LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-PROC-217057

Hierarchical Bayesian Approach to Locating Seismic Events

G. Johannesson, S. C. Myers, W. G. Hanley

November 14, 2005

2005 Joint Statistical Meetings
Minneapolis, MN, United States
August 7, 2005 through August 11, 2005

Disclaimer

Hierarchical Bayesian Approach to Locating Seismic Events

Gardar Johannesson, Stephen C. Myers, and William G. Hanley
Lawrence Livermore National Laboratory

Abstract

We propose a hierarchical Bayesian model for conducting inference on the location of multiple seismic events (earthquakes) given data on the arrival of various seismic phases to sensor locations. The model explicitly accounts for the uncertainty associated with a theoretical seismic-wave travel-time model used along with the uncertainty of the arrival data. Posterior inferences is carried out using Markov chain Monte Carlo (MCMC).

1. Introduction

Locating a seismic event (earthquake) is almost without exception the first step in seismic processing. Determination of event location is, for example a prerequisite to the calculation of event magnitude. And in the case of damaging earthquakes, event location is crucial for directing emergency responders.

The location parameters of a single seismic event consists of the hypocenter (latitude, longitude, and depth) and the origin time. Our main goal will be to conduct inference on those parameters given (1) the observed arrival-time of seismic phases, (2) a theoretical travel-time model, and (3) some prior knowledge (e.g., on location, travel-time model uncertainty, etc.). In particular, we are interested in drawing inference about events that are not well defined as a result of any combination of: small amount of (noise-corrupted) data, little prior information, and/or an unestablished regional travel-time model.

The arrival-time data consists of observed arrival times of various seismic waves from multiple events to multiple stations. Each seismic event generates different types of seismic waves, phases, that propagate from the event origin along different paths. The arrival of these phases is then recorded by a seismometer at various seismic stations. The "picking" of the arrival-times, and which phase they belong to, is a mix of art and science. As such, its error process is not well defined or understood; some phases are easier to pick than others, some people are better (more consistent) at picking than others, etc. Seismic bulletins (databases) are universally contaminated with incorrect phase assignments and large arrival-time outliers. Hence, an important ingredient to an accurate seismic locator is a good characterization and modeling of the this error process.

The observed arrival-times are linked to the seismic events using a seismic-wave travel-time model. A (regional) travel-time model for a given wave-phase is typically a function of the angular distance between the event and the station with a correction for the vertical depth of the event and the elevation of the station. However, the travel-time model is just an approximation to the true travel-time. Systematic errors in the travel-time model can therefore greatly affect the accuracy of a seismic locator.

In addition to the observed arrival-time data and the theoretical travel-time model, we have (potentially) available some prior knowledge about the quality of the data, the model, and the location of some of the events. For example, we might have prior knowledge about the vertical depth of some of the events, while for other events we might have prior knowledge of their lat-long position. Events for which we have strong prior knowledge might not be of particular interest themselves, however, they could be very valuable in "constraining" the analysis of events that are not as well defined *a priori*.

The most commonly used event location algorithms process one event at a time. These methods minimize a loss function to determine the optimal location. The loss function is based on the difference between observed and predicted arrival times of various seismic phases at established sensors. These methods are adequate for determining initial location estimates, but location errors on the order of tens of kilometers are common (Bondár et al., 2004).

Simultaneous determination of many event locations (multiple-event location) can help to identify and correct travel-time prediction and seismic analysis errors. Douglas (1967) introduced the multiple-event location method, and subsequent efforts (Dewey, 1972; Jordan & Sverdrup, 1981; Pavlis & Booker, 1983; Rodi & Toksöz, 2000) are variations on the original theme. Multiple event methods typically solve for each event location, as well as a scalar travel-time correction for each station/phase. The problem is typically formulated as a linear approximation, which allows the application of matrix inversion methods as solvers. The full set of equations is typically singular and constraints are needed to invert the matrix. Fixing the location of one event or assuming a zero mean population of station/phase correction are the most common constraints in practice. More recently, Waldhauser & Ellsworth (2000) used a travel-time residual differencing (i.e. double difference) procedure to implicitly account for station/phase corrections. Most importantly, the double-difference method implements an ad hoc spatial correlation function that can account for changes in the station/phase correction as a function of event position. While each of the current multiple-event methods has unique merits, the full knowledge of the multiple-event problem is not currently utilized. As a result of underutilized prior knowledge, extensive data culling is a prerequisite to maintain inversion stability.

We propose a hierarchical Bayesian model that explicitly models the error-processes associated with the arrival-data and the travel-time model. The Bayesian approach also provides a natural framework to take advantage of prior information. As our approach is probabilistic, it yields posterior distributions of event locations, and other model parameters (travel-time corrections, etc.).

1.1 Notation

In what follows, we assume I number of events with data observed at J number of stations, where the type of arrivals are restricted to the phase-set $\mathcal{W} = \{w_1, \ldots, w_N\}$ (e.g., $\mathcal{W} = \{\mathrm{Pn}, \mathrm{Pg}, \mathrm{Lg}\}$).

The event location parameters are:

$\mathbf{x}_i \equiv (x_i, y_i, z_i) =$ the *hypocenter* (location) of the i-th event.

$o_i \equiv$ the *origin time* of the i-th event.

The station data is given by:

$\mathbf{s}_j \equiv$ the *location* of the j-th station.

$n_{ij} \equiv$ the number of recorded arrivals from the i-th event to the j-th station.

$a_{ijk} \equiv$ the k-th observed *arrival-time* from the i-th event to the j-th station.

$w_{ijk} \equiv$ the *phase-label* assigned to the a_{ijk} arrival-time, $w_{ijk} \in \mathcal{W}$.

The phase labels $\{w_{ijk}\}$ are treated as data with potential errors. That is, it might be the case that the arrival-time a_{ijk} is *not* the arrival of the w_{ijk} phase — it might even not correspond to the arrival of any of the N phases in $\mathcal{W}$. To account for both possible mislabeling of phases and potential "outliers", let

$W_{ijk} \equiv$ the *true phase-label* associated with the arrival-time a_{ijk}, $W_{ijk} \in \mathcal{W}^+ \equiv \mathcal{W} \cup \{0\}$.

Note we have introduced an additional "phase-label", 0, such that if $W_{ijk} = 0$, then a_{ijk} is not assumed to be an arrival of any of the phases in $\mathcal{W}$.

To conduct inference on the event parameters given the station data, a theoretical seismic phases travel-time model is used (e.g., the IASP91 model of Kennett & Engdahl (1991)). Denote by

$F^w(\mathbf{x}, \mathbf{s}) \equiv$ the *model-predicted travel-time* of phase w from event location $\mathbf{x}$ to station location $\mathbf{s}$, and let,

$F^w_{ij} \equiv F^w(\mathbf{x}_i, \mathbf{s}_j)$.

However, the model-predicted travel-time is only an approximation of the true travel-time of each phase, we therefore explicitly define,

$T^w(\mathbf{x}, \mathbf{s}) \equiv$ the *true travel-time* of phase w from event location $\mathbf{x}$ to station location $\mathbf{s}$, and let,

$T^w_{ij} \equiv T^w(\mathbf{x}_i, \mathbf{s}_j)$.

One can think of the true travel-time as related to the model-predicted travel-time via

$$T^w(\mathbf{x}, \mathbf{s}) = F^w(\mathbf{x}, \mathbf{s}) + (\text{model error}), \qquad (1)$$

where the size of the model-error component depends on the accuracy of the travel-time model. Given the true travel-time and the origin time, we denote the arrival-time of phase w from the i-th event to the j-th station by

$$A^w_{ij} \equiv o_i + T^w_{ij}.$$

When $W_{ijk} \neq 0$ (i.e., a_{ijk} is a "good" arrival observation), let

$$T_{ijk} \equiv T^{W_{ijk}}_{ij} \quad \text{and} \quad A_{ijk} \equiv o_i + T_{ijk},$$

be the true (the expected) travel-time and arrival-time, respectively, associated with the observed arrival-time a_{ijk}. The observed arrival-times $\{a_{ijk}\}$ are thought to be related to the expected arrival-times $\{A_{ijk}\}$ via

$$a_{ijk} = A_{ijk} + (\text{measurement error}), \qquad (2)$$

where the measurement error captures, among other things, the "picking" error associated with the $\{a_{ijk}\}$.

We will refer to a subset of parameters by simply dropping one or more sub/super scripts. For example,

$$\mathbf{x} \equiv \{\mathbf{x}_i\}, \; \mathbf{a} \equiv \{a_{ijk}\}, \text{ and } \mathbf{T} \equiv \{T^w_{ij}\}.$$

2. Hierarchical Bayesian Statistical Modeling Approach

While defining our notation in the previous section, we introduced informally the concept of measurement and model errors; see (2) and (1). We shall now better formulate these concepts (along with phase-assignment errors) and formally take advantage of any prior knowledge, particularly about the seismic location parameters.

The approach we take is to model the seismic data via a three-stage hierarchical Bayesian model:

Data-Model: A conditional distribution of $\mathbf{a}$ and $\mathbf{w}$ given $\mathbf{A}$ and $\mathbf{W}$:

$$p(\mathbf{a}, \mathbf{w} \mid \mathbf{A}, \mathbf{W}, \boldsymbol{\sigma}) = p(\mathbf{a}, \mathbf{w} \mid \mathbf{o}, \mathbf{T}, \mathbf{W}, \boldsymbol{\sigma}) \quad (3)$$

where $\boldsymbol{\sigma}$ is a vector of parameters (if any) associated with the data model.

Process-Model: A conditional distribution of $\mathbf{T}$ and $\mathbf{W}$ given the model-predicted travel-time table $\mathbf{F}$:

$$p(\mathbf{T}, \mathbf{W} \mid \mathbf{F}, \boldsymbol{\tau}) = p(\mathbf{T}, \mathbf{W} \mid \mathbf{x}, \boldsymbol{\tau}) \quad (4)$$

where $\boldsymbol{\tau}$ is a vector of possible additional parameters. Recall that $\mathbf{F} = \{F^w(\mathbf{x}_i, \mathbf{s}_j)\}$.

Parameter-Model: A prior distribution of $\mathbf{x}$, $\mathbf{o}$, $\boldsymbol{\sigma}$, and $\boldsymbol{\tau}$:

$$p(\mathbf{x}, \mathbf{o}, \boldsymbol{\sigma}, \boldsymbol{\tau}) = p(\mathbf{x}, \mathbf{o}) p(\boldsymbol{\sigma}) p(\boldsymbol{\tau}). \quad (5)$$

The second expression assumes independence between $\{\mathbf{x}, \mathbf{o}\}$, $\boldsymbol{\sigma}$, and $\boldsymbol{\tau}$.

The joint posterior distribution of all the parameters involved is given by

$$\begin{aligned} p(\mathbf{x}, \mathbf{o}, \mathbf{T}, \mathbf{W}, \boldsymbol{\sigma}, \boldsymbol{\tau} \mid \mathbf{a}, \mathbf{w}) &\propto p(\mathbf{a}, \mathbf{w} \mid \mathbf{o}, \mathbf{T}, \mathbf{W}, \boldsymbol{\sigma}) \\ &\times p(\mathbf{T}, \mathbf{W} \mid \mathbf{x}, \boldsymbol{\tau}) p(\mathbf{x}, \mathbf{o}) p(\boldsymbol{\sigma}) p(\boldsymbol{\tau}). \end{aligned} \quad (6)$$

Hence, in addition to provide posterior inference about the event-origin parameters, $\mathbf{x}$ and $\mathbf{o}$, the posterior distribution (6) also carries information about, for example, the actual (not model predicted) travel-times, $\mathbf{T}$.

We now discuss the three modeling stages in more details.

2.1 The Data Model

The data-model describes (probabilistically) the difference between the observed and the expected arrival-times and phase assignments; see (3). This model can alternatively be written as,

$$\begin{aligned} &p(\mathbf{a}, \mathbf{w} \mid \mathbf{o}, \mathbf{T}, \mathbf{W}, \boldsymbol{\sigma}) \\ &\quad = p(\mathbf{w} \mid \mathbf{a}, \mathbf{o}, \mathbf{T}, \mathbf{W}, \boldsymbol{\sigma}_w) p(\mathbf{a} \mid \mathbf{o}, \mathbf{T}, \mathbf{W}, \boldsymbol{\sigma}_a), \end{aligned} \quad (7)$$

where $\boldsymbol{\sigma} = \{\boldsymbol{\sigma}_a, \boldsymbol{\sigma}_w\}$. Hence, a model can be specified for the phase assignments that is conditional on the observed arrival-times (in addition to the expected arrival-times and phase assignments) and then a separate model can be specified for the observed arrival-times. This partly reflects how the arrival-times and the phases are picked in practice, where an arrival-time can be picked without knowing the phase, but the picked arrival-time can be valuable in deciding on the phase label.

The Arrival-Time Data Model

Consider first the case when all the arrival-time observations are assumed to be good phase-picks; that is, $W_{ijk} \neq 0$. We assume that,

$$a_{ijk} - A_{ijk} \sim \mathrm{Gau}(b_{ijW_{ijk}}, V_{ijW_{ijk}}),$$

where $\mathrm{Gau}(b_{ijW_{ijk}}, V_{ijW_{ijk}})$ denotes a Gaussian distribution with mean $b_{ijW_{ijk}}$ and variance $V_{ijW_{ijk}}$.

We currently take the "bias" parameters $\{b_{ijw} : w \in \mathcal{W}\}$ to be

$$b_{ijw} = b_j,$$

to capture small-scale station-specific variation. The $\{b_j\}$ are assumed to be independently Gaussian distributed *a priori*, with an inverse-gamma distributed variance; that is,

$$b_j \sim \mathrm{Gau}(0, V_b), \text{ with } V_b^{-1} \sim \mathrm{Gam}(h_1^b, h_2^b),$$

where $\mathrm{Gam}(h_1^b, h_2^b)$ denotes a gamma distribution with mean h_1^b/h_2^b and variance $h_1^b/(h_2^b)^2$, with both h_1^b and h_2^b specified in advance.

Our initial model for the variance parameters $\{V_{ijw} : i = 1, \ldots, I,\ j = 1, \ldots, J,\ w \in \mathcal{W}\}$ is

$$V_{ijw}^{-1} = \phi_{ijw} = \phi_{1,w}\phi_{2,j}\phi_{3,i},$$

which captures phase, station, and event specific variation in the precision of the arrival data given the expected arrival times. The phase-specific scaling parameters $\{\phi_{1,w}\}$ are given a gamma prior distribution;

$$\phi_{1,w} \sim \mathrm{Gam}(h_{11,w}^V, h_{12,w}^V),$$

with $\{h_{11,w}^V, h_{12,w}^V\}$ specified in advance, allowing for a different prior for each $\phi_{1,w}$. The $\{\phi_{2,j}\}$ and $\{\phi_{3,i}\}$ are treated slightly different. We assume that

$$\begin{aligned} &\phi_{2,j} \sim \mathrm{Gam}(1 + r_2, r_2), \text{ with} \\ &r_2^{-1} \sim \mathrm{Gam}(h_{21}^V, h_{22}^V), \end{aligned} \quad (8)$$

with h_{21}^V and h_{22}^V given. Note that a gamma distribution with shape equal to $1 + r_2$ and rate r_2 has a mode at 1 for all values of r_2. Since r_2 is not assumed known a priori, the $\{\phi_{2,j}\}$ are *marginally* correlated as they all provide some information about the value of r_2. This is not the case for the $\{\phi_{1,w}\}$. The event-specific precision parameters $\{\phi_{3,i}\}$ are modeled analogously to the $\{\phi_{2,j}\}$;

$$\begin{aligned} &\phi_{3,i} \sim \text{Gam}(1 + r_3, r_3), \text{ with} \\ &r_3^{-1} \sim \text{Gam}(h_{31}^V, h_{32}^V). \end{aligned} \tag{9}$$

In the case when $W_{ijk} = 0$, that is when a_{ijk} is not assumed to be a valid observation of any of the N phases considered, a_{ijk} is given a broad (non-informative) distribution. We take,

$$a_{ijk} \sim \text{Gau}(\mu_{ijk}, \sigma_0^2),$$

where $\mu_{ijk} = A_{ijw_{ijk}}$ (the expected arrival-time given the observed phase-label) and σ_0^2 is a large variance parameter specified in advance.

The Phase-Assignment Data Model

The phase-assignment model part of (7), $p(\mathbf{w} \mid \mathbf{a}, \mathbf{o}, \mathbf{T}, \mathbf{W}, \boldsymbol{\sigma}_w)$, is not only a phase-assignment model, but rather group-assignment model. This is because we allow W_{ijk} to assign observations to the "does not belong to any known phase" category by putting $W_{ijk} = 0$. This can be particularly useful to guard against "outliers" which can arise under a number of circumstances.

The phase-assignment model is a discrete distribution over the vector of possible phase assignments that one could expect to see for a given event-station pair. For example, if the set of possible phase-labels (indicators) $\mathcal{W} = \{1, 2, 3\}$ and the i-th event results in two picked arrivals at the j-th station ($n_{ij} = 2$), the possible configurations for the observed phase-vector $\mathbf{w}_{ij} = (w_{ij1}, w_{ij2})$ are given by

$$\mathcal{W}_2 \equiv \{(1,2), (2,1), (1,3), (3,1), (2,3), (3,2)\}.$$

On the other hand, the possible configurations for $\mathbf{W}_{ij} = (W_{ij1}, W_{ij2})$ are given by

$$\begin{aligned} \mathcal{W}_2^+ \equiv \{&(1,2), (2,1), (1,3), (3,1), (2,3), (3,2), \\ &(1,0), (2,0), (3,0), (0,1), (0,2), (0,3), (0,0)\}. \end{aligned}$$

Currently, we assume a very simple phase-assignment model given by

$$p(\mathbf{w} \mid \mathbf{a}, \mathbf{o}, \mathbf{T}, \mathbf{W}, \boldsymbol{\sigma}_w) = \prod_{i,j} p(\mathbf{w}_{ij} \mid \mathbf{W}_{ij}),$$

where

$$p(\mathbf{w}_{ij} \mid \mathbf{W}_{ij}) = \begin{cases} \rho_{ij} & \text{if } \mathbf{w}_{ij} = \mathbf{W}_{ij}, \\ \frac{1-\rho_{ij}}{|\mathcal{W}_{n_{ij}}|-1} & \text{otherwise,} \end{cases}$$

with $\rho_{ij} \in [0, 1]$, a pre-specified phase-picking accuracy probability.

The above model barely scratches the surface of taking advantage of the observed phase assignments. However, we expect most of the data information to be carried in the arrival times. For example, if a Pg arrival is wrongly labeled as a Lg arrival, we expect the discrepancy in the observed and the predicted arrival times to flag that error.

2.2 The Process Model

Recall the process model in (4). As in the case of the data model, it is natural to factor the process model as,

$$p(\mathbf{T}, \mathbf{W} \mid \mathbf{F}, \boldsymbol{\tau}) = p(\mathbf{W} \mid \mathbf{T}, \boldsymbol{\tau}_W) p(\mathbf{T} \mid \mathbf{F}, \boldsymbol{\tau}_T),$$

where $\boldsymbol{\tau} = \{\boldsymbol{\tau}_W, \boldsymbol{\tau}_T\}$.

The Phase-Assignment Process Model

The (prior) phase-assignment model $p(\mathbf{W} \mid \mathbf{T}, \boldsymbol{\tau}_W)$ can be taken to be rather simple. Note that this model simply assigns probabilities to the different possible phase configurations that each $\mathbf{W}_{ij}$ can take *prior* to observing any data, but conditional on the expected travel-times. Without looking at the data one does not know much about what phases were observed. We assume a uniform (non-informative) prior for $\mathbf{W}$,

$$p(\mathbf{W} \mid \mathbf{T}, \mathbf{F}, \boldsymbol{\tau}_W) = \prod_{i,j} p(\mathbf{W}_{ij}),$$

where $p(\mathbf{W}_{ij}) = 1/|\mathcal{W}_{n_{ij}}^+|$ and $|\mathcal{W}_{n_{ij}}^+|$ is the number of different phase configurations $\mathbf{W}_{ij}$ can assume.

The Travel-Time Process Model

We assume the following model for the expected travel-times (conditional on the predicted travel-times),

$$T^w(\mathbf{x}_i, \mathbf{s}_j) = \alpha^w(\mathbf{x}_i, \mathbf{s}_j) + \beta^w(\mathbf{x}_i, \mathbf{s}_j) F^w(\mathbf{x}_i, \mathbf{s}_j),$$

where $\alpha^w(\cdot, \cdot)$ and $\beta^w(\cdot, \cdot)$ are stochastic phase-specific travel-time shift and scaling model-correction terms, respectively. We currently take

$$\alpha^w(\mathbf{x}_i, \mathbf{s}_j) = \alpha_0^w \text{ and } \beta^w(\mathbf{x}_i, \mathbf{s}_j) = \beta_0^w,$$

where both α_0^w and β_0^w are assigned Gaussian prior distributions. The additive term α_0^w can be thought of as correcting for variation in regional crust thickness while β_0^w can be thought as correcting for variation in travel speed.

Obvious extensions are to allow the two correction terms to vary with location. For example, assume that

$$\alpha^w(\mathbf{x}_i, \mathbf{s}_j) = \alpha_0^w + \alpha_1^w(\mathbf{x}_i) + \alpha_1^w(\mathbf{s}_j),$$

where $\alpha_1^w(\cdot)$ is given a spatially correlated Gaussian prior.

2.3 The Prior Knowledge

We take the prior distribution for the origin location $\mathbf{x}_i$ to be multivariate Gaussian where the event depth is log-transformed (i.e., $(x_i, y_i, \log z_i)' \sim$ Gaussian). Similarly, the origin-times $\{o_i\}$ are assigned independent Gaussian prior distributions. In most cases we have vague prior information about the event origin parameters. However, there are events that have well established origins that we would like to include in our analysis to provide, indirectly, information about other parameters of the model (e.g., travel-time correction parameters).

3. Posterior Inference

Realizations from the joint posterior distribution (6) are generated using a Markov chain Monte Carlo (MCMC) algorithm with a mix of Gibbs and Metropolis transition steps (e.g., Gelman et al., 2004, p. 292). At each iteration, the parameters are sampled in the following order:

1. Origin times $\{o_i\}$, each individually.
2. Origin locations $\{\mathbf{x}_i\}$, each individually.
3. Phase-assignments $\{\mathbf{W}_{ij}\}$, each individually.
4. Travel-time shift and scaling parameters $\{\alpha_0^w, \beta_0^w\}$, each individually.
5. Station-specific variations $\{b_j\}$, each individually.
6. Precision parameters $\{\phi_{1,w}, \phi_{2,j}, \phi_{3,i}\}$, each individually.
7. Precision variation parameters $\{r_2, r_3\}$, each individually.

Given the current model, we note that

$$a_{ijk} = o_i + \alpha_0^w + \beta_0^w F_{ij}^w + b_j + \varepsilon_{ijw},\ w = W_{ijk}, \quad (10)$$

where $\varepsilon_{ijw} \sim \text{Gau}(0, V_{ijw})$. Note that the parameters $\{o_i\}$, $\{\alpha_0^w\}$, and $\{b_j\}$ all yield a shift in the expected arrival time and as such are highly correlated, and therefore difficult to sample efficiently on an individual basis, as outlined above. To improve chain mixing, the following two transition kernels are also applied:

8. Origin times $\{o_i\}$, travel-time shifts $\{\alpha_0^w\}$, and station-specific variations $\{b_j\}$, all sampled jointly.
9. Travel-time shift and scaling parameters $\{\alpha_0^w, \beta_0^w\}$, all sampled jointly.

We now give some more details on how each of these transition steps are carried out.

3.1 Sampling Origin Times, Travel-Time Parameters, and Station-Specific Variations

The origin times $\{o_i\}$, the travel-time correction parameters $\{\alpha_0^w, \beta_0^w\}$, and the station-specific variation parameters $\{b_j\}$ all have a Gaussian prior distribution. And they all interact with the data through the additive model (10), with Gaussian errors. Hence, using standard Gaussian-Gaussian conjugate results (e.g., Gelman et al., 2004, p. 46), the full conditional distribution of each of these parameters is Gaussian. We therefore adopt a Gibbs-update for the individual transitions needed in transition steps 1, 4, and 5.

In addition to the individual parameters having a Gaussian full conditional distribution, the joint distribution of any collection of these parameters has a multivariate Gaussian distribution (e.g., Gelman et al., 2004, p. 578). Hence, for transition steps number 7 and 8, we also adopt Gibbs-updates. However, as these updates involve multivariate Gaussian distributions, and hence matrix computations (including QR decompositions), they are considerably more computationally involved than the single parameter updates. These joint updates are therefore not carried out in every iteration.

3.2 Sampling Origin Locations

Note that origin location parameters $\{\mathbf{x}_i\}$ interact with the observed data and other model parameters only through the travel-time model $F^w(\cdot,\cdot)$, which is a physics-based simulation computer code. This does not leave many options available to construct a transition kernel for the origin locations. We adopt a Metropolis multivariate Gaussian random-walk to sample each $\mathbf{x}_i = (x_i, y_i, z_i)'$, where z_i is the event depth log-transformed. Hence, given the r-th realization $\mathbf{x}_i^{(r)}$, a new proposal $\mathbf{x}_i^*$ is generated from $\text{Gau}(\mathbf{x}_i^{(r)}, \Sigma_i)$, where Σ_i is given. The new proposal is either accepted or rejected as the $(r+1)$-th realization according to the Metropolis acceptance ratio (e.g., Gelman et al., 2004, p. 289).

3.3 Sampling Phase Assignments

The phase-assignment vector $\mathbf{W}_{ij}$ takes possible values in the finite set $\mathcal{W}_{n_{ij}}^+$. This allows us to tabulate the full conditional distribution of $\mathbf{W}_{ij}$,

$$p(\mathbf{W}_{ij} = w \,|\, \mathbf{a}, \mathbf{w}, \cdots), \text{ for } w \in \mathcal{W}_{n_{ij}}^+,,$$

where $\cdots$ denotes all the remaining parameters in the model. The transition step 3 is therefore carried out using Gibbs-updates.

3.4 Sampling Precision Parameters

Since the errors in (10) are Gaussian and the precision parameters $\{\phi_{1,w}, \phi_{2,j}, \phi_{3,i}\}$ have gamma prior distributions, one can show that (e.g., Gelman et al., 2004, p. 50)

$$p(\phi \,|\, \mathbf{a}, \mathbf{w}, \cdots) \text{ is Gamma,}$$

where ϕ is any of $\{\phi_{1,w}, \phi_{2,j}, \phi_{3,i}\}$. Hence, we use Gibbs-updates in transition step 6.

3.5 Sampling Precision Variation Parameters

The precision rate parameters r_2 and r_3, of (8) and (9), respectively, do not yield full conditional distributions that are readily available for sampling. However, these full conditional distributions can be evaluated up to a constant of proportionality and in addition can be shown to be unimodal. This makes it relatively easy to implement a slice-sampler (Neal, 2003) for both r_2 and r_3. This is the approach we take in transition step 7.

4. Application: Nevada Nuclear Test-Site Events

We selected nine nuclear explosion tests from the Nevada test site (NTS) database (Walter et al., 2003) along with a (subset) of arrival data from nine stations recording the seismic activity following these tests; see Figure 1. The arrival data was limited to three phases, $\mathcal{W} = \{\text{Pn}, \text{Pg}, \text{Lg}\}$, yielding a total of 128 arrival times for the nine events, with the number of arrivals observed from each event ranging from 7 (reported by 3 stations) to 23 (reported by 8 stations).

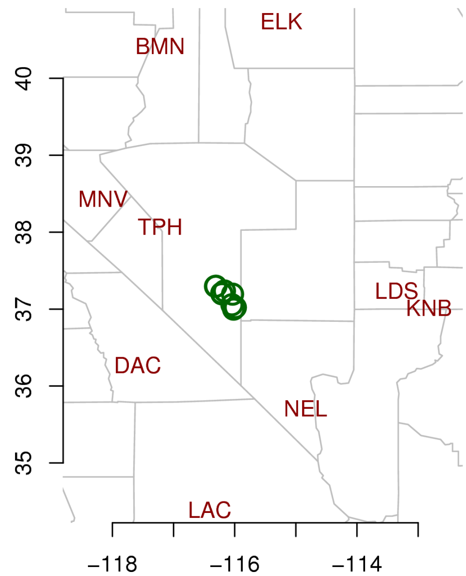

Figure 1: The location of the nine events (circles) and the station considered (three letter id).

Analysis of this arrival data was carried out under different prior and model assumptions, including using corrupted arrival data. We present the result from one of these analysis, where: (1) informative prior was assigned to the origin of two of the events, (2) the arrival data was corrupted, and (3) the model carried out phase/outlier identification. This represents a scenario one could expect, where there are some events that have well established origin (ground truth events) and the arrival data is potentially corrupted with outliers and misidentified phases.

The NTS arrival data was corrupted in two ways: (1) Pn/Pg phase-labels were exchanged in the arrival data for four event-station pairs and (2) 30sec was added to the recorded arrival-time of the Pg phase for four event-station pairs. In both cases, arrival data from the same three events were corrupted (events 591069, 628994, and 576701). The original number of arrivals for these three events was 7, 12, and 17, respectively, but after corrupting the data there are only 4, 9, and 11 "good" arrivals for these events, respectively.

The travel-time model used was the global IASP91 model, which is known to produce biased travel-times for this particular region (Anderson & Myers, 2005).

In sampling from the posterior distribution, 10 MCMC chains were run in parallel, each of length 4,000, with the first 2,000 iterations used for burn-in and tuning of the Metropolis random-walk event-location proposal distributions. Of the 2,000 iterations remaining from each chain, every 10th iteration was stored, yielding a sample of 2,000 iterations (200 samples from each chain).

Figure 2 shows the lat-long posterior distributions for the nine events. Note that each event is drawn at a different scale. The events with corrupted data are the last event in the first row (event 576701) and the first and last events in the second row (events 591069 and 628994). The true lat-long locations of the events is well captured by the posterior distributions, with some of the events within the 50% posterior probability contours and all within the 95% contours. Including the three events with corrupted data.

The corrupted data was identified by the model. This is shown in Table 1 for the four cases where Pn/Pg phase-labels were switched. In all of these cases, the highest posterior probability was put on the correct phase-assignment vector. A similar success was seen for the four cases were 30sec was added to the Pg arrivals.

Analysis of the posterior realizations of the travel-time correction parameters showed significant variation among the phases, both in terms of shift parameters $\{\alpha_0^w\}$ and scaling parameters $\{\beta_0^w\}$, with the Lg phase showing the strongest deviation from the IASP91 travel-time model. The posterior distributions of the phase-specific precision scaling parameters $\{\phi_{1,w}\}$ indicated that Pn arrivals had the greatest influence (greatest precision), followed by Pg, and then Lg. There was also notable difference in the station-specific precision parameters $\{\phi_{2,j}\}$. But there was less variation among the event-specific precision parameters $\{\phi_{3,i}\}$.

These initial results are very promising. With corrupted data, we are both able to construct realistic posterior distributions for the event lat-long locations (corrupted and non-corrupted events) and also identify the corrupted data. The travel-time correction model is relatively simple and future efforts aim at a more realistic travel-time correction model with spatially-varying travel-time corrections. Additional improvements under investigation are event-specific precision parameters that vary with travel-time length, to mirror how event magnitude might impact the arrival "picking" accuracy.

```
Event 591069, Station ELK:
     O[Lg] O[Pg] O[Pn]   Prob CumProb
  1     Lg    Pn     Pg 0.9825  0.9825
  2     Lg    Pn     NA 0.0070  0.9895
  3     NA    Pn     Pg 0.0060  0.9955
  4     Lg    NA     Pg 0.0045  1.0000
Event 628994, Station KNB:
    O[Lg] O[Pn] O[Pg]    Prob CumProb
  1     Lg    Pg     Pn 0.9645  0.9645
  2     NA    Pg     Pn 0.0175  0.9820
  3     Lg    NA     Pn 0.0110  0.9930
  4     Lg    Pg     NA 0.0070  1.0000
Event 576701, Station BMN:
    O[Lg] O[Pg] O[Pn]    Prob CumProb
  1     Lg    Pn     Pg 0.9135  0.9135
  2     NA    Pn     Pg 0.0735  0.9870
  3     Lg    Pn     NA 0.0095  0.9965
  4     Lg    NA     Pg 0.0030  0.9995
Event 576701, Station NEL:
    O[Lg] O[Pg] O[Pn]    Prob CumProb
  1     Lg    Pn     Pg 0.9700  0.9700
  2     Lg    NA     Pg 0.0110  0.9810
  3     NA    Pn     Pg 0.0095  0.9905
  4     Lg    Pn     NA 0.0085  0.9990
  5     NA    NA     Pg 0.0005  0.9995
```

Table 1: Identifying the four cases where Pn/Pg phase labels were switched in the corrupted data. The first line in each table shows the observed (called) phase labels, with Lg correct, but Pn and Pg switched. These are then followed with the phase-label vectors with the highest posterior probability. The phase label `NA` denotes an outlier.

5. Acknowledgment

This work was performed under the auspices of the US Department of Energy by the Lawrence Livermore National Laboratory under contract number W-7405-ENG-48.

References

Anderson, M. & Myers, S. C. (2005). Assessment of regional-distance location calibration using a multiple event location algorithm. *The Bulletin of the Seismological Society of America*. In perparation.

Bondár, I., Myers, S. C., Engdahl, E. R., & Bergman, E. A. (2004). Epicentre accuracy based on seismic network criteria. *Geophysical Journal International*, *156*, 483–496.

Dewey, J. (1972). Seismicity and tectonics of western venezuela. *The Bulletin of the Seismological Society of America*, *62*, 1711–1751.

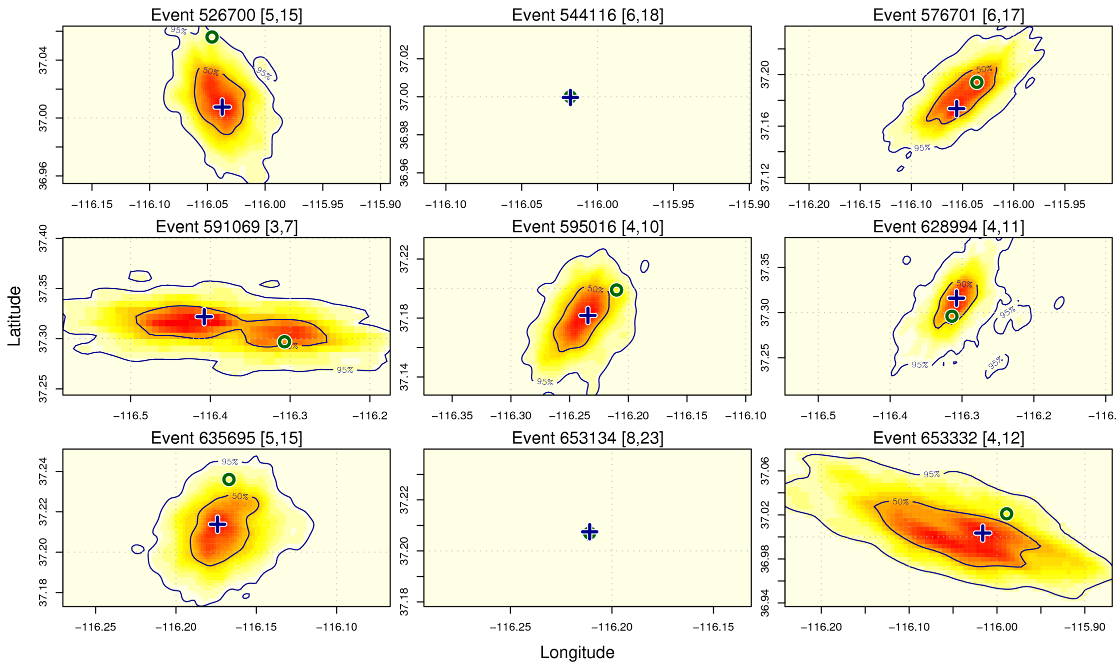

Figure 2: Lat-long posterior probability distributions for the nine events. Circles show the true location of the events, while crosses show the posterior model. 50% and 95% posterior probability contours are drawn for each event.

Douglas, A. (1967). Joint epicentre determination. *Nature*, *215*, 47–48.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian Data Analysis* (2nd ed.). Boca Raton, Florida: Hapman and Hall/CRC.

Jordan, T. H. & Sverdrup, K. A. (1981). Teleseismic location techniques and their application to earthquake clusters in the south-central pacific. *The Bulletin of the Seismological Society of America*, *71*, 1105–1130.

Kennett, B. L. N. & Engdahl, E. R. (1991). Traveltimes for global earthquake locations and phase identification. *Geophysical Journal International*, *105*(429–465).

Neal, R. M. (2003). Slice sampling (with discussion). *Annals of Statistics*, *31*, 705–767.

Pavlis, G. L. & Booker, J. R. (1983). Progressive multiple event location (pmel). *The Bulletin of the Seismological Society of America*, *73*, 1753–1777.

Rodi, W. & Toksöz, M. N. (2000). Grid-search techniques for seismic event location. In *The Proceedings of the 22nd Annual DoD/DOE Seismic Research Symposium*, New Orleans. Defense Threat Reduction Agency (DTRA).

Waldhauser, F. & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the North Hayward fault, California. *The Bulletin of the Seismological Society of America*, *90*, 1353–1368.

Walter, W. R., Smith, K. D., O'Boyle, J. L., Hauk, T. F., Ryall, F., Ruppert, S. D., Myers, S. C., Anderson, M., & Dodge, D. A. (2003). Improving the fundametal understanding of regional seismic signal processing with a unique Western United States dataset. In *Proceedings of the 25th Seismic Research Review*, Tucson, AZ.

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under contract W-7405-Eng-48.