...,u(T)} Pr(u(1), ..., u(T) \,|\, z(1), ..., z(T)) \qquad (3)$$

In the continuous model, this estimate corresponds to the trajectory of the sensor, and is implemented for this specific case by the well-known Kalman smoother [16, 23].

The ADG representation is so useful because there exists a complete computational framework for solving the inference problem for any graphical model [15]. The usual methods for time series analysis are special cases within this framework, e.g. the forward-backward and Viterbi algorithms for hidden Markov models [26], and filtering and smoothing operations in Kalman filters [18]. Furthermore, the computational complexity of all these methods scales linearly with T, the number of samples, so the computations are tractable even for very long series.

Learning models While the structure of the graph tells which variables depend on each other, the *distributional parameters* quantify how the dependence works via the probabilities (1) and (2). As mentioned, either discrete or conditionally Gaussian distributions are used here, and these free parameters are generically denoted θ. The second fundamental problem of finding these models then arises: estimating the distributional parameters from data, which is only partially known via $z(1), ..., z(T)$. The standard criterion is maximum-likelihood,

$$\hat{\theta} = \arg\max_{\theta} Pr(z(1), ..., z(T); \theta) \qquad (4)$$

where $Pr(\cdot; \theta)$ indicates that the probability depends on the parameters θ. Intuitively, this means varying θ to best explain the observation z.

Since G contains unobserved variables, we use the Expectation-Maximization (EM) algorithm to estimate the parameters [7, 19]. EM is a general technique for iteratively computing maximum likelihood estimates of parameters in the presence of hidden variables $u(t)$. This method is well established in this context, e.g. for its application to Kalman filtering see [24]. It is stable and guaranteed to increase the likelihood at every iteration.

In the context of hidden-variable models, each iteration of EM contains a step (the "E step") which estimates the model's hidden states; this is the MAP estimate described above. This is followed by an "M step" that determines parameter values for that combination of observed and (estimated) hidden data. The parameters are then updated and the cycle continues until convergence. Similar to the case of the inference algorithms outlined above, the scaling is linear in sequence length T for hidden Markov and Kalman filter models. Typically, to determine on the order of tens of parameters from time series on the order of hundreds of samples, the EM algorithm will require on the order of hundreds of iterations to complete. This takes a few minutes on a modern workstation.

Checking models We use two main methods to validate our results. The first uses residuals and goodness of fit tests to check fitted models, mainly for internal consistency. Specifically, in order to fit a certain observation sequence, a corresponding most-probable sequence of hidden variables are inferred from (3). This sequence is governed by a model telling how it evolves, and its adherence to the model can be tested. For instance, the continuous-state Kalman filter variables must evolve according to an autoregressive, Brownian-type additive feedback with given statistics. In the event of a mis-specified model, the hidden variable sequence will need to be distorted to fit the data that was actually observed, and the residual tests will detect this. (For an example, see figure 6.)

The second main method of checking results is cross-validation [25], which can be used for end-to-end model testing. In general, cross-validation works by splitting the available data into two parts, one for training and one for testing. The training set is used to select a model, and the model fit is evaluated on the test set. This split-train-test sequence is repeated several times for different splits to gauge model robustness. This is done partly qualitatively, but also by evaluating the fit, $Pr(\text{data}_s \,|\, \text{model}_s)$ for each split s. This provides a way to judge the relative performance of two different models, when one has greater complexity than the other.

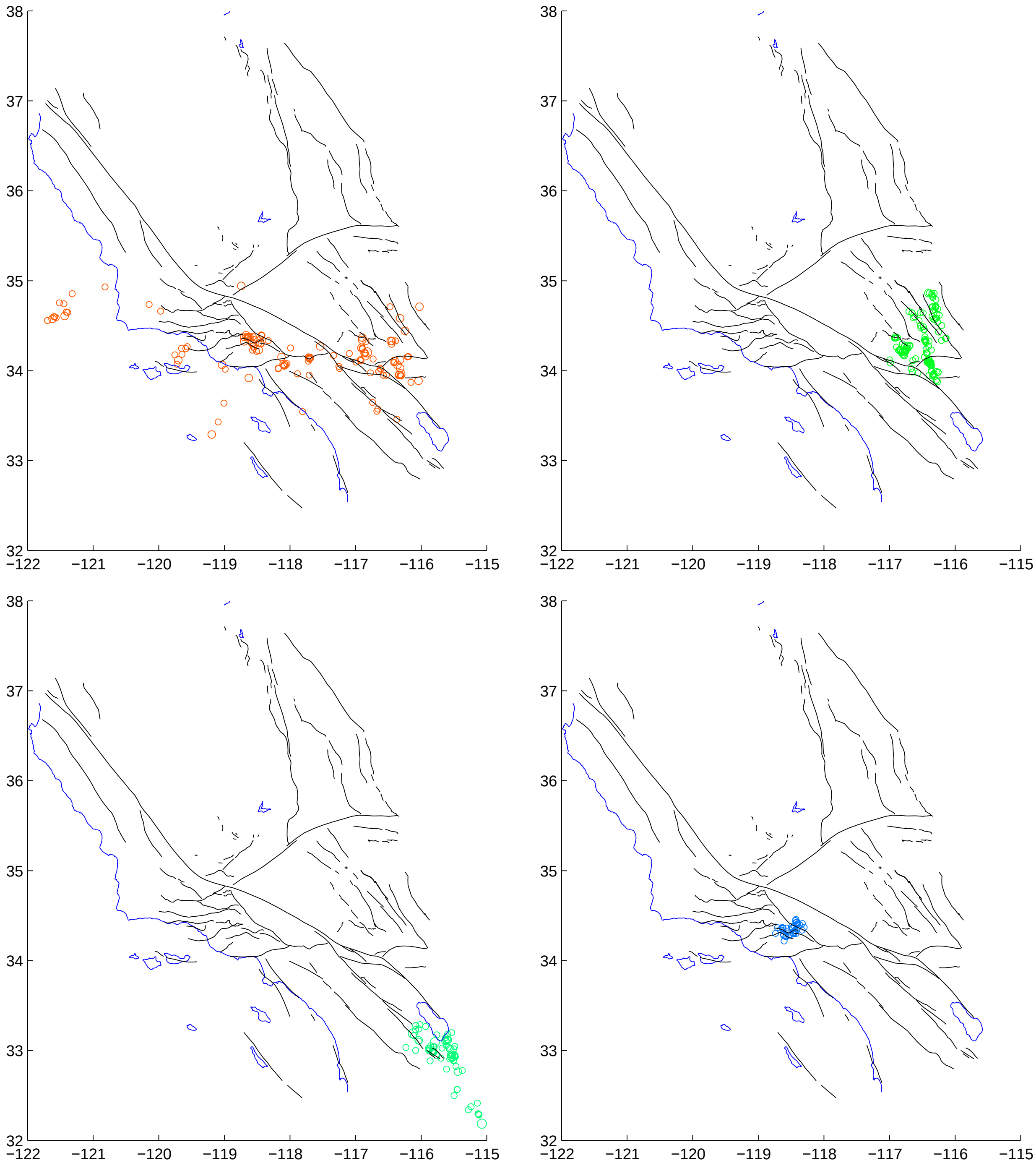


Figure 3: Preliminary HMM analysis results for SCEC catalog seismicity data. Upper left: the class of Transverse Range events; upper right: the class of Hector Mine and Landers earthquake aftershocks; bottom left: the class of Salton Sea swarm events; bottom right: the class of Northridge earthquake aftershocks.

The next subsections describe how these inference and model training procedures work for seismic record and GPS sensor analysis.

2.3.2 Seismic Record Analysis

Numerous seismic records exist that contain, at a minimum, the location, magnitude, and time of earthquakes that occur within a particular region over some period of time. These records are a key source of information for understanding patterns of earthquake behavior, fault interactions, and the underlying stresses in the Earth's crust.

We propose to conduct an analysis of seismic record data using hidden Markov models (HMMs) [1, 2, 3, 4, 22], which as noted above, are a special case of our general dynamical system modeling framework. Preliminary work performed utilizing HMMs for analysis of the Southern California seismic record freely available from the Southern California Earthquake Center (SCEC) led to some very encouraging early results [13, 12].

Preliminary Results The SCEC catalog covers several decades of events, but for these experiments we considered only events between 1960 January 1 and 1999 December 31. For this time period the catalog is considered complete for events of magnitude three and above. Directly from the record, we extracted four attributes for each event: latitude, longitude, depth, and magnitude. We also calculated two derived attributes based on the event times: the time to the next event, and the time since the previous event.

Example results from a preliminary experiment for a seventeen state HMM are presented in figure 3. For this experiment, we considered all events of magnitude greater than four. Circles indicate the location of earthquakes; circle size corresponds to (floored) magnitude. Black lines represent major faults, and blue lines represent coastlines and bodies of water.

In the figure, we see that data was grouped into meaningful classes and the technique did associate earthquakes with underlying physical processes. The four classes shown represent clusters of aftershocks for the Hector Mine, Landers, and Northridge earthquakes, Transverse Range events, and Salton Sea area swarm events [17].

Methodology These results were obtained as follows. We view a seismic record as a time series, with each event representing a point in that time series. We used an HMM with the graphical form of figure 2: a single discrete internal state variable u controlling a six-dimensional observable z. The characteristics of the events become the observed values $z(1), ..., z(T)$, and we are interested in the hidden values $u(1), ..., u(T)$. Each of the $N = 17$ discrete states $u(t) = i$ is associated with a set of N transition probabilities $A_i = \{a_{i1}, \ldots, a_{ij}, \ldots, a_{iN}\}$, and a probability distribution of observable outputs b_i. The transition probabilities determine the probability of the next state given the current one.

Given this structure, we train the HMM: calculate the model parameters $\theta = (A, b)$ that satisfy (4). Then, given these model parameters, we define the optimal state sequence via (3) as that in which the state at each point in time is most likely. The trained HMM is used to determine the optimal state for each point in the time series; the events are classified according to their associated HMM states. Often, the HMM is trained once with a subset of the data, and the resulting model fixed and used to determine the hidden states for new data.

First, given the state sequence, we can classify the events in a seismic record according to their type. The properties of each class are described by the output probabilities of the associated state. Second, given the trained model, we may analyze the pattern of interactions between different classes of events. These dynamic interactions between types of events are revealed through inspection of the transition probability matrix A. For example, a high probability of moving from one class to another implies a potential causal link between the two classes of events.

Application and Development These ideas were tried on a pilot basis for this data, but the more thorough analysis we propose should reveal more about the nature and dynamics of seismic events in Southern California. Also, we propose to use the same methods on the Northern California seismicity record to look for common patterns and test the generality of the underlying method.

The HMM method in this experiment classified events into categories, but the number of states $N = 17$ was chosen empirically. We propose to equip the developed HMM analyzer to *automatically determine the number of event classes* in the data with the idea of *cross-validation* presented in the preceding section, an idea used successfully in another geophysical application [27]. Cross-validation requires splitting the data into two sets, for training and testing. The simplest way is to train the model on contiguous subsets of the data, and test it on a disjoint contiguous subset. Alternatively, it may be more informative to remove the time element for the purposes of state-number determination, and first model the data using a non-dynamical finite mixture model (FMM). In the FMM case, there is no dynamical state dependence — the underlying state of each observation is chosen independently. Because the FMM discards the time information, cross-validation can be performed by splitting the aggregated time samples into training and testing sets which need not be contiguous. This will give us another means to estimate the number of states based on the distribution of the observations in the parameter space.

Our method will not only require minimal or no parameter tuning, but will also be able to return consistent results across experiments; this frees the investigator from having to examine numerous candidate results. The expectation-maximization (EM) method used for optimization of the HMM parameters can be sensitive to the initial parameter estimates used to begin the E-M iteration. This means that with a standard implementation, the results may vary from experiment to experiment on the same data, depending on the initialization method. This is due to the existence of many local optima in the likelihood function, and for many applications, the degree of variance in the results is low. However, it is desirable to extend the applicability of these methods to situations where the data are noisy, sparse, or incomplete because of limitations of the data collection equipment. We propose to address this kind of data by employing a technique known as deterministic annealing [30].

The deterministic annealing approach makes use of ideas from statistical mechanics to break the main optimization problem down into into a series of sub-problems. Each sub-problem is characterized by a computational temperature parameter which modifies the maximum likelihood optimization criterion: the higher the temperature, the fewer local optima are present. The first sub-problem is solved at a very high temperature, so that the distributions of the output and transition probabilities turn out to be close to uniform. Then the next sub-problem is solved at a slightly lower temperature, so that the distributions of output and transition probabilities are slightly closer to that of the original problem, this time using the previous sub-problem's solution as the initial condition. This continues until the last sub-problem is completely "cool" and is, in fact, identical to the original problem. Annealing regularizes the parameter estimates and helps avoid non-optimal solutions: at the beginning of the process, all the components of the solution are similar. As the temperature decreases, the influence of each observation is gradually localized. Because each new global maximum is close to the old one as the temperature decreases, the solution is able to track it across temperature steps. At the same time, finer structure, closer to that of the true model, gradually emerges in the fitted parameters.

2.3.3 Continuous Trajectory Modeling

In a second major thrust of this proposal, we will detect patterns and new features in GPS data coming from SCIGN, a new network of over 250 GPS sensors [14] that is analyzed to define the relative position of the sensors to millimeter accuracy. In this section we outline several avenues of scientific research built around the understanding offered by hidden-variable models.

Preliminary Results An application of the HMM methods to this data is shown in figure 4. In this case, the observed variables z are the three-dimensional sensor position through time, and the hidden variable is a five-state mode.

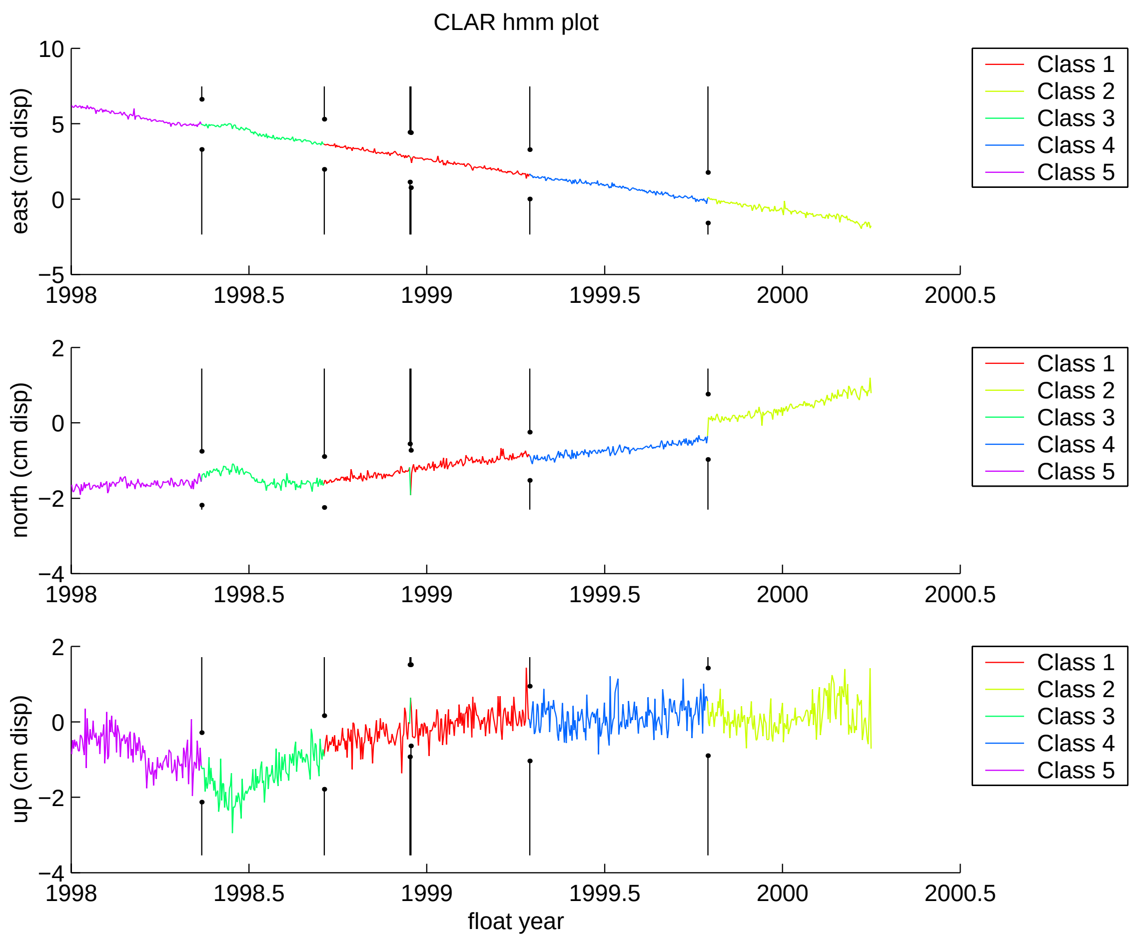


Figure 4: HMM segmentation of a GPS signal recorded in Claremont, California. Note that the HMM has distinguished between different modes of the signal, including behavior in 1998 resulting from local ground water pumping, and the 1999 Hector Mine earthquake

When there is a shift in the observed displacements, it shows up as a changed mode $u(t)$. This provides one principled way to detect changes in the pattern of output exhibited by a sensor. For instance, both the 1998 water-pumping signals and the 1999 Hector mine earthquake are associated with a change of mode.

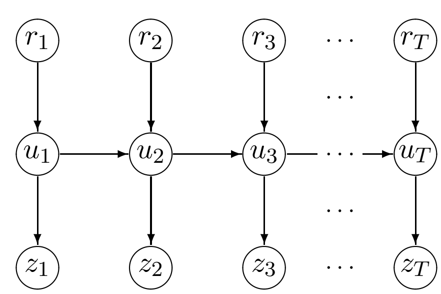


Figure 5: Kalman filter with exogenous inputs

One difficulty with the pure HMM method of change detection is that it assumes homogeneity of the time series *between* state changes. For example, over the course of a red interval of figure 4, a single type of output is observed, governed by $P(z(t) \,|\, u(t) = \text{red})$. This is approximately true over moderate time scales, but over long periods the output will drift and a new discrete state will be introduced to allow for the evolving behavior of outputs. This drift shows up as residuals having systematic trends (see section 2.3.1), and tells us the model is incomplete.

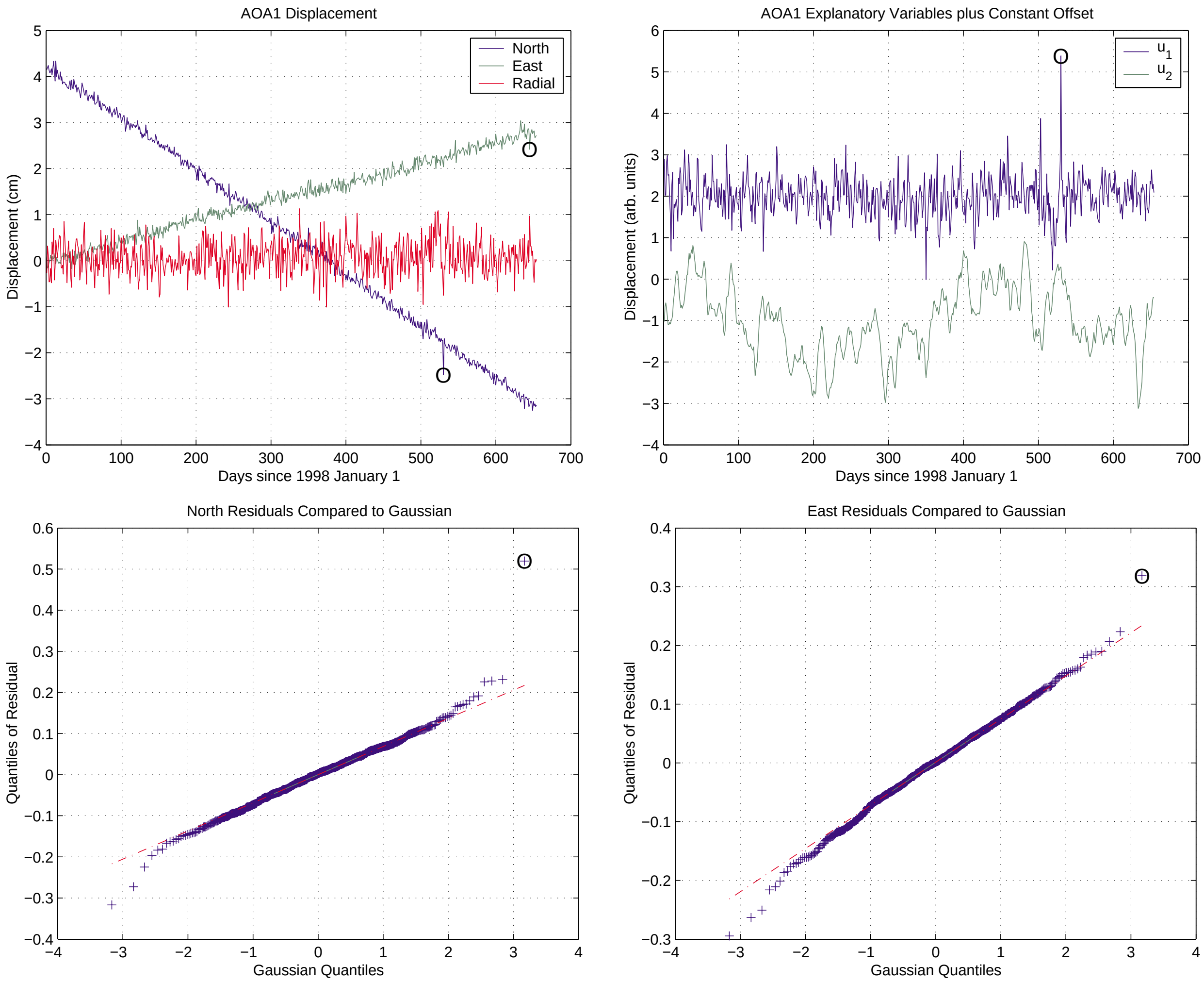


Figure 6: Kalman filter models for continuous data from SCIGN station AOA1; analysis shows slowly varying cyclic sensors motion and identifies two outlying points. The displacements (top left) are modeled over time by estimating the two hidden variables (top right). One of these variables is a high-frequency Brownian-type process which models fast variations in the position, while the other hidden variable models slower trends at the time scale of months to years. Residuals (north at bottom left, east at bottom right) show two deviations from Gaussianity at times 530 and 645 (circled).

Methodology This *pattern of continuous sensor motion* can be modeled by a continuous state-space model, or Kalman filter. As we see below, there are several ways to do this, depending on the information we want to extract from the data. We initially propose a model of the form of figure 5 for sensor trajectories — in essence, a known forcing term r influences a hidden state u which in turn produces outputs z.

The smoothly changing sensor positions are captured via a continuous hidden variable $u(t) \in R^p$. For example, $u(t)$ might contain the true sensor position and velocity, neither of which is measured directly. In general, $u(t)$ is an embedding of the local physical process in a p-dimensional phase space.

The Kalman filter models the evolution of u

and the observations z as the recurrence

$$\begin{aligned} u(t+1) &= Au(t) + Br(t) + e_u(t) \\ z(t) &= Cx(t) + Dr(t) + e_z(t) \quad . \end{aligned} \tag{5}$$

The crucial point is that, as shown in figure 5, $u(t+1)$ depends only on $u(t)$, the external forcing $r(t)$, and an independent error term $e_u(t)$. Similarly, the observation $z(t)$ depends only on the state, the forcing term at that instant, and another error term $e_z(t)$. (Both noise terms are independent across time steps.) This simple dependence permits the computational simplicity of the Kalman filter, which updates $u(t)$ incrementally based on the corresponding values of z and r. This ability to predict new states, and to measure the deviation of an observed state from its expectation, is useful as an indicator of novel geophysical activity: see figure 6.

The ability to model sensor trajectories is not of use unless the parameters $\theta = (A, B, C, D)$ can be learned for a given sensor via (4), which for this problem means solving

$$\max_{A,B,C,D} \log Pr(z(1), ..., z(T); A, B, C, D) \quad . \tag{6}$$

This is just a Gaussian probability distribution but it depends on the matrix parameters in a complicated way. Complex gradient descent schemes for (6) were known [6, ch. 7] but did not reach wide use; more recently, simpler and more flexible EM algorithms were introduced for the problem [24]. These methods and their extensions have had great recent success in tracking visual objects [20], speech recognition [8], and other applications. The simplicity of the EM methods hinges on their use of the Kalman smoother as a *subroutine* to estimate the hidden variables u based on the observations; these estimates are in turn used to update the parameters.

Preliminary Results Figure 6 shows an example of a SCIGN trajectory that is modeled by a two-state Kalman model, where matrix parameters A, C, and D were learned from data using (6); B was taken to be zero. The forcing term $r(t)$ has two components: one equal to unity for all time to allow for an arbitrary position offset, and the other increasing linearly with time to account for a constant velocity. The coefficients of D thus represent the sensor velocity. Just under 100 EM iterations were required for convergence.

The top left panel shows the original displacements for 654 days from 1998 January 1 to the Hector Mine quake, and the top right is the two explanatory variables. It is apparent that the first variable models high-frequency displacement changes, on the order of one to two days, while the other variable models longer trends, which appear to have at least two superimposed periodicities. The residual, the difference between the predicted observation and the actual one, is very close to a white Gaussian sequence, as shown by the close adherence of the two lower quantile plots to a linear pattern. This means first of all that Kalman model fit to the data is self-consistent with its assumptions. Two outliers are indicated as circles in these plots; they correspond to small disturbances at day 530 (large North residual, also big enough to affect the upper right plot) and day 645 (large East residual) in the original time series.

This finding amplifies the use of these models as sensor-specific *indicators of unusual behavior* — this behavior shows up as a non-Gaussian residual on some days. It is also apparent that the estimated model can be used with (5) to *fill in missing values* consistent with the expected sequence behavior. This translates into the hidden-variable framework as finding the values ($z(t_1), ..., z(t_2)$, say) that are most probable given the other data, $z(1), ..., z(t_1 - 1)$ and $z(t_2 + 1), ..., z(T)$. This would be useful as a preprocessing step for methods that require contiguous data sets.

Research Initiatives One difficulty with the pure HMM method of change detection is that it assumes homogeneity of the time series *between* state changes, as seen in figure 4. This limitation can be fixed by using a mixed continuous/discrete system state $u(t) = [u^1(t)u^2(t)]$, where u^1 is the integer mode and $u^2 \in R^p$ is the continuous state. Now a combination Kalman filter/forward-backward algorithm can be used to estimate the unknown parameters [10] using the same EM algorithm methodology. This flexibility indicates the power and generality of the dynamical modeling methods we will apply to the SCIGN GPS data.

Another analysis suggested by these results is the *identification of tropospheric delay errors* by identifying correlations in the residuals across sensors. Such delays should affect nearby sensors in the same way, by perturbing their values away from the model prediction. A spatially covarying position residual would be one indicator of such a delay artifact.

We also propose to *determine microblock constituents* using SCIGN trajectory models. In the GPS setting, sensors across crust boundaries will have different motion patterns; we expect these patterns to cluster into similarly-moving groups at the spatial level of microblocks. GPS stations on the same block will by definition rotate coherently with the block, while stations on different blocks will present different motions. These similar motion patterns will show up as similar state-space models (A, B, C, D), which provides one way to group sensors. Alternatively, in a more sophisticated setup, a block tag would be another (integer) hidden variable. The value of this hidden variable would be learned by analyzing a group of sensors motions together. The cross-validated model selection methods, moreover, would provide an objective count of the number of distinct microblocks.

As a related example, it is anticipated that station velocities change with time throughout the earthquake cycle. We would like to test the theory that there exists a hidden *continuous* mode variable which indicates the relative time within the earthquake cycle of a given site and fault. This dependence can be captured by introducing such a variable, influencing the hidden state, modeling the time since the last energy release. This mode variable could be inferred due to its influence on the station velocity, aiding the assessment of earthquake hazard. This could all be done within the framework presented above.

2.3.4 Project Vision and Extensions

These methods will allow the geoscience community to take advantage of emerging technologies in the data mining and statistical modeling communities. Two modes of operation suggest themselves. In the first, the scientist wants to find instances of something whose features are known and recognizable. For example, one might try to find all instances of strike-slip faulting in an interferogram. This mode of operation is probably quite familiar in concept to most geoscientists. In the other mode, the scientist may ask to find any interesting features, patterns, or modes that might be in the data without specifying a priori how they will be recognized: this is the "discovery" setting. When a pattern or feature is identified, the scientist is presented with a challenge to explain it.

This type of exploring is likely to lead to breakthroughs. It is also at risk of turning up several dry leads for every breakthrough. For this reason, the technologies we advocate above are automated and easy so that it is possible to sort quickly through many of them. The models are flexible and extensible so that several alternative hypotheses can be formulated and tested against each other by examining residuals and cross-validation. This generality is possible due to recent advances in the technology used to infer explanatory variables from data in a statistical setting. These models provide a way to use advances in computer cost and power not just to yield more data, but to give scientists the means to explain it.

This project will take place in the context of the General Earthquake Model (GEM) project, funded largely through NASA's HPCC program. The GEM project is developing a web-based problem solving environment for modeling long period earthquake processes. It will include:

- A database system for handling both real and simulated data.
- Fully three-dimensional finite element code with adaptive mesh generator capable of running on workstations and supercomputers for carrying out earthquake simulations.

- Inversion algorithms and assimilation codes for constraining the models and simulations with data.

- A collaborative portal (object broker) for allowing for seamless communication between codes, reference models, and data.

- Visualization codes for interpretation of data and models.

We intend to integrate the results of this project into the GEM effort, and through it the larger geophysical research community including SCEC. This will provide a mechanism to make the tools we develop available to the general public, and will also allow us access to many datasets through the GEM database system. It will also allow us access to simulation tools with which we will be able to test and/or train the algorithms which we will be developing under this project.

3 References

[1] L. E. Baum. An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes. *Inequalities*, 3:1–8, 1972.

[2] L. E. Baum and J. A. Egon. An inequality with applications to statistical estimation for probabilistic functions of a Markov process and to a model for ecology. *Bull. Amer. Math. Soc.*, 73:360–363, 1967.

[3] L. E. Baum and T. Petrie. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Stat.*, 37:1554–1563, 1966.

[4] L. E. Baum, T. Petrie, G. Soules, and H. Weiss. A maximization technique occuring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.*, 41(1):164–171, 1970.

[5] M. Burl, L. Asker, P. Smyth, U. Fayyad, P. Perona, J. Aubele, and L. Crumpler. Learning to recognize volcanoes on Venus. *Machine Learning*, 1998.

[6] P. E. Caines. *Linear Stochastic Systems*. Wiley, 1988.

[7] A. D. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, B-39:1–37, 1977.

[8] V. Digalakis, J. R. Rohlicek, and M. Ostendorf. ML estimation of a stochastic linear system with the EM algorithm and its application to speech recognition. *IEEE Trans. Speech Audio Proc.*, 1(4):431–442, 1993.

[9] U. M. Fayyad and P. Smyth. Cataloging and mining massive datasets for science data analysis. *J. Comput. Graph. Statist.*, 8(3):589–610, 1999.

[10] Z. Ghahramani and G. Hinton. Variational learning for switching state-space models. *Neural Computation*, 12(4):963–996, 2000.

[11] Z. Ghahramani and S. T. Roweis. Learning nonlinear dynamical systems using an EM algorithm. In *Adv. in Neural Inform. Process. Syst.*, volume 11, pages 599–605. MIT, 1999.

[12] R. Granat and A. Donnellan. A hidden Markov model based tool for geophysical data exploration. In *PAGEOPH special issue (ACES meeting)*, 2001. Hakone, Japan.

[13] R. A. Granat and A. Donnellan. Deterministic annealing hidden markov models for geophysical data exploration. *Proc. 3rd Workshop Sci. Datamining*, pages 1–8, 2001.

[14] K. W. Hudnut, Y. Bock, F. Webb, and B. Young. The Southern California integrated GPS network (SCIGN). *Proc. Plate Boundary Observatory Workshop*, 1999.

[15] F. V. Jensen, S. L. Lauritzen, and K. G. Olesen. Bayesian updating in recursive graphical models by local computations. *Computat. Statist. Quarterly*, 4:269–282, 1990.

[16] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME, Series D*, 82:35–45, 1960.

[17] P. W. Kasameyer, L. W. Younker, and J. M. Hanson. Development and application of a hydrothermal model for the Salton Sea geothermal field, California. *Geo. Soc. Amer. Bull.*, 95(10):1242–1252, 1984.

[18] B. C. Levy, A. Benveniste, and R. Nikoukhah. High-level primitives for recursive maximum likelihood estimation. *IEEE Trans. Automatic Control*, 41(8):1125–1145, 1996.

[19] G. J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions*. Wiley, 1997.

[20] B. North, A. Blake, M. Isard, and J. Rittscher. Learning and classification of complex dynamics. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(9):1016–1034, 2000.

[21] J. Pearl. *Probabilistic reasoning in intelligent systems*. Morgan-Kaufman, 1988.

[22] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE*, 77(2):257–286, 1989.

[23] H. E. Rauch. Solutions to the linear smoothing problem. *IEEE Trans. Auto. Control*, 8:371–372, 1963.

[24] R. H. Shumway and D. S. Stoffer. An approach to time series smoothing and forecasting using the EM algorithm. *J. Time Series Anal.*, 3(4):253–264, 1982.

[25] P. Smyth. Model selection for probabilistic clustering using cross-validated likelihood. *Statistics and Computing*, 10(1):63–72, 2000.

[26] P. Smyth, D. Heckerman, and M. I. Jordan. Probabilistic independence networks for hidden Markov probability models. *Neural Computation*, 9(2):227–269, 1997.

[27] P. Smyth, K. Ide, and M. Ghil. Multiple regimes in Northern hemisphere height fields via mixture model clustering. *Jour. Atmos. Sci.*, 56(21):3704–3723, 1999.

[28] D. J. Spiegelhalter. Fast algorithms for probabilistic reasoning in influence diagrams, with applications in genetics and expert systems. In R. M. Oliver and J. Q. Smith, editors, *Influence diagrams, belief nets, and Decision Analysis*, pages 361–384. Wiley, 1990.

[29] M. Turmon, J. Pap, and S. Mukhtar. Statistical pattern recognition for labeling solar active regions: Application to SoHO/MDI imagery. *Astrophysical Journal*, Mar. 2002.

[30] N. Ueda and R. Nakano. Deterministic annealing EM algorithm. *Neural Networks*, 11(2):271–282, 1998.

4 Management Plan

We anticipate that work will begin in September 2002. The task proposal calls for about 0.7 work-years for both years of the project.

4.1 Investigator Contributions

The teaming of machine learning experts with practicing geophysicists provides a strong basis for accomplishing the research goals as well as for developing knowledge discovery software for other investigators to use.

M. Turmon Lead the investigation, perform Kalman filter analysis, and help design and implement modeling software.

R. Granat Perform HMM analysis, design and implement general model inference mechanism.

K. Hurst As a practicing geophysicist, will provide guidance concerning interpretation of patterns and statistics derived from this investigation. As Project Manager of the GEM project will oversee the integration of the tools from this project into the GEM environment.

A. Donnellan As Principal Investigator of the NASA/HPCC GEM project and as a practicing geophysicist, will provide guidance concerning interpretation of patterns and statistics derived from this project.

4.2 Milestones and Deliverables

Milestones are more detailed for the first year of the project. In the second year we will expand our analysis as detailed in section 2.3.4, and also seek out other data sources to illustrate the generality of our tools.

Y01, Q1 Event classification for Southern California earthquake record using HMM technology.

Y01, Q2 SCIGN modeling and initial clustering (say, 20 stations) from SCIGN 3.0 data.

Y01, Q3 Analysis of multiple GPS sensors using HMM to identify signal modes that occur within a microblock or across multiple microblocks using data that was cleaned using continuous-state SCIGN models from Q2.

Y01, Q4 SCIGN clustering using 100 stations; mixed continuous/discrete model for identification of behavior modes.

Y02, Q1 Report results regarding identification of SCIGN modes and seismicity at the AGU fall meeting.

Y02, Q2 Event classification for Northern California earthquake record using HMM technology to show generalization of methodology.

Y02, Q4 Distribute software via GEM portal. Analysis of results reported in journal publication.

Deliverables are the following:

- Analysis results: In the first year, we will perform and report analysis of SCIGN modes and seismicity activity clusters. These results will be reported in conferences, like AGU, and in publications.
- Software: The software we develop for modeling seismicity patterns and sensor trajectories will be released independently on the GEM web site for use by the scientific community on the GEM web site. It will also be integrated with GEM and released as part of that project's regular schedule. The software will be written in C, allowing relatively flexible use, and re-use in other applications and for other datasets.

4.3 Facilities and Equipment

Seismicity data will be obtained from the Southern California Earthquake Center (SCEC) and from the Northern California Earthquake Data Center (NCEDC). GPS sensor position data will be obtained as part of the regular SCIGN distribution program. Other future data sources include the Japanese GPS network, and the Plate Boundary Observatory network, as they become available.

All software development for this project will be carried out in the Data Understanding Systems group at JPL using two Sun Microsystems Ultra 60/300 graphics workstations with 1 GB of RAM each. Data analysis will use these systems, and larger runs may use either of two, 32-node, gigahertz-class Beowulf clusters as necessary. Medium-term and long-term storage is available on a 218 GB RAID-5 disk system and multiple tape drives connected to these workstations. Because of this, equipment procurement does not figure into our budget. Software available includes the Matlab, S-plus, Mathematica, and C languages, as well as extensive modeling tools for time series analysis and probabilistic computations.

Time Series Pattern Recognition for Classifying Seismicity Records and Geodetic Sensor Trajectories
(SENH Proposal)

Year 1 Budget Summary

For Period From October 2002 to September 2003	A	NASA USE ONLY B	C
1. Direct Labor (salaries, wages, and fringe benefits)	$90.9		
2. Other Direct Costs:			
a. Subcontracts	$6.7		
b. Consultants	$0.0		
c. Equipment	$0.0		
d. Supplies	$0.0		
e. Travel	$0.0		
f. Other			
1. MPS and ADC	$41.4		
2. Services	$0.0		
3. Indirect Costs*	$9.2		
4. Other Applicable Costs			
1. Award Fee	$2.1		
2. Government Co-I			
5. **SUBTOTAL--Estimated Costs**	$150.3		
6. Less Proposed Cost Sharing (if any)			
7. Carryover Funds (if any)			
a. Anticipated amount :			
b. Amount used to reduce budget			
8. **Total Estimated Costs**	$150.3		XXXXXXX
9. **APPROVED BUDGET**	XXXXXXX	XXXXXXX	

***Facilities and Administrative Costs**

Time Series Pattern Recognition for Classifying Seismicity Records and Geodetic Sensor Trajectories
(SENH Proposal)

Year 2 Budget Summary

For Period From October 2003 to September 2004	A	NASA USE ONLY B	C
1. Direct Labor (salaries, wages, and fringe benefits)	$93.5		
2. Other Direct Costs:			
a. Subcontracts	$6.7		
b. Consultants	$0.0		
c. Equipment	$0.0		
d. Supplies	$0.0		
e. Travel	$0.0		
f. Other (MPS & ADC)			
1. MPS & ADC	$43.9		
2. Services	$0.0		
3. Indirect Costs*	$10.9		
4. Other Applicable Costs			
1. Award Fee	$2.0		
2. Government Co-I			
5. **SUBTOTAL--Estimated Costs**	$157.0		
6. Less Proposed Cost Sharing (if any)			
7. Carryover Funds (if any)			
a. Anticipated amount :			
b. Amount used to reduce budget			
8. **Total Estimated Costs**	$157.0		XXXXXXX
9. **APPROVED BUDGET**	XXXXXXX	XXXXXXX	

***Facilities and Administrative Costs**

JPL

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, California 91109

NASA Proposal Cost Plan (FY 2002 - FY 2005)

Rates effective 10/10/2001

PROPOSAL TITLE: **Time Series Pattern Recognition for Classifying Seismicity Records and Geodetic Sensor Trajectories**
(SENH Proposal)

DATE: 7-Feb-02

A. DIRECT COMPENSATION

	PRODUCTIVE								WEEKLY	FY 2002	FY 2003	FY 2004	FY 2005
	WORKHOURS				WORKYEARS (X.XX)				SALARY	NASA New Start Inflation Index			
	FY 2002	FY 2003	FY 2004	FY 2005	FY 2002	FY 2003	FY 2004	FY 2005	$	1.000	1.028	1.057	1.086
JPL LABOR CLASSIFICATION	0.85100	0.85200	0.85200	0.85200	**2,080**	**2,080**	**2,080**	**2,080**		$K	$K	$K	$K
Principal Investigator	-	248.1	248.1	-		0.14	0.14		2,236	-	14.3	14.7	-
ISCS Senior	-	708.9	708.9	-		0.40	0.40		1,646	-	30.0	30.8	-
Line Manager I	-	177.2	177.2	-		0.10	0.10		2,278	-	10.4	10.7	-
Line Manager II	-	88.6	88.6	-		0.05	0.05		2,750	-	6.3	6.4	-
	-	-	-	-						-	-	-	-
JPL PRODUCTIVE WORKHOURS	-	1,222.8	1,222.8	-	-	0.69	0.69	-					
	0.958	0.958	0.958	0.958									
SUBCONTRACTOR WORKHOURS	-	-	-	-									
TOTAL WORKHOURS AND WORKYEARS A1	-	1,222.8	1,222.8	-	-	0.69	0.69	-					
LABOR COST SUBTOTAL									A2	-	60.9	62.6	-
APPLIED FRINGES COSTS (% X A2)					49.3%	49.3%	49.3%	49.3%	A3	-	30.0	30.9	-
JPL DIRECT COMPENSATION									TOTAL A	-	90.9	93.5	-

B. TRAVEL

DESTINATION	No.Trips	Cost/Trip		FY 2002	FY 2003	FY 2004	FY 2005
1							
2							
3							
4							
			TOTAL TRAVEL B	-	-	-	-

C. SERVICE

(LIST BY TYPE, e.g., COMPUTING, DOCUMENTATION AND PUBLICATION, etc.)		FY 2002	FY 2003	FY 2004	FY 2005
1					
2					
3					
4					
	TOTAL SERVICES C	-	-	-	-

D. PROCUREMENTS

	(LIST BY TYPE, e.g. CONTRACT LABOR, CONSULTANTS, MATERIALS & SUPPLIES, EQUIP., etc.)		FY 2002	FY 2003	FY 2004	FY 2005
D2 CONTRACTS/	1 System Administration			3.0	3.0	
CHARGEBACKS	2 UCS--Desktop/Networking Services and Telephone			3.7	3.7	-
		TOTAL CONTRACTS D1	-	6.7	6.7	-
D2 CONSULTANTS	3					
	4					
		TOTAL CONSULTANTS D2	-	-	-	-
		TOTAL CONTRACTS AND CONSULTANTS D1+D2	-	6.7	6.7	-
D3. PO's EQUIPMENT	1					
	2					
		TOTAL PO - EQUIPMENT D3	-	-	-	-
D4 SUPPLIES	3					
	4					
		TOTAL PO - SUPPLIES D4	-	-	-	-
		TOTAL PO D3+D4	-	-	-	-
		TOTAL PROCUREMENTS D	-	6.7	6.7	-

E. FACILITIES

(LIST AS EITHER NEW CONSTRUCTION OR MODIFICATION - IF C OF F FUNDED, DO NOT INCLUDE)		FY 2002	FY 2003	FY 2004	FY 2005
1					
2					
	TOTAL FACILITIES E	-	-	-	-

		FY 2002	FY 2003	FY 2004	FY 2005		FY 2002	FY 2003	FY 2004	FY 2005
F. MULTIPLE PROGRAM SUPPORT (MPS) (MEP, SESPD, TAPD, TMOD, or None)	($/HR X A1)	$ 11.09	$ 11.95	$ 13.53	$ 14.57	F	-	14.6	16.5	-
G. TOTAL DIRECT COST						(SUM A THROUGH F) G	-	112.2	116.7	-
H. ALLOCATED DIRECT COSTS (ADC)										
H1. APPLIED ENG & SCI ADC	($/HR X A1)	$20.88	$21.76	$22.22	$22.42	H1	-	26.6	27.2	-
H2. APPLIED CONTRACTS ADC	(% X (D1+D2))	2.82%	2.62%	2.27%	2.47%	H2	-	0.2	0.2	-
H3. APPLIED PURCHASE ORDERS (PO) ADC	(% X (D3+D4))	11.31%	11.83%	10.60%	9.48%	H3	-	-	-	-
I. COSTS SUBTOTAL						(SUM A THROUGH H) I	-	139.0	144.1	-
I1. APPLIED GENERAL ADC	(% X I)	6.21%	6.65%	7.60%	7.88%	I1	-	9.2	10.9	-
J. TOTAL JPL COSTS						(I + I1) J	-	148.3	155.0	-
K. AWARD FEE	(% X J)	1.50%	1.40%	1.30%	1.30%	K	-	2.1	2.0	-
L. GRAND TOTAL COSTS						(J + K) L	-	150.3	157.0	-

Cost Estimation Rates and Factors/JPL Cost Accumulation System 10/10/01

The NASA prime contract NAS7-1407 is a Cost Reimbursable Award Fee type instrument. All costsincurred are billed to the Government on a 100% reimbursable basis. The costs to be charged for the proposed work must be consistent with contractual provisions and established procedures for costing under the current contract between NASA and Caltech. All charges developed at the Laboratory, including JPL applied burdens, are billed to the Government as direct charges at the rates in effect at the time the work is accomplished. Government audit is performed on a continuing basis by a Defense Contract Audit Agency team in residence.

Distributive Cost Rates and Factors

JPL defines distributive costs as cost elements not directly identifiable with a specific sponsor funded task. Major classifications under this category include: (1) Allocated Direct Costs; (2) Multiple Program Support; (3) Employee Benefits; and (4) Service Centers.

1. Labor Fringe Rates - Employee Benefits

   The accumulation process applies JPL employee benefits on a composite basis. Cost estimates will apply benefits as a percentage of total straight-time wages, salaries, and overtime. Functions and activities covered by this rate include paid leave, vacations, and other benefits including retirement plans, group insurance plans, and tuition reimbursements.

2. Desktop and Network Service (DNS)/Telephone

   Desktop and Network Service (DNS) links each personal computer to an individual employee or accountable contractor. The DNS billing process occurs monthly. Time charges generated by an individual during the month are the basis of the cost allocation to specific projects or tasks. The billing process allocates charges to an account worked on a proportional basis unless otherwise requested. This same process is used collect costs for telephone billing.

3. Multiple Program Support (MPS) Factors

   The MPS Rate applies costs for program management and technical infrastructure. Cost estimates and application will apply the composite rate to all program direct hours charged to projects managed by JPL.

4. Allocated Direct Cost Rates (ADC)

   ADC rates contain cost elements benefiting multiple work efforts, including Project Direct, MPS, and Support and Services activities. Rate applications for cost estimates are specific to the given category as stated below.

a. Engineering and Science (formally Labor ADC). Applied as a rate per JPL and accountable contractor straight-time and overtime workhours on Project Direct, MPS, and Service Center activities.

b. Procurement

(1) Subcontract. Applied as a percentage of contracted invoice costs related to subcontract defined articles or activities.

(2) Purchase Order. Applied as a percentage of direct procurement invoice costs related to purchase order defined articles or activities.

c. General. Applied as a percentage of all costs except General pool costs on direct projects.

5. R&D Program Charge. JPL applies the R&D Program Charge to cost estimates for reimbursable research and development funding (appropriations) to the extent the task utilizes such funds. JPL defines all non-government funding as R&D funds. Cost estimates will include this charge to all appropriate new activities, and when revisions to the scope of work change the baseline estimated total cost.

Cost estimates will include the R&D Program Charge by applying the rate (percentage which is annually determined) to the sum of total direct and allocated direct costs for the cost estimates that are $250,000 or more.

This is a charge taken by JPL. It is the responsibility of the Contract Management Office to determine the appropriate R&D Program Charge for each new cost estimate including any revisions involving reimbursable R&D funding.

6. Award Fee

Sponsors placing funds on contract contribute a percentage of task order dollars to the award fee. The local NASA Management Office determines the rate (percentage) annually. NASA applies the rate (percentage) to the funding available net of the CAAS charge.

7. NASA G&A. NASA applies this charge to reimbursable work. The basis for this rate is total reimbursable cost (including DRDF and Award Fee) less CAAS charges.

8. NASA Contract Administration and Audit Services (CAAS) Charge

NASA applies this charge. NASA sets the actual charge through application of the published schedule to the total reimbursable dollar value shown on the Resource Authority Warrant (#506A or #822). This is a non-refundable amount based upon the total agreement value (task plan cost estimate) and will be charged on each order. If the total reimbursable estimate value is increased, the CAAS charge will be recalculated and the customer charged accordingly.

All reimbursable efforts whose total estimated costs exceed $250,000 will be subject to Award Fee and R&D charges as described above.

Current and Pending Support

Dr. Michael Turmon

Current Support

Agency	Title	FY02 Award	Total Award	Period	% Effort
NASA Code S	Startool: Pattern Recognition for Solar Imagery	$150K	$450K	FY00–FY02	0.2 WY
NASA Code S (AISRP)	Statistical Object Identification, Tracking, and Analysis	$130K	$400K	FY01–FY03	0.14 WY
NASA Code S (LWS)	Evolution of Magnetic Fields and Variations of Solar Irradiance	$20K	$40K	FY01-FY02	0.08 WY

Pending Support

Agency	Title	FY03 Award	Total Award	Period	% Effort
NASA		— This proposal —			

This proposal has not been submitted in response to any other announcements of opportunity.

Current and Pending Support

(See GPG Section II.D.8 for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Robert Granat	Other agencies (including NSF) to which this proposal has been/will be submitted. None

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title: Time series pattern recognition for classifying seismicity records and geodetic sensor trajectories

Source of Support: NASA

Total Award Amount: $307K   Total Award Period Covered: 02-04

Location of Project: JPL

Person-Months Per Year Committed to the Project. 5   Cal: Acad: Sumr:

*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

NSF Form 1239 (10/99)   USE ADDITIONAL SHEETS AS NECESSARY

5

Current and Pending Support

(See GPG Section II.D.8 for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Kenneth Hurst

Other agencies (including NSF) to which this proposal has been/will be submitted.
None

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title: Detection and analysis of time-dependent Earthquake and Volcanic processes using space geodesy

In collaboration with Paul Segall at Stanford (lead) and Jeff McGuire at WHOI

Source of Support: NASA

Total Award Amount: $81.1K (JPL portion) Total Award Period Covered: 02-05

Location of Project: JPL

Person-Months Per Year Committed to the Project. 1.3 Cal: Acad: Sumr:

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:Understanding sub-surface volcanic processes using continuous differential microgravity and continuous

Source of Support: NASA

Total Award Amount: $292k Total Award Period Covered: 02-05

Location of Project: JPL, UCLB, Open University

Person-Months Per Year Committed to the Project. 2.3 Cal: Acad: Sumr:

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:Time series pattern recognition for classifying seismicity records and geodetic sensor trajectories

Source of Support: NASA

Total Award Amount: $307K Total Award Period Covered: 02-04

Location of Project: JPL

Person-Months Per Year Committed to the Project. 1 Cal: Acad: Sumr:

*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

NSF Form 1239 (10/99) USE ADDITIONAL SHEETS AS NECESSARY

5

Current and Pending Support: Andrea Donnellan

Investigator: Andrea Donnellan

Other agencies to which this proposal has been submitted: None

---

Support: [X] Current [ ] Pending [ ] Submission Planned in Near Future [ ] * Transfer of Support

Project/Proposal Title: Development of Fully 3-D Finite Element Model for Crustal Deformation Data

Source of Support: NASA/SENH

Award Amount (or Annual Rate): $570,000 Period Covered: 12/99 to 11/02

Location of Project: Southern California

Person-Months Committed to the Project. Cal.: 1 mos Acad: Summ:

Support: [X] Current [ ] Pending [ ] Submission Planned in Near Future [ ] * Transfer of Support

Project/Proposal Title: Numerical Simulations for Active Tectonic Processes: Increasing Interoperability and Performance

Source of Support: NASA/HPCC

Award Amount (or Annual Rate): $2,208,000 Period Covered: 2/02 to 9/04

Location of Project: Southern California

Person-Months Committed to the Project. Cal.: 2 mos Acad: Summ:

Support: [X] Current [ ] Pending [ ] Submission Planned in Near Future [ ] * Transfer of Support

Project/Proposal Title: Automated Data Analysis of Geodetic Sensor Networks

Source of Support: NASA/CETDP

Award Amount (or Annual Rate): $291,000 Period Covered: 11/99 to 9/02

Location of Project: Southern California

Person-Months Committed to the Project. Cal.: 0.5mos Acad: Summ:

Support: [ ] Current [X] Pending [ ] Submission Planned in Near Future [ ] * Transfer of Support

Project/Proposal Title: GPS measurement of isostatic rebound and tectonic def. in Marie Byrd Land, West Antarctica

Source of Support: NASA

Award Amount (or Annual Rate): $300,000 Period Covered: 7/02 to 9/05

Location of Project: Antarctica

Person-Months Committed to the Project. Cal.: 3 mos Acad: Summ:

Support: [ ] Current [ ] Pending [ ] Submission Planned in Near Future [ ] * Transfer of Support

Project/Proposal Title:

Source of Support:

Award Amount (or Annual Rate): Period Covered:

Location of Project:

Person-Months Committed to the Project. Cal.: Acad: Summ:

Support: [ ] Current [ ] Pending [ ] Submission Planned in Near Future [ ] * Transfer of Support

Project/Proposal Title:

Source of Support:

Award Amount (or Annual Rate): Period Covered:

Location of Project:

Person-Months Committed to the Project. (average per year) Cal.: Acad: Summ:

---

* If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

Michael Turmon

M/S 126-347; Jet Propulsion Laboratory; Pasadena, CA 91109
+1 818 393 5370
turmon@jpl.nasa.gov

Education	PhD 1995	Cornell University, Ithaca, NY
	BSEE 1987, BSCS 1987, MSEE 1990	Washington University, St. Louis, MO

Experience

Principal Member Technical Staff, 2001–present
Senior Member Technical Staff, 1995–2001
Jet Propulsion Laboratory

Principal investigator for solar feature identification project, which has developed an automatic, scientist-trainable mechanism for objectively identifying solar active regions from multimode imagery. The resulting software is now in use by solar physicists at UCLA and Big Bear Solar Observatory.
Funded co-investigator on the SoHO spacecraft; supervise development of solar-feature identification software to be incorporated into the standard data pipeline for that mission.
Co-Investigator of CNES PICARD spacecraft; develop methods for robust identification of solar activity.

Visiting Scientist, June 1995–August 1995 — Jet Propulsion Laboratory
Developed new pattern recognition techniques for Magellan SAR images of Venus.

Research Assistant, 1990–1995 — Cornell University
Investigated learning in neural networks.

Research Assistant, 1987–1990 — Washington University
Developed statistical techniques for direction-of-arrival estimation and radar imaging.

Honors and Achievements

Presidential Early Career Award (PECASE): 2000
NASA Exceptional Achievement Medal: 1999
Co-Investigator of CNES PICARD satellite: 1999–present
Best Paper by a Young Researcher, IASC Compstat conference: 1998
Co-Investigator, NASA/ESA SoHO spacecraft: 1996–1998
Invited to apply for AT&T PhD Fellowship: 1992
NSF Graduate Fellow: 1987–1990

Professional Service

Member IEEE (Computer and Information Theory Societies)
Member Institute for Mathematical Statistics

Participated in technical reviews in the NSF IDM review panel, spring 1999, fall 2000.

Program committee memberships:

- NIPS-00 (workshop on Software Support for Bayesian Analysis Systems)
- CVPR-99 (Computer Vision and Pattern Recognition)

KDD-98 (Knowledge Discovery in Databases)
NIPS-96 (Neural Information Processing Systems)
WCNN-95 (World Conference on Neural Networks)

Reviews for *IEEE Transactions on Neural Networks*, *IEEE Transactions on Signal Processing*, and *Machine Learning*.

Relevant Publications

M. Turmon Statistical pattern recognition for labeling solar active regions: Application to SoHO/MDI imagery. *Astrophysical Journal*. March, 2002. With J. Pap and S. Mukhtar.

A language for probabilistic modeling of scientific data. In *Proc. Second Conf. Highly Structured Stochastic Systems*, pages 298–300, Pavia, Italy, 1999. With L. Ramsey, E. Mjolsness and V. Gluzman.

Analysis of the SoHO/VIRGO total and spectral irradiances based on the SoHO/MDI images. Presented at IUGG-99 (International Union of Geodesy and Geophysics), July 1999. With J.M. Pap, M. Anklin, R. Bogart, L. Floyd, C. Fröhlich, F. Varadi and Ch. Wehrli.

Automatically finding solar active regions using SoHO/MDI photograms and magnetograms. In *Proc. SoHO 6/GONG '98 Workshop on Structure and Dynamics of the Sun*, pages 979–984, 1998. With J. M. Pap and S. Mukhtar.

Representing solar active regions with triangulations. In *Proc. Compstat-98*, pages 473–478, 1998. With S. Mukhtar.

Recognizing chromospheric objects via Markov chain Monte Carlo. In *Proc. IEEE ICIP-1997*, pages III, 320–323, 1997. With S. Mukhtar.

Bayesian inference for identifying solar active regions. In H. Mannila D. Heckerman and D. Pregibon, editors, *Proc. Third Conf. on Knowledge Discovery and Data Mining*. MIT Press, 1997. With S. Mukhtar and J. Pap.

Segmenting chromospheric images with Markov random fields. In G. J. Babu and E. D. Feigelson, editors, *Statistical Challenges in Modern Astronomy II*, pages 408–411. Springer, 1997. With J. Pap.

Generalization in feedforward neural networks. In *IEEE 1995 International Symposium on Information Theory*, 1995. With T. L. Fine. Long paper.

Sample size requirements for feedforward neural networks. In G. Tesauro et al., editor, *Neural Information Processing Systems 7*, pages 327–334. Morgan-Kauffman, 1995. With T. L. Fine.

Sample size requirements of feedforward neural network classifiers. In *IEEE 1993 International Symposium on Information Theory*, 1993. With T. L. Fine.

Maximum-likelihood estimation of constrained means and Toeplitz covariances with application to direction-finding. *IEEE Trans. on Signal Processing*, 42(5):1074–1086, May 1994. With M. I. Miller.

Performance evaluation of maximum-likelihood Toeplitz covariance estimates generated using the expectation maximization algorithm. In *Proc. Fourth ASSP Workshop on Spectrum Estimation and Modeling*, pages 182–185, August 1988. With M. I. Miller, D. L. Snyder, and J. A. O'Sullivan.

Efficient implementation of the EM algorithm for Toeplitz covariance estimation. In *Proc. Twenty-second Annual Conference on Information Sciences and Systems*. Princeton University, March 1988. With D. R. Fuhrmann and M. I. Miller.

The application of maximum-entropy aand maximum-likelihood for spectral estimation. In *IEEE 1986 International Symposium on Information Theory*, October 1986. With M. I. Miller and D. L. Snyder.

Robert Granat

M/S 126-347; Jet Propulsion Laboratory; Pasadena, CA 91109
+1 818 393 5353
granat@aig.jpl.nasa.gov

Education	PhDEE expected 2002	University of California, Los Angeles, CA
	MSEE 1998	University of California, Los Angeles, CA
	BSCNS 1992	California Institute of Technology, Pasadena, CA

Experience

Senior Member Technical Staff, 1996–present
Jet Propulsion Laboratory

Principal investigator for real time earthquake location and imaging project, which is developing crustal imaging techniques based on Kirchoff integration and pattern recognition for rapid identification and localization of seismic events.
Task manager, intelligent systems project: develop machine learning tools for automated analysis of geodetic sensor networks.
Key research and development roles for a number of projects: atmospheric time series data analysis, large scale machine learning, automated spectral analysis, fault tolerant computing, and sub-pixel change detection in images.

Research Assistant, 1997–1998 — University of California, Los Angeles

Developed a neural network based real time method for imaging dynamic ionospheric electron densities, for use with relocatable over-the-horizon radar (ROTHR).

Research Assistant, 1994–1996 — California Institute of Technology

Implemented and tested a model independent motion and structure recovery algorithm for use on video sequences.
Researched a laser induced fluorescence (LIF) based method for detecting impurities in large tokamaks.

Honors and Achievements

NASA "Level B" Award: 2001
NASA Achievement Award: 1997

Professional Service

Member IEEE
Member American Geophysical Union

Relevant Publications

A hidden Markov model based tool for geophysical data exploration. In *PAGEOPH special issue, ACES meeting, Hakone, Japan*, 2001. With A. Donnellan.

Estimating dynamic ionospheric changes without a priori models. In *Radio Science*, 35(2):341–349, 2000. With H. Na.

A neural network approach to imaging ionospheric motion. Master's thesis, University of California, Los Angeles, 1998.

Kenneth J. Hurst

ADDRESS:
Mail Stop 126-347; Jet Propulsion Laboratory; 4800 Oak Grove Dr.; Pasadena, CA 91109;
Phone: [818] 354-6637; fax: [818] 393-4965; email: kenneth.j.hurst@jpl.nasa.gov

EDUCATION:

Columbia University, New York, New York	PhD Geology, October, 1987
Title of thesis: "The Measurement of Vertical Crustal Deformation"	
Columbia University, New York, New York	MPh Geology, 1987
	MA Geology, 1982
Earlham College, Richmond Indiana	BA Geology, 1980
	BA Physics, 1980

EMPLOYMENT EXPERIENCES:

March 2001-present: Group Supervisor, Data Understanding Systems Group, Exploration Systems Autonomy Section, Jet Propulsion Laboratory
1/2000-12/2001: Part-time Faculty, California State University, Northridge.
1997-2001: Senior Member of the Technical Staff, Satellite Geodesy and Geodynamics Group, Tracking Systems and Applications Section, Jet Propulsion Laboratory.
1990-1997: Member of the Technical Staff, Jet Propulsion Laboratory
1988 - 1989: Post-doctoral Research Scientist - LDGO, Columbia University
1988: Research Assistant - LDGO, Columbia University
1987 -1988: Research Associate - University of Colorado
1987: Professional Research Assistant - University of Colorado
1983 - 1987: Graduate Research Assistant (Seismology) - Columbia University

PROFESSIONAL AFFILIATIONS:

Geological Society of America
American Geophysical Union
Seismological Society of America
National Association of Geoscience Teachers

AWARDS:

Co-recipient of NASA Group Achievement award
NASA Award for Excellence, April 1996.
Co-recipient of runner-up award in NASA software of the year competition 1995.
Co-recipient of NASA Group Achievement award for Global Positioning System Geodesy Development, Jet Propulsion Laboratory, 1994.
Co-recipient of NASA Group Achievement award for development of the GIPSY-OASIS GPS analysis software Jet Propulsion Laboratory 1992.
Departmental honors in Geology, Earlham College 1980. Deformation in the Shumagin Seismic Gap, Alaska, *Geophys. Res. Let., 14*, 1234-1237, 1987.
All college honors, Earlham College 1980.
Anna Hubbard Grave, Thomas Clarkson Grave honorary scholarship for graduate work in science, Earlham College 1980.

SELECTED PEER REVIEWED PUBLICATIONS

Hurst, K., D. Argus, A. Donnellan, M. Heflin, D. Jefferson, G. Lyzenga, The co- and immediate post-seismic geodetic signature of the Oct 16 1999 Hector Mine earthquake. *Geophys Res. Let 2733-2736, 2000*

Argus, Donald, M. Heflin, A. Donnellan, F. Webb, D. Dong, K. Hurst, D. Jefferson, G. Lyzenga, M. Watkins, J. Zumberge, Shortening and thickening of metropolitan Los Angeles measured and inferred by using geodesy, *Geology, 703-706*, 1999.

Fox, G., K. Hurst, A. Donnellan, and J. Parker,Introducing a New Paradigm for Computational

Earth Science - A web-object-based approach to Earthquake Simulations, *Physics of Earthquakes* edited by John Rundle, DonaldTurcotte and William Klein, published by AGU pp 219-245, 2000.

MClusky, S., S. Balassanian, A. Barka, C. Demir, S. Ergintav, I. Georgiev, O. Gurkan, M. Hamburger, K. Hurst, H. Kahle, K. Kastens, G. Kekelidze, R. King, V. Kotzev, O. Lenk, S. Mahmoud, A. Mishkin, M. Nadariya, A. Ouzounis, D. Paradissis, Y. Peter, M. Prilepin, R. Reilinger, I. Sanli, H. Seeger, A. Tealeb, M. N. Toksoz, and G. Veis. GPS constraints on plate kinematics and dynamics in the Eastern Mediterranean and Caucasus,*J. Geophys. Res., 105*,5695-5720, 2000

Heflin, Michael, D. Dong, A. Donnellan, K. Hurst, D. Jefferson, M. Watkins, F. Webb, J Zumberge, D. Dauger, G. Lyzenga, Rate change observed at JPLM after the Northridge earthquake, *Geophysical Res. Let.* 93-96, 1998.

Kahle, Hans, C. Straub, R. Reilinger, S. McClusky, R. King, K. Hurst, G. Veis, K. Kastens, P. Cross, The strain rate field in the eastern Mediterranean region, estimated by repeated GPS measurements, *Tectonophysics*, 1997.

Geoffrey Blewitt, Michael B. Heflin, Kenneth J. Hurst, David C. Jefferson, Frank H. Webb, and James F. Zumberge, Absolute far-field displacements from the June 28, 1992, Landers earthquake sequence,*Nature, 361*, 340-342, Jan 28, 1993.

Lundgren, Paul, Susan Kornreich Wolf, Marino Protti, and Kenneth J. Hurst, GPS measurements of crustal deformation following the 22 April 1991, Valle de la Estrella, Costa Rica Earthquake,*Geophys. Res. Let. 20*,407-410, 1993.

Heflin, M., G. Blewitt, W. Bertiger, A. Freedman, K. Hurst, S. Lichten, U. Lindqwister, R. Malla, Y. Vigue, F. Webb, T. Yunck, and J. Zumberge, Global geodesy using GPS without fiducial sites *Geoph. Res. Let.*, *19*, 131-134, 1992.

Beavan, J., R. Bilham, K. Hudnut, and K. Hurst, Techniques and results of crustal deformation measurement using sea-level gauges, leveling and extensometers,*in Crustal Deformation and Earthquakes, ed. Wu Bing, Seismological Press*, 302-319, 1988.

Bilham, R. G., and K. Hurst, Relationships between fault zone deformation and segment obliquity on the San Andreas Fault, California, *in Crustal Deformation and Earthquakes, ed. Wu Bing, Seismological Press,* 510-524, 1988.

Hurst, K. J., and J. Beavan, Improved sea level monitors for measuring vertical crustal

Ware, R. H., C. Rocken, and K, Hurst, A global positioning system baseline determination including bias fixing and Water Vapor Radiometer corrections, *J. Geophys. Res.*, *91*, 9183-9192, 1986.

Hurst, K. J. and R. Bilham, Hydrostatic levels in precision geodesy and crustal deformation measurement, *J. Geophys. Res.*, *91*, 9202-9216, 1986.

Beavan, J., K. Hurst, R. Bilham, and L. Shengold, A densely spaced array of sea level monitors for the detection of vertical crustal deformation in the Shumagin seismic gap, Alaska, *J. Geophys. Res.*, *91*, 9067-9080, 1986.

Bilham, R., J. Beavan, K. Evans, and K. Hurst, Crustal deformation metrology at Lamont-Doherty Geological Observatory, *Earthq. Predict. Res.*, *3*,391-411, 1985.

Beavan, J., R. Bilham, and K. Hurst, Coherent tilt signals observed in the Shumagin seismic gap: Detection of time-dependent subduction at depth?, *J. Geophys. Res., 89*, 4478-4492, 1984.

Hurst, K, K. VanZant, and C. Hurst, A new driving mechanism for hand coring, *Palynology, 5*, 81-84, 1981.

ANDREA DONNELLAN CURRICULUM VITAE

Education

Ph.D., Geophysics, California Institute of Technology (1991)
M.S., Geophysics, California Institute of Technology (1988)
B.S., Geology, Ohio State University, *with honors and distinction in geology* (1986)
Pursuing M.S., Computer Science, University of Southern California (present)

Professional Experience

Deputy Manager, Exploration Systems Autonomy (2000–present)
Supervisor, Data Understanding Systems Group, Jet Propulsion Laboratory (1999–2001)
Research Professor, Department of Earth Sciences, University of Southern California (1999–present)
Research Scientist, Satellite Geodesy and Geodynamics Systems Group, Jet Propulsion Lab. (1997–1999)
Member of Technical Staff, Satellite Geodesy and Geodynamics Sys. Group, Jet Prop. Lab. (1993–1997)
Visiting Associate, Seismological Laboratory, California Institute of Technology, (1995–1996)
National Research Council Resident Research Associate, NASA Goddard Space Flight Center (1991–1993)
Graduate Research Assistant, California Institute of Technology (1986–1991)
Research Assistant, Institute of Polar Studies, Ohio State University (1983–1986)
Geochemistry Group, Sohio Research and Development (1985)
Thin Section Laboratory Technician, Ohio State University (1983)

Selected Professional Activities

NASA Solid Earth Science Working Group for NASA HQ (2000-present)
US Rep. to the Int. Sci. Board of the APEC Cooperation on Earthquake Simulations (2000–present)
American Geophysical Union (AGU) nonlinear geophysics committee (2000–present)
JPL Science and Technology Management Council (2001–present)
JPL SESPD Science Advisory Group to Charles Elachi (1994–2001)
Plate Boundary Observatory steering committee (1999–present)
General Earthquake Models (GEM) program planning and other committees (1998–present)
Convenor NSF/NASA Sponsored Autonomous Systems in Extreme Environments Workshop (1999)
AGU 2000 and 2001 Spring Meeting program committee geodesy section chair (2000–2001)
AGU Geodesy representative for education and outreach (1999)
Southern California Earthquake Center (SCEC) Crustal Deformation Working Group (1993–present)
S. Cal. Integrated GPS Network (SCIGN) Coord. Board (SCEC rep: 1994–1998; NASA rep: 1999–2001)
UNAVCO Field Operations Working Group, Chair (1995–1997)
Development Oversight of SCEC GPS Educational Modules (1996–present)
Panel member National Earthquake Hazards Reduction Program External Research Program (1994–1999)

Selected Awards

JPL Lew Allen Award for Excellence (2000)
Southern California Earthquake Center Outreach Award for Education (1998)
Presidential Early Career Award for Scientists and Engineers (1996)
National Research Council Postdoctoral Fellowship (1991–1993)

Selected Publications

Granat, R., and A. Donnellan, Deterministic annealing hidden Markov models for geophysical data exploration, PAGEOPH, in press.

Donnellan, A., J. Parker, and G. Peltzer, Combined GPS and InSAR models of postseismic deformation from the Northridge earthquake, PAGEOPH, in press.

Hurst, K.J., D. Argus, A. Donnellan, M.B. Heflin, D. Jefferson, G.A. Lyzenga, J.W. Parker, F.H. Webb, J.F. Zumberge, The Co- and Immediate Post-seismic geodetic signature of the 1999 Hector Mine Earthquake, Geophys. Res. Lett., 27, 2733–2736, 2000.

Fox, G.C., K. Hurst, A. Donnellan, and J. W. Parker, Introducing a New Paradigm for Computational Earth Science—A Web-Object-Based Approach to Earthquake Simulations, Geocomplexity and the Physics of Earthquakes, American Geophysical Union Monograph, 2000.

Donnellan, A. and G. A. Lyzenga, Fault afterslip and upper crustal relaxation following the Northridge earthquake, *J. Geophys. Res.*, **103**, 21,285–21,297, 1998.

Donnellan, A. and F.H. Webb, Geodetic observations of the M 5.1 January 29, 1994 Northridge aftershock, *Geophys. Res. Lett.*, **25**, 667–670, 1998.