IMPROVED EVENT LOCATION UNCERTAINTY ESTIMATES

István Bondár, Keith McLaughlin, and Hans Israelsson

Science Applications International Corporation

Sponsored by Air Force Research Laboratory

Contract No. FA8718-05-C-0018

ABSTRACT

While many studies aimed to reduce location bias by introducing improved travel-time corrections, little effort was devoted to the complete estimation of location uncertainties, despite the fact that formal error ellipses are often overly optimistic. Since most location algorithms assume that the observations are independent, correlated systematic errors due to similar ray paths inevitably result in underestimated location uncertainties. Furthermore, the tails of real seismic data distributions are heavier than Gaussian. The main objectives of this project are to develop, test and validate methodologies to estimate location uncertainties in the presence of correlated, systematic and non-Gaussian errors. Particular attention will be paid to robust and transportable models of a travel-time covariance matrix.

The characterization of the full covariance matrix will separate and estimate non-Gaussian, heavy-tailed distributions of measurement and model errors and take into account the correlation due to systematic errors. We will characterize measurement errors as a function of signal parameters, such as phase, distance and amplitude. To achieve this goal we will perform fully controlled experiments by using known signals, scaled down to several magnitude levels and embedded in clean noise (Kohl et al., 2004). We will estimate the correlation structure in the data using various variogram models. To estimate non-linear dependence structures in the data that are not captured by the full covariance matrix, we will apply the theory of copulas, a quickly growing field of statistics to describe tail dependence. Based on the copula theory we will develop a hypothesis test, independent of formal uncertainties, to assess the reliability of the error ellipses obtained from the classical approach using the full covariance matrix.

During the first year of the project preliminary methodologies will be developed, tested and demonstrated on a limited set of event clusters. For validation purposes we use event clusters with GT0-2 events. Our primary choice is therefore the Nevada Test Site (NTS) where an abundance of GT0 nuclear explosions, well-recorded in all distance ranges, is available. The NTS cluster allows us to simulate sparse, unbalanced networks by using subsets of stations. We will also use the Lop Nor, China nuclear explosions (well-recorded teleseismically but with sparse regional networks) and the Lubin, Poland mining explosions (recorded by dense but unbalanced regional networks) to demonstrate the applicability of our methodologies.

OBJECTIVES

The objectives of this project are to develop methodologies to estimate location uncertainties in the presence of correlated, systematic model errors; to characterize measurement errors as a function of signal parameters such as phase and signal-to-noise ratio; and to describe the total error budget in the case of non-linear, non-Gaussian dependence structure. The improved understanding of the complete error budget will be applied to non-linear location estimators to make location programs more robust in the presence of correlated errors and outliers. The resulting error budgets will lead to more robust estimates of location uncertainty. A hypothesis test (independent of formal uncertainty estimates) will be developed to assess the reliability of location uncertainty estimates.

RESEARCH ACCOMPLISHED

The assumption of independent error processes prevails in most modern location algorithms, despite the fact that the problem arising from inadequate representation of systematic bias has been known to seismologists since the advent of modern instrumental seismology. A classic example is the Longshot nuclear explosion (29 October 1965, Amchitka). Herrin and Taggart (1968) showed that a large number of arrivals traveling along similar ray paths through an unmodeled oceanic subducting slab introduced location bias. If unrecognized, correlated systematic errors result in unrealistic error ellipses with degraded coverage (true locations do not lie within the ellipses) and introduce location bias. To further illustrate our motivation to consider the correlation structure in the data, we performed a constrained bootstrapping (Yang et al., 2004) experiment on the 7 October 1994 Lop Nor, China nuclear explosion. The explosion is considered GT1 (Fisk, 2002) and recorded by some 600 stations at teleseismic distances. As Figure 1a indicates, the station distribution is far from azimuthally uniform; and is dominated by the networks in California, Japan and Europe. Figure 1b shows the trajectory of the mislocation vector with increasing number of stations. As more and more stations contribute to the solution the location is driven away from the GT1 location. Since the location algorithm does not account for correlated travel-times along similar ray paths, the relative importance of the Californian and European stations steadily increases, resulting in ever more increasing location bias. As the information carried by the network geometry is exhausted relatively early, adding more stations merely increases data redundancy and increases bias. Furthermore, as shown in Figure 1c, the area of the 90% coverage ellipse monotonically decreases with increasing number of stations. This is because the off-diagonal elements of the covariance matrix are ignored assuming independent errors. Hence, it is guaranteed that the error ellipse will not cover the true location once a sufficiently large number of correlated systematic errors contribute to the solution.

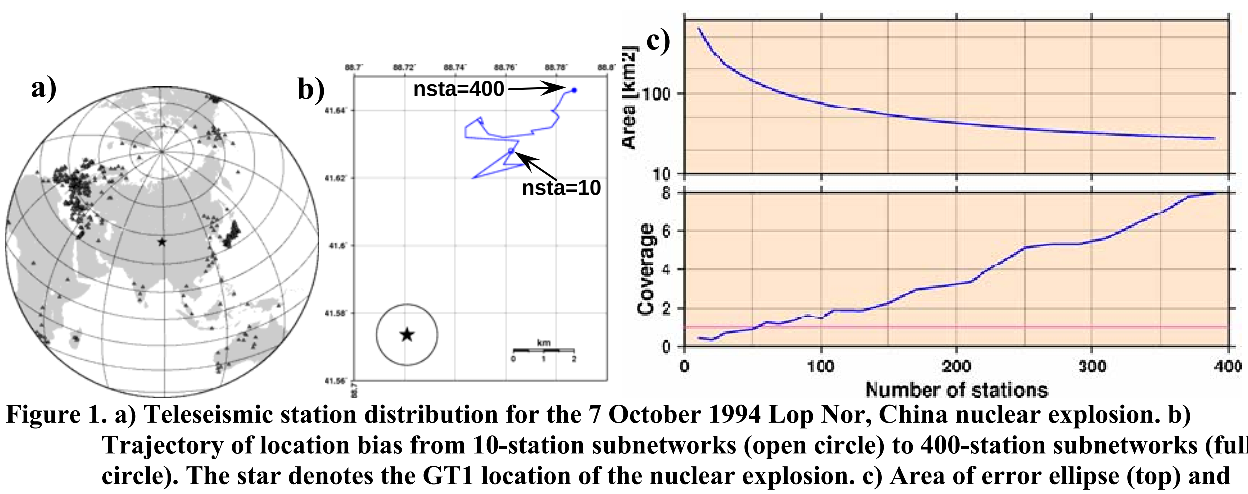

Figure 1. a) Teleseismic station distribution for the 7 October 1994 Lop Nor, China nuclear explosion. b) Trajectory of location bias from 10-station subnetworks (open circle) to 400-station subnetworks (full circle). The star denotes the GT1 location of the nuclear explosion. c) Area of error ellipse (top) and actual ellipse coverage (bottom) with respect the number of stations. If the coverage parameter is larger than one, the true location falls outside the error ellipse.

In this project we focus on the treatment of correlated errors, with non-Gaussian, non-zero-mean, heavy-tailed skewed distributions of reading errors. We will employ variogram analysis of observed residuals to estimate the correlation structure in the data. We will estimate the full covariance matrix by fitting variogram models to empirical station-station variograms using fixed ground truth events and event clusters, as well as event-event variograms for fixed stations. Note that estimating the correlation structure through variogram analysis is essentially the same process that is used to construct empirical travel-time correction surfaces by kriging (Schultz and Myers, 1998; Myers and Schultz, 2000; Rodi, 2003). We will retain the information derived for the correlation structure offered by the variogram analysis to construct a full covariance matrix.

Chang et al (1983) have shown that incorporating the full covariance matrix in the location algorithm is straightforward. The correlation structure implies that linear combinations of station residuals may exist. This can be taken into account by diagonalizing the covariance matrix, thus reducing the dimensionality of the problem. The estimated location error ellipses then necessarily become larger, reflecting the reduction in the equivalent number of uncorrelated observations. We apply a methodology developed by McLaughlin et al. (1988) that transforms the empirical correlation matrix, derived from variogram analysis, so that it becomes positive definite, with positive eigenvalues and unit diagonal elements.

Data sets

To validate the methodologies developed during the course of the project we will rely on high quality, GT0-2 event clusters. Our primary choice is the Nevada Test Site (NTS) data set (Figure 2), which contains a large number of GT0 events, well recorded in all distance ranges. The NTS data set will allow us to perform Monte Carlo and bootstrap experiments using subsets of events and stations to investigate the effect of systematic errors due to unbalanced networks.

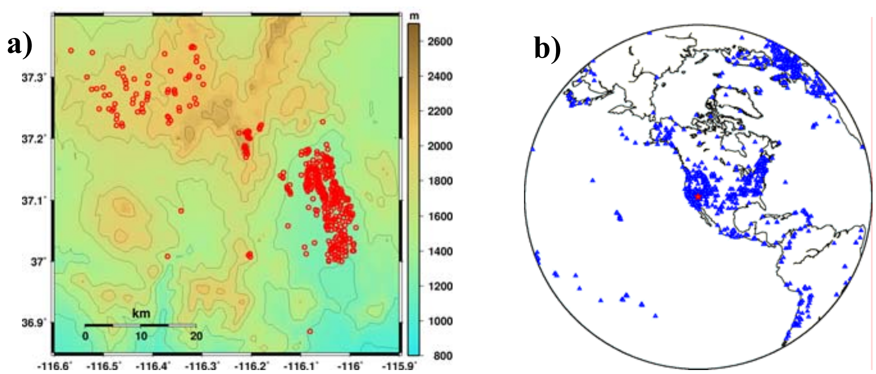

Figure 2. Nevada Test Site (NTS) data set. a) 401 GT0 underground nuclear explosions at Pahute Mesa and Yucca Flat. b) Station distribution in the 0-90° distance range.

The second data set we identified is the GT1-2 underground nuclear explosions at the Lop Nor Test Site (Figure 3). These events are well-recorded at teleseismic distances, but by only a sparse regional network.

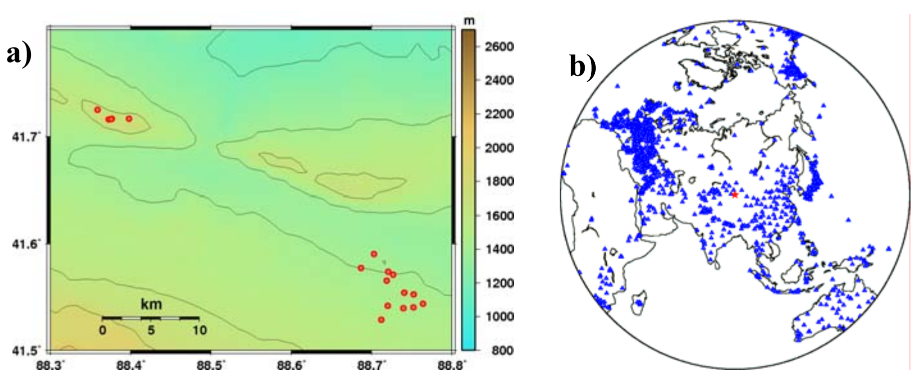

Figure 3. Lop Nor Test Site data set. a) 17 GT1-2 underground nuclear explosions. b) Station distribution in the 0-90° distance range.

Our third data set constitutes the Lubin, Poland GT1-2 mine events (Figure 4). These are recorded by a sparse teleseismic network, and with a dense, but heavily unbalanced, regional network.

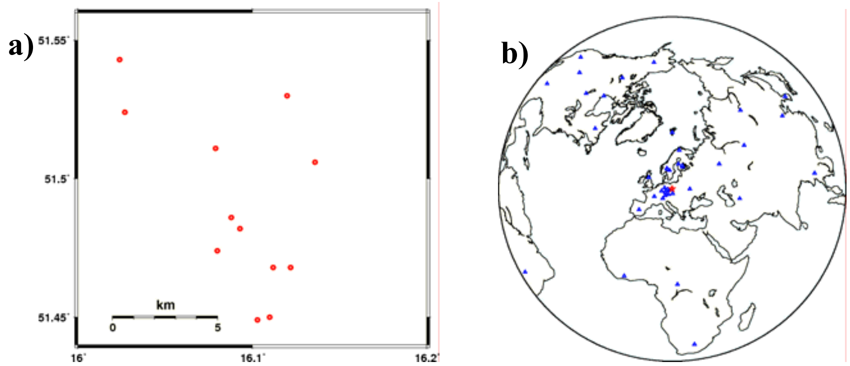

Figure 4. Lubin, Poland data set. a) 12 GT1-2 mine events. b) Station distribution in the 0-90° distance range.

<u>Preliminary analysis of measurement errors</u>

Errors of arrival times are usually assumed to be Gaussian in seismic event location algorithms. It has long been realized, however, that distributions of picking errors are skewed; for example, errors in arrival times of weak signals, picked by both seismic analysts and automatic algorithms, are frequently biased late (Buland, 1986). With improved travel time models being developed, accurate descriptions of picking error distributions and their dependence on signal characteristics, such as SNR and dominant frequency, become more important for location error estimates. Douglas et al. (2005) emphasize the effect of SNR on measurement errors.

The lack of "true" onset times makes the estimation of the error distribution difficult. Consistency of independent readings of seismic analysts is sometimes used as a baseline or “true” onset. However, such "true" onset times based on analyst consistency cannot escape the element of subjectivity in manual readings. Moreover, standard errors of manual picks of about 0.2 s for impulsive phases have been reported (Leonard, 2000). In this project we will use seismic events with ground truth (GT0-2) locations and origin times as well as controlled experiments as a basis for estimating statistical characteristics of picking errors.

An example of using ground truth information is shown by the box plot in Figure 5, which shows the errors in Pn arrival time picks at the station PRI (Priest, CA) from GT0 underground nuclear explosions at the Yucca Flat, Nevada Test Site. The bias in the picking errors gradually increases with decreasing magnitudes and becomes 1s or larger below mb=4.5. Notice also the increased scatter in the errors as the magnitudes become smaller, which indicates the increasing uncertainty in analyst picks as clean onsets fade away with decreasing signal-to-noise ratio.

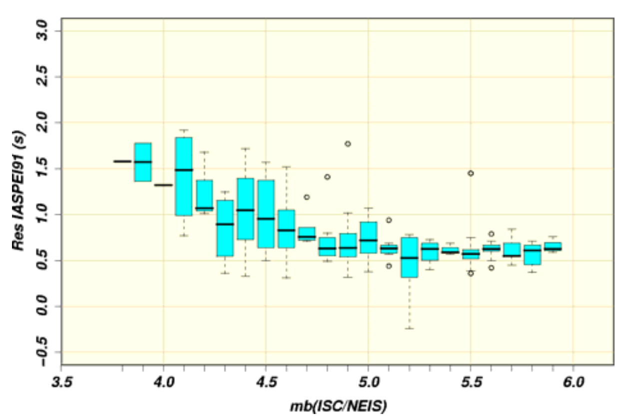

Figure 5. Picking errors of Pn arrival times at the station PRI from GT0 underground nuclear explosions at Yucca Flat, Nevada Test Site as a function of magnitude. Residuals are calculated relative to the IASPEI91 predictions from the known origin times.

Unfortunately, GT0 event clusters are not always at our disposal to investigate the effect of decreasing SNR on arrival time picks. To overcome this problem, we follow the methodology of Kohl et al (2004, 2005) which uses known signals scaled to various magnitude levels and embedded in clean background noise. Since we know exactly where the embedded signals are and that they are not contaminated by other signals (hence the notion of clean noise), the procedure allows us to design controlled experiments.

An example of using signals embedded at known times in clean background noise for estimating picking error characteristics is given in Figure 6. The example shows picking errors of an automatic algorithm (DFX) for P signals at FINES from large underground nuclear explosions at the Lop Nor Test Site, scaled down to varying sizes and embedded in clean noise. Figure 6a shows how the bias or lateness of the picks sets in for SNR around 6-7 and continues to increase with decreasing SNR, much like the effect illustrated in Figure 5. The QQ plot (quantiles of observed picking errors plotted against Gaussian quantiles) in Figure 6b indicates that picking errors of signals at various SNR ranges exhibit varying means and variances (Rodi, 2004) and deviate from the normal distribution even at high SNR levels.

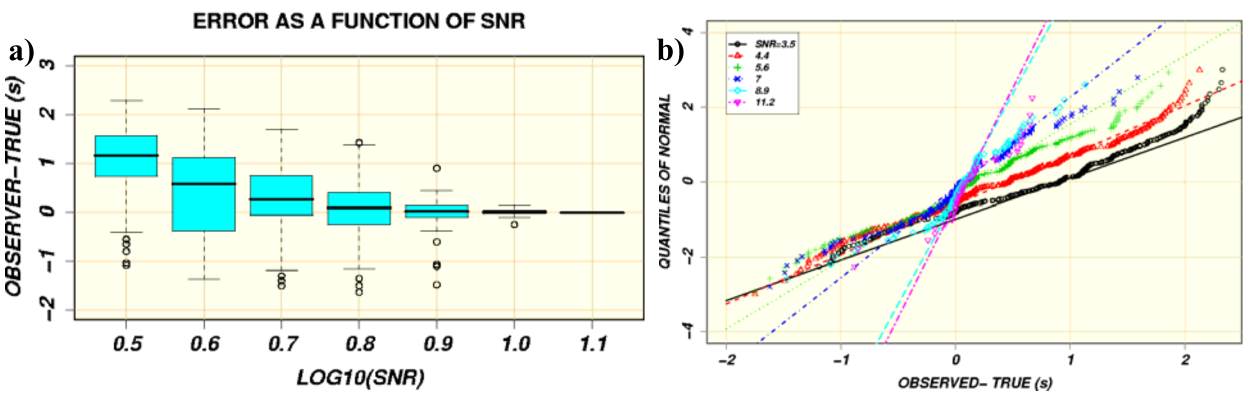

Figure 6. Picking errors of an automatic algorithm obtained for signals, scaled down to various amplitudes and embedded in clean background noise, at the FINES array. a) Reading errors become biased as SNR decreases. b) Reading errors at various SNR levels. If the errors were Gaussian, the observed picking errors (symbols) would align with the normal quantiles (lines).

We will use the methodology outlined above to obtain improved models of observational errors with respect to phase and to account for the effect of signal-to-noise ratio on the residuals when constructing the full covariance matrix.

Non-linear dependence structure

Estimating the full covariance matrix will allow us to account for correlated systematic errors in regions where the bias is unknown (uncalibrated). However, the linear Gaussian approach has its limitations. The full covariance matrix approach implies that the observations are described by a multivariate Gaussian distribution, which can only account for linear correlation structures. Non-linear dependence structures may exist in the data that are not captured by the correlation matrix. Every location algorithm, either linearized or non-linear, minimizes a misfit function, which is typically expressed as the sum of powers of weighted residuals. The inherent assumption is that the likelihood function can be written as the product of individual probability density functions of the observations – that is, the observations are independent. If the observations are dependent, the joint distribution is no longer the mere product of the marginal distributions; and the negative logarithm of the likelihood function can no longer be written as the simple sum of powers of weighted residuals.

Constructing the likelihood (i.e. the joint probability density) function in the general case often proves to be very difficult, and this is exactly why location algorithms make the somewhat unsupported assumption of independent error processes. Sklar's theorem (1959) offers a way to construct the joint distribution function of continuous multivariate random variables. The theorem states that if H is an n-dimensional joint distribution function with marginal cumulative distributions $F_1,\ldots,F_n$, then there exists a unique copula function C such that $H(x_1,\ldots,x_n) = C(F_1(x_1),\ldots,F_n(x_n))$, where $u_i = F_i(x_i)$ denotes the probability integral transformations of x_i. Thus, the copula is the joint cumulative distribution function of the order statistics of the univariate marginal distributions.

The converse of Sklar's theorem is also true, and it implies that we can link together univariate distributions of any type with any copula in order to get a valid multivariate distribution. If F_i^{-1} denotes the inverse of the marginal distribution functions, then there exists a unique copula such that $C(u_1,\dots,u_n) = H(F_1^{-1}(u_1),\dots,F_n^{-1}(u_n))$. The separation of the dependence structure from the marginals is apparent in the form of the likelihood function:

$L(x_1,\dots,x_n;p) = c(F_1(x_1),\dots,F_n(x_n);\vartheta)\prod_i f_i(x_i;p)$, where $c(u_1,\dots,u_n;\vartheta) = \dfrac{\partial^n C(u_1,\dots,u_n;\vartheta)}{\partial u_1 \dots \partial u_n}$ denotes the copula density function; p and ϑ stand for the model and copula parameters, respectively. Hence, a copula is a function that joins or 'couples' a multivariate distribution function to its one-dimensional marginal distribution functions. For a detailed discussion of copulas see Joe, (1997) and Nelsen (1999). The basic idea behind the copula formalism is to separate dependence and marginal behavior between elements of multivariate random vectors.

Using Sklar's theorem, one can construct multivariate distributions with arbitrary margins. For simplicity, we consider bivariate distributions. A great many examples of copulas can be found in the literature and most of the copulas are members of families with one or more real parameters. When the joint multivariate distribution is Gaussian with a covariance matrix Σ, the likelihood function can be written as

$$L(x_1,\dots,x_n;p) = \frac{1}{(2\pi)^{n/2}\sqrt{\det\Sigma}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)} = c^G(\Phi_1(x_1),\dots,\Phi_n(x_n);\Sigma)\prod_i f_i(x_i;p)$$

Hence, the copula formalism offers a way to develop a hypothesis test: if the best fitting copula to the data is the Gaussian copula, then the full covariance matrix adequately describes the dependence structure and provides a reliable estimate for the location uncertainty.

Variogram models with copulas

To illustrate the power of the copula approach, we apply the copula formalism to derive variograms for the NTS data set. Copulas of the form $C(u,v) = \varphi^{-1}(\varphi(u)+\varphi(v))$ are called Archimedean copulas, where φ is a convex, decreasing function with domain (0,1] and range [0,∞) such that $\varphi(1) = 0$. The function φ is called the generator function, which uniquely determines an Archimedean copula. Table 1 lists the most frequently used one-parameter Archimedean copulas. For a more complete set of Archimedean copulas see Nelsen (1999).

Table 1. One-parameter Archimedean copulas

Copula	$C(u,v)$	$\varphi(t)$	α	Limits	Kendall's τ
Clayton	$\max\left(\left[u^{-\alpha}+v^{-\alpha}-1\right]^{-1/\alpha},0\right)$	$\dfrac{(t^{-\alpha}-1)}{\alpha}$	$(0,\infty)$	$C_0=\Pi$ $C_\infty=M$	$\dfrac{\alpha}{\alpha+2}$
Ali-Mikhai-Haq	$\dfrac{uv}{1-\alpha(1-u)(1-v)}$	$\ln\dfrac{1-\alpha(1-t)}{t}$	$[-1,1)$	$C_0=\Pi$	$\dfrac{3\alpha-2}{3\alpha}-\dfrac{2(1-\alpha)^2}{3\alpha^2}\ln(1-\alpha)$
Gumbel	$\exp\left(-\left[(-\ln u)^\alpha+(-\ln v)^\alpha\right]^{1/\alpha}\right)$	$(-\ln t)^\alpha$	$[1,\infty)$	$C_1=\Pi$ $C_\infty=M$	$1-\dfrac{1}{\alpha}$
Frank	$-\dfrac{1}{\alpha}\ln\left(1+\dfrac{(e^{-\alpha u}-1)(e^{-\alpha u}-1)}{e^{\alpha}-1}\right)$	$-\ln\dfrac{e^{-\alpha t}-1}{e^{-\alpha}-1}$	$(-\infty,\infty)\setminus\{0\}$	$C_{-\infty}=W$ $C_0=\Pi$ $C_\infty=M$	$1-\dfrac{4}{\alpha}\left[D_1(-\alpha)-1\right]$ $D_1(-\alpha)=\dfrac{\alpha}{2}+\dfrac{1}{\alpha}\int_0^\alpha \dfrac{t}{e^t-1}dt$
Joe	$1-\left[(1-u)^\alpha+(1-v)^\alpha-(1-u)^\alpha(1-v)^\alpha\right]^{1/\alpha}$	$-\ln\left[1-(1-t)^\alpha\right]$	$[1,\infty)$	$C_1=\Pi$ $C_\infty=M$	No closed form

Since φ is a function of the copula parameter α, identifying φ is equivalent to identifying the Archimedean copula itself. Genest and Rivest (1993) described a procedure to identify the form of φ from a sample of bivariate observations. The procedure is based on generating the intermediate (unobserved) random variable $\omega_i = F(x_i, y_i)$ that has a distribution function $K(t) = P(\omega_i \le t)$. Thus, $K(t)$ is the cumulative distribution function of the pseudo-observations ω_i, or in other words, the multivariate probability integral transformation of $F(x,y)$ (Genest and Rivest, 2001; Genest et al., 2002; Nelsen et al., 2003). This distribution function is related to the generator of an Archimedean copula through the expression $K(t) = t - \varphi(t)/\varphi'(t)$.

Thus, to identify the best fitting copula, we

1. Estimate Kendall's τ from the sample by the non-parametric estimate
$$\tau_n = \frac{2}{n(n-1)} \sum_{i<j} sign\left((x_i - x_j)(y_i - y_j)\right)$$
2. Construct a non-parametric estimate of $K(t)$
   a. $\omega_i = \frac{\sum_j 1(x_j < x_i, y_j < y_i)}{n-1}, \quad i = 1, \ldots, n$
   b. $K_n(t) = \frac{\sum_i 1(\omega_i \le t)}{n}, \quad 0 < t < 1$
3. Construct a parametric estimate of $K_\varphi(t)$
   a. Use τ_n to get an initial estimate of α_n
   b. Use α_n to estimate $\varphi_n(t)$
   c. Use $\varphi_n(t)$ to estimate $K_\varphi(t)$ using the relationship $K_\varphi(t) = t - \varphi(t)/\varphi'(t)$
   d. Refine α so that it minimizes $\sum \left| K_n(t) - K_\varphi(t) \right|$

Repeat step 3 for several choices of φ and select the best fitting copula.

The copula formalism offers an elegant way to construct the conditional probability distributions and derive quantile regression curves of y subject x (Frees and Valdez, 1998). The p-th quantile regression curve is defined as $y_p = F_2^{-1}(v_p)$ where v_p is the solution of the equation $C(v_p \mid u) = \frac{\partial C(u, v_p)}{\partial u} = p$. Setting p to 0.5 yields the median regression curve of y subject to x.

Because of the well-known local upper-mantle velocity heterogeneity at the NTS site (e.g. Cormier, 1987; Lynnes and Lay, 1988) we treat Pahute Mesa and Yucca Flat separately. We use robust statistics (*smad*) to estimate the variance of residual differences as a function of station separation for fixed events. Using the copula framework allows us to derive strictly data-driven models of variograms, (i.e. we are not forcing any a priori models, such as the commonly used exponential or spherical models), which still yield closed formulas. We define the variogram as the median regression curve of *smad* with respect to station separation, and we derive the median regression from the best fitting copula.

Figure 7 shows the Pn median regression curves for Pahute Mesa and Yucca Flat. Note that in both cases the best fitting copula is identified as the Clayton copula, with a slightly different parameter: 0.56 and 0.35, respectively. The difference in the copula parameters for Pahute Mesa and Yucca Flat may also account for the SNR dependence in the residuals, as signals from Yucca Flat explosions have typically lower SNR than those from Pahute Mesa.

For the Clayton copula, solving $C(v_p \mid u) = u^{-\alpha-1}\left(u^{-\alpha} + v_p^{-\alpha} - 1\right)^{-(\alpha+1)/\alpha} = p$ for v_p we obtain
$$v_p = \left(1 + u^{-\alpha}(p^{-\alpha/(\alpha+1)} - 1)\right)^{-1/\alpha}.$$
The Clayton copula exhibits lower tail dependence, conveniently describing the fact that with decreasing station separation the residual differences become increasingly correlated.

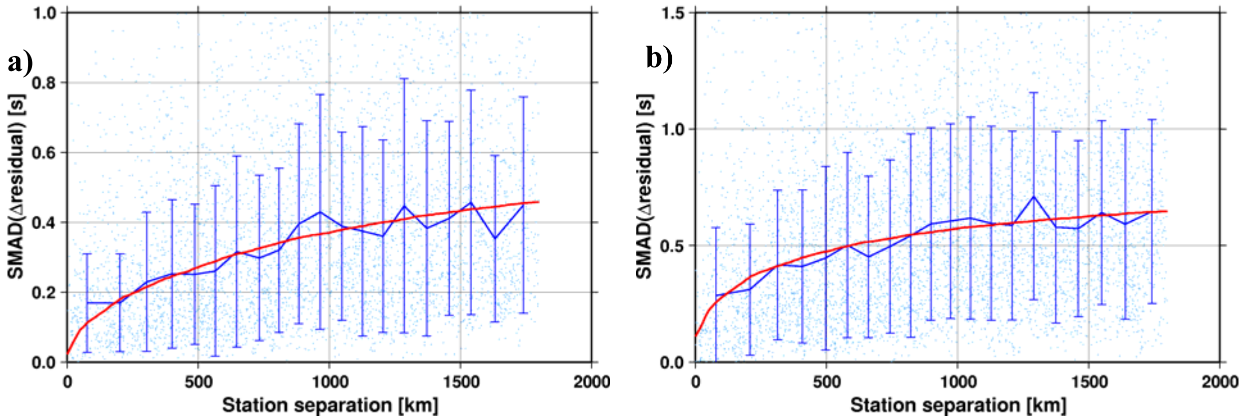

Figure 7. Variogram estimation for Pn phases at Pahute Mesa (a) and Yucca Flat (b). The thick blue line connects the median smad values at every 5-percentile worth of data; the red line shows the median regression curve derived from the best fitting copula.

CONCLUSIONS AND RECOMMENDATIONS

We have identified the data sets we will use to test and validate the methodologies developed in the course of the project. These include the NTS GT0 and Lop Nor GT1-2 underground nuclear explosions, as well as the GT2 mining events in Lubin, Poland.

We have developed preliminary methodologies to obtain improved models of reading errors and deriving the full covariance matrix from variograms. We have also developed a method to transform the empirical covariance matrix so that it becomes a positive definite matrix.

We have developed a data-driven methodology, based on copula theory, to obtain robust estimates of variogram models.

During the first year of the project we concentrate our efforts to develop, test and validate methodologies, and demonstrate their applicability on a limited set of event clusters.

REFERENCES

Buland, R. (1986). Uniform reduction error analysis, *Bull. Seism. Soc. Am.* 76: 217-230.

Chang, A.C., R.H. Shumway, R.R. Blandford and B.W. Barker (1983). Two methods to improve location estimates – preliminary results, *Bull. Seism. Soc. Am*. 73: 281-295.

Cormier, V.F. (1987). Focusing and defocusing teleseismic P waves by known three-dimensional structure beneath Pahute Mesa, Nevada Test Site, *Bull. Seism. Soc. Am.*, 77, 1688-1703.

Douglas, A., J.B. Young, D. Bowers and P.D. Marshall (2005). An analysis of P travel times for Nevada Test Site explosions recorded at regional distances, *Bull. Seism. Soc. Am.* 95: 941-950.

Fisk, M. (2002). Accurate locations of nuclear explosions at the Lop Nor test site using alignment of seismograms and IKONOS satellite imagery, *Bull. Seism. Soc. Am.* 92: 2911-2925.

Frees, E.W. and E.A. Valdez (1998). Understanding relationships using copulas, *North Am. Actuarial J*. 2: 1-25.

Genest, C., L-P. Rivest (1993). Statistical inference procedures for bivariate Archimedean copulas, *J. Am. Stat. Assoc.* 88: 1034-1043.

Genest, C. and L-P. Rivest (2001). On the multivariate probability integral transformation, *Statistics & Probability Letters* 53: 391-399.

Genest, C., J-F. Quessy and B. Rémillard (2002). Test of serial independence based on Kendall's process, *Canadian J. Statistics* 30: 1-21.

Herrin, E., and J. Taggart (1968). Source bias in epicenter determination, *Bull. Seism. Soc. Am.* 58: 1791-1796.

Joe, H. (1997). *Multivariate models and dependence concepts*, Chapman and Hall, London.

Kohl, B., T. J. Bennett, I. Bondár, B. Barker, W. Nagy, C. Reasoner, and J. Hanson (2004), Development of a network data set for evaluating detection and network processing performance, in *Proceedings of the 26th Seismic Research Review: Trends in Nuclear Explosion Monitoring*, LA-UR-04-5801, Vol. 2, pp. 725-734.

Kohl, B., T.J. Bennett, I. Bondár, B. Barker, W. Nagy, C. Reasoner, H. Israelsson and P. Piraino (2005). Development of a network data set for evaluating detection and network processing performance, in *Proceeding of the 27th Seismic Research Review: Trends in Nuclear Explosion Monitoring*, Palm Springs, CA.

Leonard, M. (2000). Comparison of manual and automatic onset time picking, *Bull. Seism. Soc. Am.* 90: 1384-1390.

Lynnes, C.S. and T. Lay (1988). Analysis of amplitude and travel-time anomalies for short-period P-waves from NTS explosions, *Geophys. J. Int.* 92: 431-443.

McLaughlin, K.L., R.H. Shumway and T.W. McElfresh (1988). Determination of event magnitudes with correlated data and censoring: A maximum likelihood approach, *Geophys. J. Int.* 95: 31-44.

Myers, S.C. and C.A. Schultz (2000). Improving sparse network seismic location with Bayesian kriging and teleseismically constrained calibration events, *Bull. Seism. Soc. Am.* 90: 199-211.

Nelsen, R.B., J.J. Quesada-Molina, J.A. Rodriguez-Lallena and M. Ubeda-Flores (2003). Kendall distribution functions, *Statistics & Probability Letters* 65: 263-268.

Nelsen, R.B. (1999). *An introduction to copulas*, Lecture Notes in Statistics 139, Springer Verlag, New York.

Rodi, W., S.C. Myers and C.A. Schultz (2003). Grid-search location methods for ground-truth collection from local and regional seismic networks, in *Proceedings of the 25th Seismic Research Review - Nuclear Explosion Monitoring: Building the Knowledge Base,* LA-UR-03-6029, Vol. 1, pp. 311-319.

Rodi, W.L. (2004), Error analysis in the joint event location/seismic calibration inverse problem, in *Proceedings of the 26th Seismic Research Review: Trends in Nuclear Explosion Monitoring*, LA-UR-04-5801, Vol. 1, pp. 307-316.

Schultz, C.A., S.C. Myers, J. Hipp and C.J. Young (1998), Nonstationary Bayesian kriging: A predictive technique to generate spatial corrections for seismic detection, location and identification, *Bull. Seism. Soc. Am.* 88: 1275-1288.

Sklar, A. (1959), Fonctions de repartition à *n* dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris* 8: 229-231.

Yang, X., I. Bondár, J. Bhattacharyya, M. Ritzwoller, N. Shapiro, M. Antolik, G. Ekström, H. Israelsson and K. McLaughlin (2004), Validation of regional and teleseismic travel-time models by relocating GT events, *Bull. Seism. Soc. Am.* 94: 897-919.