BAYESLOC MULTIPLE-EVENT LOCATION APPLIED TO A GLOBAL DATA SET

Stephen C. Myers, Gardar Johannesson, and Nathan A. Simmons

Lawrence Livermore National Laboratory

Sponsored by the National Nuclear Security Administration

Award No. DE-AC52-07NA27344/LL10-DM4TT-NDD02

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 that includes 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 is to a data set consisting of all well-recorded events in the Middle East and the most well-recorded events in 5° bins globally. This sampling strategy is designed to provide the ray coverage needed to determine lithospheric-scale 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 arrivals of P,Pn, pP, sP, and PcP recorded at 4606 stations. Relocated epicenters are an average of 16 km from bulletin locations. Although epicenter priors are not used, epicenters are found to be within 5.6 km from those that are known to within 1 km (e.g., peaceful nuclear explosions). Much of the improvement in location accuracy is attributed to dynamic assessment of data precision, which factors into data weights. Location accuracy will improve when location priors are used. For arrivals labeled *P*, *Pn*, 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, Pn, PcP*, and depth phases (*pP* and *sP*), respectively. Labels found to be incorrect, but not erroneous, were reassigned to another phase label. *P* and *Pn* residual standard deviation with respect to *ak135* travel times are dramatically reduced from 3.45 seconds to 1.01 seconds. Simmons et al. (2010, these Proceedings) use the Bayesloc-processed data set in a global tomographic study, which reduces residual standard deviation to 0.53 seconds (consistent with pick error). This result suggests that the dominant contribution to global bulletins residuals is location and picks errors, not travel time prediction errors due to 3D structure. Modeling the whole multiple-event system results in exceedingly accurate locations and an internally consistent data set that is ideal for tomography and other travel time calibration studies.

OBJECTIVES

Production of high-quality data sets for calibration of seismic travel times remains a painstaking and costly endeavor. The value of meticulous data analysis is beyond reproach, but the number and coverage of events with accurate locations and carefully measured seismic-phase arrival times is not sufficient for either development of high-fidelity (3-dimensional) models or comprehensive empirical calibration. Although regional and global bulletins of seismic data provide excellent data coverage, these databases are contaminated by inaccurate and spurious entries.

This project adapts the Bayesloc method (Myers et al., 2007, 2009) of multiple-event location for application to global data sets. Using the updated Bayesloc algorithm, we simultaneously relocate events, assess pick precision, estimate path-specific travel time corrections, and probabilistically assess phase labels to produce an accurate and consistent global bulletin. Bayesloc allows prior constraints on any aspect of the multiple-event system, enabling direct utilization of previous work that statistically characterizes the accuracy of event hypocenters and picks (e.g., Bondár et al., 2004; Bondár and McLaughlin, 2009). Simultaneous analysis if the whole data set allows robust estimation of travel time corrections, probabilistic phase labels (including outlier identification), and assessment of pick precision. Travel-time corrections mitigate location bias and probabilistic assessment of pick error optimizes data weighting. The posteriori bulletin is, on the whole, the most accurate set of global locations available with self-consistent arrival-time picks and phase labels. Therefore, the posteriori bulletin is ideal for tomographic studies.

RESEARCH ACCOMPLISHED

Bayesloc

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

$$p(o,x,T,W,\delta,\phi \mid a,w) = p(a \mid o,T,W,\phi)\, p(T(x) \mid F(x),\delta,W)\, p(W \mid w,a,T(x)) \\ p(x,o)\, p(\phi)\, p(\delta)\, p(W \mid w) / p(a), \quad (1)$$

where *o* represents event origin times, *x* represents event locations, *T* is the collection of travel times from each event to each station for each phase (model prediction plus correction), *W* is the collection of all phase labels, ϕ is the collection of arrival-time precision parameters, δ is a collection of travel-time corrections, *a* and *w* are the collection of arrival times and phase labels. Equation 1 decomposes the inversion of arrival time data to solve for the components of the multiple-event system (left-hand side of equation 1), into a collection of "forward" problems and priors (right-hand side of equation 1). Specifically, the first term (right-hand side) computes the probability of observing the collection of arrivals given a set of hypocenters, travel times, phase labels, and pick precisions. The second term computes the probability of all travel times, given a model-based prediction, a collection of corrections, and the phase labels. The third term computes the probability of the true phase labels given a set of input phase labels, the observed arrivals, and the corrected travel times. The fourth, fifth, sixth, and seventh terms are prior constraints on hypocenters, arrival-time measurement precision, travel-time corrections, and input phase labels, respectively. The denominator is the probability over all arrival data, which serves as normalization. Analytical expressions for each term in equation 1 are provided in Myers et al. (2007, 2009).

Bayesloc uses the MCMC method to sample the joint probability of the multiple-event system (Gelman et al., 2004). Sampling the probability function is accomplished by starting with an initial configuration of the system, then proposing a new configuration that is consistent with prior information. The proposal process is random in the absence of prior information. The probability of each multiple-event configuration is computed using the forward calculations afforded by equation 1. A proposed configuration is always accepted as the new "state" of the system if the probability is greater than the current state. If the probability of the proposed state is lower than the current state, then the new state is accepted at a rate specified by the ratio of the probability for the proposed state and the current state. The process of proposing and accepting/rejecting configurations is continued until adequate sampling of the joint probability density is achieved (typically 10,000 to 20,000 samples). Graphical examination of the MCMC samples can be used to assess the non-parametric probability density, or an analytical form (e.g., Gaussian) may be used to summarize the MCMC samples. For example, the mean or mode of *lat* and *lon* samples for given event may be used to compute an epicenter, and *lat* and *lon* covariance estimates may be used to estimate an uncertainty ellipse.

Modifications to Bayesloc

Myers et al. (2007, 2009) implement simple adjustments to the travel-time curve and a zero-mean collection of station terms to robustly correct for gross prediction errors. That formulation is suitable for event clusters, where a station term can adequately capture deviations from the travel-time curve. In order to apply Bayesloc to a broad-area data set, the travel-time corrections must be path specific. As such, we recast the Bayesloc travel-time correction as,

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

where δ is the travel-time correction, T is the corrected travel time, F is the model-based travel time, and α terms are static corrections. β is an adjustment to the slope of the travel-time curve with x and s representing event and station positions, respectively. The double bars indicate a norm giving event-station distance. Subscripts i, j, and W are indices for event, station, and phase, respectively. We further constrain the collection of station and event terms (α_i, α_j) to be zero mean 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}) are constrained to be zero mean with respect to their station and event terms. Decomposition of terms in this way allows robust determination of station and event corrections, and refinement of event-phase and station-phase corrections is possible if sufficient data are available.

Data Set

For the first application of Bayesloc to a broad-area data set we have gathered an extensive list of events throughout the Middle East that is complemented by approximately 5° sampling (~550 km spacing) of global events. This sampling strategy is designed to provide the ray coverage needed to determine lithospheric-scale 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. Simmons et al. (2010, these Proceedings) report on the tomography study using the data resulting from Bayesloc relocation and processing.

We relocate 5401 events using 878,535 *P*, *Pn*, *pP*, *sP*, and *PcP* arrivals recorded at 4606 stations (Figure 1). Table 1 lists the number of arrivals for each phase. Event locations and arrival times for Middle East events are from the LLNL database (Ruppert et al., 2005), which is a compilation of global and regional bulletins as well as ~20,000 travel-time measurements at regional stations (*Pn*) made by LLNL staff. The global *P*-wave data set is from Engdahl et al. (1998), referred to hereafter as the EHB data set. Event selection from the EHB data set is based on the number of associated *P*-phases. After ordering events by the number of P-phases, the event having the most *P*-phases is selected, and all remaining events within 5° epicenter distance from the selected event are removed from the list. The procedure is repeated until the list is exhausted. In order to preserve the depth sampling afforded by the EHB bulletin, geographic event sampling is conducted in depth bins with lower bounds of 35 km, 75 km, 150km, 300 km, 450 km, and 700 km.

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%

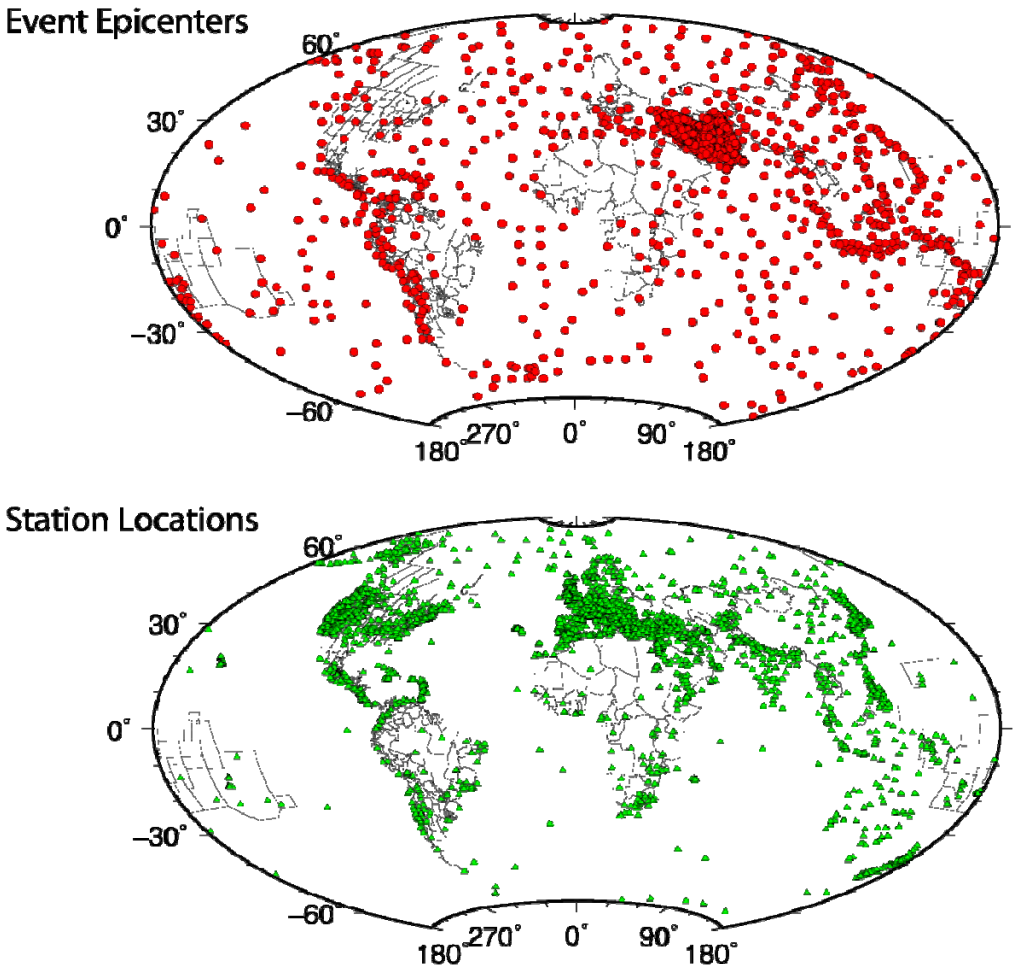

Figure 1. Event epicenters and station locations. All well-recorded events in the Middle East are considered and global event sampling is approximately 5°. Global sampling is performed independently in event depth intervals down to 700 km (see text). The resulting data set provides horizontal and vertical ray coverage through the Middle East, which is used by Simmons et al. (2010, these Proceedings) in a tomographic study.

Bayesloc Relocation

The joint posteriori distribution is determined using 4 Markov Chains. For this data set the starting position is the station location with the earliest pick. Starting depths were set to 15 km, except for events with EHB locations greater than 100 km, in which case the EHB depth was used as the starting depth. EHB depths were determined by scrutinizing depth phases, including depth phases that pass through the oceanic water column, suggesting that event depths in subduction zones are well constrained. Therefore, in addition to starting the MCMC chain at the input depth, we place 5 km standard deviation prior constraints on EHB depths greater than 70km.

Prior constraints on travel-time corrections are based on prediction errors assessed by Engdahl et al. (1998) and our own experience (Flanagan et al., 2007; Myers et al., 2010). We place tight prior constraints on the shift to the travel-time curve for teleseismic phases (*P*, *pP*, *sP*, *PcP*), because the absolute travel time for these phases is established using nuclear explosions with known origin times, as well as events recorded by local networks with well-constrained origin times. A prior with standard deviation of ±5 seconds is imposed on the regional *Pn*-phase. These constraints force the average travel time to be consistent with the model used for teleseismic travel-time prediction (*ak135*; Kennett et al, 1995). Myers et al. (2007) demonstrate that adjustments to the slope of the travel-time curves are robustly determined with large data sets and are, therefore, not needed in this case. As such, we allow phase velocity for all phases to vary by a factor of 3. Priors on the standard deviation of event, station, event-phase, and station-phase corrections are uninformative, but recall that the collection of each parameter must be

zero mean. Priors 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.

Relocation Results

The results presented here are averages of the last 12,000 of 15,000 MCMC samples. The first 3,000 samples ("burn in") are used to find the neighborhood of the mode of the posteriori distribution and to adapt MCMC sampling. As such the, first 3,000 samples are not necessarily representative of the joint probability density. The samples are non-parametric in nature, but we summarize many of the samples by assuming a Gaussian distribution. The Gaussian assumption allows hypocenters to be represented using conventional parameters, including epicenter ellipses.

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, and the average difference between reference epicenters and Bayesloc epicenters is 5.6 km. If this level of accuracy holds for the events with less well-known locations, this suggests that the Bayesloc locations are substantially more accurate than the bulletin location given that the average epicenter shift is ~16 km. However, more testing is needed to further quantify Bayesloc epicenter accuracy.

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 even. Residual travel times at European stations with respect to the known location, which is based on satellite imagery (Albright et al., 1998), are negative (fast). The predominance of European stations with biased negative residuals results in a northward bias in the location when the ak135 model is used (mislocation of 10.1 km). Because the prediction errors are biased (not completely random), the resulting epicenter error ellipse does not cover the true location (Figure 3). 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. Perhaps 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.

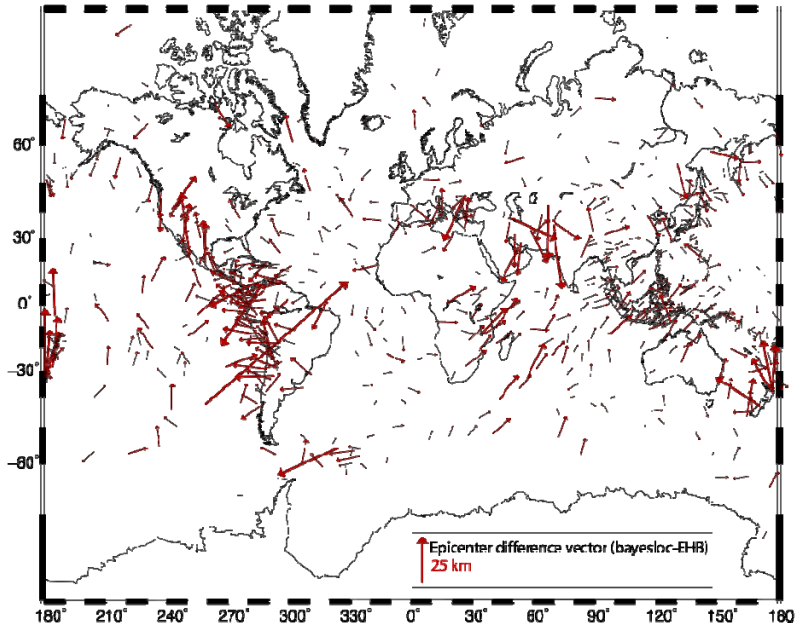

Figure 2. Epicenter relocation vectors. The tail of each vector is at the starting location based on the EHB and LLNL bulletins. Vector length is scaled by the magnitude of the epicenter shift (see inset scale), and vector orientation is in the direction of epicenter shift.

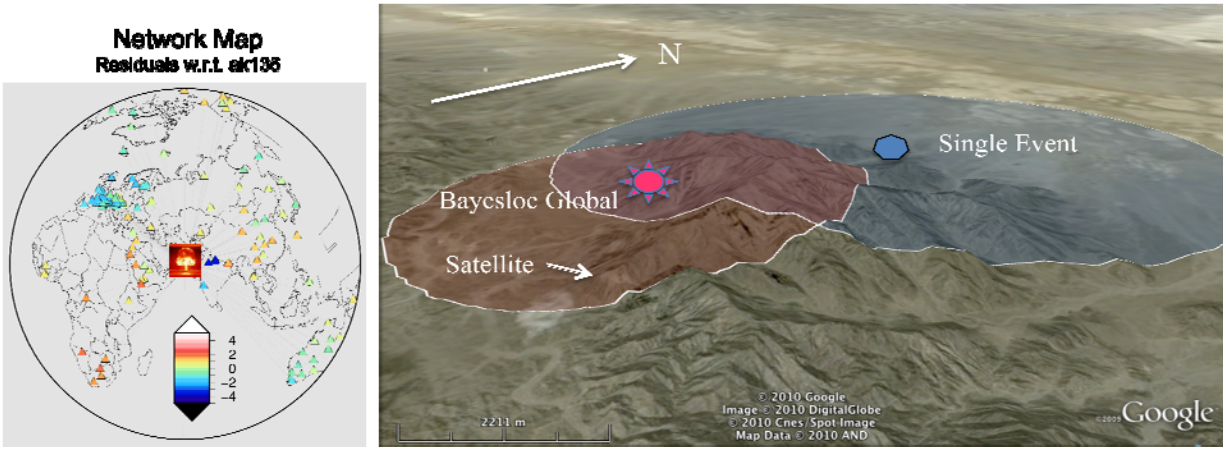

Figure 3. Single-event and Bayesloc global relocation of the May 28, 1998 Pakistan nuclear test. Bayesloc mislocation is 4.5 km with epicenter error ellipse area of 70 km^2. Single event mislocation is 10.1 km with epicenter error ellipse area of 207 km^2. The satellite location is from Albright et al.,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

We imposed tight priors on the shift of travel-time curves for teleseismic phases (*P*, *PcP*, *pP*, *sP*). Loose priors were used on the slope (travel time/distance) of the teleseismic travel-time curves but posteriori changes to the travel-time slope were insignificant. A loose prior ±5 seconds was used for the regional phase (*Pn*) travel times, and the posteriori shift of the *Pn* curve is 0.42 seconds. The slope of the *Pn* curve changes significantly and equates to *Pn* phase velocity of 8.16 km/s compared to 8.05 km/s in ak135. In summary, the *Pn* phase travels faster horizontally through the upper mantle than the global average, but *Pn* travel times are delayed relative to the global average due to thick crust.

Figure 4 is a Gaussian representation of the posteriori data precision for each phase. In this case prior information on the pick error was not used, resulting in a pure data-driven assessment of the precision for each phase. The errors are relative to the corrected travel times, but we did include the 0.42 second shift in the *Pn* travel-time prediction to show the significance of the shift with respect to the overall Pn distribution. *P* is found to be most precisely timed, followed by *Pn, pP, PcP, and sP*. Summary of the posteriori pick error is provided in Table 1.

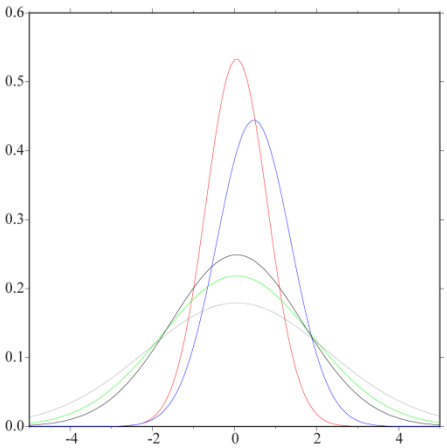

Figure 4. Gaussian representation of posteriori measurement error distribution. *P*=red, *Pn*=blue, *PcP*=green, *pP*=black, and *sP*=grey. The 0.42 second shift of *Pn* travel prediction is included. Standard deviation of measurement error for each pick is also tabulated in Table 1.

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 assessed by counting the number of times a label is accepted. 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 about 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 *Pn* phases, are not likely to be erroneous, but the rate of erroneous data entries for later-arriving phase – *pP*, *sP*, *and PcP* – is from 1% to 2%. The difference between columns 5 and 6 is the rate of misidentified phases, i.e., valid arrivals with the wrong phase assignment. For example, approximately 3% of reported *P* phases are mislabeled. Detailed examination of the Bayesloc output finds that if a *P* phase is relabeled it is most likely to be relabeled as a depth phase. The depth phase *pP* is also commonly relabeled as either another depth phase or as *P*. Figure 5 shows an example of phase relabeling with waveforms added to substantiate the Bayesloc result. Clearly the removal of one *P* phase is correct, and relabeling the *pP* arrival to *sP* appears reasonable given the apparent arrival of the true *pP* phase that precedes the relabeled phase.

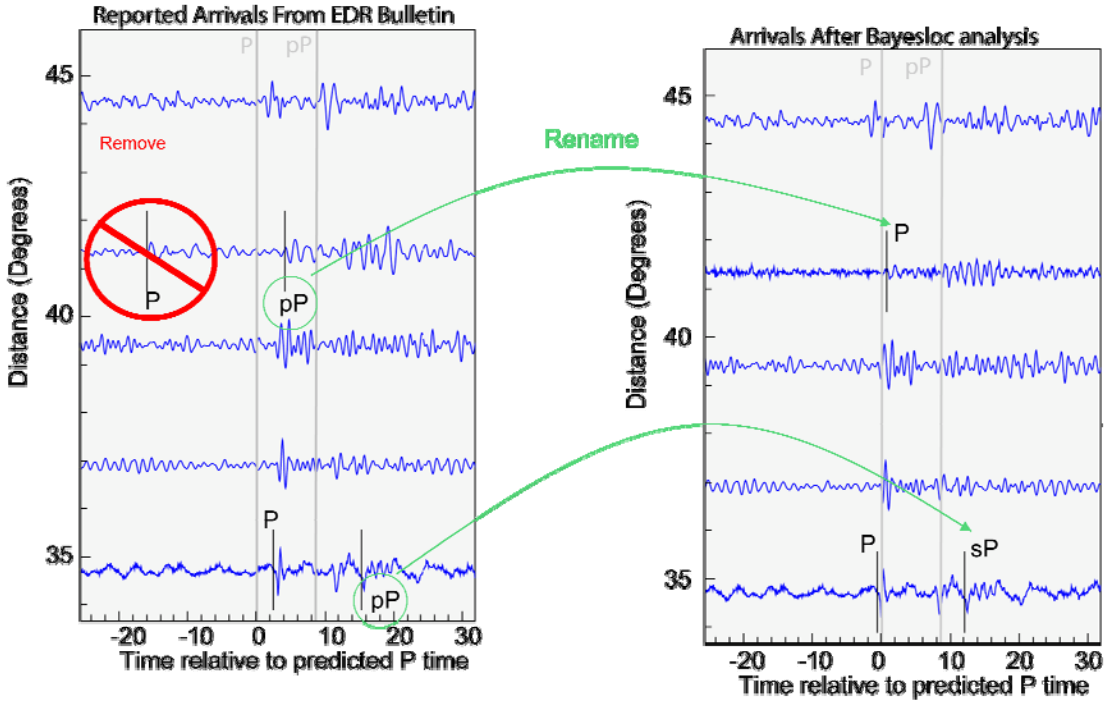

Figure 5. Example of phase re-labeling. In this case, posteriori labels have probability greater than 0.9. Validation of phase relabeling using waveform data suggests that posteriori labels are correct. The data shown here are a small subset of the data for this particular event.

Figure 6 plots posteriori precision (1/variance) for the 3 components of the Bayesloc error model. Low precision indicates that no configuration of the multiple-event system could be found to precisely fit the data for the tested station, phase, or event. High precision indicates precise data fit. The precision of each arrival-time datum is the product of station, phase, and event terms. Posteriori precision is dominated by pick (measurement) error, but also includes other errors that are not accounted for in the travel-time correction. The *P* phase is found to be the most precise, followed by *Pn*, *PcP*, *pP*, and *sP* (also see Table 1). Stations show the largest variability in precision: arrival time data are very consistent at some stations and inconsistent at others. Likewise, arrival times are more consistently fit by corrected travel-time predictions for some events. A possible reason for this is that some events (e.g., explosions) are more impulsive, as noted by Bondár et al. (2004). Variable data precision is used to weight the

penalty for data misfit. Therefore, Bayesloc relocations and travel-time corrections are weighted to fit high-precision stations and events by design.

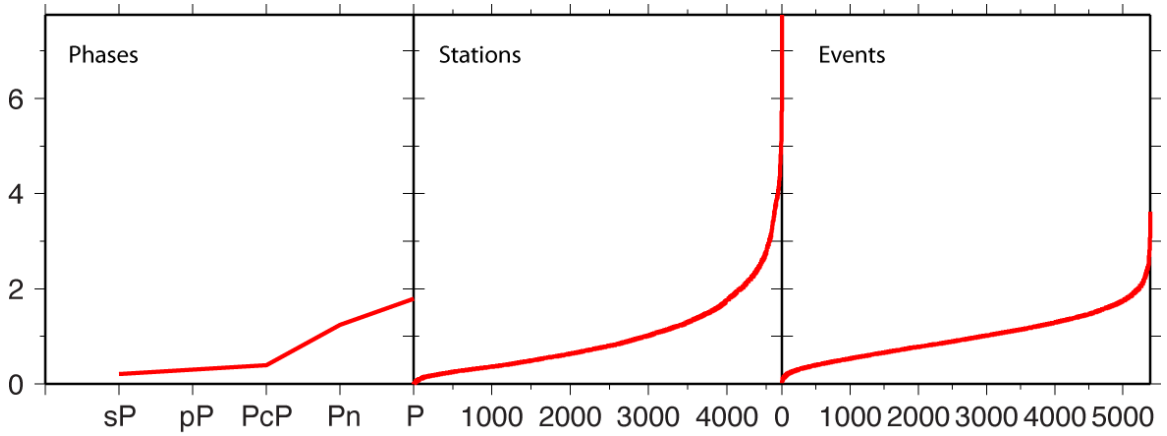

Figure 6. Posteriori precision (1/variance) for phases, stations, and events. Phases, stations, and events are ordered from least precise to most precise. See text for discussion.

Bayesloc posteriori residual standard deviation for the first-arriving *P* wave (*P* and *Pn*) is more than a factor of 3 smaller than input residual standard deviation. Figure 7 shows the density of residual occurrence as a function of distance for input data and Bayesloc output. Approximately 4% of the data are removed since the phases were not confidently determined. The result in Figure 7 suggests that the dominant contribution to global bulletin residuals is location and picks errors, rather than the effects of 3D velocity heterogeneity. At teleseismic distances, the residual distribution shows a slight negative trend, as well as patterns within the body of the residual distribution. *Pn* residuals exhibit a distinct negative trend, consistent with the Bayesloc correction to the *Pn* travel-time curve. The *Pn* distribution after trend removal is slightly larger than the distribution for teleseismic *P*, because of increased pick error and extreme lateral heterogeneity in the Middle East region. Simmons et al. (2010, these Proceedings) use the Bayesloc output as input to 3D tomography. With respect to the 3D velocity model, residuals are 0 mean and standard deviation is reduced to 0.53 seconds. Moreover, the residual trends seen in Figure 7 are removed when travel times are predicted on the basis of the new global tomography model.

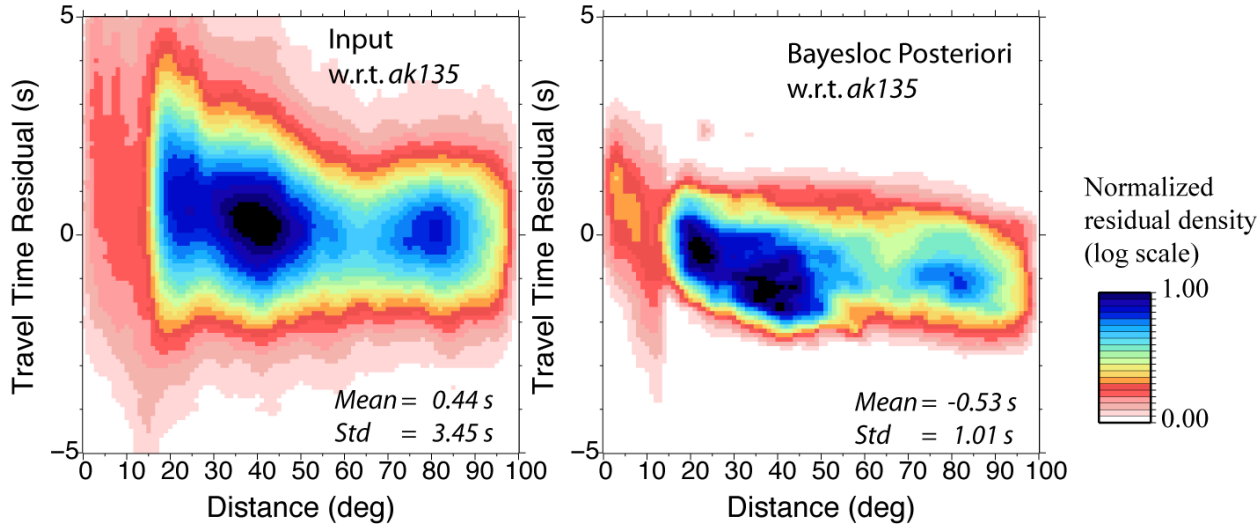

Figure 7. Input and output (posteriori) residual occurrence (density). See text for details.

CONCLUSIONS AND RECOMMENDATIONS

We have made modifications to the Bayesloc multiple-event algorithm that enable application to a global arrival time data set. The first application is to a data set with approximately 5° event spacing globally, to which we add all well-recorded events in the Middle East. Bayesloc relocated epicenters are, on average, 16 km from global bulletin locations (Figure 2). Measured against events that are known to within 1 km, Bayesloc epicenter errors are 5.6 km on average. For the *P* phase: 92% of the bulletin phase labels are confidently determined to be correct; 4% have a significant probability of being improperly labeled, but the input label is the most likely choice; 3.4% are most likely mislabeled; and 0.6% are confidently determined to be erroneous. Analysis for other phases is listed in Table 1. Data precision is decomposed into event, station, and phase components. *P*-phase arrivals are found to be the most precise, followed by *Pn*, *pP*, *PcP*, and *sP*. The most variability in precision is seen for stations, reinforcing common knowledge that some stations are very good and others are not. The range in precision for events is intermediate between stations and phases (Figure 6). Residual standard deviation for the *P* and *Pn* phases is reduced from 3.45 seconds to 1.01 seconds (Figure 7). Reduction in residual magnitude includes identification of outliers and arrivals with ambiguous phase labels. Simmons et al. (2010, these Proceedings) use the posteriori Bayesloc data set for global-scale tomography and demonstrate how the data consistency yields markedly detailed velocity structure beneath the Middle East region. After tomography, residual standard deviation for *P* and *Pn* is 0.53 seconds, which is consistent with Bayesloc assessment of pick error.

ACKNOWLEDGEMENTS

We thank our LLNL colleagues for day-to-day interactions and support. Thanks also to Bill Rodi for numerous, wide-ranging conversations on the location problem.

REFERENCES

Albright, D., Gay, C., and Pabian, F. (1998). New details emerge on Pakistan's nuclear test site, Earth Observation Magazine, December.

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

Bondár, I., and K. L. McLaughlin (2009). A new ground truth data set for seismic studies, *Seismol. Res. Lett.* 80: 3.

Flanagan, M. P., S. C. Myers, and K. D. Koper (2007). Regional travel-time uncertainty and seismic location improvement using a 3Dimensional *a priori* velocity model, *Bull. Seismol. Soc. Am.* 97:804–825.

Engdahl, E. R., R. van der Hilst, and R. Buland (1998). Global Teleseismic Earthquake Relocation with Improved Travel Times and Procedures for Depth Determination, *Bull. Seismol. Soc. Am.* 88: 722–743.

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

Kennett, B.J.N., E.R. Engdahl and R. Buland (1995). Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int,* 122: 108–124.

Ruppert, S., D., Dodge, A. Elliott, M. Ganzberger, T. Hauk and E. Matzel (2005). Enhancing seismic calibration research through software automation and scientific information management, in *Proceedings of the 27th Seismic Research Review: Ground-Based Nuclear Explosion Monitoring Technologies*, LA-UR-05-6407, Vol. 2, pp. 937–945.

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

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.