LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-452494

Global-scale P-wave tomography optimized for prediction of teleseismic and regional travel times for Middle East events: 1. Data set development

S. C. Myers, G. Johannesson, N. A. Simmons

September 1, 2010

Journal of Geophysical Research

Disclaimer

Global-scale P-wave tomography optimized for prediction of teleseismic and regional travel times for Middle East events: 1. Data set development

S.C. Myers[1], G. Johannesson[2] & N.A. Simmons[1]

[1]Atmospheric, Earth and Energy Division, Lawrence Livermore National Laboratory
[2]Systems and Decision Sciences, Lawrence Livermore National Laboratory

Corresponding author:

Stephen Myers
L-046
Lawrence Livermore National Laboratory
Livermore, CA 94550
(925) 423-4988
smyers@llnl.gov

For submission to *Journal of Geophysical Research – Solid Earth*

Prepared by LLNL under Contract DE-AC52-07NA27344

Abstract

We extend the *Bayesloc* seismic multiple-event location algorithm for application to global arrival time data sets. *Bayesloc* is a formulation of the joint probability distribution across multiple-event location parameters, including hypocenters, travel time corrections, pick precision, and phase labels. Stochastic priors may be used to constrain any of the *Bayesloc* parameters. Markov Chain Monte Carlo (MCMC) sampling is used to draw samples from the joint probability distribution, and the posteriori samples are summarized to infer conventional location parameters such as the hypocenter. The first application of the broad-area *Bayesloc* algorithm is to a data set consisting of all well-recorded events in the Middle East and the most well-recorded events with 5° spatial sampling globally. This sampling strategy is designed to provide the ray coverage needed to determine lithospheric-scale P-wave velocity structure in the Middle East using the complementary ray geometry provided by regional (sub-horizontal) and teleseismic (sub-vertical) ray paths, and to determine a consistent – albeit lower resolution – image of global mantle structure. The data set consists of 5401 events and 878,535 *P*, P_n, *pP*, *sP*, and *PcP* arrivals recorded at 4606 stations. Relocated epicenters are an average of 16 km from bulletin locations. Epicenters are found to be within 5.6 km on average from epicenters that are known to within 1 km (e.g. nuclear explosions). For arrivals labeled *P*, P_n, and *PcP*, ~92%, ~90%, and 96% are properly labeled with probability > 0.9, respectively. Incorrect phase labels are found to be erroneous at rates of 0.6%, 0.2%, 1.6%, and 2.5% for *P*, P_n, *PcP*, and depth phases (*pP* and *sP*), respectively. Labels found to be incorrect, but not erroneous, were reassigned to another phase label. *P* and P_n residual standard deviation with respect to *ak135* travel times are dramatically reduced from 3.45 seconds to 1.01 seconds. The differences between travel time residuals for nearly reciprocal ray paths are significantly reduced from the

input event locations, suggesting that *Bayesloc* relocation improves data set consistency. This results also suggest that the dominant contribution to travel time residuals calculated from information provided in global bulletins is location and picks errors, not travel time prediction errors due to 3-D structure. Modeling the whole multiple-event system results in accurate locations and an internally consistent data set that is ideal for tomography and other travel time calibration studies. *Simmons et al.* (companion paper) use the *Bayesloc*-processed data set to develop a 3-D tomographic image, which further reduces residual standard deviation to 0.50 seconds.

1. Introduction

Production of accurate global and regional seismic bulletins for use in Earth model development and empirical travel time calibration remains a painstaking and costly endeavor. The International Seismic Centre (ISC) produces the most comprehensive bulletin of seismic event information, containing reconciled arrival time measurements (picks) and event hypocenters based on data contributions from the National Earthquake Information Center (NEIC) and numerous regional- and local-network operators. As such, the ISC bulletin is the summary of an impressive international effort to analyze a massive stream of global seismic data. Despite the wealth of information contained within the ISC bulletin, direct use of the ISC bulletin (or constituent bulletins) in travel time tomography has been discounted. A number of investigations find that tomographic signal – the component of a travel time residual attributed to deviation from a reference Earth model – is obscured by errant picks and substantial hypocenter errors [e.g. *Grand* 1990; *Gundmundsson et al*. 1990; *Creager and Boyed* 1992; *Roehm et al.* 1999; *Husen et al*. 2009].

Recognizing both the wealth of information in the ISC bulletin, as well as data quality issues, *Engdahl et al.* [1998] developed and applied methods to identify and remove pick outliers and to correct phase-naming mistakes. *Engdahl et al.* [1998] also made use of core phases (e.g. *PKP*) and secondary phases (*pP*, *sP*, *pwP*) to relocate and improve hypocenter accuracy. The resulting bulletin (EHB hereafter) remains the most utilized data set for the development of global P-wave tomography models [e.g *van der Hilst et al*. 1997; *Kennett et al.* 1998; *Bijwaard et al.* 1998; *Boschi and Dziewonski* 1999; *Kárason and van der Hilst* 2000; *Antolik et al*. 2003; *Montelli et al*. 2004]. *Li et al.* [2008] make use of teleseismic and regional EHB data to develop, arguably, the most comprehensive image of mantle P-wave structure to date. The EHB bulletin

85 continues to evolve as detailed analysis of individual events and multiple-event analysis of event
86 clusters (e.g. aftershock sequences) are completed [e.g. *Bondár et al*. 2008, 2009]. However,
87 inaccuracies and inconsistencies are inevitable for bulletins that are based on single-event
88 locations; because additions and improvements are generally made one event at a time.
89 Other approaches use waveform cross correlation to make exceedingly precise arrival-
90 time measurements. Recorded signals can be cross correlated with synthetic seismograms [e.g.
91 *Grand* 1994, 2002; *Simmons et al.* 2006, 2007], or many waveforms for a single event can be
92 cross correlated to obtain relative arrival times [e.g *Masters et al*. 2000; *Bolton and Masters*
93 2001; *Ritsema and van Heijst* 2002; *Houser et al.* 2008]. Despite producing exceedingly precise
94 measurements, application of correlation methods is currently limited to long-period, teleseismic
95 signals. Unlike teleseismic waveforms, widespread waveform similarity for recorded regional
96 waveforms is uncommon, which severely limits the applicability of empirical waveform
97 correlation. Matching synthetic and empirical regional waveforms is similarly challenging;
98 because, region-specific models – or fully 3-dimensional models – are needed in order to
99 adequately match observed regional waveforms. As a result, waveform methods require further
100 development before a self-consistent regional and teleseismic data can be produced.
101 In this study we continue the effort to improve the accuracy and consistency of bulletin
102 data, by adapting the *Bayesloc* method [*Myers et al*. 2007, 2009] of multiple-event location for
103 application to global seismicity. *Bayesloc* is a formulation of the joint probability function that
104 spans hypocenters, travel-time corrections, pick precision, and phase labels. Previous versions of
105 *Bayesloc* were tailored for application to event clusters (e.g. aftershock sequences), with travel-
106 time correction and pick precision formulations that are designed for robustness. Application to
107 a global data set requires a more general formulation. Specifically, the formulation must enable

datum-specific assessment of travel time correction and pick precision, while maintaining robustness. The updated *Bayesloc* algorithm enables production of an accurate and consistent global/regional bulletin by simultaneously relocating events, assessing pick precision, a more generalized travel time correction, and probabilistically assessing phase labels. Assessment of phase labels includes the possibility that an arrival-time datum is erroneous. *Bayesloc* allows prior constraints on any aspect of the multiple-event system, enabling directly utilization of previous work that statistically characterizes the accuracy of event hypocenters and picks [e.g. *Bondár et al*. 2004; *Bondár and McLaughlin* 2009]. The use of prior information helps to mitigate regional location bias and improve outlier identification. The first application is to a data set of all well-recorded events in the Middle East and an even sampling of the best-recorded events sampling the globe.

2. Method

2.1 Multi-Event Relocation

Multiple-event methods simultaneously invert arrival times for many events to determine both event locations and a set of travel-time corrections. Travel-time corrections typically take the form of station/phase terms [e.g. *Douglas* 1967], which restricts applicability to instances where travel time prediction errors at a station and for a given phase are approximately equal for all events (e.g. event clusters). Multiple event methods are known for generating precise relative locations, but loss of location accuracy is inherent to the unconstrained station/phase formulation, resulting in consistent location bias [*Douglas* 1967; *Jordan and Sverdrup* 1981; *Pavlis and Booker* 1983]. *Waldhauser and Ellsworth* [2000] mitigate the limitations of the station phase term by solving for event locations based on residual differences (a.k.a. double

131 difference) and allowing the residual difference to vary slowly with event location. *Menke and*
132 *Schaff* [2004] show that double difference methods can resolve absolute location if network
133 coverage is outstanding.

134 The *Bayesloc* method [*Myers et al.* 2007] parameterizes the travel time correction as an
135 adjustment to the travel time curve for each phase, with the addition of station terms with a zero-
136 mean prior. Correcting the travel time curve mitigates gross travel time prediction errors, and
137 zero-mean station-phase terms maintains resolution of absolute location while accounting for
138 path-specific travel time variations. This formulation is robust to poor network configuration,
139 allowing broader application of the multiple-event method.

140 In addition to precise locations, multiple-event relocation can be used to identify and
141 remove outlier data [e.g. *Engdahl and Bergman* 2001]. In single-event algorithms, the event
142 location can move significantly to accommodate an outlier datum; because, the tolerance for
143 overall arrival time misfit is relatively large. In multiple-event locators, the use of a self-
144 consistent set of travel time correction parameters significantly reduces the tolerance for arrival
145 time misfit. Reduced tolerance for data misfit reduces the distance that an event location can
146 move to better fit an outlier datum, which results in a larger residual for outlier data. Further, an
147 outlier residual is more confidently identified in the context of many events recorded at a
148 common station.

149

150 **2.2 *Bayesloc* modified for a global data set**

151 *Bayesloc* is a statistical model of the multiple event system, including event locations,
152 travel time corrections, assessments of arrival-time measurement (pick) precision, and phase
153 labels. The overarching statistical model is

154

155 $$p(o,x,T,W,\sigma,\tau \mid a,w) = p(a \mid o,T,W,\sigma)\,p(T(x) \mid F(x),\tau)\; p(x,o)\,p(\sigma)\,p(\tau)\,p(W \mid w)/\,p(a)\,, \qquad (1)$$

156

157 where *o* represents event origin times, *x* represents event locations, *T* is the collection of travel
158 times from each event to each station for each phase (model-based prediction plus correction), *W*
159 is the collection of all phase labels, τ is a collection of travel time correction parameters, σ is the
160 collection of arrival-time precision parameters, *a* and *w* are the collection of arrival times and
161 input phase labels. Equation 1 decomposes the inversion of arrival time data to solve for the
162 components of the multiple-event system (left-hand side of equation 1), into a collection of
163 "forward" problems and prior constraints (right-hand side of equation 1). Specifically, the first
164 term (right-hand side) computes the probability of observing the collection of arrivals given a set
165 of hypocenters, travel times, phase labels, and pick precisions. The second term computes the
166 probability of all travel times, given a model-based prediction (event location is implicit) and a
167 collection of correction parameters. The third, fourth, fifth, and sixth terms are prior constraints
168 on hypocenters, arrival-time measurement precision parameters, travel time correction
169 parameters, and input phase labels, respectively. The denominator is the probability over all
170 arrival data, which serves as normalization. Analytical expressions for each term in equation 1
171 are provided in *Myers et al.* [2007, 2009].

172 *Bayesloc* uses the Markov-Chain Monte Carlo (MCMC) method to sample the joint
173 probability of the multiple-event system [see *Gelman et al.* 2004]. Sampling the probability
174 function is accomplished by starting with an initial configuration of the system, then randomly
175 proposing a new configuration using specifically designed proposal distributions (See Myers et
176 al., 2007, 2009). The probability of the current and proposed multiple-event configuration is

177 computed using the forward calculations afforded by Equation 1. A proposed configuration is
178 always accepted as the new "state" of the system if the probability is greater than the current
179 state. If the probability of the proposed state is lower than the current state, then the new state is
180 accepted at a rate specified by the ratio of the probability for the proposed state and the current
181 state. Clearly, configurations of event locations and travel time corrections that reduce arrival
182 time residuals are higher probability configurations and are, therefore, preferred by the MCMC
183 algorithm. Likewise, fitting residuals with representative arrival-time measurement precision
184 (1/variance) parameters maximizes probability by including the most data within the high
185 probability portion of the distribution.

186 The process of accepting/rejecting proposed configurations is continued until adequate
187 sampling of the joint probability density is achieved (typically 10,000 to 20,000 iterations).
188 Graphical examination of the MCMC samples can be used to assess the non-parametric
189 probability density or an analytical form (e.g. Gaussian) may be used to summarize the MCMC
190 samples. For example, the mean or mode of latitude and longitude samples for given event may
191 be used as an epicenter estimate, and an epicenter uncertainty ellipse can be computed from the
192 latitude and longitude covariance.

193 *Myers et al.* [2007, 2009] implement simple adjustments to the travel time curve and a
194 collection of station terms with a prior mean of zero to robustly correct for gross prediction
195 errors. The travel time curve adjustment formulation is suitable for event clusters, where a
196 station term can adequately capture deviations from the adjusted travel-time curve. However,
197 travel time corrections must be path specific in order to apply *Bayesloc* to a broad-area data set.
198 As such, we recast the *Bayesloc* travel time correction as,

199 $$\delta_{ijW} = T_{ijW} - F_{ijW} = \alpha_W + \alpha_i + \alpha_j + \alpha_{iW} + \alpha_{jW} + \beta_W \left\| x_i - s_j \right\|, \quad (2)$$

where δ is the travel time correction, T is the corrected travel time, F is the model-based travel time, and β is an adjustment to the slope of the travel time curve. The double bars in the β term indicate event-station distance and the variables x and s represent event and station positions, respectively. The α terms are static corrections attributed to a station (j), event (i), phase (W) or combination thereof. Realizations for each category of α term (α_i, α_j, α_{iw}, α_{jw}) are drawn from a zero-mean Gaussian distribution with unknown variance. The variance of each α term is estimated throughout Bayesloc sampling. In effect, station and event terms (α_i, α_j) are small corrections with respect to the corrected travel time curve ($F_w + \alpha_w, + b_w||x_i - s_j||$) and the station-phase and event-phase terms (α_{iw}, α_{jw}) can further refine the travel time correction. Decomposition of terms in this way allows robust determination of station and event corrections, with resolution of event-phase and station-phase corrections if sufficient data are available. In the absence of sufficient data α terms will tend towards zero.

3. Data Set

3.1. Event selection and data sources

For the first application of *Bayesloc* to a broad-area data set we have gathered a list of 5,401 events throughout the Middle East and well-recorded events that provide global epicenter coverage. Arrival time data and event locations are drawn from the EHB bulletin [*Engdahl et al.* 1998] and a subset of the Lawrence Livermore National Laboratory (LLNL) database [see *Ruppert et al*. 2005]. The LLNL database contains a compilation of global and regional bulletins as well as ~20,000 travel time measurements at regional stations (P_n) made by LLNL staff. The data set contains over 4,000 events in the Middle East region. Outside of the Middle East, the best-recorded events with epicenter spacing of ~5° achieve broad global sampling. Sampling of

223 the global data set is accomplished by selecting the event with the most teleseismic P-phase
224 arrivals and removing all other events within ~5° arc distance of the selected event. The process
225 is repeated until the event list is exhausted. In order to preserve the depth sampling afforded by
226 the EHB bulletin, geographic event sampling is conducted in depth bins with lower bounds of 35
227 km, 75 km, 150km, 300 km, 450 km, and 700 km. All together, 878,535 *P*, P_n, *pP*, *sP*, and *PcP*
228 arrivals recorded at 4,606 stations comprise the data set (Figure 1). Table 1 lists the number of
229 arrivals for each phase.
230 The event sampling strategy is designed to provide the ray coverage needed to determine
231 lithospheric-scale structure in the Middle East using the complementary ray geometry provided
232 by regional (sub-horizontal) and teleseismic (sub-vertical) ray paths, and to determine a
233 consistent – albeit lower resolution – image of global mantle structure. *Simmons et al.* (this
234 volume) report on the tomography study using the *Bayesloc* bulletin developed in this study.
235
236 **3.2 *Bayesloc* Relocation**
237 The *Bayesloc* joint posteriori distribution for the Middle East/Global data set is
238 determined using 4 Markov Chains. The results presented here are averages of the last 12,000 of
239 15,000 MCMC samples for each chain. The first 3,000 samples ("burn in") were used to find the
240 neighborhood of the mode of the posteriori distribution and to adapt MCMC sampling. As such,
241 the first 3,000 samples are not necessarily representative of the joint probability density. Chain
242 mixing – using the parameter configuration of another chain as a proposed configuration – was
243 used to test for local minima. In order to better sample event locations, the starting position for
244 each chain was randomly perturbed from the location of the station having the earliest arrival
245 time pick for that event. Starting depths were set to 15 km, except for depths greater than 70 km

246 as listed in the EHB bulletin, in which case the EHB depth was used as the starting depth. In
247 addition to using the EHB depths as the starting positions, we place a prior constraint on event
248 depth of ±5 km (standard deviation) for EHB locations with depths greater than 70 km. Our
249 justification is that EHB depth-phase data were scrutinized for consistency and the effect of slow
250 wave propagation through the water column was accounted for when the surface bounce point
251 was determined to be in the ocean.

252 The *ak135* model [*Kennett et al*. 1995] was used for all base-model travel time
253 predictions. The *Bayesloc* travel time correction model includes a shift and change of slope of
254 each model-based travel time curve. *Myers et al.* [2007] demonstrate that *Bayesloc* robustly
255 determines the slope of the travel time curve, so we use non-informative priors on the slope
256 correction (β in Equation 2). However, resolution of the shift of the travel time curve (α_w in
257 Equation 2) requires prior constraints on either the shift itself or prior constraints on some event
258 origin times. Because the *ak135* model was developed and validated using the limited set of
259 nuclear explosions with known origin times and using locally recorded events, for which the
260 origin time is well constrained [*Kennett et al.* 1995], the absolute travel time for teleseismic
261 phases cannot be improved in this study. Therefore, we place tight constraints on the shift of
262 teleseismic travel time curves (*P*, *pP*, *sP*, *PcP*). Unlike teleseismic phases, *ak135* travel time
263 errors for the regional P_n phase are exceedingly non-stationary, with regional biases of many
264 seconds. We allow *Bayesloc* to model the regional P_n travel time bias in the Middle East by
265 placing a loose prior on the shift of the travel time curve of ±5 seconds (standard deviation).

266 In addition to travel time curve corrections, the *Bayesloc* travel time correction model
267 includes event, station, event-phase, and station-phase parameters. Prior constraints on the
268 standard deviation of these parameters are uninformative, but a prior mean value of zero for the

collection of each parameter category is imposed. Prior constraints on the measurement precision are also uninformative, so data weighting is entirely determined by adapting precision parameters to fit data distributions during the MCMC sampling. Last, MCMC phase label sampling is set so that if an alternative phase labels is tested, than there is an equal chance that the label will be one of the other phases under consideration or the "erroneous" label.

4. Results

4.1. Epicenter shifts

Figure 2 shows epicenter shifts for all of the global events and a representative sampling of events in the Middle East. Epicenters shift by 16 km on average relative to the input bulletin locations, and regional trends in the vector directions are evident. Eight of the events in the global data set are listed in the IASPEI Reference Event List [*Bondár and McLaughlin* 2009] with location accuracy of 1 km or better. The average difference between reference epicenters and *Bayesloc* epicenters is 5.6 km. This result suggests that the *Bayesloc* locations are substantially more accurate than the bulletin location given that the average epicenter shift is far larger than the error with respect to known locations.

The shift in position from the starting (single-event) epicenter to the *Bayesloc* epicenter is not random (Figure 2). The consistency and magnitude of epicenter shifts for events in and near subduction zones are particularly strong. The direction of epicenter shifts in subduction zones tends to be toward the subduction trench, which is consistent with the expected shift due to a slab of fast oceanic lithosphere dipping under the overriding plate [*Creager and Boyed* 1992]. Outside of subduction zones, epicenters tend to shift northward in the Pacific basin, eastward in Africa, and southward in the Middle East. Little shift is observed throughout northern Asia,

because most of the events are nuclear explosions for which the known location is listed in the EHB bulletin.

Location Example

Figure 3 shows the location of the May 28, 1998 Pakistan nuclear explosion and *Bayesloc* location predictions. The event was well recorded, but station sampling is not geographically uniform. Residual travel times at European stations are early (negative residual) with respect to the known location [*Albright et al.* 1998] and the *ak135* [*Kennett et al.* 1995] travel times. The predominance of European stations with negative residuals results in a northward epicenter bias when the *ak135* model is used for travel time prediction (mislocation of 10.1 km). Because the prediction errors are not random, the resulting epicenter error ellipse does not cover the true location. *Bayesloc* travel time corrections mitigate travel time prediction bias, resulting in an epicenter error of 4.5 km. Modeling all components of the location system, including pick and model error, results in a reduction of the epicenter error ellipse area from 207 km^2 to 70 km^2. More importantly, the *Bayesloc* error ellipse covers the known location because the marginal probability of the event location integrates over the joint probability of all other multiple-event parameters.

4.2. *Bayesloc* Posteriori Travel times

Travel Time Curves

The posteriori shift of the P_n travel time curve is 0.42 seconds, and the slope of the P_n travel time curve changes significantly. The posteriori slope of the P_n travel time curve equates to a phase velocity of 8.16 km/s compared to the starting value of 8.05 km/s for *ak135*. The

315 delay in P_n travel time is consistent with the thick crust of the Iranian Plateau and the decrease in
316 slope of the P_n curve suggests faster upper mantle velocity than the global average. Despite loose
317 priors on the slope of the *P*, *PcP*, *pP*, and *sP* travel time curves, the posteriori curves where not
318 significantly altered in the *Bayesloc* inversion.
319
320 *Data Precision*
321 Application of *Bayesloc* to global data sets with more events and arrivals further
322 improves the robustness of data precision modeling. The precision for each datum is the product
323 of an event, a station, and a phase term. Posteriori precision is dominated by pick (measurement)
324 error, but also includes other errors that may not be accounted for in the travel time correction.
325 Priors on the pick error were non-informative, resulting in a purely data-driven assessment of
326 precision. Figure 4 summarizes posteriori precision for the 3 components of the *Bayesloc* error
327 model. Low precision indicates that no configuration of the multiple-event system could be
328 found to fit the subset of data for the tested station, phase, or event.
329 The largest variability in precision is found for stations: arrival time data are very
330 consistent at some stations and inconsistent at others. The variability in precision equates to
331 station standard deviation ranging from 0.36 seconds to 23.62 seconds. The large variation in
332 station precision is likely a combination of inconsistent arrival time measurement practices,
333 station timing errors, and/or errors in the station location. Likewise, arrival times are more
334 consistently fit for some events than for others, with the variability in event standard deviation
335 ranging from 0.59 seconds to 6.26 seconds. A possible reason for variations in data fit between
336 events is that some events (e.g. explosions) generate more impulsive signals, as noted by *Bondár*
337 *et al.* [2004].

Figure 5 is a Gaussian representation of the posteriori data precision for each phase. Precision is determined relative to the corrected travel times, but we do include the 0.42 s shift in the P_n travel time prediction to show the significance of the shift with respect to the overall P_n distribution. *P* is found to be most precisely picked, followed by P_n, *pP*, *PcP,* and *sP*. This result is consistent with common knowledge that first arrivals (*P* and P_n), which do not arrive in the code of earlier phases, are most precisely picked, and the later phases, which generally arrive soon after the *P*-phase, are less precisely picked. A summary of the posteriori pick uncertainty is listed in Table 1.

Phase Labels

MCMC sampling includes testing alternate phase labels for each arrival datum. The phase labels that increase overall probability are more likely to be accepted, and posteriori probability is measured by the frequency of each label in the MCMC sample. Posteriori summary statistics for each phase are listed in Table 1. The 4th column of Table 1 lists the percentage of instances where the input and posteriori phase label agree and the posteriori probability of the phase label is greater than 0.9. The results suggest that input phase labels are correct, with high confidence, in approximately 90% of the instances for this data set. The 5th column lists the percentage of instances where the input and most likely posteriori phase label agree. The 6th column lists that percentage of instances where the posteriori phase label was deemed "erroneous" (i.e. the provided arrival time did not match the timing for any of the phases considered in this study).

The results also suggest that the first-arrival, *P* and P_n phases, are not likely to be erroneous, but the rate of erroneous data entries for later-arriving phase – *pP*, *sP*, and *PcP* – is

361 approximately 1% to 2%. The difference between columns 5 and 6 is the rate of phase
362 misidentification, i.e. valid arrivals with the wrong phase assignment. For example,
363 approximately 3% of reported *P* phases are mislabeled. Detailed examination of the *Bayesloc*
364 output finds that if a *P*-phase is relabeled it is most likely to be relabeled as a depth phase. This
365 can occur when the *P*-phase is on a node of the focal sphere and a depth phase is not nodal. The
366 depth phase *pP* is also commonly relabeled as either another depth phase or as *P*. Figure 6
367 shows an example of phase relabeling with waveforms added to substantiate the *Bayesloc* result.
368 Clearly the removal of one *P*-phase is correct, and relabeling the *pP* arrival to *sP* is sensible
369 given the apparent arrival of the true *pP* phase that precedes the relabeled phase.
370
371 *Travel Time Residuals*
372 *Bayesloc* reduces P and Pn residual standard deviation by more than a factor of 3, from
373 approximately 3.45 s to 1.01 s. The *Bayesloc* data set excludes ~4% of input data that are not
374 labeled as *P* or P_n with posteriori probability exceeding 0.9, but only a small portion of the
375 observed reduction in standard deviation can be attributed directly to data removal. Simply
376 removing 4% of the data comprising the tails of a Gaussian distribution would reduce the
377 standard deviation from 3.45 to approximately 3.00 seconds. In fact, the observed reduction in
378 residual standard deviation would require removing approximately 60% of the data comprising
379 the tails a Gaussian distribution. Reduction of residual standard deviation by a factor 3 through
380 *Bayesloc* relocation suggests that the dominant contribution to global bulletin residuals is
381 location and picks errors, rather than the effects of 3-D velocity heterogeneity. This finding is
382 consistent with the findings of studies referenced in the introduction.

383 Plotting *ak135* travel time residual density as a function of event-station distance (Figure
384 7) clarifies residual trends. At teleseismic distances, the residual distribution shows a slight
385 negative trend. Patterns within the body of the residual distribution at between 20° and 50° also
386 emerge in the *Bayesloc* residuals. P_n residuals exhibit a distinct negative trend, consistent with
387 the *Bayesloc* correction to the P_n travel time curve for the Middle East. The P_n distribution after
388 trend removal is slightly larger than the distribution for teleseismic *P*, because of increased pick
389 error and significant velocity heterogeneity in the Middle East region. *Simmons et al.*
390 (companion paper) use the *Bayesloc* output as input to 3-D tomography. With respect to the 3-D
391 velocity model, *Bayesloc* residuals become 0 mean with standard deviation of 0.50 seconds.
392 Moreover, the residual trends as seen in Figure 7 are removed when travel times are predicted on
393 the basis of the new global tomography model.
394
395 *Data Consistency Measured by Reciprocal-Path Travel Time Residuals*
396 Data set consistency can be measured by differencing travel times for reciprocal paths.
397 To the extent that reciprocal travel times disagree, residuals cannot be attributed to velocity
398 model errors and the data set is less desirable for tomography. A truly reciprocal path would
399 require two events that occur at the location of two seismic stations. For global data sets that are
400 comprised primarily of earthquakes, reciprocal pairs would be rare indeed, primarily because the
401 events would have to occur at the surface. We can, however, identify station-event pairs that are
402 reciprocal within some spatial tolerance. We do not expect the travel times to be identical, but
403 the residuals with respect to a reference model (*ak135* in this case) should be highly correlated,
404 because of significant overlap in the ray path sensitivity kernels [*Rodi and Myers* 2007].

We identify 150 teleseismic paths with event-station distance between 15° and 92°, and 65 regional paths with event-station distance between 5° and 15°. The maximum distance (tolerance) between events and stations forming the ends of a reciprocal path is set to 1° for teleseismic paths and 0.7° for regional paths. This difference between ray end points constitutes a smaller percentage of the overall path for teleseismic than for regional paths, so we may expect better agreement for teleseismic reciprocal paths. Decreasing the tolerance for regional reciprocal paths results in too few paths for meaningful analysis. Only events with depth less than 20 km are considered.

Figure 8 shows reciprocal paths and the differences between residual travel times for each reciprocal pair. Teleseismic reciprocal paths provide good sampling in the Middle East and a few paths at the Eastern Pacific Rim. Regional-distance reciprocal paths are entirely in the Middle East for this data set. The standard deviation of residual differences for input teleseismic reciprocal paths is 3.1 s, revealing considerable data set inconsistency. The standard deviation of residual differences for *Bayesloc* teleseismic reciprocal paths is reduced to 0.8 s, suggesting vastly improved data set consistency. The number of path pairs is dramatically reduced at regional distances, but the results also show improved data set consistency, with the standard deviation of residual differences reduced from 1.7 seconds (input) to 1.0 seconds for *Bayesloc*. It should be noted that although reciprocity tests reveal substantial improvement in data consistency, the *Bayesloc* procedure did not include reciprocity constraints, and improved agreement between travel-time residuals for reciprocal paths is primarily an indication of improved event locations.

4. Discussion and Conclusions

Global bulletin arrival times provide the seismic community with a wealth of information. However, use of bulletin data for the purposes of determining phase travel times is limited, due to event hypocenter errors and contamination by a small percentage of erroneous data. Hypocenter errors are the result of both erroneous data and inaccurate travel time predictions for radially symmetric Earth models. Unfortunately, making bulletin data useful for tomographic and other travel time calibration studies is not as straightforward as identifying and removing large-residual arrival time data; because, single-event locators move the event location to a position that minimizes all travel time residuals, including an erroneous datum. Therefore, an erroneous datum may not be evident.

Modifications to the *Bayesloc* joint probability formulation of the entire multiple-event system afford simultaneous reanalysis of all arrival time data. Through MCMC hypothesis testing, *Bayesloc* captures the range of data, travel time, and hypocenter configurations that peak overall probability. During MCMC sampling, *Bayesloc* continually explores how the entire multiple event system responds to the hypothesis that any given datum is incorrectly labeled or that the datum is erroneous. Therefore, an early assessment that a datum is erroneous does not preclude reintroduction of the datum under more favorable parameter configurations. This underscores the probabilistic nature of the *Bayesloc* result and that ambiguity in determining a phase label, or any other multiple-event parameter, is propagated to the uncertainty of all other parameters. This gives us confidence in the meaningfulness of high posteriori probability data labels and in the posteriori assessment of measurement precisions, travel time corrections, and event hypocenter probability regions.

Improved data consistency is exemplified by comparing residuals for nearly reciprocal paths (Figure 8). Differencing residuals for reciprocal paths reveals significant travel time inconsistencies for the input (single-event location) data set. Because reciprocal inconsistencies cannot be attributed to the model error, reciprocal travel time inconsistency is a measure of the degree to which a tomographic algorithm cannot fit the data. Recent studies demonstrate that data inconsistencies are not sufficiently mitigated by data averaging [*Husen et al.* 2009; *Diehl et al.* 2009], and data inconsistencies resulting from event mislocation are particularly problematic for tomography [*Creager and Boyed* 1992].

The *Bayesloc* result presented here is used in a companion paper [*Simmons et al*. 201x] to simultaneously image detailed lithospheric structure in the Middle East as well as broad structure throughout the mantle. The tomographic study reinforces the accuracy and consistency of the *Bayesloc* results presented here, as the residual distribution is further reduced by more than a factor of 2 (variance reduction of 75%) to 0.50 seconds. Lastly, lithospheric structure in the Middle East is clearly imaged, and the magnitudes of velocity anomalies are significantly larger than previous studies, despite the low variance of the input data set.

Acknowledgments

Funding for this project is provided by the Ground-Based Nuclear Detonation Detection program (Leslie Casey program manager) within NNSA/NA22. We thank our LLNL colleagues for day-to-day interactions and support and we thank the broader community for motivation. Thanks also to Bill Rodi for frequently reminders about the power of travel time reciprocity, as well as wide-ranging conversations on the location problem.

References

Albright, D., Gay, C. & Pabian, F. (1998), New details emerge on Pakistan's nuclear test site, EOM. http://www.modernag.com/Common/Archives/DecemberJan/gay.htm.

Antolik, M., Gu, Y., Ekstrom, G. & Dziewonski, A. (2003), J362D28: a new joint model of compressional and shear velocity in the Earth's mantle, *Geophys. J. Int.,* **153,** 443–466.

Bijwaard, H., W. Spakman and E. R. Engdahl (1998), Closing the gap between regional and global travel time tomography, *J. Geophys. Res*., *103(B12)*, 30,055-30,078.

Bolton, H. and G. Masters (2001), Travel times of P and S from the global digital networks: Implications for the relative variation of P and S in the mantle, *J. Geophys. Res*., *106*, 13527-13540.

Bondár, I., S.C. Myers, E.R. Engdahl, E.A. Bergman (2004), Epicenter accuracy based on seismic network criteria, *Geophys. Jour. Int*., **156,** 483–496.

Bondár, I., E. Bergman, E.R. Engdahl, B. Kohl, Y-L. Kung, and K. McLaughlin (2008). A hybrid multiple event location technique to obtain ground truth event locations, *Geophys. J. Int.*, **175**, 185-201, doi: 10.1111/j.1365-246X.2008.03867.x.

Bondár, I., K.L. McLaughlin (2009), A New Ground Truth Data Set For Seismic Studies, Seismological Research Letters Volume 80, Number 3, May/Jun.

Boschi, L. and A. M. Dziewonski (1999), High- and low-resolution images of the Earth's mantle: implications of different approaches to tomographic modeling, *J. Geophys. Res*., *104*(B11), 25567-25594.

Creager, K.C., and T.M. Boyed (1992). Effects of earthquake mislocation on estimates of velocity structure, *Phys. Earth and Planet. Int*., **75**, 63-76.

493 Diehl, T., E. Kissling, S. Husen, and F. Aldersons (2009), Consistent phase picking for regional
494 tomography models:Application to the greaterAlpine region: *Geophys. J. Int*., **176**, 542–
495 554.
496 Douglas, A. (1967), Joint epicentre determination. *Nature* **215**, 47–48.
497 Engdahl, E.R., R. van der Hilst, and R. Buland (1998), Global Teleseismic Earthquake
498 Relocation with Improved Travel Times and Procedures for Depth Determination, *Bull.*
499 *Seism.. Soc. Am.,* **88**, 722-743.
500 Engdahl, E.R., and E.A. Bergman (2001), Validation and generation of reference events by
501 cluster analysis, *Proceedings*, 23rd Annual DoD/DOE Seismic Research Review, Jackson
502 Hole, WY.
503 Flanagan, M.P., S.C. Myers, K.D. Koper (2007), Regional travel-time uncertainty and seismic
504 location improvement using a 3-Dimensional *a priori* velocity model, *Bull. Seismol. Soc.*
505 *Am*., **97**,804-825.
506 Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. (2004), Bayesian Data Analysis (2nd ed.).
507 Boca Raton, Florida: Hapman and Hall/CRC.
508 Grand, S. P. (1990), A possible station bias in travel time measurements reported to ISC:
509 Geophysical Research Letters, 17, 17–20.
510 Grand, S.P. (1994), Mantle shear structure beneath the Americas and surrounding oceans, *J.*
511 *Geophys. Res.*, *99*, 11591-11621.
512 Grand, S. P. (2002), Mantle shear-wave tomography and the fate of subducted slabs, *Phil. Trans.*
513 *R. Soc. Lond. A*, *360*(1800), 2475-2491.

514 Gudmundsson, O., J.H.Davies, andR.W.Clayton (1990), Stochastic analysis of global traveltime
515 data—Mantle heterogeneity and random errors in the ISC data: Geophysical Journal
516 International, 102, 25–43.
517 Houser, C., G. Masters, P. Shearer and G. Laske (2008), Shear and compressional velocity
518 models of the mantle from cluster analysis of long-period waveformes, *Geophys. J. Int*.,
519 *174*, 195-212.
520 Husen, S., T. Diehl, and E. Kissling (2009), The effects of data quality in local earthquake
521 tomography: Application to the Alpine region, *Geophysics*, **74**, no. 6, WCB71-WCB79.
522 Jordan, T. H., and K. A. Sverdrup (1981), Teleseismic location techniques and their application
523 to earthquake clusters in the south-central Pacific, *Bull. Seism. Soc. Am.* **71,** 1105–1130.
524 Kárason, H. & R. D. van der Hilst (2000), Constraints on mantle convection from seismic
525 tomography, in: M. R. Richards, M. R. Gordon, R. D. van der Hilst (Eds.), The History
526 and Dynamics of Global Plate Motion, AGU, Washington, DC, pp. 277-288.
527 Kennett, B.L.N., E.R. Engdahl and R. Buland (1995), Constraints on seismic velocities in the
528 Earth from traveltimes, *Geophys. J. Int,* **122**, 108-124.
529 Kennett, B.L.N., S. Widiyantoro and R.D. van der Hilst (1998), Joint seismic tomography for
530 bulk sound and shear wave speed in the Earth's mantle, *J. Geophys. Res*., **103**(B6),
531 12469-12493.
532 Li, C., R. D. van der Hilst, E. R. Engdahl and S. Burdick (2008), A new global model for P wave
533 speed variations in Earth's mantle, *Geochem. Geophys. Geosys*., *9*(5), doi:
534 10.1029/2007GC001806.

Masters, G., G. Laske, H. Bolton & A. M. Dziewonski (2000), The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: implications for chemical and thermal structure, *in Earth's Deep Interior: Mineral Physics and Tomography from the Atomic to the Global Scale*, edited by S. -I. Karato et al., pp. 63-87, AGU, Washington, DC.

Menke, W., and D. Schaff (2004), Absolute earthquake locations with differential data, *Bull. Seismol. Soc. Am*., **94**, 2254-2264.

Montelli, R., G. Nolet, F. A. Dahlen, G. Masters, E. R. Engdahl and S.-H. Hung (2004), Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, 303, 338-343.

Myers, S.C., G. Johannesson, and W. Hanley (2007) A Bayesian hierarchical method for multiple-event seismic location, *Geophys. J. Int*., **171**, 1049-1063.

Myers, S.C., G. Johannesson, and W. Hanley (2009), Incorporation of probabilistic seismic phase labels into a Bayesian multiple-event seismic locator, *Geophys. J. Int*., **177**, 193-204.

Myers, S.C., M. L. Begnaud, S. Ballard, M. E. Pasyanos, W. S. Phillips, A. L. Ramirez, M. S. Antolik, K. D. Hutchenson, J. Dwyer, and C. A. Rowe, and G. S. Wagner (2009). A crust and upper mantle model of Eurasia and North Africa for Pn travel time calculation, *Bull. Seismol. Soc. Am*., **100**, 640-656.

Pavlis, G. L., and J. R. Booker (1983), Progressive multiple event location (PMEL), *Bull. Seim. Soc. Am.* **73,** 1753–1777.

Ritsema, J. and H. van Heijst (2002), Constraints on the correlation of P- and S-wave velocity heterogeneity in the mantle form P, PP, PPP, and PKPab traveltimes, *Geophys. J. Int*., 149, 482-489.

558 Rodi, W. L., and S. C. Myers (2007), Modeling travel-time correlations based on seismicity
559 kernels and correlated velocity anomalies, in 29th Monitoring Research Review: Ground-
560 Based Nuclear Explosion Monitoring Technologies, Denver, Colorado, 463–471.
561 Roehm, A. H. E., J. Trampert, H. Paulssen, and R. K. Snieder (1999), Bias in reported seismic
562 arrival times deduced from the ISC Bulletin: Geophysical Journal International, 137,
563 163–174.
564 Ruppert, S., D., Dodge, A. Elliott, M. Ganzberger, T. Hauk and E. Matzel (2005). Enhancing
565 seismic calibration research through software automation and scientific information
566 management, in *Proceedings of the 27th Seismic Research Review: Ground-Based*
567 *Nuclear Explosion Monitoring Technologies,* LA-UR-05-6407, Vol. 2.
568 Simmons, N. A., A. M. Forte and S. P. Grand (2006), Constraining mantle flow with seismic and
569 geodynamic data: a joint approach, *Earth Planet. Sci. Lett. 246*, 109–124.
570 Simmons, N. A., A. M. Forte and S. P. Grand (2007), Thermochemical structure and dynamics
571 of the African superplume, *Geophys. Res. Lett., 34*(2), L02301, doi:
572 10.1029/2006GL028009.
573 Simmons, N. A., S. C. Myers and G. Johannesson (2010a), Global-scale P-wave tomography
574 optimized for prediction of teleseismic and regional travel times for Middle East events:
575 2. Tomographic inversion, *J. Geophys. Res*., (companion paper).
576 van der Hilst, R. D., S. Widiyantoro and E. R. Engdahl (1997), Evidence for deep mantle
577 circulation from global tomography, *Nature*, *386*, 578-584.
578 Waldhauser, F., and W. L. Ellsworth (2000). A double-difference earthquake location algorithm:
579 method and application to the Northern Hayward Fault, CA, *Bull. Seism. Soc. Am.* **90,**
580 1353–1368.

581

582 **Figure Captions**

583

584 Figure 1. Event epicenters and station locations. All well-recorded events in the Middle East are
585 retained and global event sampling is approximately 5°. Global sampling is performed
586 independently in event depth intervals down to 700 km (see text). The resulting data set provides
587 horizontal and vertical ray coverage through the Middle East, which is used by Simmons et al.
588 (this volume) in a tomographic study.

589

590 Figure 2. Epicenter relocation vectors. The tail of each vector is at the starting location based on
591 the EHB and LLNL bulletins. Vector length is scaled by the magnitude of the epicenter shift
592 (see inset scale), and vector orientation is in the direction of epicenter shift.

593

594 Figure 3. Single-event and *Bayesloc* global relocation of the May 28, 1998 Pakistan nuclear test.
595 *Bayesloc* mislocation is 4.5 km with epicenter error ellipse area of 70 km^2. Single event
596 mislocation is 10.1 km with epicenter error ellipse area of 207 km^2. The satellite location is from
597 Albright et al., 1998.

598

599 Figure 4. Posteriori precision (1/variance) for phase, station, and event terms. Data precision is
600 the product of these terms. Phases, stations, and events are ordered from least precise to most
601 precise. See text for discussion.

602

603 Figure 5. Gaussian representation of posteriori measurement error distribution for each seismic
604 phase. The 0.42 second shift of P_n travel prediction is included. Standard deviation of
605 measurement error for each pick is also tabulated in Table 1.
606
607 Figure 6. Example validation of phase relabeling based comparison with waveform data. In this
608 case, posteriori labels have probability greater than 0.9. The data shown here are a small subset
609 of the data for this particular event.
610
611 Figure 7. Input and output (Bayesloc posteriori) residual occurrence as a function of event-
612 station distance. Input and output residuals are computed using the *ak135* model. Color indicates
613 the density of residual occurrence on a normalized log scale. See text for discussion.
614
615 Figure 8. Residual difference for approximately reciprocal paths. Solid lines are forward paths
616 and dashed lines are reciprocal paths. The choice of which path is forward and which is
617 reciprocal is arbitrary. *Bayesloc* reciprocal travel times are significantly more consistent than the
618 input data set. All event depths are 20 km or less to minimize path differences near stations.
619 Teleseismic path lengths are 15° to 95° and regional path lengths are 5° to 15°. Maximum
620 separation between events and stations defining reciprocal-path end points is 1° for teleseismic
621 paths and 0.7° for regional paths. See for discussion.

622

623 Table 1. Number of picks for each event and summary of posteriori assessment of phase labels.

Phase	Number of picks	Estimated standard deviation	Phase label retained with prob.>0.9	Input label is most probable	Most probably erroneous.
P	817,552	0.74 s	92%	96%	0.6%
Pn	42,327	0.90 s	90%	98%	0.2%
pP	10,524	1.60 s	90%	95%	2.1%
sP	4,992	2.22 s	92%	97%	2.6%
PcP	3,140	1.83 s	96%	98%	1.6%

624
625
626 Table 2. Prior constraints (standard deviation). Priors not listed are uninformative (broad).

	Travel Time Curves		Event Depth	Correct Phase Label input
	Shift	Slope		
P	10^{-6} s	5 s/degree		0.9
Pn	5 s	5 s/degree		0.9
pP	10^{-6} s	5 s/degree		0.9
sP	10^{-6} s	5 s/degree		0.9
PcP	10^{-6} s	5 s/degree		0.9
EHB events with depth>70 km			5 km	

627
628

Figure 1.

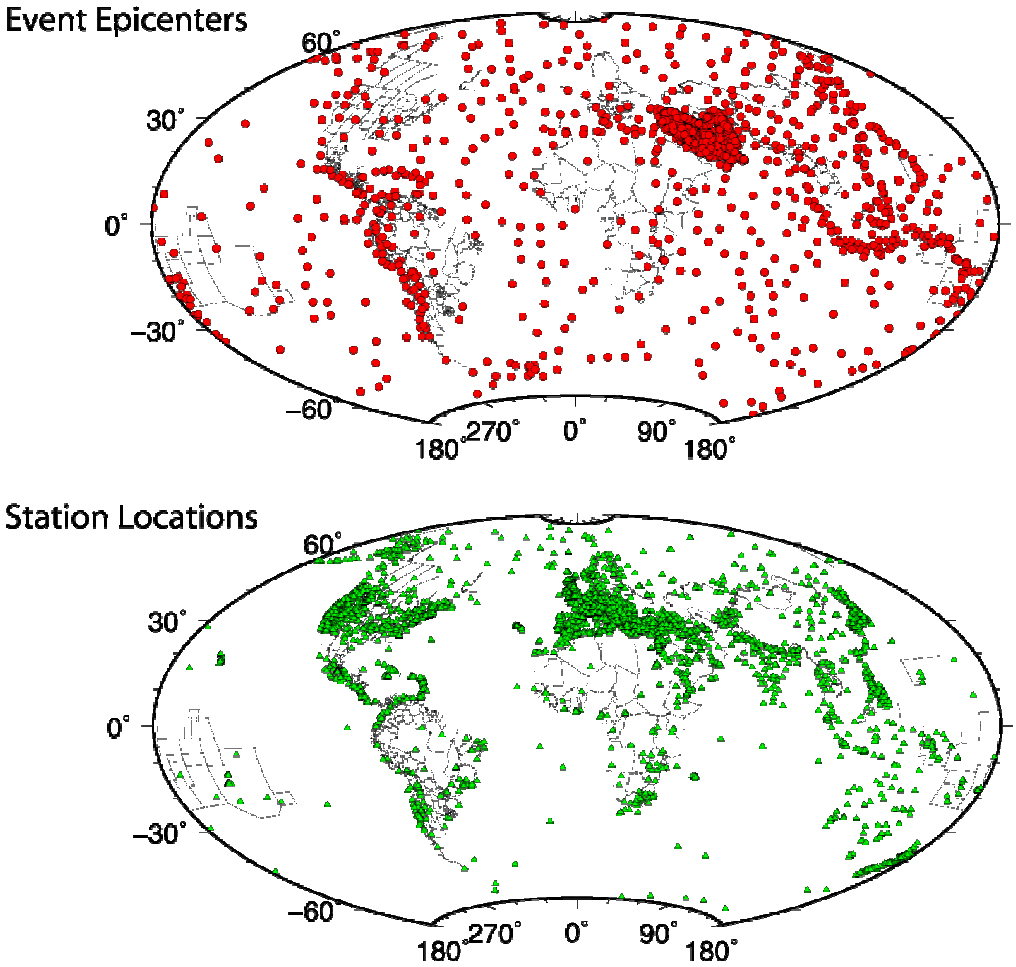

Event Epicenters
60°
30°
0°
−30°
−60°
180°
270°
0°
90°
180°
Station Locations
60°
30°
0°
−30°
−60°
180°
270°
0°
90°
180°

Figure 2.

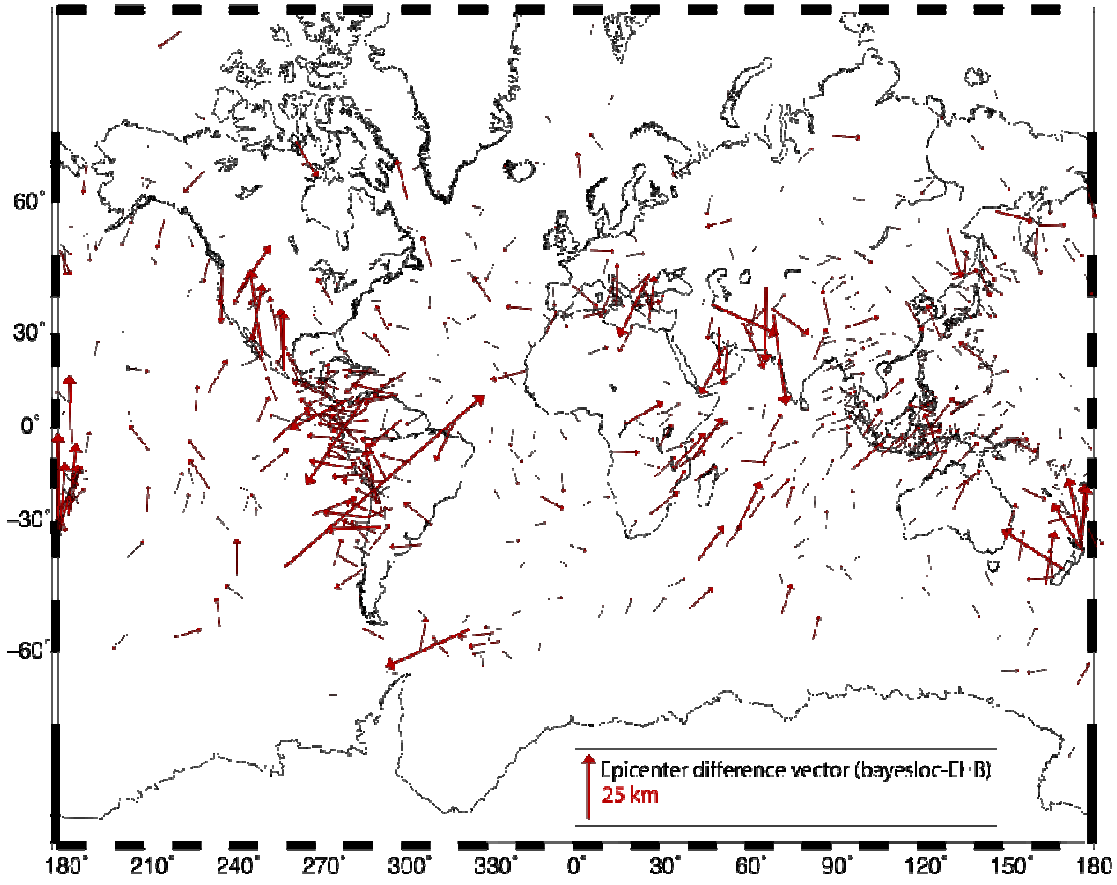

60°
30°
0°
-30°
-60°
180°
210°
240°
270°
300°
330°
0°
30°
60°
90°
120°
150°
180
Epicenter difference vector (bayesloc-EHB)
25 km

Figure 3

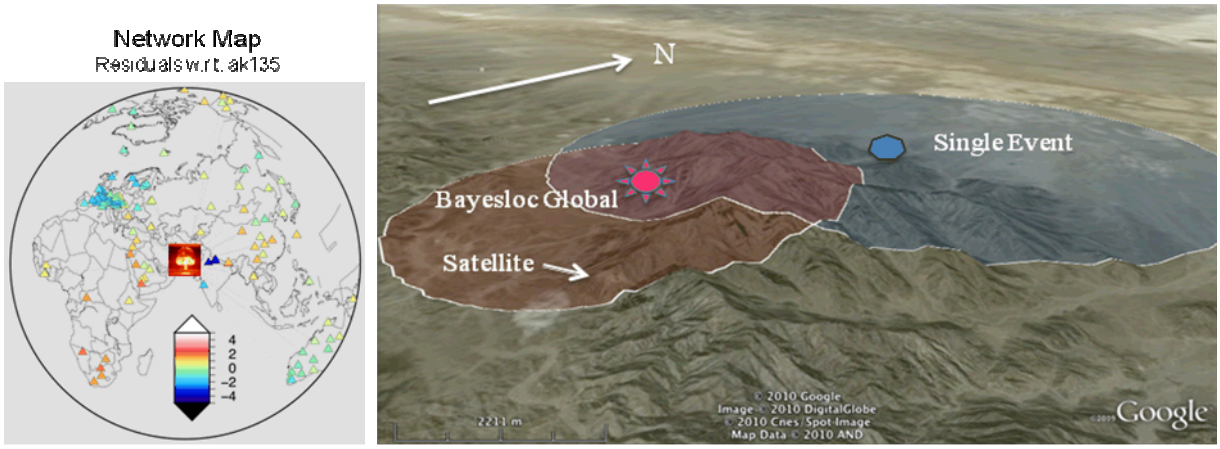

Network Map
Residuals w.r.t. ak135
4
2
0
-2
-4
N
Single Event
Bayesloc Global
Satellite
© 2010 Google
Image © 2010 DigitalGlobe
© 2010 Cnes / Spot Image
Map Data © 2010 AND
2211 m
Google

Figure 4

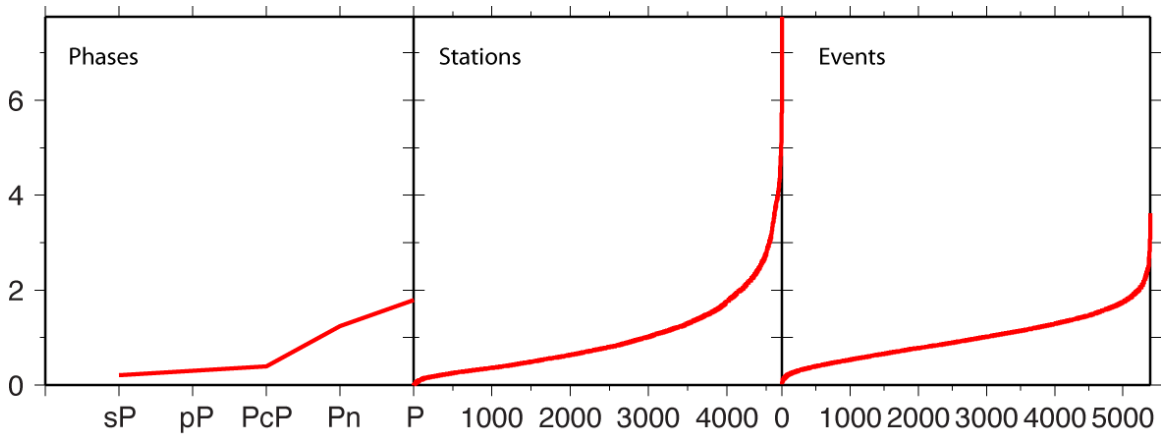

Phases
Stations
Events
6
4
2
0
sP
pP
PcP
Pn
P
1000
2000
3000
4000
0
1000
2000
3000
4000
5000

Figure 5

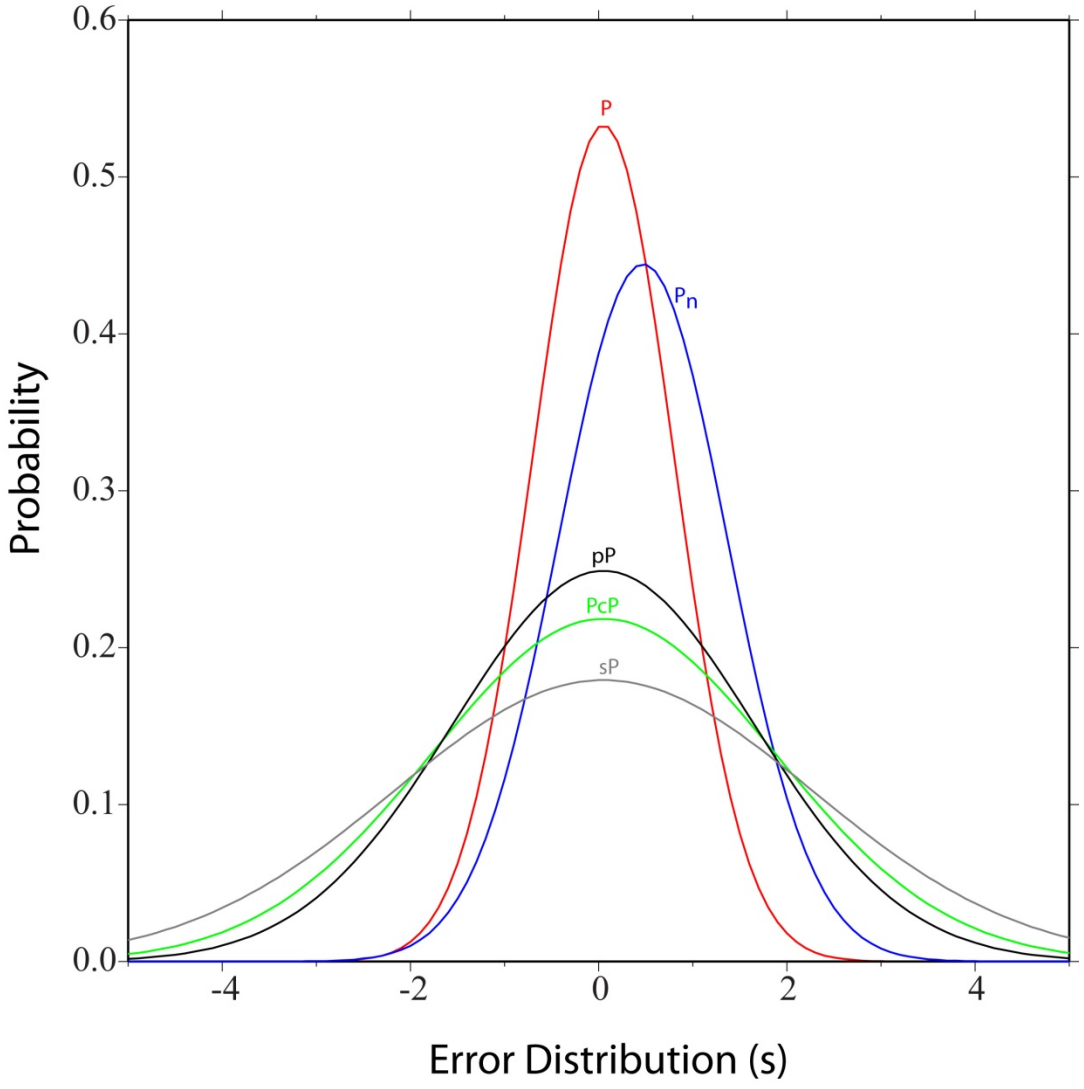
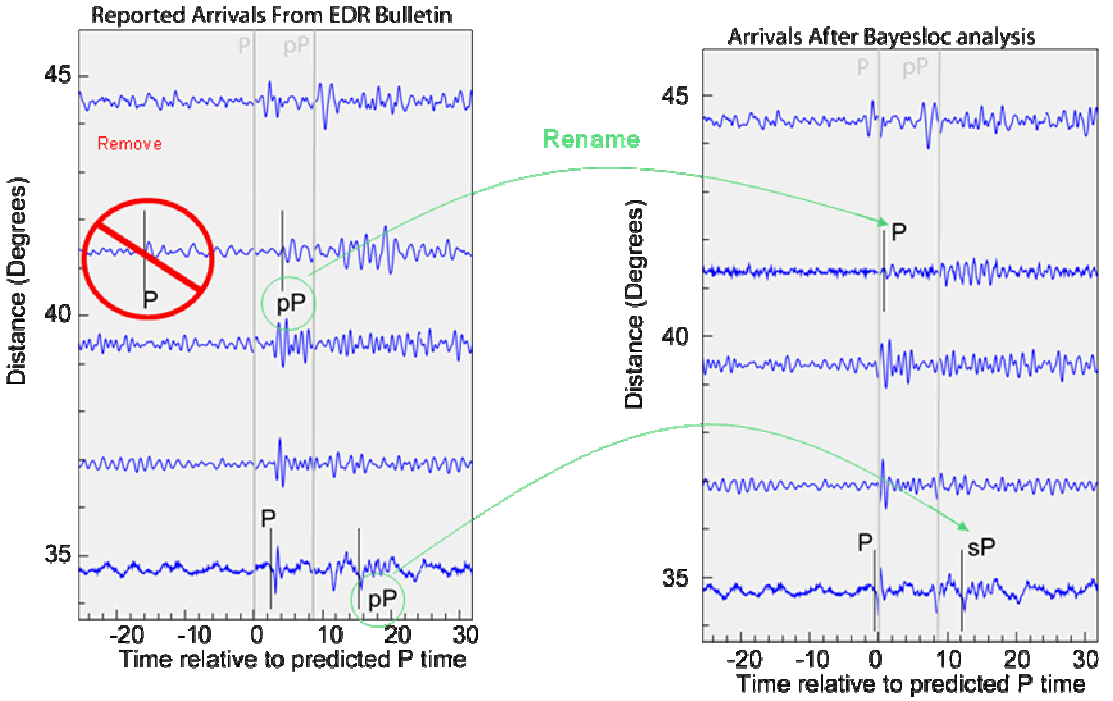
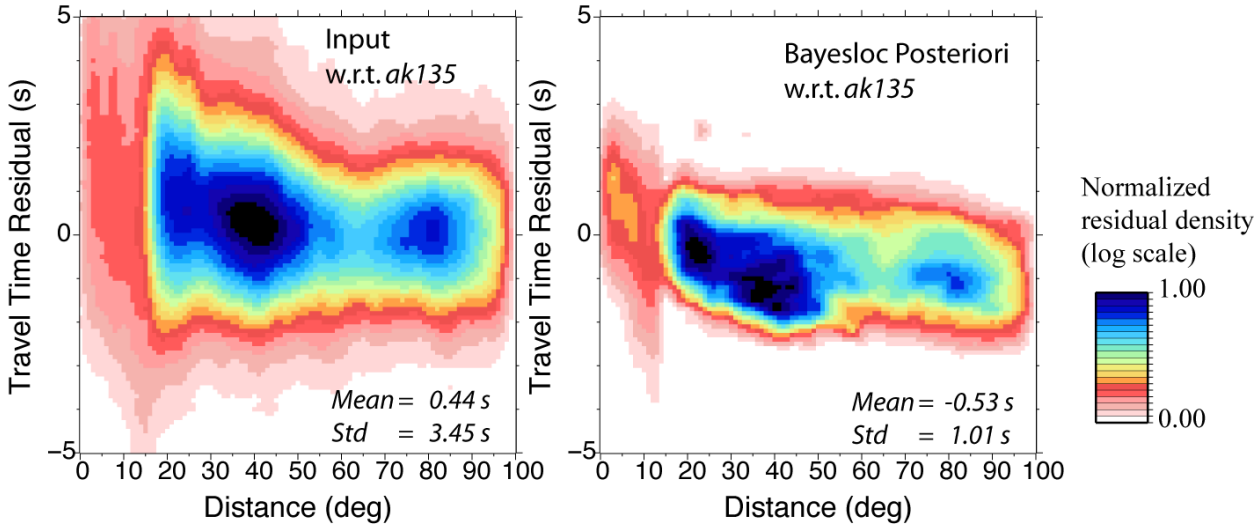

0.6
0.5
0.4
0.3
0.2
0.1
0.0
Probability
P
Pn
pP
PcP
sP
-4
-2
0
2
4
Error Distribution (s)

Figure 6

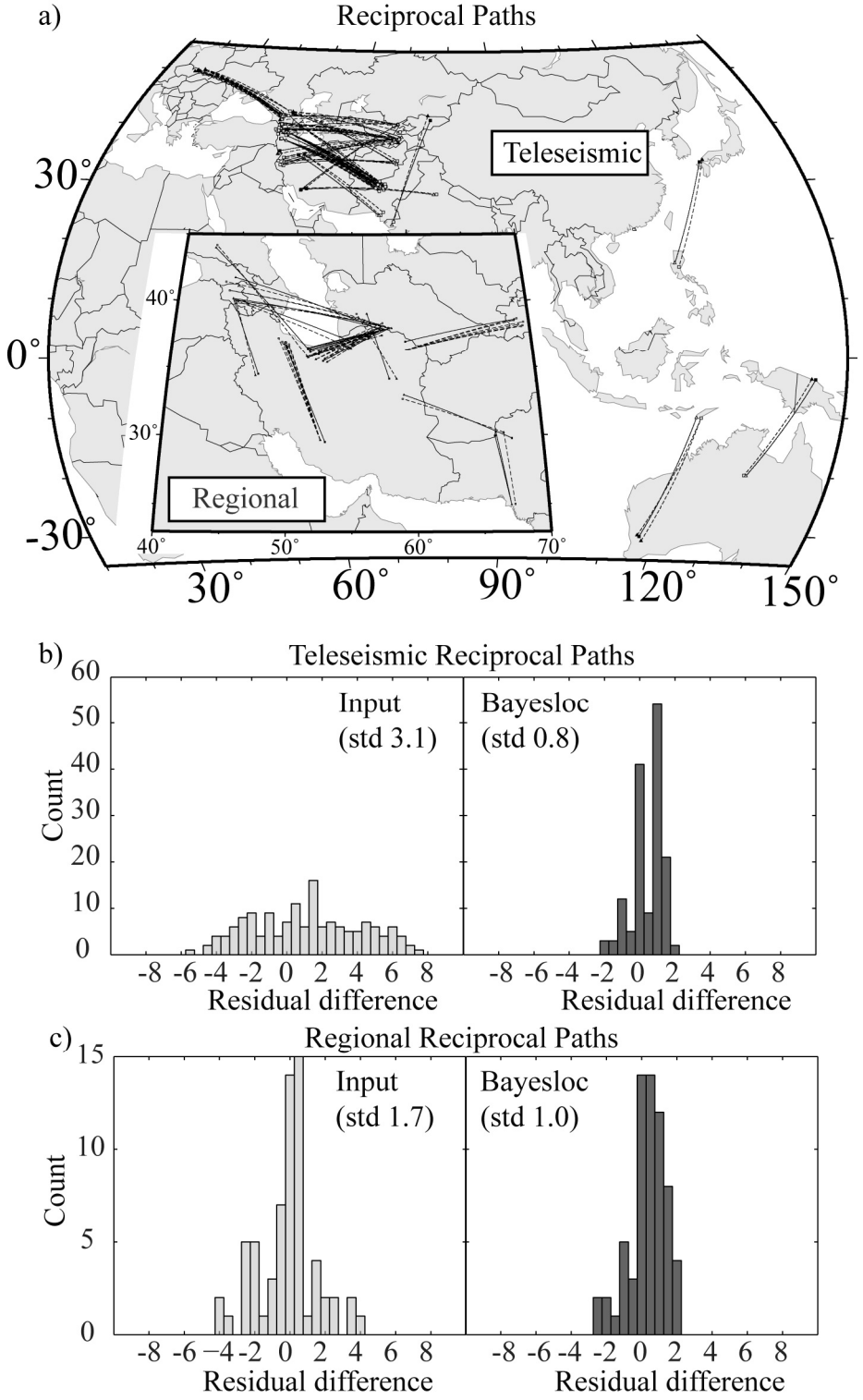


Waveforms without picks are added to show the accuracy of Bayesloc phase determinations.

Figure 7


Input
w.r.t. ak135
Mean = 0.44 s
Std = 3.45 s
Bayesloc Posteriori
w.r.t. ak135
Mean = -0.53 s
Std = 1.01 s
Travel Time Residual (s)
Distance (deg)
5
0
−5
0 10 20 30 40 50 60 70 80 90 100
Normalized residual density (log scale)
1.00
0.00

690
691 Figure 8
692
693
694



695