# IMPROVED SEISMIC EVENT LOCATION USING A DAMPED LEAST SQUARES ALGORITHM

Sanford Ballard and Paul C. Reeves

Sandia National Laboratories

Sponsored by National Nuclear Security Administration
Office of Nonproliferation Research and Engineering
Office of Defense Nuclear Nonproliferation

Contract No. DE-AC04-94-AL85000

## ABSTRACT

Perhaps the most important seismic parameter needed to distinguish explosions from natural earthquakes is the depth of the seismic event. Unfortunately, without seismic observations obtained quite close to the seismic event epicenter, this parameter is difficult to resolve precisely because of the inherent coupling between event depth and origin time. In this paper we will explore techniques to improve the location resolution of the iterative linear least squares location algorithm, which is currently the standard technique for locating seismic events.

To locate seismic events we are using a new code called LOCOO (LOCator Object Oriented). LOCOO uses an iterative least squares inversion procedure that employs singular value decomposition (SVD) to solve the associated matrix equations (Geiger, 1910). This algorithm is essentially identical to the one implemented in LocSAT and EVLoc, which are standard codes that are widely used to locate global and regional seismic activity.

For poorly constrained events, i.e., events that were observed by relatively few, distant stations, the location frequently oscillates during implementation of the iterative least squares algorithm, between two competing solutions, neither of which is the best fit to the available data in a least squares sense. In order to converge, the system must be damped such that a stable solution can be achieved that is near the point in solution space characterized by the minimum sum squared residuals. The details of how such damping is applied are critical to obtaining the best possible solution. We have developed an evolutionary scheme for selecting the proper amount of damping to apply such that a stable solution is achieved. Application of this algorithm in test cases indicates that with this code, we are able to resolve depth much more frequently than existing codes, without having to resort to arbitrarily fixing depth, and that the solutions obtained often display significantly lower residuals.

## OBJECTIVE

The most widely used algorithm for estimating seismic event hypocenters and origin times, and the confidence bounds on those quantities, is the linear least squares inversion algorithm solved using singular value decomposition. This technique is described in detail by Menke (1989) and Lay and Wallace (1995) and is the basis for seismic event location applications and a next-generation application called LOCOO under development at Sandia National Laboratories. In this paper we review the mathematical basis of the algorithm, discuss the major assumptions made in its derivation, and explore the utility of using damping to improve the performance of the location algorithm in cases where some of the assumptions inherent in the algorithm have been violated.

## RESEARCH ACCOMPLISHED

### Formulation of the Inverse Problem

The seismic event location problem can be described as follows: We are presented with *N* observations of seismic arrival times, station-to-event azimuth, and horizontal slowness, which we denote as a vector **d** of length *N*, and we wish to determine the location in time and space of the seismic event which produced the observations. The location is described by a vector **m**, of length $M = 4$, which contains the event latitude, longitude, depth and origin time. In order to determine **m**, $N$ must be $\geq M$ and we need an operator **F**, which relates the event location to the observations:

$$\mathbf{F(m) = d} \tag{1}$$

For the seismic event location problem, **F** is an earth model as defined by the geometry of the Earth and a set of travel timetables with associated corrections. Equation 1 is a statement of the forward problem: given an Earth model and a seismic event location, we can use it to calculate a set of observations. The problem we wish to solve is the inverse problem: given an Earth model, **F**, and a set of observations **d**, find **m**, the location of the seismic event that produced the observations. Since **F** is non-linear, we cannot solve the inverse problem directly. The standard approach is to start with an initial estimate of the event location, $\mathbf{m}_0$, and try to deduce an appropriate perturbation to that location, $\delta\mathbf{m}$, such that

$$\mathbf{m}_0 + \delta\mathbf{m} = \mathbf{m} \tag{2}$$

In other words, find the change in event location which will move the event location from the initial estimate of the event location to the event location that, when operated on by the Earth model, will produce the observations. Combining Equations 1 and 2 we have

$$\mathbf{F}(\mathbf{m}_0 + \delta\mathbf{m}) = \mathbf{d} \tag{3}$$

To solve Equation 3 for $\delta\mathbf{m}$, we first linearize it by expanding the left hand side into a Taylor series

$$\mathbf{F}(\mathbf{m}_0 + \delta\mathbf{m}) = \mathbf{F}(\mathbf{m}_0) + \left[\frac{\partial\mathbf{F}(\mathbf{m}_0)}{\partial\mathbf{m}_0}\right]\delta\mathbf{m} + \left[\frac{\partial^2\mathbf{F}(\mathbf{m}_0)}{\partial\mathbf{m}_0^2}\right]\frac{\delta\mathbf{m}^2}{2!} + \left[\frac{\partial^3\mathbf{F}(\mathbf{m}_0)}{\partial\mathbf{m}_0^3}\right]\frac{\delta\mathbf{m}^3}{3!} + \ldots \tag{4}$$

Ignoring all terms of order 2 and greater and substituting the remainder into Equation 3 yields

$$\left[\frac{\partial\mathbf{F}(\mathbf{m}_0)}{\partial\mathbf{m}_0}\right]\delta\mathbf{m} \approx \mathbf{d} - \mathbf{F}(\mathbf{m}_0) \tag{5}$$

At this point we want to consider our *a priori* uncertainty information. Associated with each observation $d_i$ and prediction $F_i$ are uncertainty estimates $\sigma_{d,i}$ and $\sigma_{F,i}$, respectively, which combine to yield an uncertainty associated with each residual $\sigma_{r,i}$ given by

$$\sigma_{r,i} = \sqrt{\sigma_{d,i}^2 + \sigma_{F,i}^2} \qquad (6)$$

This *a priori* uncertainty information is introduced into the analysis by creating an $N \times N$ diagonal matrix $\boldsymbol{\sigma}$, the diagonal elements of which contain the values $\sigma_{r,i}$ as defined in Equation 6. We multiply both sides of Equation 5 by $\boldsymbol{\sigma}^{-1}$, which becomes

$$\boldsymbol{\sigma}^{-1} \cdot \left( \frac{\partial \mathbf{F}(\mathbf{m}_0)}{\partial \mathbf{m}_0} \right) \delta\mathbf{m} \approx \boldsymbol{\sigma}^{-1} \cdot \left[ \mathbf{d} - \mathbf{F}(\mathbf{m}_0) \right] \qquad (7)$$

Let us examine the terms of this equation. The right hand side is a vector of *N* weighted residuals, defined as the differences between the observations, **d**, and the predictions of those observations derived from the Earth model, **F**, applied to the estimate of the event location, $\mathbf{m}_0$. The weighting factors are the combined measurement and model uncertainties. Their inclusion gives greater weight in the analysis to residuals that have smaller uncertainties, and it renders all of the residuals unitless. The left hand side consists of three terms. The first is an $N \times M$ matrix of weighted partial derivatives of the *N* model predictions with respect to the *M* components of the estimated event location. Each element of this matrix describes the amount by which one of the predictions will increase as a result of a positive change in one of the *M* components of the estimated event location. When the predictions increase, the residuals decrease so the left hand side is a recipe for decreasing the residuals. The second term on the left side of Equation 7, $\delta\mathbf{m}$, is a vector of length *M* containing finite changes in each of the *M* components of the event location.

Note that the right and left hand sides of Equation 7 are related by an "approximately equal" symbol. This indicates that Equation 7 is only an approximation since the high order terms in Equation 4 were omitted. Equation 7 will be correct only to the extent that, in regions of the model parameter space near $\mathbf{m}_0$, changes in weighted residuals are directly proportional to weighted changes in model parameters. In this context, "near" means within a distance characterized by $\delta\mathbf{m}$.

For convenience of notation in all that follows, let us define a matrix **A** of weighted partial derivatives

$$\mathbf{A} = \boldsymbol{\sigma}^{-1} \cdot \frac{\partial \mathbf{F}(\mathbf{m}_0)}{\partial \mathbf{m}_0} \qquad (8)$$

and a vector **r** of weighted residuals

$$\mathbf{r} = \boldsymbol{\sigma}^{-1} \cdot \left[ \mathbf{d} - \mathbf{F}(\mathbf{m}_0) \right] \qquad (9)$$

so that Equation 7 becomes

$$\mathbf{A} \cdot \delta\mathbf{m} \approx \mathbf{r} \qquad (10)$$

Note that Equation 10 does not lead directly to a solution to our original problem, i.e., it does not yield an estimate of the value of $\delta\mathbf{m}$ that leads to **m**. Rather, it tells us only how small changes in a given event location will affect residuals between observed and predicted seismic quantities. What we seek is the particular value of $\delta\mathbf{m}$ which, when added to $\mathbf{m}_0$, **will** lead to the minimum of the sum squared weighted residuals, $\chi^2_{\text{min}}$, where the sum squared weighted residuals, $\chi^2$, is defined as

$$\chi^2 = \sum_{i=1}^{N} \left( r_i - \sum_{j=1}^{M} A_{ij} \delta m_j \right)^2 \qquad (11)$$

$\chi^2_{\text{min}}$ is found by taking the derivative of $\chi^2$ with respect to each of the model parameters and setting them equal to 0:

$$\frac{\partial \chi^2}{\partial(\delta m_k)} = -2\sum_{i=1}^{N}\left(r_i - \sum_{j=1}^{M} A_{ij}\delta m_j\right)A_{ik} = 0\,; \quad k = 1, M \tag{12}$$

Reorganizing terms, this becomes

$$\sum_{i=1}^{N}\left(\sum_{j=1}^{M} A_{ij}\delta m_j\right)A_{ik} = \sum_{i=1}^{N} r_i A_{ik}\,; \quad k = 1, M \tag{13}$$

or equivalently

$$\mathbf{A}^T \cdot \mathbf{A} \cdot \delta\mathbf{m} = \mathbf{A}^T \cdot \mathbf{r} \tag{14}$$

Now, $\mathbf{A}^T \cdot \mathbf{A}$ is an $M \times M$ symmetric matrix so, if it is not singular, it will have an inverse and we can write

$$\delta\mathbf{m} = \left[\mathbf{A}^T \cdot \mathbf{A}\right]^{-1} \cdot \mathbf{A}^T \cdot \mathbf{r} \tag{15}$$

It remains to actually solve Equation 15. Menke (1989) and Lay and Wallace (1995) show that singular value decomposition (SVD) of **A** produces exactly the same least squares solution as Equation 15. SVD decomposes **A** into a set of matrices **U**, **W** and $\mathbf{V}^\mathbf{T}$ where **U** is an $N \times M$ column-orthogonal matrix and $\mathbf{V}^\mathbf{T}$ is the transpose of an $M \times M$ orthogonal matrix. **U** and **V** are orthogonal in the sense that their columns are orthonormal. The columns of **V** are $M$-dimensional vectors that describe the principal axes of the error ellipsoid of the solution vector $\delta\mathbf{m}$. **W** is an $M \times M$ diagonal matrix containing the so-called singular values of matrix **A**. The diagonal elements of **W** are all zero or positive (Press et al., 2002).

Replacing **A** by its decomposition we obtain

$$\mathbf{U} \bullet \mathbf{W} \bullet \mathbf{V}^\mathrm{T} \bullet \delta\mathbf{m} = \mathbf{r} \tag{16}$$

Now, since **U** and $\mathbf{V}^\mathrm{T}$ are orthogonal, their inverses are equal to their transposes. And since **W** is diagonal, its inverse is equal to the diagonal matrix whose elements are the reciprocals of the elements **W**. This allows us to write the solution to Equation 16 as

$$\delta\mathbf{m} = \mathbf{V} \cdot \mathbf{W}^{-1} \cdot \mathbf{U}^\mathrm{T} \cdot \mathbf{r} \tag{17}$$

or equivalently

$$\delta m_k = \sum_{i=1}^{M}\left(\frac{\mathbf{U}_i \cdot \mathbf{r}}{w_i}\right)\mathbf{V}_{k,i}\ ; k = 1, M \tag{18}$$

where the subscripts $i$ on matrices **U** and **V** indicate column numbers and $w_i$ indicates the $i$'th diagonal element of **W**. Equations 17 and 18 say that each element of $\delta\mathbf{m}$ is a linear combination of the columns of **V**, with coefficients obtained by forming the dot products of the columns of **U** with the vector of residuals, **r**, scaled by the singular values.

We have, finally, the solution that we are after: the value of $\delta\mathbf{m}$ which, when added to the initial event location estimate $\mathbf{m}_0$, produces **m**. When **m** is operated on by our earth model, **F**, (Equation 3), a set of predictions is produced that agrees as well as possible with the observations, in a least squares sense.

Alas, the last statement is almost, but not quite, true. It would be true if Equation 5 were exactly equivalent to Equation 3, which it is not because we linearized Equation 3 by ignoring all high order terms in Equation 4. However, if the conditions outlined in the second paragraph following Equation 5 are approximately satisfied, then $\delta\mathbf{m}$ will lead to a solution **m**, which is closer to $\chi^2_{\min}$ than was $\mathbf{m}_0$. If Equation 3 is nearly linear in the model parameters, then **m** will be much closer than $\mathbf{m}_0$. If we accept the improved location as the initial estimate for another application of Equations 2-18, then the solution will be improved further still. If we continue to iteratively apply Equations 2-18 in this manner, we will ultimately reach a situation where the residuals are very close to the minimum we seek and the solution ceases to change significantly. At that point we can conclude that we have reached convergence and accept the solution as the final solution.

## Poorly Constrained Events

Up to this point we have assumed that the solution to our problem was well constrained by the available data. What if that is not the case? For example, if the data set **d** consists only of azimuth observations from a large number of widely distributed seismic stations, it will be possible to accurately constrain the latitude and longitude of the event but it will not be possible to resolve the depth or the origin time. In matrix **A**, all of the elements in the columns corresponding to the derivatives with respect to depth and origin time will contain zeroes. In this case, **A** is singular and the singular values that correspond to depth and origin time will be equal to zero.

Another situation in which problems can arise is when travel time observations are available from a cluster of stations, but all the stations are far from the event. In this case, there will be a trade off between the depth of the event and the origin time. Increasing the origin time of the event, without significantly impacting the residuals, can compensate for an increase in the estimated depth of the event. In this case, **A** is not singular but rather ill-conditioned. The singular values corresponding to depth and/or origin time will be much smaller than those corresponding to latitude and longitude. The degree to which **A** is ill-conditioned is characterized by the ratio of the largest to the smallest singular values, a quantity that is referred to as the condition number.

One of the significant advantages of the SVD algorithm is that after computing the SVD of **A**, we can examine **W** to assess potential difficulties and take steps to alleviate them before finally computing $\delta\mathbf{m}$ using Equation 18 (Press et al., 2002).

If **A** is singular, then one or more of the singular values will be equal to zero and the condition number will be infinite. The difficulty will be encountered when we attempt to compute Equation 18 since we will be attempting to divide by zero. In this case we must take action before computing Equation 18 or our algorithm will generate an exception and fail.

In cases where the condition number is extremely large, round off errors in the computer may come in to play, yielding a solution with wildly large components that send our event location off into regions that make no physical sense. Fortunately, this situation appears to be rare in overdetermined seismic event location problems. More common is for the condition number to be moderately large, in which case it is possible to solve for all of the location parameters, but it may not be particularly useful to do so. In the second example described above where a trade-off between depth and origin time exists, the uncertainties on the depth and origin time of the event may be extremely large (larger than the dimensions of the Earth, for example) rendering the results meaningless.

When we encounter a situation with a large condition number our best option is to change the smallest singular value to infinity. This causes the algorithm to simply not solve for the parameter that corresponds to the manipulated singular value. In the two examples above, if we set the singular value corresponding to depth to infinity, we would simply give up trying to solve for depth. In Equation 18, where $w_i$ appears in the denominator, this means that the element of $\delta\mathbf{m}$ corresponding to depth will be zero and depth will not deviate from its initial estimate.

## Oscillating Solutions

Another difficulty that is frequently encountered when implementing the iterative linear least squares seismic event location algorithm described above is for the solution vector $\delta\mathbf{m}$ to oscillate, resulting in seismic event locations **m** that alternate between two or more significantly different locations. A set of seismic observations that generates this behavior is presented in Table 1 and the resulting oscillatory behavior is illustrated by the green curves in Figure 1. These six observations were drawn randomly from a global set of hand-picked travel time observations recorded for the May 1998 India nuclear test. Note that not only do all four components of the event location oscillate, but so does $\chi^2$. This fact indicates that our algorithm has failed since the solution is not moving in the direction of $\chi^2_{\min}$.

To understand this behavior, we first seek the "best" solution, defined as the solution with the lowest $\chi^2$. To accomplish this we ran a series of locations with depth fixed at different levels from –24 km to 216 km. Since only 3 location parameters were being solved for, the algorithm behaved quietly, producing stable estimates of the latitude, longitude, origin time and $\chi^2_{\min}$, with the event location fixed at each different depth. Figure 2a illustrates $\chi^2_{\min}$ as a function of depth with the blue curve representing the results of the experiment just described. $\chi^2_{\min}$ occurs near a depth of 30 km, which corresponds to the depth of the base of the crust in the IASPEI travel time model. Figure 2b illustrates the "best-fit" latitude and longitude as a function of depth with the blue curve representing the results of the experiment just described. The sharp kink in the curve corresponds to the latitude and longitude of the event location obtained with the event depth fixed at 30 km.

Table 1

| Station | Latitude (degrees) | Longitude (degrees) | Phase | Observed Travel Time (seconds) | Distance (degrees) | Event-to-Station Azimuth (degrees) |
|---|---|---|---|---|---|---|
| UCH | 42.2 | 74.5 | Sn | 379.1 | 14.9 | 7.9 |
| PDY | 59.6 | 112.7 | PcS | 825.5 | 42.6 | 29.5 |
| HIA | 49.3 | 119.7 | P | 483.0 | 42.6 | 46.0 |
| KS31 | 37.4 | 127.9 | P | 522.6 | 47.8 | 63.2 |
| EIL | 29.7 | 35.0 | P | 395.1 | 32.3 | 282.8 |
| OBN | 55.1 | 36.6 | P | 441.4 | 37.6 | 327.1 |

The kink in the blue curve in Figure 2b suggests that a plausible cause of the failure of the location algorithm is that the assumption of linear behavior has been violated. Recall that in the derivation of the location algorithm we assumed that the derivative of the predicted seismic observations with respect to event location was approximately linear (Equations 4 and 5). The kink in Figure 2b indicates quite clearly that this is not the case and that the particular combination of observations that we have available in this location problem has led to a situation in which the combined derivatives of travel time with respect to latitude are highly non-linear at the crust-mantle boundary. When our location $\mathbf{m}_0$ is below the crust-mantle boundary, the derivatives and residuals combine to produce a change in location $\delta\mathbf{m}$ that moves the event location far above the crust-mantle boundary to a location where, due to non-linearity, $\chi^2$ is actually larger than it was at the previous location. At this new location, the derivatives and residuals combine to produce a change in location $\delta\mathbf{m}$ that moves the event location back to where it was before the most recent change in location was imposed.

In the particular example we are considering, the problem seems to be related to the inclusion of an observation of an Sn phase at station UCH. If this observation is omitted the situation is ameliorated

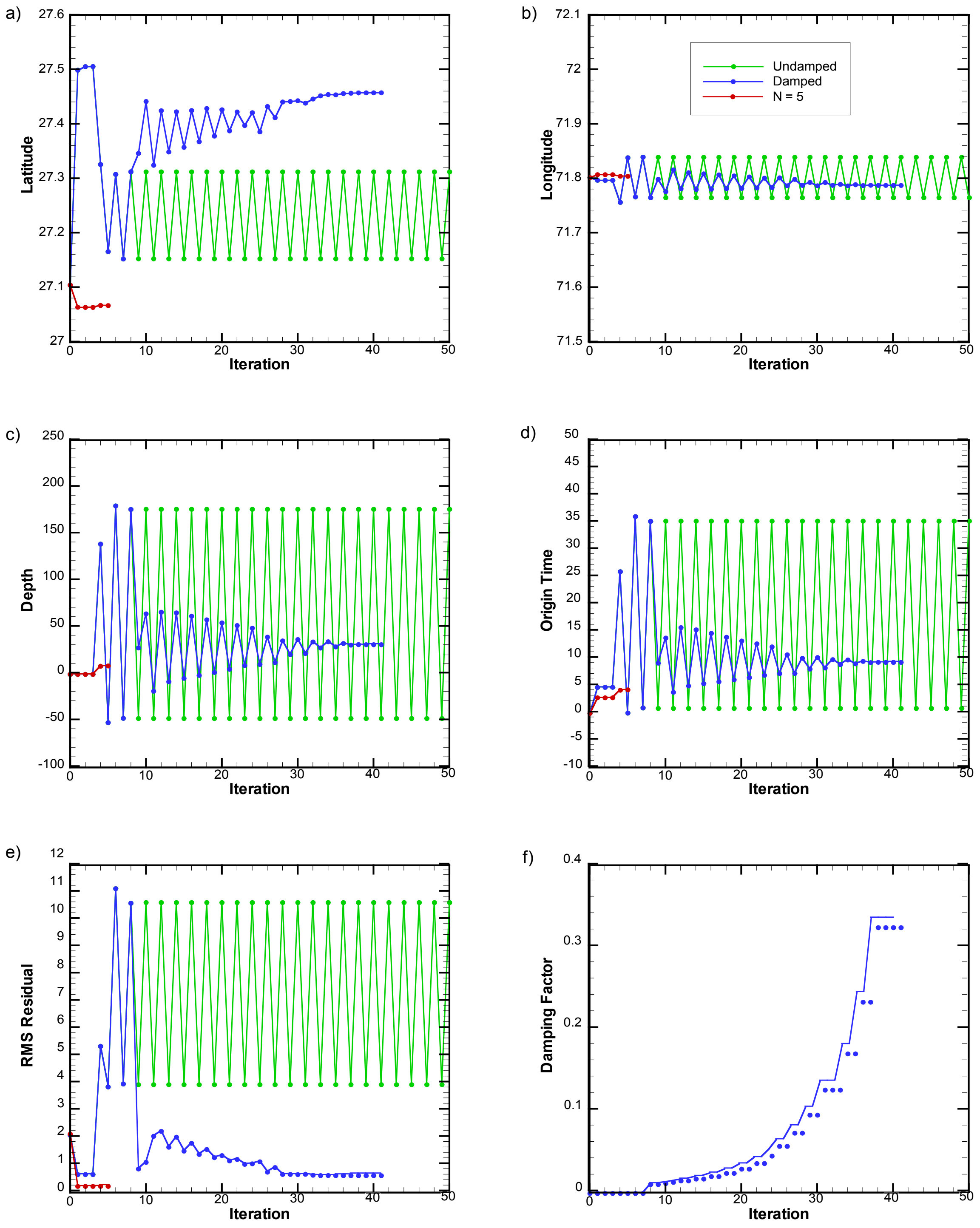

**Figure 1 – The seismic event location, root mean square weighted residual and applied damping factor as a function of iteration number. The green curve illustrates the evolution of the undamped solution, which never converged, while the blue curve illustrates the damped solution. The red curve portrays the evolution of the solution obtained when the Sn phase was omitted from the data set.**

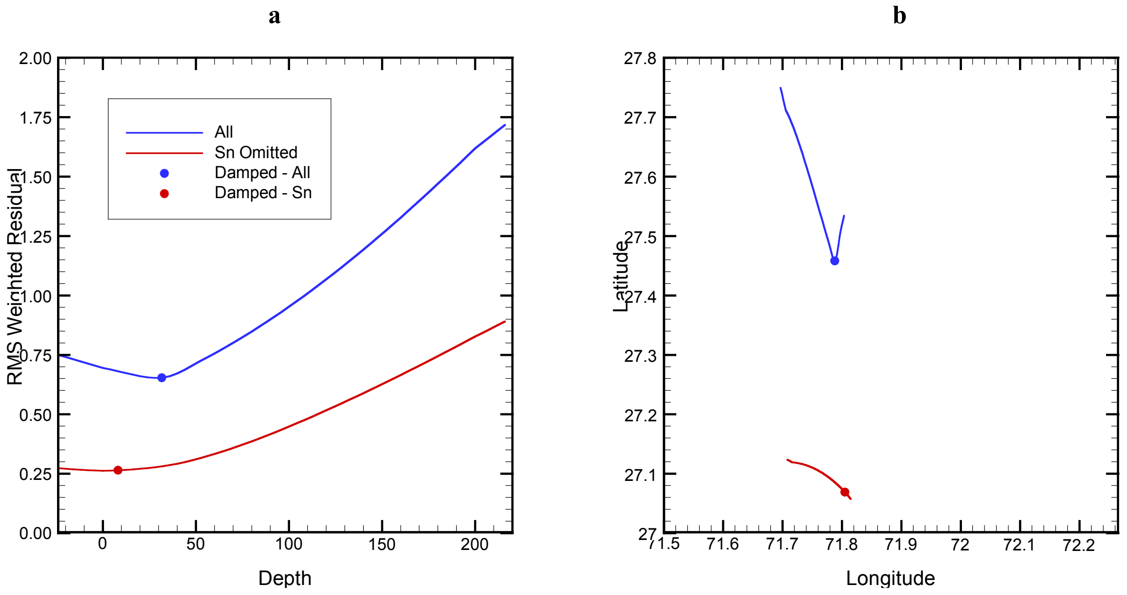

**Figure 2 – a) Curves illustrate the root mean square weighted residual as a function of depth with depth held fixed at different levels. b) The curves illustrate the projection of the best-fit latitude and longitude onto a common horizontal plane when the depth was held fixed at different levels. In both figures, the symbols represent the position of the best-fit damped solution when depth was not held fixed. Blue curves and symbols represent solutions when all 6 observations were included, Red curves and symbols represent solutions when the Sn observation was omitted from the analysis.**

substantially, as illustrated by the red curves in Figures 1 and 2. In Figure 2b, the kink in the latitude-longitude curve has disappeared indicating that the severe non-linearity, which was apparently introduced by the Sn observation, has disappeared. In Figure 1, the solution moves very quickly to convergence without exhibiting any oscillatory behavior.

While the problem can be resolved by discarding observations, that is a highly undesirable solution to the problem. A better alternative is to damp the solution vector such that the location changes by smaller amounts at each iteration, ultimately converging on the best-fit solution. This damping must be properly imposed however, such that the solution continues to move toward $\chi^2_{\min}$. The appropriate course is to impose an additional constraint on the system, over and above the constraint that we seek to minimize $\chi^2$.

The additional constraint is that we wish to also minimize the length of the solution vector (Menke, 1989). These two constraints compete with other so by adding the minimum length constraint, we weaken the $\chi^2_{\min}$ constraint. Total minimization of the solution length alone would be accomplished by setting all the singular values to infinity, in which case the solution vector would consist of all zeroes and its length would be zero. We combine the $\chi^2_{\min}$ and minimum length constraints by adding a constant damping factor, $\lambda$ to each element of **W**. Equation 17 and 18 become

$$\delta\mathbf{m} = \mathbf{V}\cdot(\mathbf{W}+\lambda\mathbf{I})^{-1}\cdot\mathbf{U}^{\mathrm{T}}\cdot\mathbf{r} \qquad (19)$$

and

$$\delta\mathbf{m} = \sum_{i=1}^{M}\left(\frac{\mathbf{U}_i\cdot\mathbf{r}}{w_i+\lambda}\right)\mathbf{V}_i \qquad (20)$$

Since $\lambda$ appears in the denominator of Equation 20, it acts to reduce the magnitude of the changes in the location parameters in each iteration. So long as $\lambda$ is not too large, the minimum $\chi^2$ constraint will dominate and the solution will continue to move toward $\square^2_{\text{min}}$. This is precisely the behavior needed to overcome the oscillations induced by non-linearity. The trick is to impose sufficient damping to overcome the oscillations but not to impose so much that the solution stops moving before it reaches $\square^2_{\text{min}}$.

The algorithm we have implemented to accomplish this begins by initializing the applied damping factor, $\square$, to zero at the start of the location algorithm. If, after some minimum number of iterations, $\square^2$ is observed to increase over its value at the conclusion of the previous iteration, the applied damping factor is set to some small initial damping factor, $\square_0$, which is typically some small fraction of the largest singular value. After each successive iteration, if $\square^2$ is observed to increase over its value at the conclusion of the previous iteration, then the applied damping factor is multiplied by an additional factor, $\square$, which consists of two components, $\square_{const}$ and $\square_{factor}$, with

$$\square = \square_{const} + \square_{factor} \frac{iteration}{\max iterations} \tag{21}$$

where *iteration* is the current iteration number and *maxiteration* is the user-specified maximum number of iterations that the algorithm is allowed to implement before aborting. Typically, both $\square_{const}$ and $\square_{factor}$ are set to equal 1. In this case, $\square$ grows very slowly early during a run but then grows very rapidly as the maximum number of allowed iterations is approached. Experience indicates that this algorithm allows the solution to reach convergence before the damping factor gets to be so large that the solution no longer moves.

Application of this algorithm to the example problem is illustrated by the blue curve in Figure 1. The damping factors imposed as a function of iteration number are illustrated in Figure 1f. The maximum number of iterations allowed was set at 100. Oscillations are detected after just a few iterations and the damping factor begins to increase. After about 36 iterations, the solution has stabilized and oscillations, while still occurring with very small amplitude, are no longer discernible at the scale of the illustration in Figure 1. Finally, after 42 iterations, the algorithm converges. The fact that the algorithm converged on the "correct" solution is illustrated in Figure 2b, where the small symbols represent the solutions obtained with the damped algorithm just described. In all cases, the damped algorithm achieved solutions characterized by the minimum of $\square^2$ as determined by the multiple fixed depth runs.

It is interesting to note that the $\square^2_{\text{min}}$ solution occurs very near the base of the crust at 30 km depth even though the event that produced the observations is known to have occurred at the earth's surface. Note also that the solutions obtained when the observation that induced the non-linearity was omitted occur at a much shallower depth than the non-linear solutions. One might speculate that the non-linearity induced by the particular combination of observations used in the analysis somehow pulled $\square^2_{\text{min}}$ to the point in the model space where the non-linearity was most pronounced. If this is the case, then the minimum sum squared weighted residual solution, whether obtained by the damped linear least squares algorithm or by any other solution algorithm that seeks to minimize $\square^2$, might be misleading. The importance of this hypothesis must be kept in perspective, however, given that the uncertainty on the depth in the example presented is large (± 550 km), even compared to the amplitude of the oscillations encountered. This issue is the subject of continuing investigation.

## CONCLUSION AND RECOMMENDATION

In this paper, we have reviewed the mathematical basis of the linear least squares inversion algorithm, which is the most widely used algorithm for estimating seismic event hypocenters and origin times. We have placed particular attention on the assumption of linearity of the predictions of the observations with respect to the seismic event location, showing how the assumption arises in the derivation and discussing the implications of the assumption. Finally, using a sample data set drawn from the full set of observations generated by the May 1998 India nuclear test, we explored the benefits of using damping in the linear least squares inversion technique to improve its performance in situations where the linearity assumption is violated.

## REFERENCES

Lay, T. and T. C. Wallace (1995), *Modern Global Seismology*, Academic Press.

Menke, W. (1989), *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press.

Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery (2002), *Numerical Recipes in C++, The Art of Scientific Computing*, 2nd Edition, Cambridge University Press.