LAWRENCE LIVERMORE NATIONAL LABORATORY

LLNL-TR-658716

# A Bayesian Framework for Locating Seismic Events Using Absolute Arrival Time Data along with Back Azimuth and Slowness Observations

G. Johannesson, S. C. Myers

August 12, 2014

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

# A Bayesian Framework for Locating Seismic Events Using Absolute Arrival Time Data along with Back Azimuth and Slowness Observations

Gardar Johannesson and Stephen C. Myers

Lawrence Livermore National Laboratory

Version: August 13, 2014

LLNL-TR-658716

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

**Abstract**

We outline a Bayesian framework to blend together observed seismic arrival times, back azimuth, and slowness data to locate seismic events. Previously we have developed a Bayesian seismic event locator, BayesLoc, that uses seismic arrival times of multiple phases to locate seismic activity. The approach taken to incorporate back azimuth and slowness data very much reflects the treatment the arrival time data in the current version of BayesLoc; by introducing statistical "bias" corrections and "precision" factors. There is a work in progress to extend the arrival time only BayesLoc program to take advantage of possible back azimuth and slowness data following the template outlined in this document.

# 1 Introduction

The problem of locating seismic events can be viewed as a classical inversion problem: given a forward model and observed output data, invert for unknown seismic origin input parameters to the forward model that "best" explain the observed output data. A particular successful approach to such inversion/calibration problems is a Bayesian approach, which treats the unknown inputs/parameters as stochastic and therefore yields a probabilistic solution; see for example Kaipio and Somersalo (2004) and Tarantola (2004) for a general introduction to the topic.. Bayesian methods have been applied successfully to locate seismic events, as for example in BayesLoc (Myers et al., 2007, 2009; Johannesson et al., 2010).

One of the major strength of the Bayesian approach is its probabilistic characterization of uncertainties in the estimated seismic source parameters, which can be critical for downstream inference/decision making (e.g., the probability that a given seismic event is within a given region or simply a map of the region that is deemed to contain the event with say 90% probability). Another, and often overlooked, strength of Bayesian inversion is the general statistical approach to describe all the sources of errors in the whole process, from the unknown seismic source parameters to the observed output, which includes how the errors in the observations are treated and how a bias in the forward model is handled (e.g. travel times of various seismic phases). This is particularly important when the noise (the errors) in the observed data, say the picked absolute arrival times of various seismic phases at a collection of stations, is far from being homogeneous, with some station producing more accurate observations than others and with some phases being more easily "picked" than others. Similarly, the assumed travel-time model has biases which varies with phases and stations, all of which can be described with statistical models.

In short, a successful Bayesian seismic inversion doesn't only yield a probabilistic characterization of the unknown seismic origin parameters, but also a probabilistic correction to the assumed travel-time model and correctly "weights" the observed arrival times to reflect the variation in accuracy across stations and phases. This is exactly what BayesLoc does when locating multiple seismic events given observed arrival times.

The general Bayesian approach used to "correct" the underling travel time model and "weight" the observed arrival data differently can also be applied to take advantage of a totally new source of seismic observations that related to the seismic origin parameters, for example back azimuth and slowness (Rost and Thomas, 2002). As for the arrival time data, biases in the "forward model" for the back azimuth and slowness are estimated, within the Bayesian framework, and the accuracy of the back azimuth/slowness data is characterized. As a result, the estimated seismic origin parameters are both impacted by the arrival time data and the back azimuth/slowness data, but the "weight" of each data source reflects the fidelity of the forward model and the accuracy of the observed data, both of which are estimated simultaneously along with the unknown source parameters.

What follows is a brief overview of the BayesLoc model for absolute arrival time data (Myers et al., 2007, 2009; Johannesson et al., 2010). We then outline how the current BayesLoc model can be augmented with back azimuth and slowness data, to yield stronger inference on the unknown seismic origin parameters.

## 2 BayesLoc for Arrival Time Data from a Cluster of Events

Starting with the (absolute) arrival time data, let

$a_{wij} \equiv$ the arrival time of phase $w$ from the $i$-th event to the $j$-th seismic station.

We note that not all combination of $wij$ have observed arrival times and are simply treated as "missing" (to simplify notation). The forward model in this case is the predicted travel time for each phase of interested from the unknown event locations to the known seismic stations,

$F_w(\mathbf{s}_i, \mathbf{r}_j) = F_{wij} \equiv$ the predicted travel time of phase $w$ from the $i$-th event to the $j$-th seismic station,

where

$\mathbf{s}_i = (x_i^s, y_i^s, z_i^s) \equiv$ the location of the $i$-th event and
$\mathbf{r}_j = (x_j^r, y_j^r, z_j^r) \equiv$ the location of the $j$-th seismic station.

The predicted arrival time is then given by $o_i + F_{wij}$, where

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

## 2.1 Bayesian Formulation

At the core of the Bayesian formulation of the seismic location problem in BayesLoc are conditional and prior probability distribution functions (PDFs), which given data, yields the posterior PDF of all the unknowns; see Gelman et al. (2013) for an overview of building hierarchical Bayesian models for data. The core conditional PDF is that of the observed arrival time data, given a particular realization of the unknown seismic origin parameters, along with all possible travel time correction parameters and arrival time residuals precision parameters. The core prior PDFs are then those that express our prior knowledge about the seismic origin parameters (lat, long, depth, and time), along with prior PDFs for possible travel-time corrections and precision parameters, to be outlined below.

To start with, the observed arrival time data is assumed to be independently Gaussian distributed;

$$p(a_{wij} \mid o_i, \mathbf{s}_i, \delta_{wij}, \kappa_{wij}) = \mathrm{Gau}(a_{wij} \mid o_i + F_w(\mathbf{s}_i, \mathbf{r}_j) + \delta_{wij}, \kappa_{wij}^{-2}), \qquad (1)$$

where $p(\cdot \mid \cdots)$ denotes a conditional PDF, $\mathrm{Gau}(y \mid \mu, \sigma^2)$ denotes the Gaussian PDF (for $y$) with mean $\mu$ and variance $\sigma^2$, $\delta_{wij}$ is a travel time correction to the assumed travel time model, and $\kappa_{wij}$ is the precision of the Gaussian PDF, equal to the inverse variance.

The travel time correction applied in BayesLoc for event clusters is given by an additive model of various correction factors;

$$\delta_{wij} = \beta_{1,w} + \beta_{2,w} D_{ij} + \beta_{3,j} + \beta_{4,wj}, \qquad (2)$$

where $\beta_{1,w}$ is a phase-specific shift in the travel time curve, while $\beta_{2,w}$ capture event-to-station distance related deviation, with $D_{ij} = \|\mathbf{s}_i - \mathbf{r}_j\|_g$ as the geodesic distance (in degrees) between the $i$-the event and the $j$-th station The $\beta_{3,j}$ and $\beta_{3,wj}$ are somewhat different and capture station and station-phase specific travel time corrections. In the Bayesian formulation, the $\beta$'s are all assumed unknown and stochastic with the prior distribution for $\beta_{1,w}$ and $\beta_{2,w}$ taken as

$$p(\beta_{1,w}) = \mathrm{Gau}(\beta_{1,w} \mid M_{\beta_1 w}, V_{\beta_1 w}), \quad \text{for } w = 1, \ldots, n^{\mathrm{ph}}, \text{ and}$$
$$p(\beta_{2,w}) = \mathrm{Gau}(\beta_{2,w} \mid M_{\beta_2 w}, V_{\beta_2 w}), \quad \text{for } w = 1, \ldots, n^{\mathrm{ph}},$$

where the means and the variances, the $M$'s and the $V$'s, are assumed known. On the other hand, the $\beta_{3,j}$'s and the $\beta_{4,wj}$'s are treated slightly differently where we assume;

$$p(\beta_{3,j} \,|\, \tau_{\beta_3}) = \mathrm{Gau}(\beta_{3,j} \,|\, 0, \tau_{\beta_3}^{-2}), \quad \text{for } j = 1, \ldots, n^{\mathrm{sta}}, \text{ and}$$
$$p(\beta_{4,wj} \,|\, \tau_{\beta_4 w}) = \mathrm{Gau}(\beta_{4,wj} \,|\, 0, \tau_{\beta_4 w}^{-2}), \quad \text{for } w = 1, \ldots, n^{\mathrm{ph}},\ j = 1, \ldots, n^{\mathrm{sta}},$$

where further the $\tau$'s are assumed unknown and stochastic. The prior PDFs fo the $\tau$'s are taken as;

$$p(\tau_{\beta_3}) = \mathrm{Gam}(\tau_{\beta_3} \,|\, A_{\beta_3}, B_{\beta_3}) \quad \text{and}$$
$$p(\tau_{\beta_4 w}) = \mathrm{Gam}(\tau_{\beta_4,w} \,|\, A_{\beta_4 w}, B_{\beta_4 w}), \quad \text{for } w = 1, \ldots, n^{\mathrm{ph}},$$

where $\mathrm{Gam}(y \,|\, \alpha, \beta)$ denotes a Gamma PDF (for $y$) with shape parameter $\alpha$ and scale parameter $\beta$ (i.e., a Gamma PDF with mean $\alpha/\beta$ and variance $\alpha/\beta^2$).

**Few remarks**. The $\beta_1$'s and $\beta_2$'s capture large-scale corrections to the assumed travel time model (e.g., ak135), while the $\beta_3$'s and $\beta_4$'s capture small-scale shifts at each station (the $\beta_3$'s) and further at each station-phase (the $\beta_4$'s). Note that the distribution of the $\beta_3$'s and $\beta_4$'s is centered on zero a priori, but with unknown variance (such variables are often termed as random effects in the statistics literature). This formulation drives the $\beta_3$'s and $\beta_4$'s to concentrated as much as possible around zero (yielding higher total likelihood), with the tightness controlled by the $\tau$'s. The prior PDF for the $\tau$'s is typically taken to yield vague prior knowledge, as to let the posterior PDF of the $\tau$'s, and hence the spread of the $\beta_3$'s and the $\beta_4$'s to be drive by the need to correct the underlying travel time model at the station and station-phase level for a particular event cluster.

The Bayesian treatment of the arrival time residual precision, $\kappa_{wij}$, mirrors that of the travel time corrections, with

$$\kappa_{wij} = \kappa_{1,w}\kappa_{2,i}\kappa_{3,j}, \tag{3}$$

where we refer to the $\kappa$'s as precision factors. The precision factors are assumed unknown and stochastic with the following know prior for the phase-specific precision factor;

$$p(\kappa_{1,w}) = \mathrm{Gam}(\kappa_{1,w} \,|\, A_{\kappa_1 w}, B_{\kappa_1 w}), \quad \text{for } w = 1, \ldots, n^{\mathrm{ph}},$$

where $A_{\kappa_1 w}$ and $B_{\kappa_1 w}$ are known. We treat the $\kappa_2$'s and the $\kappa_3$'s as precision random scaling effects by assuming that

$$p(\kappa_{2,i} \,|\, \lambda_2) = \mathrm{Gam}(\kappa_{2,i} \,|\, \lambda_{\kappa_2}, \lambda_{\kappa_2}) \quad \text{for } i = 1, \ldots, n^{\mathrm{ev}}, \text{ and}$$
$$p(\kappa_{3,j} \,|\, \lambda_3) = \mathrm{Gam}(\kappa_{3,j} \,|\, \lambda_{\kappa_3}, \lambda_{\kappa_3}) \quad \text{for } j = 1, \ldots, n^{\mathrm{sta}},$$

where $\lambda_{\kappa_2}$ and $\lambda_{\kappa_3}$ are unknown and assigned a vague Gamma prior PDFs;

$$p(\lambda_{\kappa_2}) = \mathrm{Gam}(\lambda_{\kappa_2} \,|\, S_{\kappa_2}, B_{\kappa_2}) \quad \text{and} \quad p(\lambda_{\kappa_3}) = \mathrm{Gam}(\lambda_{\kappa_3} \,|\, S_{\kappa_3}, B_{\kappa_3}).$$

**Few remarks.** The $\kappa_{1,w}$'s capture the overall precision of the corrected arrival time residuals $(a_{wij} - (o_i + F_{wij} + \delta_{wij}))$ for each phase $w$. The $\kappa_2$'s and the $\kappa_3$'s can be interpreted as event- and station precision random scaling factors that are assumed to come from a Gamma PDF with mean equal to one (and variance of $1/\lambda$). The $\lambda$'s are then typically assigned a vague prior PDFs as to let the data drive the shape of the posterior PDFs (and the spread of the precision scaling factors around one).

## 2.2 Posterior Inference

The joint PDF of the arrival time data and all the unknown variables (the origin parameters plus both travel time correction and precision parameters) is given by

$$\begin{aligned} p(\mathbf{a}, \mathbf{s}, \mathbf{o}, \boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\kappa}, \boldsymbol{\lambda}) = & \; p(\mathbf{a} \,|\, \mathbf{o}, \mathbf{F}(\mathbf{s}), \boldsymbol{\beta}, \boldsymbol{\kappa}) \\ & \times p(\boldsymbol{\beta} \,|\, \boldsymbol{\tau}) p(\boldsymbol{\tau}) \\ & \times p(\boldsymbol{\kappa} \,|\, \boldsymbol{\lambda}) p(\boldsymbol{\lambda}) \\ & \times p(\mathbf{s}) p(\mathbf{o}), \end{aligned} \tag{4}$$

where

$$p(\mathbf{a} \,|\, \mathbf{o}, \mathbf{F}(\mathbf{s}), \boldsymbol{\beta}, \boldsymbol{\kappa}) = \prod_{wij} \mathrm{Gau}(a_{wij} \,|\, o_i + F_w(\mathbf{s}_i, \mathbf{r}_j) + \delta_{wij}(\boldsymbol{\beta}), \kappa_{wij}(\boldsymbol{\kappa})^{-1}),$$

and $\mathbf{a}$ is a vector containing all the arrival data, $\mathbf{o}$ a vector of all the origin times, $\mathbf{F}(\mathbf{s})$ a vector of all the predicted travel times (for a given collection of origin locations, $\mathbf{s}$), $\boldsymbol{\beta}$ is a vector of all the travel time correction parameters (and $\delta_{wij}(\boldsymbol{\beta})$ the resulting travel time correction), and $\boldsymbol{\kappa}$ is a vector of all the precision factors (and $\kappa_{wij}(\boldsymbol{\kappa})$ the resulting precision). The remaining PDFs in (4) are similarly defined as the product of the individual PDFs, with relevant variables grouped into vectors. We can simplify notation considerable by letting

$$\boldsymbol{\alpha} = (\boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\kappa}, \boldsymbol{\lambda}),$$

that is grouping all the travel time correction and the precision parameters into a single vector. The joint (prior) PDF of $\boldsymbol{\alpha}$ is then given by

$$p(\boldsymbol{\alpha}) = p(\boldsymbol{\beta} \,|\, \boldsymbol{\tau}) p(\boldsymbol{\tau}) p(\boldsymbol{\kappa} \,|\, \boldsymbol{\lambda}) p(\boldsymbol{\lambda})$$

and the joint PDF of (4) becomes

$$p(\mathbf{a}, \mathbf{s}, \mathbf{o}, \boldsymbol{\alpha}) = p(\mathbf{a} \,|\, \mathbf{o}, \mathbf{F}(\mathbf{s}), \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) p(\mathbf{s}) p(\mathbf{o}). \tag{5}$$

The posterior PDF of interest, which conditions on the observed arrival time data, is then simply proportional to the joint PDF,

$$\begin{aligned} p(\mathbf{s}, \mathbf{o}, \boldsymbol{\alpha} \,|\, \mathbf{a}) = p(\mathbf{a}, \mathbf{s}, \mathbf{o}, \boldsymbol{\alpha})/p(\mathbf{a}) \propto p(\mathbf{a}, \mathbf{s}, \mathbf{o}, \boldsymbol{\alpha}) \\ p(\mathbf{a} \,|\, \mathbf{o}, \mathbf{F}(\mathbf{s}), \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) p(\mathbf{s}) p(\mathbf{o}). \end{aligned} \tag{6}$$

Markov chain Monte Carlo (MCMC) algorithm is used to sample from the joint posterior PDF of (6) using a mixture of conditional Gibbs samplers and Metropolis-Hasting random-walk samplers (see Johannesson et al. 2010 for details and e.g. Robert and Casella 2004 for a general overview of Monte Carlo sampling techniques). The end result is a sample of realizations from (6) which can be used to estimate any summary statistics of interest (e.g., the posterior mean epicenter of each event, including summary statistics for the travel time correction parameters and the precision factors).

# 3 Incorporating Back Azimuth and Slowness Data into BayesLoc

We will take a similar approach to incorporate back azimuth and slowness observations into BayesLoc as we did with the arrival data; by formulating a joint distribution similar to the one in (4) that links the observed data to the unknown origin parameters.

Let

$\theta_{wij} \equiv$ the observed back azimuth for phase $w$, event $i$, and sta-, tion $j$

$u_{wij} = 1/v_{wij}^{\text{app}} \equiv$ the observed slowness, equal to one over the apparent velocity of the wave front at the station array, for phase $w$, event $i$, and station $j$.

As for the arrival data, not all combination of $wij$ are observed and are simply treated as missing. The forward models in this case are

$F^{\theta}(\mathbf{s}_i, \mathbf{r}_j) = F_{ij}^{\theta} \equiv$ the back azimuth (clockwise from north) from station $j$ to event $i$,

$F_w^{u}(\mathbf{s}_i, \mathbf{r}_j) = F_{wij}^{u} \equiv$ the predicted slowness for phase $w$, from event $i$ at station $j$.

Note that the predicted back azimuth is the same for all phases, but we can potentially have observations of the back azimuth derived from multiple phases for a given event-station pair.

## 3.1 Bayesian Formulation

The Bayesian formulation for including back azimuth and slowness data follows the template used to incorporate the arrival time data.

### 3.1.1 Back Azimuth

Starting with the back azimuth data, we assume that

$$p(\theta_{wij} \,|\, \mathbf{s}_i, \delta^\theta_{wij}, \kappa^\theta_{wij}) = \mathrm{vMF}(\theta_{wij} \,|\, F^\theta(\mathbf{s}_i, \mathbf{r}_j) + \delta^\theta_{wij}, \kappa^\theta_{wij}), \tag{7}$$

where $\mathrm{vMF}(\theta \,|\, \mu, \kappa)$ denotes a von-Mises-Fisher distribution (or just von-Mises in this case) for $\theta$ with mean angle $\mu$ and concentration parameter $\kappa$ (the variance of the distribution is $1 - I_1(\kappa)/I_0(\kappa)$, where $I_1(\cdot)$ and $I_0(\cdot)$ are the modified Bessel functions of order 1 and 0, respectively). The von-Mises distribution approximates a wrapped Gaussian distribution on the circle.

We take the back azimuth correction $\delta^\theta_{wij}$ to be of the form

$$\delta^\theta_{wij} = \beta^\theta_{1,j} + \beta^\theta_{2,wj},$$

which is very similar in spirit to the travel time corrections applied to the arrival data, with $\beta^\theta_1$'s capturing the overall station corrections to the predicted back azimuth to the event cluster and the $\beta^\theta_2$'s capturing station-phase interactions in the corrections. Both variables are treated as random effects, meaning that

$$\begin{aligned} p(\beta^\theta_{1,j} \,|\, \tau_{\beta^\theta_1}) &= \mathrm{vMF}(\beta^\theta_{1,j} \,|\, 0, \tau_{\beta^\theta_1}), \quad \text{for } j = 1, \ldots, n^{\mathrm{sta}}, \text{ and} \\ p(\beta^\theta_{2,wj} \,|\, \tau_{\beta^\theta_2 w}) &= \mathrm{vMF}(\beta^\theta_{2,wj} \,|\, 0, \tau_{\beta^\theta_2 w}), \quad \text{for } w = 1, \ldots, n^{\mathrm{ph}}, \ j = 1, \ldots, n^{\mathrm{sta}}. \end{aligned}$$

Hence, the collection of the station corrections is centered around zero, with the concentration parameter $\tau_{\beta^\theta_1}$ controlling how tight the population is around zero, and similarly for the station-phase corrections, with a different concentration parameter for each phase. The concentration parameters are treated as unknown and stochastic with the following known prior distributions;

$$\begin{aligned} p(\tau_{\beta^\theta_1}) &= \mathrm{Gam}(\tau_{\beta^\theta_1} \,|\, A_{\beta^\theta_1}, B_{\beta^\theta_1}) \quad \text{and} \\ p(\tau_{\beta^\theta_2 w}) &= \mathrm{Gam}(\tau_{\beta^\theta_2, w} \,|\, A_{\beta^\theta_2 w}, B_{\beta^\theta_2 w}), \quad \text{for } w = 1, \ldots, n^{\mathrm{ph}}, \end{aligned}$$

where the $A$'s and the $B$'s are known (and typically specified to yield a vague prior knowledge).

The concentration parameter $\kappa^\theta_{wij}$ of (7) is treated similarly as in the case of the arrival time data and factored into the following concentration scaling factors,

$$\kappa^\theta_{wij} = \kappa^\theta_{1,w}\kappa^\theta_{2,i}\kappa^\theta_{3,j}.$$

Each of the $\kappa^\theta_{1,w}$ is assigned a known PDF,

$$p(\kappa^\theta_{1,w}) = \mathrm{Gam}(\kappa^\theta_{1,w} \,|\, A_{\kappa^\theta_1 w}, B_{\kappa^\theta_1 w}), \quad \text{for } w = 1,\ldots,n^{\mathrm{ph}}.$$

However, the $\kappa^\theta_2$'s and the $\kappa^\theta_3$'s are treated as random scaling factors with,

$$\begin{aligned} p(\kappa^\theta_{2,i} \,|\, \lambda_{\kappa^\theta_2}) &= \mathrm{Gam}(\kappa^\theta_{2,i} \,|\, \lambda_{\kappa^\theta_2}, \lambda_{\kappa^\theta_2}) \quad \text{for } i = 1,\ldots,n^{\mathrm{ev}}, \text{ and} \\ p(\kappa^\theta_{3,j} \,|\, \lambda_{\kappa^\theta_3}) &= \mathrm{Gam}(\kappa^\theta_{3,j} \,|\, \lambda_{\kappa^\theta_3}, \lambda_{\kappa^\theta_3}) \quad \text{for } j = 1,\ldots,n^{\mathrm{sta}}, \end{aligned}$$

where $\lambda_{\kappa^\theta_2}$ and $\lambda_{\kappa^\theta_3}$ are unknown and given known Gamma prior PDFs,

$$p(\lambda_{\kappa^\theta_2}) = \mathrm{Gam}(\lambda_{\kappa^\theta_2} \,|\, S_{\kappa^\theta_2}, B_{\kappa^\theta_2}) \quad \text{and} \quad p(\lambda_{\kappa^\theta_3}) = \mathrm{Gam}(\lambda_{\kappa^\theta_3} \,|\, S_{\kappa^\theta_3}, B_{\kappa^\theta_3}).$$

### 3.1.2 Slowness

The slowness data is treated very much like the arrival time data, except with the assumption that the slowness is log-Gaussian distributed, that is,

$$p(\log u_{wij} \,|\, \mathbf{s}_i, \delta^u_{wij}, \kappa^u_{wij}) = \mathrm{Gau}(\log u_{wij} \,|\, F^u_w(\mathbf{s}_i, \mathbf{r}_j) + \delta^u_{wij}, (\kappa^u_{wij})^{-2}). \quad (8)$$

The log-slowness corrections are given by the additive model,

$$\delta^u_{wij} = \beta^u_{1,w} + \beta^u_{2,w} D_{ij} + \beta^u_{3,j} + \beta^u_{4,wj}, \quad (9)$$

which is identical in form as the travel-time correction model in (2). The correction parameters are modeled in an identical fashion as those of the travel time correction model, with the following known priors PDFs for $\beta^u_1$'s and $\beta^u_2$'s;

$$\begin{aligned} p(\beta^u_{1,w}) &= \mathrm{Gau}(\beta^u_{1,w} \,|\, M_{\beta^u_1 w}, V_{\beta^u_1 w}), \quad \text{for } w = 1,\ldots,n^{\mathrm{ph}}, \text{ and} \\ p(\beta^u_{2,w}) &= \mathrm{Gau}(\beta^u_{2,w} \,|\, M_{\beta^u_2 w}, V_{\beta^u_2 w}), \quad \text{for } w = 1,\ldots,n^{\mathrm{ph}}. \end{aligned}$$

On the other hand, the $\beta_3^u$'s and the $\beta_4^u$'s are treated as Gaussian random effects with

$$p(\beta^u_{3,j} \,|\, \tau_{\beta^u_3}) = \mathrm{Gau}(\beta^u_{3,j} \,|\, 0, (\tau_{\beta^u_3})^{-2}), \quad \text{for } j = 1, \dots, n^{\mathrm{sta}}, \text{ and}$$
$$p(\beta^u_{4,wj} \,|\, \tau_{\beta^u_4}) = \mathrm{Gau}(\beta^u_{4,wj} \,|\, 0, (\tau_{\beta^u_4 w})^{-2}), \quad \text{for } w = 1, \dots, n^{\mathrm{ph}},\ j = 1, \dots, n^{\mathrm{sta}},$$

where the $\tau$'s are assumed unknown and stochastic and given known Gamma priors PDFs;

$$p(\tau_{\beta^u_3}) = \mathrm{Gam}(\tau_{\beta^u_3} \,|\, A_{\beta^u_3}, B_{\beta^u_3}) \quad \text{and}$$
$$p(\tau_{\beta^u_4 w}) = \mathrm{Gam}(\tau_{\beta^u_4, w} \,|\, A_{\beta^u_4 w}, B_{\beta^u_4 w}), \quad \text{for } w = 1, \dots, n^{\mathrm{ph}}.$$

The Bayesian treatment of the log-slowness precision, the $\kappa^u_{wij}$ of (8), is also identical to the treatment of the arrival time precision in (3), assuming the following factorization;

$$\kappa^u_{wij} = \kappa^u_{1,w}\kappa^u_{2,i}\kappa^u_{3,j}.$$

The phase-specific precision factors are given a known Gamma priors,

$$p(\kappa^u_{1,w}) = \mathrm{Gam}(\kappa^u_{1,w} \,|\, A_{\kappa^u_1 w}, B_{\kappa^u_1 w}), \quad \text{for } w = 1, \dots, n^{\mathrm{ph}},$$

but the event- and station-specific factors are treated as Gamma random scaling factors;

$$p(\kappa^u_{2,i} \,|\, \lambda_{\kappa^u_2}) = \mathrm{Gam}(\kappa^u_{2,i} \,|\, \lambda_{\kappa^u_2}, \lambda_{\kappa^u_2}) \quad \text{for } i = 1, \dots, n^{\mathrm{ev}}, \text{ and}$$
$$p(\kappa^u_{3,j} \,|\, \lambda_{\kappa^u_3}) = \mathrm{Gam}(\kappa^u_{3,j} \,|\, \lambda_{\kappa^u_3}, \lambda_{\kappa^u_3}) \quad \text{for } j = 1, \dots, n^{\mathrm{sta}},$$

where $\lambda_{\kappa^u_2}$ and $\lambda_{\kappa^u_3}$ are unknown and given known Gamma prior PDFs;

$$p(\lambda_{\kappa^u_2}) = \mathrm{Gam}(\lambda_{\kappa^u_2} \,|\, S_{\kappa^u_2}, B_{\kappa^u_2}) \quad \text{for} \quad p(\lambda_{\kappa^u_3}) = \mathrm{Gam}(\lambda_{\kappa^u_3} \,|\, S_{\kappa^u_3}, B_{\kappa^u_3}).$$

**Some remarks**. As stated above, the log-slowness corrections are independent from the travel time corrections of (2). However, as the slowness relates to the inverse of the derivative of the travel time curve, it is not unreasonable to propagate some of the corrections applied to the travel time model to the slowness model, for example the large-scale correction in slope of the travel time curve provided by $\beta_2 D_{ij}$ in (2).

### 3.2 Posterior Inference

As with the arrival data model, to simplify notation we let

$$\boldsymbol{\alpha}^\theta = (\boldsymbol{\beta}^\theta, \boldsymbol{\tau}^\theta, \boldsymbol{\kappa}^\theta, \boldsymbol{\lambda}^\theta) \quad \text{and} \quad \boldsymbol{\alpha}^u = (\boldsymbol{\beta}^u, \boldsymbol{\tau}^u, \boldsymbol{\kappa}^u, \boldsymbol{\lambda}^u).$$

The prior PDFs for the collection of the correction and the precision parameters is then given by

$$\begin{aligned} p(\boldsymbol{\alpha}^\theta) &= p(\boldsymbol{\beta}^\theta \,|\, \boldsymbol{\tau}^\theta) p(\boldsymbol{\tau}^\theta) p(\boldsymbol{\kappa}^\theta \,|\, \boldsymbol{\lambda}^\theta) p(\boldsymbol{\lambda}^\theta) \quad \text{and} \\ p(\boldsymbol{\alpha}^u) &= p(\boldsymbol{\beta}^u \,|\, \boldsymbol{\tau}^u) p(\boldsymbol{\tau}^u) p(\boldsymbol{\kappa}^u \,|\, \boldsymbol{\lambda}^u) p(\lambda^u). \end{aligned}$$

We can then write the joint distribution of the back azimuth data and the log-slowness data as

$$\begin{aligned} p(\boldsymbol{\theta}, \mathbf{u}, \mathbf{s}, \boldsymbol{\alpha}^\theta, \boldsymbol{\alpha}^u) = p(\boldsymbol{\theta} \,|\, \mathbf{F}^\theta(\mathbf{s}), \boldsymbol{\alpha}^\theta) p(\mathbf{u} \,|\, \mathbf{F}^u(\mathbf{s}), \boldsymbol{\alpha}^u) \\ \times\, p(\boldsymbol{\alpha}^\theta) p(\boldsymbol{\alpha}^u) p(\mathbf{s}). \end{aligned} \tag{10}$$

Note that the back azimuth/slowness data does not provide any (direct) information about the origin times (the $o_i$'s). The posterior PDF with respect to the back azimuth/slowness data is then given by

$$p(\mathbf{s}, \boldsymbol{\alpha}^\theta, \boldsymbol{\alpha}^u \,|\, \boldsymbol{\theta}, \mathbf{u}) = \frac{p(\boldsymbol{\theta}, \mathbf{u}, \mathbf{s}, \boldsymbol{\alpha}^\theta, \boldsymbol{\alpha}^u)}{p(\boldsymbol{\theta}) p(\mathbf{u})} \propto p(\boldsymbol{\theta}, \mathbf{u}, \mathbf{s}, \boldsymbol{\alpha}^\theta, \boldsymbol{\alpha}^u). \tag{11}$$

This posterior PDF can be sampled using MCMC, using a combination of Gibbs sampler and Metropolis-Hasting random-walk samplers, very much along the current approach in BayesLoc for the arrival time data.

It should be obvious by now how the arrival data and the back azimuth/slowness data can be blended together to form a posterior PDF of all unknowns ($\mathbf{s}$, $\mathbf{o}$, $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}^\theta$, $\boldsymbol{\alpha}^u$) conditioning on all the data ($\mathbf{a}$, $\boldsymbol{\theta}$, $\mathbf{u}$);

$$\begin{aligned} & p(\mathbf{s}, \mathbf{o}, \boldsymbol{\alpha}, \boldsymbol{\alpha}^\theta, \boldsymbol{\alpha}^u \,|\, \mathbf{a}, \boldsymbol{\theta}, \mathbf{u}) \\ & \qquad \propto p(\mathbf{a} \,|\, \mathbf{o}, \mathbf{F}(\mathbf{s}), \boldsymbol{\alpha}) p(\boldsymbol{\theta} \,|\, \mathbf{F}^\theta(\mathbf{s}), \boldsymbol{\alpha}^\theta) p(\mathbf{u} \,|\, \mathbf{F}^u(\mathbf{s}), \boldsymbol{\alpha}^u) \\ & \qquad \times\, p(\boldsymbol{\alpha}) p(\boldsymbol{\alpha}^\theta) p(\boldsymbol{\alpha}^u) p(\mathbf{s}) p(\mathbf{o}). \end{aligned} \tag{12}$$

Regarding a MCMC sampler that samples all the unknowns. Such sampler would loop through the following sub-sampling steps:

1. Sampling $\mathbf{s}$: A Metropolis-Hasting sampler. A new proposed value for $\mathbf{s}$ would be evaluated (accepted/rejected) based on feedback from all three data sources.

2. Sampling $\mathbf{o}$: A Gibbs sampler. A new realization would only depend on the arrival time data.

3. Sampling $(\boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\kappa}, \boldsymbol{\lambda})$: A mixture of Metropolis-Hasting random-walk and Gibbs samplers: A new proposed value would only depend on the arrival time data.

4. Sampling $(\boldsymbol{\beta}^\theta, \boldsymbol{\tau}^\theta, \boldsymbol{\kappa}^\theta, \boldsymbol{\lambda}^\theta)$: A mixture of Metropolis-Hasting random-walk and Gibbs samplers: A new proposed value would only depend on the back azimuth data.

5. Sampling $(\boldsymbol{\beta}^u, \boldsymbol{\tau}^u, \boldsymbol{\kappa}^u, \boldsymbol{\lambda}^u)$: A mixture of Metropolis-Hasting random-walk and Gibbs samplers: A new proposed value would only depend on the slowness data.

# 4 Conclusion

We have outlined how the Bayesian framework can be used to blend together three sources of seismic data (arrival times, back azimuth, and slowness) to yield a more accurate event locations. The approach taken here to develop the statistical model for the back azimuth and the slowness data very much mirrors the approach taken in the current version of BayesLoc, which only uses arrival time data, by introducing statistical "bias" corrections and "precision" factors for each data source.

There is a work in progress to extend the arrival time only BayesLoc program to take advantage of possible back azimuth and slowness data following the template outlined in this document.

## References


A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian Data Analysis*. Hapman and Hall/CRC, 3nd edition, 2013.

G. Johannesson, S.C. Myers, and W. Hanley. Bayesloc: Bayesian multiple seismic-event locator for event clusters. Technical Report LLNL-SM-427937, Lawrence Livermore National Laboratory, 2010. 85 pages.

J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*. Springer, 2004.

S.C. Myers, G. Johannesson, and W. Hanley. A Bayesian hierarchical method for multiple-event seismic location. *Geophysical Journal International*, 171:1049–1063, 2007.

S.C. Myers, G. Johannesson, and W. Hanley. Incorporating of probabilistic seismic phase labels into a Bayesian multiple-event seismic locator. *Geophysical Journal International*, 177:193–204, 2009.

C. P. Robert and G. Casella. *Monte Carlo Statistical Methods (Second Edition)*. Springer, 2004.

S. Rost and C. Thomas. Array seismology: methods and applications. *Reviews of Geophysics*, 40(3):2–1 – 2–27, 2002.

A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2004.