LOCATION OF MICROEARTHQUAKES INDUCED BY HYDRAULIC FRACTURING

by

William Rodi, Yingping Li, and C.H. Cheng

Earth Resources Laboratory
Department of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

This paper examines the problem of locating microearthquakes induced by hydraulic fracturing using seismic arrival time data. In addition to the use of absolute arrival times measured for individual events, we consider the use of differential arrival times amongst a set of two or more seismic events as a means of constraining their locations relative to one another. Differential arrival times can be measured very accurately using cross-correlation techniques and are less sensitive than absolute arrival times to subsurface velocity structure. We have developed an algorithm which combines relative event location techniques with conventional absolute location techniques and applied it to a set of 19 microearthquakes recorded during a hydraulic fracturing experiment conducted as part of the Los Alamos Hot Dry Rock project. We find that the events, except for a few outliers, delineate a planar zone 30 meters in dimension, presumably a fracture plane. This example shows that the use of differential arrival times improves the accuracy of locating microearthquake clusters and that the relative locations of events within the cluster are better determined than their absolute locations. The results also suggest the need for directional data from three-component stations or better station geometry to further improve location accuracy.

INTRODUCTION

The use of seismic data to locate microearthquakes is becoming an established procedure for mapping hydraulically induced fractures (e.g., House, 1987; Fehler et al., 1987; Vinegar et al., 1992). The standard method of seismic event location, developed originally for global earthquake studies, is to fit the hypocenter and origin time of an event to the arrival times of seismic phases measured at a network of seismometers. The effectiveness of this procedure in hydrofracture monitoring, however, is limited by practical restrictions on the number and distribution of seismometers which can be deployed.

Previous applications have typically used a network with less than five receivers with at most one or two near the injection depth. Given such small networks it is important to use as much information as possible from each seismogram, including both P and S wave arrival times and polarization information when 3-component stations are available (see e.g., Moriya et al., 1993).

A difficulty which occurs in all applications of seismic source location is the dependence of the data on the properties of the propagation medium. In particular, arrival times depend on the seismic velocities of the earth as well as on the hypocenter and origin time of the source. Consequently, a source location determined from arrival times will depend on the velocity model used in fitting the data, and errors in this model will induce an error into the inferred location. One way of reducing these model-induced errors is to estimate the earth's velocity structure simultaneously with the hypocenters of a large set of events, i.e., to combine source location with tomographic imaging. Block (1991) and Block et al. (1993) applied this approach to data from a hydraulic injection experiment at the Fenton Hill, New Mexico geothermal area conducted as part of the Los Alamos Hot Dry Rock project. Fehler and Phillips (1991) describe a variation of simultaneous hypocenter-velocity inversion in which the velocity functions are allowed to change with time in the vicinity of microearthquakes as they occur.

Relative event location techniques are another way of addressing the hypocenter-velocity tradeoff. The underlying principle of this approach is that, for a given seismic phase and given receiver, the difference between the arrival times from two events close to one another is much less sensitive to the velocity structure than are the absolute arrival times from each individual event. Moreover, for events with similar location and focal mechanism, very precise differential arrival times can be measured using waveform correlation techniques. Differential arrival times contain information about the separation between the observed events, i.e., the difference between their hypocentral locations, rather than their absolute locations. The idea behind relative event location then is to determine the relative pattern of hypocenters within a cluster of events. This technique has been used in hydrofracture monitoring by Phillips et al. (1992) and Moriya et al. (1993). The procedure of Phillips et al. (1992) includes a simultaneous inversion for the relative locations of events in a cluster and the average P to S velocity ratio in the vicinity of the event cluster.

This report treats the problem of locating microearthquake clusters induced by hydraulic injection using data from a seismic monitoring network. The report is organized as follows. The next two sections outline the methodology of multiple event location with arrival time data and the measurement of differential arrival times from seismic waveforms. We then describe the application of the algorithm to an injection experiment at the Fenton Hill, New Mexico geothermal area conducted as part of the Los Alamos Hot Dry Rock experiment. The last section presents some conclusions of this study.

INVERSE PROBLEM OF EARTHQUAKE CLUSTER LOCATION WITH ARRIVAL TIMES

General Formulation

Let E denote a set of earthquakes occurring in a cluster. For each event $e \in E$ we let $\mathbf{x}_e$ denote the hypocenter of e and t_e its origin time. The centroid hypocenter and origin time of the event cluster E will be denoted $\mathbf{x}_E$ and t_E, respectively. These are defined as

$$\mathbf{x}_E = \frac{1}{|E|} \sum_{e \in E} \mathbf{x}_e$$
$$t_E = \frac{1}{|E|} \sum_{e \in E} t_e$$

where $|E|$ denotes the cardinal number of the set E (i.e., the number of events in the cluster). The relative hypocenter and origin time of event e with respect to the cluster are defined by

$$\Delta_E \mathbf{x}_e = \mathbf{x}_e - \mathbf{x}_E \tag{1}$$
$$\Delta_E t_e = t_e - t_E. \tag{2}$$

We let S be a set of seismic stations which record the earthquakes in E, and let W be a set of body wave types which can be identified in the recorded seismograms. For hydraulic fracture monitoring we typically have $W = \{\mathrm{P}, \mathrm{S}\}$. The location of the station $s \in S$ is denoted $\mathbf{x}_s$.

Associated with each wave type $w \in W$ is a function of position u_w which gives the spatial dependence of the slowness (reciprocal velocity) of w. We then define a traveltime functional F as

$$F(\mathbf{x}_e, \mathbf{x}_s, u_w) = \min_{\substack{\Gamma \\ \mathbf{x}_e, \mathbf{x}_s \in \Gamma}} \int_{\mathbf{x} \in \Gamma} dl(\mathbf{x})\, u_w(\mathbf{x}) \tag{3}$$

where $\int_{\mathbf{x} \in \Gamma} dl(\mathbf{x})$ denotes integration along a raypath Γ. We will denote the least-time raypath minimizing the integral in eq. (3) as $\Gamma(\mathbf{x}_e, \mathbf{x}_s, u_w)$.

Let T_{esw} denote the observed *absolute* arrival time of wave w at station s from event e. The standard model of arrival times is

$$T_{esw} = t_e + F(\mathbf{x}_e, \mathbf{x}_s, u_w) + n^{\mathrm{obs}}_{esw} \tag{4}$$

where n^{obs}_{esw} is a measurement error. Let $R \subseteq E$ denote a set of reference events chosen from the cluster E. The observed *differential* arrival time of wave w at station s of the

event e relative to R is defined as

$$\Delta_R T_{esw} = T_{esw} - \frac{1}{|R|} \sum_{r \in R} T_{rsw}.$$

The event e may or may not be an element of R. The model of differential times is given by

$$\begin{aligned} \Delta_R T_{esw} = t_e &- \frac{1}{|R|} \sum_{r \in R} t_r \\ &+ F(\mathbf{x}_e, \mathbf{x}_s, u_w) - \frac{1}{|R|} \sum_{r \in R} F(\mathbf{x}_r, \mathbf{x}_s, u_w) \\ &+ n^{\text{obs}}_{Resw}. \end{aligned} \tag{5}$$

Here n^{obs}_{Resw} denotes the measurement error in $\Delta_R T_{esw}$. We assume that $\Delta_R T_{esw}$ has been measured separately from the absolute arrival times T_{esw} and therefore do not write n^{obs}_{Resw} as differences amongst the n^{obs}_{esw}. Note that eq. (5) subsumes eq. (4) if we define $(1/|R|) \sum_{r \in R} = 0$ when R is the empty set.

We will assume that $\Delta_R T_{esw}$ has been measured for various reference sets $R \subseteq E$, events $e \in E$, stations $s \in S$ and wave types $e \in W$ with no restriction on the completeness of this data. Equation eq. (5), for each combination (R, e, s, w), defines an inverse problem with the $\Delta_R T_{esw}$ as the data and with $\mathbf{x}_e$ and t_e for each $e \in E$ and u_w for each $w \in W$ as the unknowns. The inverse problem is linear in the origin times t_e but nonlinear in the event hypocenters $\mathbf{x}_e$ and slowness functions u_w. We will treat the measurement errors as zero-mean, independent stochastic variables:

$$\mathrm{E}(n^{\text{obs}}_{Resw}) = 0 \tag{6}$$

$$\mathrm{Cov}(n^{\text{obs}}_{Resw} n^{\text{obs}}_{R'e's'w'}) = \begin{cases} N^{\text{obs}}_{Resw} & \text{if } R = R',\ e = e',\ s = s',\ w = w' \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Priors on Slowness and Modeling Errors

The usual practice in locating earthquakes has been to solve eq. (4) or eq. (5) for the source parameters $(\mathbf{x}_e, t_e)$ with the earth's slowness functions u_w assumed known. These functions are of course never known exactly, and errors in the assumed slowness models may bias the derived earthquake locations. The amount of bias depends on the magnitude and spatial dependence of the slowness errors, as well as the receiver distribution relative to the true event location. One way to account for incomplete knowledge of the u_w is to invert for these functions simultaneously with the hypocenters $\mathbf{x}_e$. An alternative is to invert only for the hypocenters but to account for the uncertainty in the u_w by admitting additional errors in the data (in addition to measurement errors). The statistics of these "modeling" errors are generated by treating the u_w as random

functions with known statistical properties. We outline this approach here using a linearization of the traveltime with respect to slowness.

Assume that u_w is a second-order random function with known mean and covariance operator, u_{0w} and C_w, respectively:

$$\mathrm{E}(u_w(\mathbf{x})) = u_{0w}(\mathbf{x}) \tag{8}$$

$$\mathrm{Cov}(u_w(\mathbf{x}), u_w(\mathbf{x}')) = C_w(\mathbf{x}, \mathbf{x}'). \tag{9}$$

Further assume that the u_w are uncorrelated with each other and with the observational errors:

$$\mathrm{Cov}(u_w(\mathbf{x}), u_{w'}(\mathbf{x}')) = 0, \quad w \neq w' \tag{10}$$

$$\mathrm{Cov}(u_w(\mathbf{x}), n^{\mathrm{obs}}_{Resw'}) = 0. \tag{11}$$

Eqs. (8)–eq. (11) embody a form of prior information about u_w.

If the covariances given by C_w are not too large, we may linearize the traveltime functional F in eq. (3) with respect to u_w around u_{0w}. We thus obtain the first order approximation (from Fermat's stationarity principle)

$$F(\mathbf{x}_e, \mathbf{x}_s, u_w) = F(\mathbf{x}_e, \mathbf{x}_s, u_{0w}) + \int_{\mathbf{x}\in\Gamma(\mathbf{x}_e,\mathbf{x}_s,u_{0w})} dl(\mathbf{x})\,(u_w(\mathbf{x}) - u_{0w}(\mathbf{x})). \tag{12}$$

Using eq. (12) we can rewrite eq. (5) as

$$\begin{aligned}\Delta_R T_{esw} = t_e &- \frac{1}{|R|}\sum_{r\in R} t_r \\ &+ F(\mathbf{x}_e, \mathbf{x}_s, u_{0w}) - \frac{1}{|R|}\sum_{r\in R} F(\mathbf{x}_r, \mathbf{x}_s, u_{0w}) \\ &+ \Delta_R n^{\mathrm{mod}}_{esw} + n^{\mathrm{obs}}_{Resw}.\end{aligned} \tag{13}$$

where

$$\Delta_R n^{\mathrm{mod}}_{esw} = n^{\mathrm{mod}}_{esw} - \frac{1}{|R|}\sum_{r\in R} n^{\mathrm{mod}}_{rsw} \tag{14}$$

and

$$n^{\mathrm{mod}}_{esw} = \int_{\mathbf{x}\in\Gamma(\mathbf{x}_e,\mathbf{x}_s,u_{0w})} dl(\mathbf{x})\,(u_w(\mathbf{x}) - u_{0w}(\mathbf{x})). \tag{15}$$

A simple example shows that the modeling error n^{mod} can be significant. Let the reference velocity be 5000 m/s, a typical value of P velocity in the shallow crust, and suppose the true velocity differs from this by 5% along a 100 m segment of one of the raypaths. For that raypath then the modeling error is 1 ms. This is of the same order

as the measurement errors we will encounter later with data from the Hot Dry Rock experiment.

From eq. (11) and eq. (15) we infer that the observational and modeling errors are uncorrelated:

$$\mathrm{Cov}(n^{\mathrm{mod}}_{esw}, n^{\mathrm{obs}}_{e's'w'} = 0. \tag{16}$$

Using eqs. (8)–eq. (10) we can infer the statistical moments of the (absolute) modeling error itself in eq. (15) are

$$\mathrm{E}(n^{\mathrm{mod}}_{esw}) = 0 \tag{17}$$

$$\mathrm{Cov}(n^{\mathrm{mod}}_{esw}, n^{\mathrm{mod}}_{e's'w'}) = \begin{cases} N^{\mathrm{mod}}_{ee'ss'w} & \text{if } w = w' \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

where the matrix N^{mod} is given by

$$N^{\mathrm{mod}}_{ee'ss'w} = \int_{\mathbf{x}\in\Gamma(\mathbf{x}_e,\mathbf{x}_s,u_{0w})} dl(\mathbf{x}) \int_{\mathbf{x}'\in\Gamma(\mathbf{x}_{e'},\mathbf{x}_{s'},u_{0w})} dl(\mathbf{x}')\, C_w(\mathbf{x},\mathbf{x}'). \tag{19}$$

This formula gives the covariance between the modeling errors in two raypaths: one from e to s and one from e' to s'. The degree of correlation between these errors will depend on the degree of correlation in the slowness function, as indicated by C_w, in relation to the proximity of the raypaths to one another.

Suppose that C_w has an effective correlation distance of λ, i.e.,

$$C_w(\mathbf{x},\mathbf{x}') \approx \begin{cases} (C_w(\mathbf{x},\mathbf{x})C_w(\mathbf{x}',\mathbf{x}'))^{1/2} & \text{if } |\mathbf{x}-\mathbf{x}'| \ll \lambda \\ 0 & \text{if } |\mathbf{x}-\mathbf{x}'| \gg \lambda. \end{cases}$$

Then the modeling errors for two raypaths separated by much more than λ will be uncorrelated while the errors for raypaths separated by much less than λ will be highly correlated. An instance of the latter would occur with an event cluster whose dimension was small compared to the correlation distance λ. Then the rays from two events in the cluster to a common station, s, will be separated by less than λ and the modeling errors for these two paths will be highly correlated:

$$N^{\mathrm{mod}}_{ee'ssw} \approx (N^{\mathrm{mod}}_{eessw} N^{\mathrm{mod}}_{e'e'ssw})^{1/2}. \tag{20}$$

To the extent that eq. (20) holds for all events e and e' in a reference set R, the modeling error in the differential arrival time $\Delta_R T_{esw}$ will be small. To see this, it follows from eq. (14) and eq. (18) that

$$\mathrm{Var}(\Delta_R n^{\mathrm{mod}}_{esw}) = N^{\mathrm{mod}}_{eessw} - \frac{2}{|R|}\sum_{r\in R} N^{\mathrm{mod}}_{erssw} + \frac{1}{|R|^2}\sum_{r,r'\in R} N^{\mathrm{mod}}_{rr'ssw}.$$

Since eq. (20) implies that $N^{\mathrm{mod}}_{rr'ssw}$ is approximately the same for all $r, r' \in R \cup \{e\}$ the terms on the right-hand side of the above equation will approximately cancel. A similar result was derived by Jordan and Sverdrup (1981) in the teleseismic location problem.

The formula for N^{mod} in eq. (19) presents a difficult computation and our current algorithm therefore does not use this formula. Instead it employs an *ad hoc* model of the modeling error variance that provides a similar qualitative behavior. That is we set

$$N^{\mathrm{mod}}_{ee'ss'w} = \mu(F(\mathbf{x}_e, \mathbf{x}_s, u_w)F(\mathbf{x}_{e'}, \mathbf{x}_{s'}, u_w))^{1/2} \exp\left\{-2\frac{|\mathbf{x}_e - \mathbf{x}_{e'}| + |\mathbf{x}_s - \mathbf{x}_{s'}|}{|\mathbf{x}_e - \mathbf{x}_s| + |\mathbf{x}_{e'} - \mathbf{x}_{s'}|}\right\} \quad (21)$$

where μ is a constant. What this implies is that the variance of the modeling error in a travel time is proportional to the calculated traveltime itself. Thus the modeling error is treated as a random walk in traveltime. The correlation between modeling errors for two raypaths is determined by the exponential function in eq. (21). What this function does is make the correlation depend on the proximity of the raypaths, as measured by how far apart their endpoints are from one another compared to the source-receiver separation.

Nonlinear Least Squares Estimation of Hypocenters

Equation eq. (13) defines a nonlinear inverse problem of the form

$$d = A(h) + Bt + n. \quad (22)$$

The data vector d contains the differential arrival times $\Delta_R T_{esw}$ for various events $e \in E$, stations $s \in S$, wave types $w \in W$ and reference event sets $R \subseteq E$. When $R = \emptyset$ (the empty set) $\Delta_R T_{esw}$ is an absolute arrival time. The model vector h contains the hypocentral parameters $\mathbf{x}_e$ for each event $e \in E$ while the second model vector t contains the event origin times. The noise vector n contains the sum of observational errors (n^{obs}_{Resw}) and modeling errors ($\Delta_R n^{\mathrm{mod}}_{esw}$). Finally, A is a nonlinear transformation obtained from the traveltime functionals F or differences between these functionals, while B is a linear transformation determined by the t_e and t_r terms in eq. (5).

The statistics of n can be summarized as

$$\mathrm{E}(n) = 0 \quad (23)$$

$$\mathrm{Var}(n) = N(h). \quad (24)$$

The variance matrix $N(h)$ is derivable from equations eq. (7), eq. (14), eq. (15), eq. (16) and eq. (18). The dependence on h obtains from the dependence of N^{mod} on the endpoints of the raypaths.

We solve eq. (22) in a least squares sense. Define the data misfit functional Ψ as

$$\Psi(h,t) = (d - A(h) - Bt)^T N(m)^{-1}(d - A(h) - Bt). \quad (25)$$

Then the model estimates, $\hat{h}$ and $\hat{t}$, are solutions to

$$\Psi(\hat{h},\hat{t}) = \min_{h,t} \Psi(h,t). \tag{26}$$

Since the data depend linearly on t, we can perform minimization with respect to t analytically and reduce eq. (26) to minimization over h only. For fixed h, minimization of Ψ with respect to t is achieved by

$$\hat{t}(h) = (B^T N(m)^{-1} B)^{-1} B^T N(m)^{-1} (d - A(h)).$$

Define the linear projector

$$Q = I - B(B^T N(h)^{-1} B)^{-1} B^T N(h)^{-1}$$

Then

$$\Psi(h,\hat{t}(h)) \equiv \Phi(h) = (d - A(h))^T Q^T N(m)^{-1} Q (d - A(h)) \tag{27}$$

and eq. (26) is equivalent to

$$\Phi(\hat{h}) = \min_{h} \Phi(h) \tag{28}$$

$$\hat{t} = (B^T N(m)^{-1} B)^{-1} B^T N(m)^{-1} (d - A(\hat{h})). \tag{29}$$

Our algorithm solves eq. (28) and eq. (29). It performs the minimization of Φ using the Polak-Ribiere variant of the conjugate gradient algorithm (see, e.g., Press et al., 1989).

Confidence Regions on Hypocenters

After finding $\hat{h}$ and $\hat{t}$ our multiple event location algorithm computes two types of confidence regions on the event hypocenters: a confidence region on each absolute hypocenter, $\mathbf{x}_e$, and a confidence region on each hypocenter relative to the cluster centroid, $\Delta_E \mathbf{x}_e$. These computations are based on a linearization of the traveltime functionals F at the least squares estimates, i.e.,

$$F(\mathbf{x}_e, \mathbf{x}_s, u_{0w}) = F(\hat{\mathbf{x}}_e, \mathbf{x}_s, u_{0w}) + \mathbf{p}(\hat{\mathbf{x}}_e, \mathbf{x}_s, u_{0w}) \cdot (\mathbf{x}_e - \hat{\mathbf{x}}_e). \tag{30}$$

The vector $\mathbf{p}(\hat{\mathbf{x}}_e, \mathbf{x}_s, u_{0w})$ is the slowness vector of the raypath $\Gamma(\hat{\mathbf{x}}_e, \mathbf{x}_s, u_{0w})$ at the endpoint $\mathbf{x} = \hat{\mathbf{x}}_e$. It points along the raypath at the reference hypocenter $\mathbf{x} = \hat{\mathbf{x}}_e$ away from the station. Its magnitude is the slowness at the reference hypocenter, $u_{0w}(\hat{\mathbf{x}}_e)$. In terms of the general inversion formulation, the linearization with respect to the event hypocenters can be written

$$A(h) = A(\hat{h}) + DA|_{\hat{h}}(h - \hat{h}) \tag{31}$$

where $DA|_{\hat{h}}$ denotes the Frechet derivative of A evaluated at $h = \hat{h}$.

A confidence region on h based on eq. (31) is an ellipsoid with center at the least squares solution. The equation for the confidence region is

$$(h - \hat{h})^T \operatorname{Var}(\hat{h})^{-1} (h - \hat{h}) = \text{ const.} \tag{32}$$

We set the constant on the right-hand side as outlined by Jordan and Sverdrup (1981) based on the assumption of normally distributed noise (n) and the use of a prior distribution on the noise variance. The variance of the solution, $\operatorname{Var}(\hat{h})$, is given by

$$\operatorname{Var}(\hat{h}) = \left(DA|_{\hat{h}})^T Q^T N(\hat{h})^{-1} Q DA|_{\hat{h}} \right)^{-1}.$$

A confidence region on any linear transformation of $\hat{h}$ of the form Lh is obtained in analogy with eq. (32):

$$(Lh - L\hat{h})^T \operatorname{Var}(L\hat{h})^{-1} (Lh - L\hat{h}) = \text{ const} \tag{33}$$

where

$$\operatorname{Var}(L\hat{h}) = L \operatorname{Var}(\hat{h}) L^T.$$

With appropriate choices of L we thus obtain confidence intervals on $\mathbf{x}_e$ and $\Delta_E \mathbf{x}_e$.

CORRELATION ANALYSIS FOR MEASURING DIFFERENTIAL ARRIVAL TIMES

It has been extensively observed that seismic waveforms generated from closely spaced earthquakes generally have almost identical signal character at a given station when the spatial separations between the events are much shorter than the distance from the events to the observing station (e.g., Geller and Mueller, 1980; Frankel, 1982; Pechmann and Kanamori, 1982; Tsujiura, 1983; Poupinet et al., 1984; Ito, 1985; 1990; Motoya and Abe, 1985; Logan et al., 1985; Scherbaum and Wendler, 1986; Spudich and Bostwick, 1987; Console and Giovambattista, 1987; Fremont and Malone, 1987; Thorbjarnardottir and Pechmann, 1987; James and Savage, 1990; Pechmann and Thorbjarnardottir, 1990; Deichmann and Garcia-Fernandez, 1992; Phillips et al., 1992; Gupta and Davis, 1992). The striking similarity among the waveforms from the similar events of "an earthquake family" (Tsujiura, 1983) can be attributed to the events having similar hypocenter locations and focal mechanisms and to the seismic waves passing through almost identical propagation paths. Seismologists generally use a cross-correlation analysis technique either in the frequency domain (e.g., Poupinet et al., 1984; Ito, 1985) or in the time domain (e.g., Pechmann and Kanamori, 1982; Frankel, 1982) to quantitatively characterize the degree of the similarity of seismic waveforms from a cluster of earthquakes

close in space and to read the arrival times of P and S waves in an accurate and objective manner.

The waveform correlation analysis for waveforms of two similar events in the frequency domain was first applied by Nakamura (1978) to analyze data of moonquakes. The method has been discussed in detail by Poupinet et al. (1984) and Ito (1985). Their methods involve calculating cross-spectra of a moving window. Let us use two real time series $x(t)$ and $y(t)$ to represent the seismograms of a reference (or master) event and a slave event, respectively, truncated to some time window $[t_1, t_2]$. The cross-spectrum of x and y is defined by

$$S_{xy}(f) = X(f)Y^*(f)$$

where f is temporal frequency, $X(f)$ and $Y(f)$ are the Fourier transforms of $x(t)$ and $y(t)$, respectively, and * denotes complex conjugate. The coherence between $x(t)$ and $y(t)$ is defined by

$$C_{xy}(f) = \frac{|S_{xy}(f)|}{|S_{xx}(f)|^{1/2}|S_{yy}(f)|^{1/2}}$$

and is a measure of the similarity between $x(t)$ and $y(t)$. The coherence ranges from zero to one, with one indicating total similarity and zero indicating no similarity between waveforms. Given that the waveforms $x(t)$ and $y(t)$ are similar, the delay time between them, τ_{xy}, can be obtained by fitting a straight line through the phase spectrum of S_{xy} with zero intercept. The slope of this line is equated to $2\pi\tau_{xy}$. A weighting function is generally used in fitting the phase spectrum (Poupinet et al., 1984; Ito, 1985; Gupta and Davis, 1992).

The cross-spectral analysis method has been widely applied to obtain accurate relative locations for moonquakes (Nakamura, 1978), microearthquake doublets along the San Andreas Fault in California (Poupinet et al., 1984), microearthquake swarms and clusters in Japan (Ito, 1985, 1990), in Germany (Scherbaum and Wendler, 1986), in Morgan Hill, California (Spudich and Bostwick, 1987), for earthquake swarms associated with volcanic activity at Mount St. Helens, Washington (Fremont and Malone, 1987) and at Kilauea, Hawaii (James and Savage, 1990), for underground nuclear explosions at Nevada Test Site (Gupta and Davis, 1992), and for a group of acoustic emission doublets at a geothermal site in Japan (Moriya et al., 1993).

Both Poupinet et al. (1984) and Ito (1990) claimed that the relative hypocenter locations of doublets or microearthquakes of a swarm can be determined with a spatial resolution of the order of 10 meters with a local seismic network. Fremont and Malone (1987) applied the cross-spectral analysis technique to 23 explosions with very precisely known locations to verify this relative location method. Using waveform data recorded at seismic stations of a local network with typical epicentral distances of about several tens of kilometers, they found that the explosions can be relocated with an accuracy

better than 20 meters for events within 250 meters to the reference explosion. Then they applied the method to an earthquake swarm of 40 events in the Mount St. Helens area and concluded that the cluster of events were located within a volume, approximately 30 m in diameter, beneath a lava dome. Cross-spectral analysis was also applied to regional seismic waveform data from Yucca Flat explosions with precisely known locations, showing that the relative relocations of the events are accurate to within about 1 km. Moriya et al. (1993) constructed a cross-spectral matrix using three-component seismic waveforms recorded by one triaxial geophone to obtain the time delay and differences of azimuths and inclinations between two events of a doublet. They relocate a group of acoustic emission doublets about 150 m away from the seismometer with an accuracy of about 1 meter, suggesting that three component data can provide additional information for a high precision relative location procedure.

An alternative way to characterize the waveform similarity of an earthquake cluster is to perform the cross-correlation analysis in the time domain. A time domain approach was used by Pechmann and Kanamori (1982) for preshocks and aftershocks of the 1979 Imperial valley, California, $M_s = 6.6$ earthquake, and by Frankel (1982) for foreshocks of a $M = 4.8$ event in the Virgin Islands. To summarize their method, let two real time series $x(t)$ and $y(t)$ denote the seismograms for a master event and a slave event, respectively. The two time series are assumed to have a mean value of zero. For a given time window $[t_1, t_2]$ we can calculate the cross-covariance function between x and y as

$$s_{xy}(t) = \int_{t_1}^{t_2-t} x(t')y(t'+t)\,dt\,.$$

Normalizing, we obtain the cross-correlation function

$$c_{xy}(t) = \frac{s_{xy}(t)}{(s_{xx}(0)s_{yy}(0))^{1/2}}.$$

The values of c_{xy} range from -1 to 1. When $c_{xy}(t) = 1$ then x and y have identical waveform shape except for a constant scaling and time shift t, i.e., $y(t') = \text{const}\ x(t'-t)$. In general, the time of the maximum peak of $c_{xy}(t)$ will reflect the time lag between the two signals within the specified time window $[t_1, t_2]$. The time domain analysis is somewhat simpler than frequency domain analysis since Fourier transforms are avoided as are phase unwrapping and fitting.

Many researchers have used the time domain cross-correlation analysis technique to characterize the waveform similarity of a cluster of similar earthquakes and to obtain better relative locations for the microearthquake clusters in California (Pechmann and Kanamori, 1982), in the Virgin Islands (Frankel, 1982), in central Italy (Console and Di Giovambattista, 1987), in Utah (Pechmann and Thorbjarnardottir, 1990), and in northern Switzerland and in many other European countries (see Deichmann and Garcia-Fernandez, 1992 for a review). The time domain correlation method is also used to obtain better relocations for mining blasts in Utah (Thorbjarnardottir and

Pechmann, 1987), in central Sweden (Israelsson, 1990), and in Estonia, in Karelian, and on the Kolo Peninsula (Riviere-Barbier and Grant, 1992), as well as for hydrofracturing events in New Mexico (Phillips et al., 1992).

Frankel (1982) applied the technique to waveform data of several pairs of microearthquakes collected from a local seismic network in the vicinity of the Virgin Islands and found that the waveforms of the events in each pair had high correlation (larger than 0.7) and that the earthquakes in each pair were located within about 110 m of each other. Using a similar method, Pechmann and Kanamori (1982) calculated cross-correlation functions for all possible pairs within a cluster of earthquakes in California and constructed cross-correlation matrixes with all possible correlation pairs, which were determined using filtered and unfiltered waveforms, to analyze the similarity between the clustered events. Their results indicated that an average peak correlation between the waveforms of eight preshocks in a volume with radius of about 300 m was about 0.74, much larger than that (0.23) of five aftershocks in the same area. Thorbjarnardottir and Pechmann (1987) tested the method by applying it to seismograms from 14 mining blasts with known locations, which are recorded at four stations of a local seismic network in the vicinity of Salt Lake City, Utah. Based on the so called quarter wavelength hypothesis (Geller and Mueller, 1980), Thorbjarnardottir and Pechmann (1987) estimated the separations between three blasts with high correlation coefficients to be about 100 meters, in good agreement with their true locations. Pechmann and Thorbjarnardottir (1990) used this method to study two earthquake clusters and concluded that foreshock clusters were concentrated within a very small volume of less than 80 m in extent while aftershock clusters were located within larger volumes of several hundred meters in diameter. Deichmann and Garcia-Fernandez (1992) analyzed waveform correlation for the seismograms from two microearthquake swarms, consisting of 83 events, recorded at a local network in northern Switzerland and obtained relative hypocentral locations for the events with an accuracy of about 20–30 meters. Using these locations they revealed rupture processes along the fault planes with dimension of several hundred meters.

Israelsson (1990) used the waveform correlation technique for seismograms of 137 regional events recorded at the NORESS array to relocate the events with high precision. He measured quantitatively the waveform similarity based on both relative amplitude and polarization using three-component data. Relative relocation results indicated that all events concentrated within an area of 4x4 km. The location errors range from 10 to 500 meters.

Table 1: Seismic Stations in the Los Alamos Injection Experiment

Station	North (m)	East (m)	Depth (m)
EE-1	-480.82	-562.26	2854.6
EE-3	-182.44	-234.42	3301.73
GT-1	1976.63	-229.82	816.86
PC-1	-954.44	613.53	570.59

LOCATION OF A MICROEARTHQUAKE CLUSTER AT THE LOS ALAMOS HOT DRY ROCK GEOTHERMAL SITE

Background

The Los Alamos Hot Dry Rock (HDR) geothermal reservoir is located at Fenton Hill, New Mexico near the boundary between the Colorado Plateau and the Rio Grande Rift (see Figure 1). Several hydraulic fracturing experiments were conducted in the reservoir to create fractures between two deep wells. In these experiments a huge volume of water was injected into boreholes under a high pumping pressure of about 48 MPa. During each injection experiment, seismometers in the nearby boreholes were used to monitor the microearthquakes induced by the injection. During one of these experiments, Experiment 2032 in December 1983, approximately 21,600 m^3 of water was injected into a wellbore at a depth of 3463 meters below the surface. During the 61 hours of water injection, more than 10,000 microearthquakes induced by hydraulic fracturing were recorded by a few borehole seismic stations. The seismic data recorded during Experiment 2032 have been analyzed extensively by scientists at the Los Alamos National Laboratory (Fehler et al., 1987; House 1987; Fehler and Phillips, 1991; Phillips et al., 1992) and at M.I.T. (Block 1991; Block et al., 1993).

We obtained waveform data from Los Alamos National Laboratory for 19 microearthquakes recorded during Experiment 2032. The events occurred during the time period between 8:00 a.m. and 9:30 p.m. on 7 December 1983. Waveforms from these events were obtained from four seismic stations, labeled EE-1, EE-2, GT-1 and PC-1. In this section we show the results of applying our data processing and multiple event location algorithms to the data from these 19 events.

Measurement of Arrival Times

The geometry of the four borehole seismic stations is shown in Figure 2. The station coordinates in a Cartesian coordinate system are listed in Table 1 (see Block, 1991, for

details). The depths are referenced to about 2650 meters above sea level. Stations EE-1, EE-3 and GT-1 are situated at depths greater than 700 meters below the surface within the Precambrian basement rock. Station PC-1 was located about 150 to 200 meters above the basement within a cavernous limestone of the Magdalena Group. Extremely few S waves could be observed at station PC-1, probably due to high attenuation in the metamorphic and sedimentary rock that exists above the basement rock.

Fehler and Phillips (1991) discuss the characteristics of the waveform data recorded during Experiment 2032. The data were recorded on an analog tape recorder and later digitized at a sampling rate of 5000 samples per second. Spectral analysis of the recorded waveforms indicated that corner frequencies for typical microearthquakes are about 400 Hz, corresponding to wavelengths of 16 m and 9 m for P and S waves, respectively. Stations EE-1 and EE-3 were equipped with triaxial geophones but stations GT-1 and PC-1 only had vertical component seismometers. In this study, we analyzed vertical component waveforms only. The best waveform data were recorded at stations EE-1 and GT-1. Seismic waveforms recorded at station EE-3 generally were clipped for larger events, especially for S waves, owing to its small distance from the hydrofracture zone. In contrast, only the larger microearthquakes could be recorded at station PC-1. Because of these limitations, the data contain usable P and S signals at all four stations only for eight of the 19 events analyzed. The remaining 11 events yielded signals only at three stations, most with both P and S waves but some with only one of the two phases. Table 2 shows station-by-station which phases were picked to obtain absolute arrival times for each event.

The top frame of Figure 3 shows vertical component waveforms from six induced microearthquakes recorded at station EE-1. The time axes shown are referenced to a calculated origin time determined from preliminary picks of P and S arrival times and a velocity model with $V_p = 5920$ and $V_s = 3500$ m/s (House, 1987). We see from Figure 3 that the P and S waveforms show a striking similarity from event to event, suggesting that the hypocentral locations and focal mechanisms of the events are very similar. The similarity of the waveforms is more clearly seen in the bottom frame of Figure 3, in which the seismograms were filtered using a fourth-order Butterworth lowpass filter with cutoff frequency of 500 Hz. The lowpass filtering reduces high frequency noise significantly. Unfiltered and filtered waveforms of the same events recorded at station GT-1 are displayed in Figure 4. Here we also see great similarity in the waveforms from different events. However, comparison of Figures 3 and 4 shows that the waveforms from any given event at the two different stations are quite dissimilar, reflecting quite different propagation effects along different paths.

In this study we located the 19 events based on a combination of absolute and differential arrival times of the observable P and S phases at the four stations. Absolute arrival times were obtained by visually picking first motion from the waveforms. Differential times were obtained by cross-correlation analysis in the time domain, as described

Table 2: Coverage of Absolute Arrival Times

	Phases picked (by station)			
Event	EE-1	EE-3	GT-1	PC-1
2499	P S	P S	P S	P S
2583	P S	P S	P S	—
2718	P S	P S	P S	—
2727	P S	P S	P S	P S
2728	P S	P S	P S	P S
2732	P S	P S	P S	—
2748	P S	P S	P S	P S
2753	P S	P S	P S	—
2754	P S	P S	P S	—
2755	P S	P S	P S	—
2764	P S	P S	S	—
2765	P S	P S	P S	—
2770	P S	P S	P S	P S
2774	P S	P S	P S	P S
2782	P S	P S	P S	P S
2787	P S	P S	P S	—
2788	P S	P S	P S	P S
2793	P S	P S	P S	—
2794	P S	P S	P S	—

in an earlier section. However, for severely clipped waveforms at station EE-3 we resorted to subtracting first motion picks as a way of augmenting the data set of differential times.

The cross-correlation method determines arrival time differences between a pair of events. Given a cluster of 19 events there are 171 possible event pairs. We determined differential arrival times for a subset of 28 of the possible pairs, i.e., Event 2728 paired with each of the other 18 events, and Event 2583 paired with 10 other events. The two reference events chosen, 2728 and 2583, had good quality P and S waveforms at at least three stations.

An example of the data processing for differential arrival times using Event 2728 as a reference is shown in Figure 5. The top two traces are the seismograms at station EE-1 for the reference event (2728, top trace) and another event (2753, second trace). From the 2728 seismogram, a 30 ms segment including the P wave was chosen and cross-correlated with each complete seismogram. The resulting correlation functions yield the third (for 2728) and fourth (for 2753) traces in the figure. The first wave packet in the third trace (labeled PA-11) is essentially the auto-correlation of the P wave from

event 2728, while the first wave packet of the fourth trace (labeled PC-12) is the cross-correlation between the P wave from event 2728 and the P wave from event 2753. The time difference between these packets, as measured from the maximum peak, yields the differential P wave arrival time between the two events. The analogous results for the S wave are shown in the bottom two traces (labeled SA-11 and SC-12) of Figure 5. In this case, the second wave packets (i.e., in the S wave window) are used to obtain the differential arrival time for S. In this example, the cross-correlation coefficients for the P and S wave windows are 0.85 and 0.89, respectively, indicating relatively high similarity between the waveforms of events 2728 and 2753. In this example the correlation analysis provides the differential P arrival time with high accuracy, perhaps 0.2 ms. The time difference between the two peaks of the correlation functions is about 3.8 ms.

Bandpass filters are often used in time-domain waveform correlation analysis for increasing the signal-to-noise ratio and for determining in which frequency band the signals are best correlated (e.g., Pechmann and Kanamori, 1982; Thorbjarnardottir and Pechmann, 1987; Israelsson, 1990; Pechmann and Thorbjarnardottir, 1990). However, in these studies the authors were mainly concerned with the effect of filtering on the correlation values but did not pay attention to its effect on the reading of differential arrival times. We now examine this latter effect in our example. The top frame of Figure 6 repeats the observed seismograms and P wave correlations from Figure 5 (traces 1–4) on an expanded time scale in a 50 ms window around the P wave. The four traces of the bottom frame of Figure 6 show the seismograms after applying a lowpass filter and then the correlation functions derived from the filtered seismograms. The filter used was a fourth-order Butterworth filter with a high cutoff frequency of 500 Hz. Comparing the top and bottom frames it is found that the filter causes a time delay of about 1 ms, but the time difference between the peak of the two correlation functions stays the same (3.8 ms) since the seismograms from both events experience the same filter delay. We also note that the cross-correlation coefficient calculated with the filtered waveforms is higher: 0.96 vs. 0.85. This shows that lowpass filtering reduces the incoherent noise significantly.

Figure 7 shows the analogous results as Figure 6 for the S wave. That is, traces 1, 2, 5 and 6 of Figure 5 are now expanded and shown within a 50 ms window around the S wave. The unfiltered (top frame) and lowpass filtered (bottom frame) results are compared. Again, the lowpass filtering is seen to cause a 1 ms time shift and to increase the S wave correlation coefficient (from 0.89 to 0.96). The filtered and unfiltered results, however, yield the same differential arrival time of 4.6 ms. We attribute the equal phase shift for filtered P and S waves to the fact that there is no significant difference between the corner frequencies of P and S waves (Fehler and Phillips, 1991) for these earthquakes.

In the previous calculation of the correlation functions, we used a single window length of 30 ms. Next we will discuss the effects of varying the window length. Figure 8 compares the cross-correlation functions between the P waves of Events 2728 and 2753

at station EE-1 obtained with five different time window lengths: 10, 20 , 30 , 40 and 50 ms. We see that the peaks of the correlation functions coincide almost exactly in time. Figure 9 shows the same comparison for the S wave correlations and again we see that the time of maximum correlation is the same for each window length. We can conclude that differential arrival times estimated with the time domain correlation method are robust with respect to the choice of window length. The correlation coefficient (peak value of the correlation function) does show some dependence on window length, however. This is illustrated in Figure 10 where the correlation coefficients extracted from Figures 8 and 9 are plotted as a function of window length. We see that the correlation coefficient decreases with window length, more so for P than for S.

We assess the accuracy of the differential arrival times picked from correlation functions based on the width of the main correlation peak. For the data processed in this study the average width between the two 3 dB points of the P wave correlation peaks was approximately 0.8 ms. For S the average is about 1.2 ms. The error in picking the correlation peaks we estimate to be no more than half of these 3 dB widths of the main correlation peak. The standard errors we actually assigned the differential arrival time estimates are 0.4 ms for P waves and 0.8 ms for S waves.

Inversion for Hypocenters

We applied our multiple event location algorithm, described in an earlier section, to a data set of absolute and differential arrival times measured for the 19 events and four stations used in this study. The data set comprises the absolute times for 2 events, Events 2728 and 2583, and the differential arrival times from some or all of the other events relative to these two reference events. Differential times from each of the other 18 events relative to 2728 were used, while times from only 10 events relative to 2583 were used. Table 3 summarizes the data set of absolute and differential arrival times. A reference event shown as "—" implies that the absolute arrival times for the event are included. For comparison, locations were also obtained from a data set containing no differential times but containing absolute arrival times from every event. Table 2 showed the coverage of absolute arrival time data by station and phase.

Table 4 lists the traveltime corrections by station and phase which we applied to the arrival time data. These corrections were obtained from W. S. Phillips (private communication) and include a recently discovered tape head delay of 11.5 ms at station PC-1. The corrections shown in the table were *subtracted* from the observed absolute arrival times. The differential times require no station corrections.

The velocity model used in the location algorithm was a homogeneous model with a P velocity of 5920 m/s and S velocity of 3500 m/s (House, 1987). The variance matrix of the modeling errors, due to errors in the velocity model, was computed according to

Table 3: Absolute/Differential Arrival Time Data Set

Event	Reference events
2499	2728
2583	—, 2728
2718	2728, 2583
2727	2728
2728	—
2732	2728, 2583
2748	2728
2753	2728, 2583
2754	2728, 2583
2755	2728, 2583
2764	2728, 2583
2765	2728, 2583
2770	2728
2774	2728
2787	2728, 2583
2788	2728
2793	2728, 2583
2794	2728, 2583

equation eq. (21). Two values of the constant μ were tried: $\mu = 0$, implying no modeling error, and $\mu = 10^{-5}$. The latter value of μ implies standard deviations of the modeling error varying between about 1 and 3 ms. Table 5 shows the modeling error standard deviation as a function of station and phase as determined by the final location of Event 2728. It varies only slightly from event to event. We note that the numbers shown are the standard deviation of the modeling error in an *absolute* arrival time. Since these errors are correlated from event to event, the modeling errors in differential arrival times are smaller (typically 0.1–0.5 ms). However, the modeling errors in the differential times

Table 4: Traveltime Corrections for the HDR Seismic Stations

	Traveltime correction (ms)	
Station	P	S
EE-1	0.6	2.3
EE-3	-0.5	-1.5
GT-1	-0.5	10.9
PC-1	28.9	44.3

Table 5: Standard Deviations of Modeling Errors With $\mu = 10^{-5}$

	Standard deviation (ms)	
Station	P	S
EE-1	0.9	1.1
EE-3	0.6	0.8
GT-1	2.4	3.1
PC-1	2.2	2.9

Table 6: Standard Deviations of Observational Errors

Type of	Standard deviation (ms)	
Datum	P	S
Absolute	1.0	1.8
Differential	0.4	0.8

are not insignificant because the raypaths from different events to a close station such as EE-3 can be significantly different. In Table 6 we show, for comparison, the typical standard deviations assigned to the observational errors. These are the values assumed when the data quality was good or average. Larger standard deviations were assigned to a few poor quality readings.

Prior to inverting the differential arrival time data, we determined the location of each event using only absolute arrival times. In doing this exercise we realized that the data for eleven of the events have a degeneracy owing to the fact that arrival times are available at only three of the four stations: EE-1, EE-3 and GT-1 (see Table 2). Arrival time data from three stations can be fit equally well by hypocentral locations on *either* side of the plane that passes through the three station locations. Solutions on both sides of this symmetry plane achieve exactly the same minimum value of the data misfit functional (Φ in eq. (27)). To make matters worse, a point on the symmetry plane is a saddle point of Φ, which the inversion algorithm might find in lieu of a minimum.

This degeneracy is illustrated in Figure 11, where two sets of locations determined from the absolute arrival times are shown in map view. The solid circles are the locations obtained when each event is given an initial guess location of -250 m north, -500 m east and 3200 m depth. The open triangles are the locations obtained when the initial guess is -400 m north, -250 m east, 3100 m depth. (Both inversions assumed zero modeling error, $\mu = 0$.) The different initial guesses yield very different hypocenter solutions for 9 of the events which are constrained by only three stations. The locations of the other 2 of these underconstrained events agree but may be saddle point solutions. The remaining 8 events, which are constrained by four stations, have the same locations in

Table 7: Final Locations of 19 Microearthquakes at Los Alamos HDR Site

Event	North (m)	East (m)	Depth (m)
2499	-320	-384	3291
2583	-331	-404	3243
2718	-318	-424	3247
2727	-331	-402	3250
2728	-330	-405	3246
2732	-325	-406	3238
2748	-308	-460	3273
2753	-322	-426	3251
2754	-327	-410	3244
2755	-358	-386	3223
2764	-333	-405	3241
2765	-309	-432	3251
2770	-366	-375	3299
2774	-332	-403	3246
2782	-324	-409	3240
2787	-320	-421	3233
2788	-332	-407	3244
2793	-320	-419	3233
2794	-331	-406	3228

both solutions.

Figure 12 shows the results of inverting the data set comprising absolute and differential arrival times, i.e., absolute arrival times for two events and differential arrival times from the other events referenced to these two events, as defined in Table 3. Two solutions are shown: one with the modeling error variance scale parameter μ set to zero (open triangles) and the other with $\mu = 10^{-5}$ (filled circles). The degeneracy in the data discussed above was resolved by using an initial guess location that led to final locations defining the smaller cluster (i.e., the initial guess yielding the filled circles in Figure 11 was used.) We see that the relative pattern of locations is very similar for the two solutions and that μ affects primarily the absolute location of the event cluster.

Figures 13–15 show orthogonal views of our final locations obtained by inverting the data set of absolute and differential arrival times. The solution shown is that obtained with $\mu = 10^{-5}$. The event locations are also tabulated in Table 7. We see that 15 of the 19 events define a small cluster with dimension of order 30 m. The map view (Figure 13) shows that the cluster delineates an approximate vertical plane striking in the northwest-southeast direction.

Table 8: Absolute Location Errors of Los Alamos HDR Events

	Epicenter confidence ellipse			
Event	Strike (° E of N	Major semi-axis (m)	Minor semi-axis (m)	Depth error (m)
2499	-60	31	22	25
2583	-68	36	21	23
2718	-73	37	22	26
2727	-65	31	23	22
2728	-65	30	21	21
2732	-71	44	22	27
2748	-77	31	23	26
2753	-72	36	22	25
2754	-70	40	22	26
2755	-67	112	22	44
2764	-68	47	22	27
2765	-76	35	22	26
2770	-61	31	23	26
2774	-65	31	22	22
2782	-67	31	22	22
2787	-74	42	22	27
2788	-66	31	23	22
2793	-74	42	22	27
2794	-71	52	22	29

Figure 16 compares, in map view, the final solution from absolute and differential times (Figure 13) with the locations determined from absolute arrival times alone. The latter, shown as open triangles, were obtained with $\mu = 10^{-5}$. We see that the use of differential arrival time data leads to a somewhat tighter cluster of locations.

Finally, Tables 8 and 9 list the parameters of the 90% confidence regions for the final event locations of Figures 13–15. Table 8 shows the confidence region on the absolute location of each event while Table 9 shows the confidence region on the location of each event relative to the centroid of the 19 event cluster. The errors in relative location, particularly the minor axis of the epicenter ellipse, are significantly smaller than the absolute location errors. This stems from the fact that modeling errors have less influence on the relative locations.

Table 9: Relative Location Errors of Los Alamos HDR Events

	Epicenter confidence ellipse			
Event	Strike (° E of N	Major semi-axis (m)	Minor semi-axis (m)	Depth error (m)
2499	-71	15	8	10
2583	-70	15	3	7
2718	-71	25	5	13
2727	-70	14	7	8
2728	-70	11	2	5
2732	-70	29	5	14
2748	-75	15	9	10
2753	-71	24	6	13
2754	-70	27	5	13
2755	-67	92	7	32
2764	-68	33	5	15
2765	-72	24	6	13
2770	-70	16	9	11
2774	-70	14	7	8
2782	-71	14	7	8
2787	-71	27	5	14
2788	-70	14	7	8
2793	-71	28	5	14
2794	-70	36	5	16

CONCLUSIONS

We have developed an algorithm for simultaneously locating earthquakes in a cluster from a data set comprising absolute arrival times of various phases to a seismic network and differential arrival times between events. A theoretical analysis shows that differential times are less sensitive to errors in the velocity model used in the location algorithm, and as a consequence that the relative locations of events within a cluster can be determined more accurately than their absolute locations. Furthermore, with waveform correlation techniques it is possible to obtain highly accurate measurements of differential arrival times.

We applied these methods to a set of data from a hydraulic fracturing experiment at the Los Alamos geothermal site in Fenton Hill, New Mexico. We processed waveform data from 19 microearthquakes recorded by four downhole stations to obtain a data set of absolute and differential arrival times. In applying our location algorithm to these data we encountered certain deficiencies in the ability of arrival data from small

networks to constrain event locations in three dimensions. These deficiencies would be remedied with additional stations or the incorporation of directional information from three-component stations.

The hypocenter locations obtained for the 19 events at Fenton Hill define a roughly planar cluster of about 30 m dimension, consistent with the events occurring on a vertical fracture. Confidence region calculations show that the relative locations of the events are constrained more than their absolute locations, as expected. The results demonstrate that the use of differential arrival time data yields a more accurate image of the microearthquake cluster than is detemined by absolute times alone. However, the improvement attributable to differential arrival time data in this example was not dramatic. One reason for this is that the absolute arrival times could be picked quite accurately from the waveform data and thus are only about a factor of two less accurate than the differential times determined with correlation analysis. A second reason is that the event location errors are controlled by poor station geometry as much as picking accuracy and velocity model accuracy, and the effects of geometry are the same for absolute and differential times.

ACKNOWLEDGEMENTS

We thank Dr. Scott Phillips and Dr. Mike Fehler for providing the seismic waveform data and other information needed to obtain our microearthquake results for the Los Alamos geothermal site

REFERENCES

Block, L.V., 1991, Joint hypocenter-velocity inversion of local earthquake arrival time data in geothermal regions, Ph.D. thesis, Massachusetts Institute of technology.

Block, L.V., C.H. Cheng, M.C. Fehler, and W.S. Phillips, 1993, Seismic imaging using microearthquakes induced by hydraulic fracturing, *Geophysics*, in press.

Console, R., and R. Di Giovambattista, 1987, Local earthquake relative location by digital records, *Phys. Earth Planet. Inter., 47,* 43–49.

Deichmann, N., and M. Garcia-Fernandez, 1992, Rupture geometry from high-precise relative hypocenter locations of microearthquake clusters, *Geophys. J. Int., 110,* 501–517.

Fehler, M., L. House and H. Kaieda, 1987, Determining planes along which earthquakes occur: method and application to earthquakes accompanying hydraulic fracturing,

J. Geophys. Res., 92, 9407–9414.

Fehler, M. and W.S. Phillips, 1991, Simultaneous inversion for Q and source parameters of microearthquakes accompanying hydraulic fracturing in granitic rock, *Bull. Seism. Soc. Am., 81,* 553–575.

Frankel, A., 1982, Precursors to a magnitude 4.8 earthquake in the Virgin Islands: spatial clustering of small earthquakes, anomalous focal mechanisms and earthquakes doublets, *Bull. Seism. Soc. Am., 72,* 1277–1294.

Fremont, M.J., and S.D. Malone, 1987, High precision relative locations of earthquakes at Mount St. Helens, Washington, *J. Geophys. Res., 92,* 10223–1023.

Geller, R.J., and C.S. Meuller, 1980, Four similar earthquakes in central California, *Geophys. Res. Lett., 7,* 821–824.

Gupta, I.N., and J.P. Davis, 1992, High-precision relative event location with cross-spectral analysis and related studies, *Proceedings of the 14th PL/DARPA Seismic Symposium,* 146–152.

Israelsson, H., 1990, Correlation of waveforms from closely spaced regional events, *Bull. Seism. Soc. Am., 80,* 2177–2193.

Ito, A., 1985, High resolution relative hypocenters of similar earthquakes by cross-spectral analysis method, *J. Phys. Earth, 33,* 279–294.

Ito, A., 1990, Earthquake swarm activity revealed from high-resolution relative hypocenters - clustering of microearthquakes, *Tectonophysics, 175,* 47–66.

James, D.E., and M.K. Savage, 1990, A search for seismic reflections from the top of the oceanic crust beneath Hawaii, *Bull. Seism. Soc. Am., 80,* 675–701.

Jordan, T.H., and K.A. Sverdrup, 1981, Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific, *Bull. Seism. Soc. Am., 71,* 1105–1130.

House, L., 1987, Locating microearthquakes induced by hydraulic fracturing in crystalline rock, *Geophys. Res. Lett., 14,* 919–921.

Logan, A., R. Evans, J. Lovell, and S. Crampin, 1985. Fault structure from accurate relocation of similar earthquakes, *EOS, Trans. AGU, 66,* 854.

Moriya, H., K. Nagano, and H. Niitsuma, 1993. Precise source location of AE doublets by spectral matrix analysis of triaxial hodogram, submitted to *Geophysics.*

Motoya, Y., and K. Abe, 1985, Waveform similarity among foreshocks and aftershocks of October 18, 1981, Eniwa, Hokkaido, earthquake, *Earthquake Predict. Res., 3,*

349–375.

Nakamura, Y., 1978, A1 moonquakes: Source distribution and mechanism, *Proc. Lunar Planet. Sci. Conf., 9,* 3589–3607.

Pechmann, J.C., and H. Kanamori, 1982, Waveforms and spectra of preshocks and aftershocks of the 1979 Imperial Valley, California, earthquake: Evidence of fault heterogeneity? *J. Geophys. Res., 87,* 10579-10597.

Pechmann, J.C., and B.S. Thorbjarnardottir, 1990, Waveform analysis of two preshock-mainshock-aftershock sequences in Utah, *Bull. Seism. Soc. Am., 80,* 519–550

Phillips, W.S., L.S. House, and M.C. Fehler, 1992, Vp/Vs and the structure of microearthquake clusters, *Seismol. Res. Lett., 63,* 56–57.

Poupinet, G., W.L. Ellsworth, and J. Frechet, 1984, Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras fault, California, *J. Geophys. Res., 89,* 5719–5731.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, 1989, *Numerical Recipes (FORTRAN Version)*, Cambridge University Press, Cambridge, U. K.

Riviere-Barbier, F., and L.T. Grant, 1992, Cluster analysis closely spaced mining blasts as a method of event location, Phillips Laboratory Final Tech. Rep. PL-TR-92-2006, Hanscom Air Force Base, MA.

Scherbaum, F., and J. Wendler, 1986, Cross spectral analysis of Swabian Jura (SW Germany) three-component microearthquake recordings, *J. Geophys., 60,* 157–166.

Spudich, P., and T. Bostwick, 1987, Studies of the seismic coda using an earthquake cluster as a deeply buried seismograph array, *J. Geophys. Res., 92,* 10526–10546.

Thorbjarnardottir, B.S., and J.C. Pechmann, 1987. Constraint on relative earthquake locations from cross-correlation of waveforms, *Bull. Seism. Soc. Am., 77,* 1626–1634.

Tsujiura, M., 1983, Characteristic frequencies for earthquake families and their tectonic implications: evidence from earthquake swarms in the Kanto District Japan, *Pure Appl. Geophys., 4,* 573–600.

Vinegar, H.J., P.B. Wills, D.C. DeMartini, J. Shlyapobersky, W.F.J. Deeg, R.G. Adair, J.C. Woerpel, J.E. Fix, and G.G. Sorrells, 1992, Active and passive seismic imaging of a hydraulic fracture in diatomite, *J. Pet. Techn., 44,* 28.

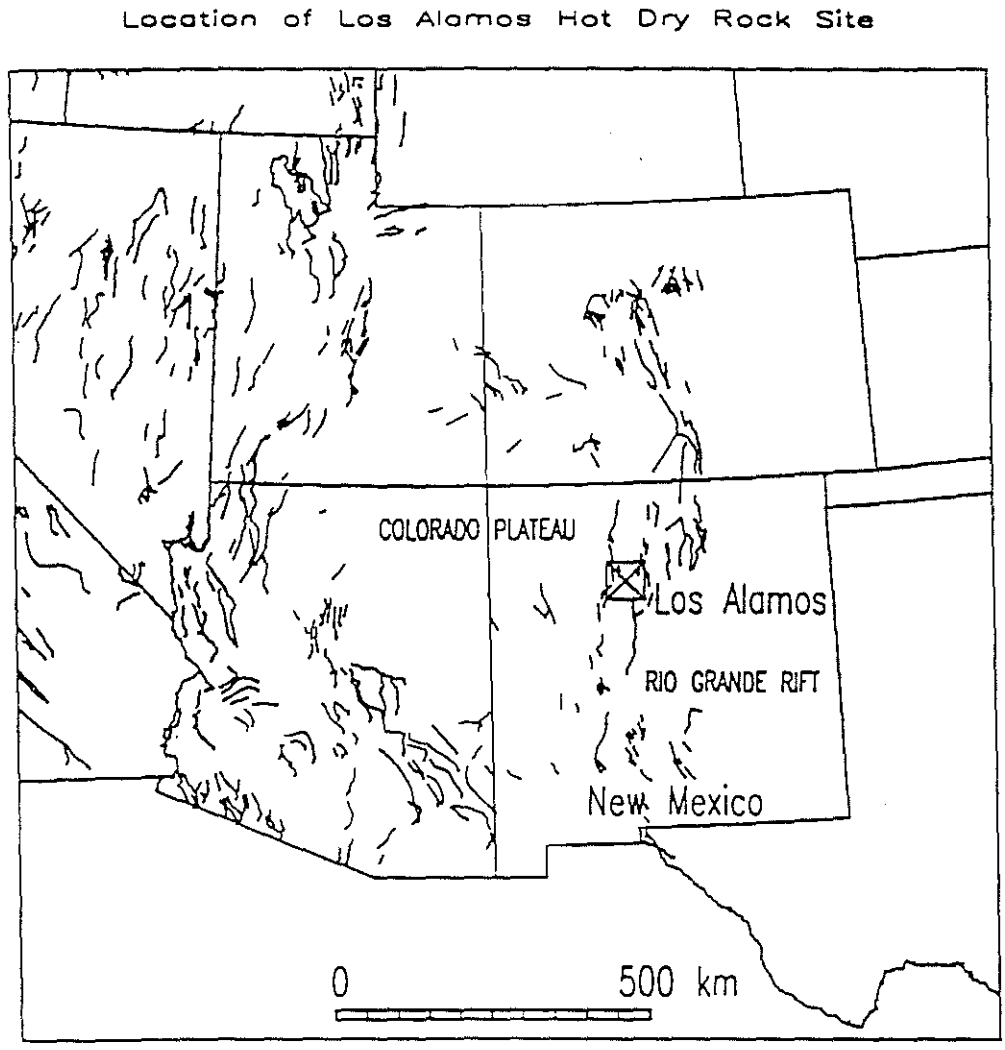

Figure 1: . Map of southeastern United States showing the major faults and the location of the Los Alamos Hot Dry Rock site in north-central New Mexico (from Block 1991).

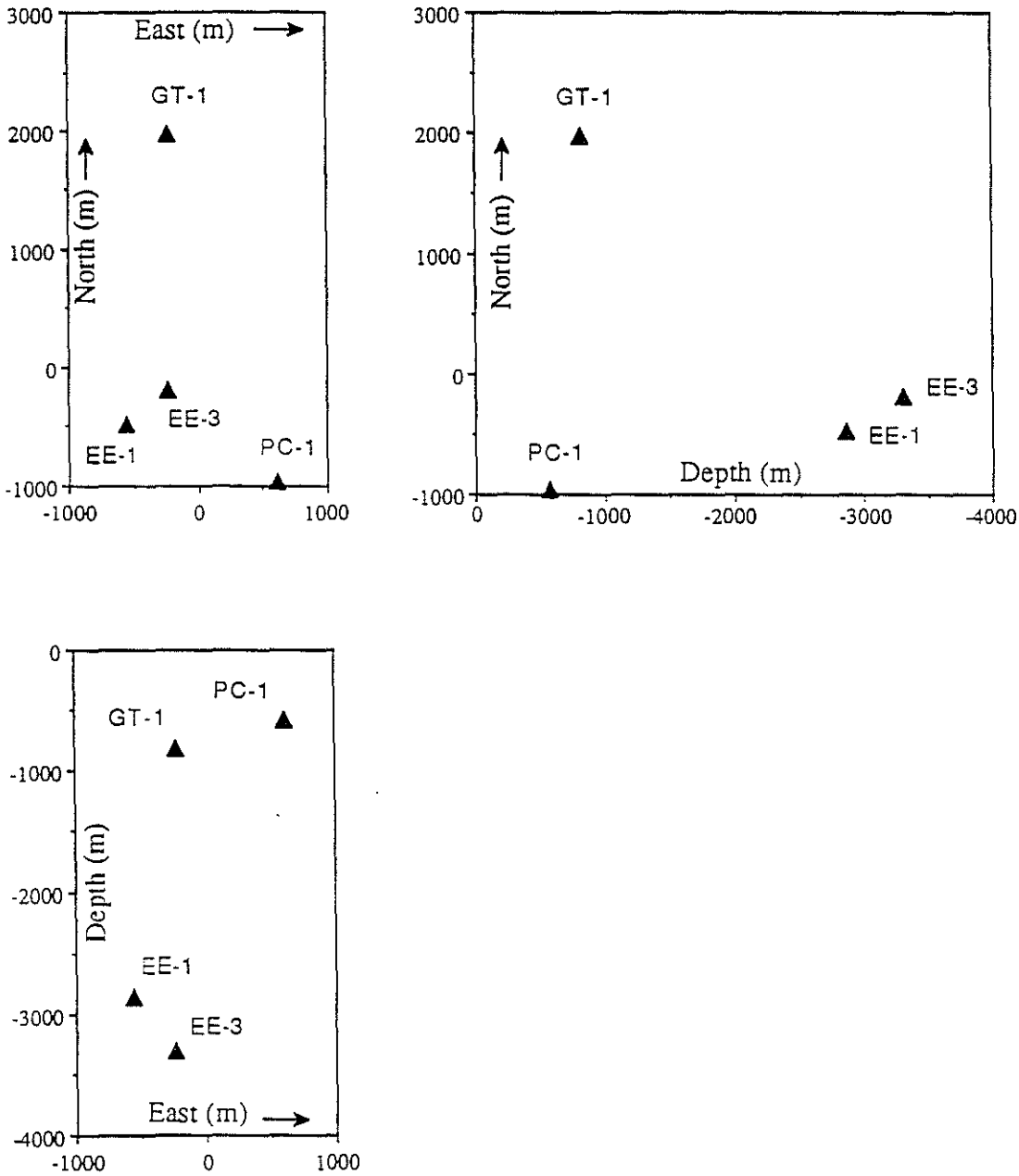

Figure 2: Orthogonal views of four seismic stations (triangles) used to monitor microearthquakes induced during the hydraulic fracturing experiment 2032 in December 1983.

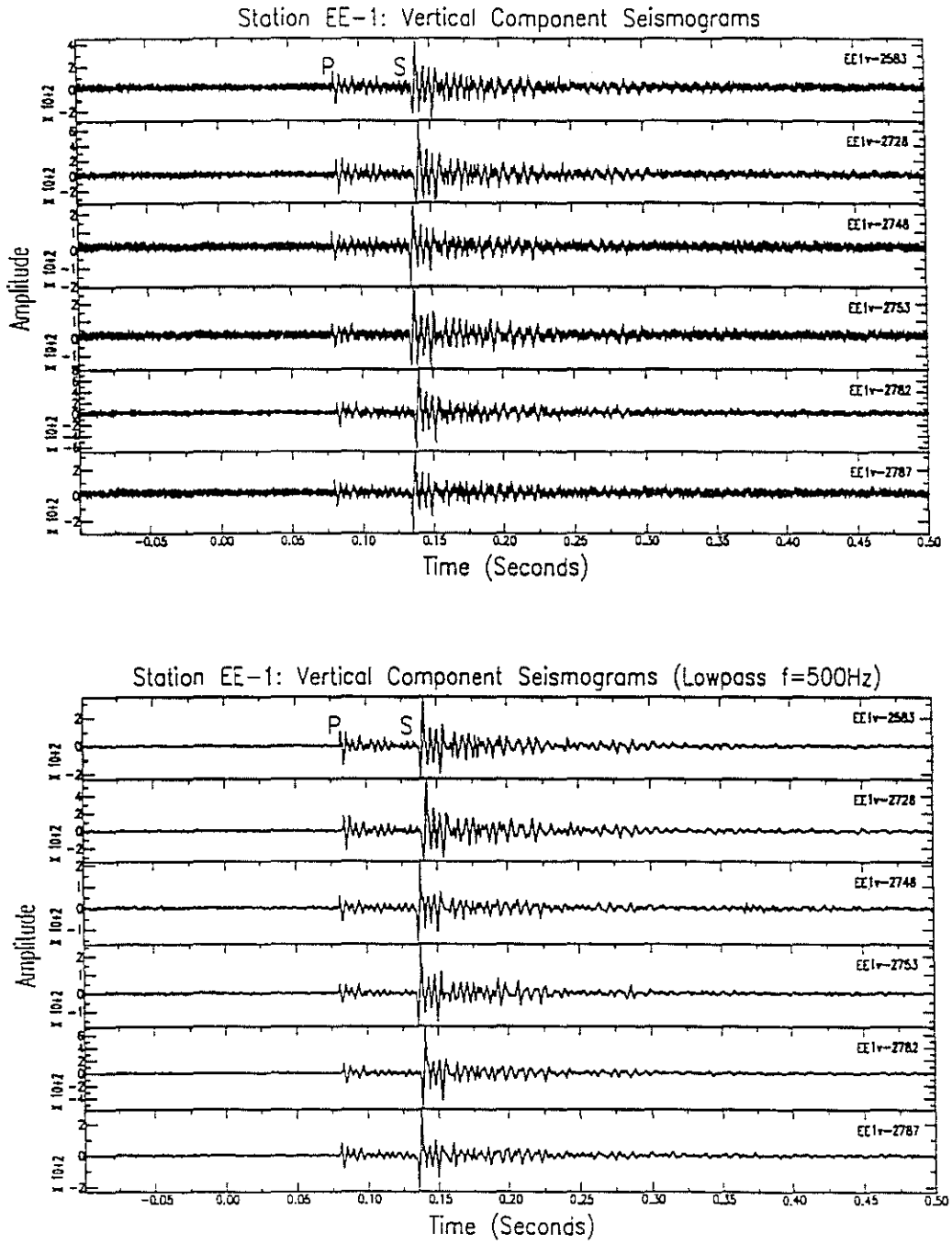

Figure 3: Unfiltered (top) and filtered (bottom) vertical component seismograms of six similar microearthquakes recorded at station EE-1. The hypocentral distance from the events to station EE-1 averages approximately 450 m.

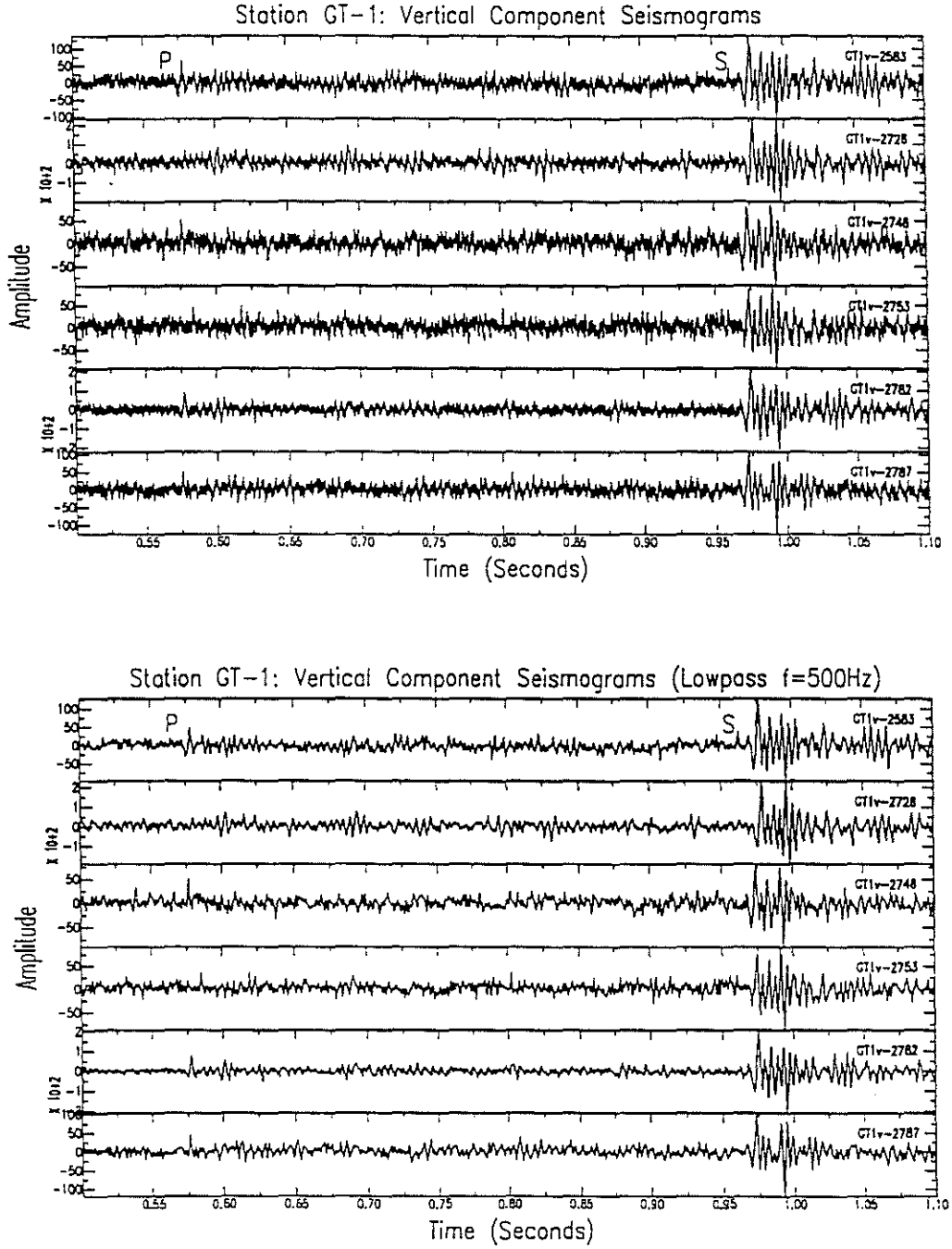

Figure 4: Unfiltered (top) and filtered (bottom) vertical component seismograms of six similar microearthquakes (same as those in Figure 3) recorded at station GT-1. The hypocentral distance from the events to station GT-1 averages approximately 3350 m.

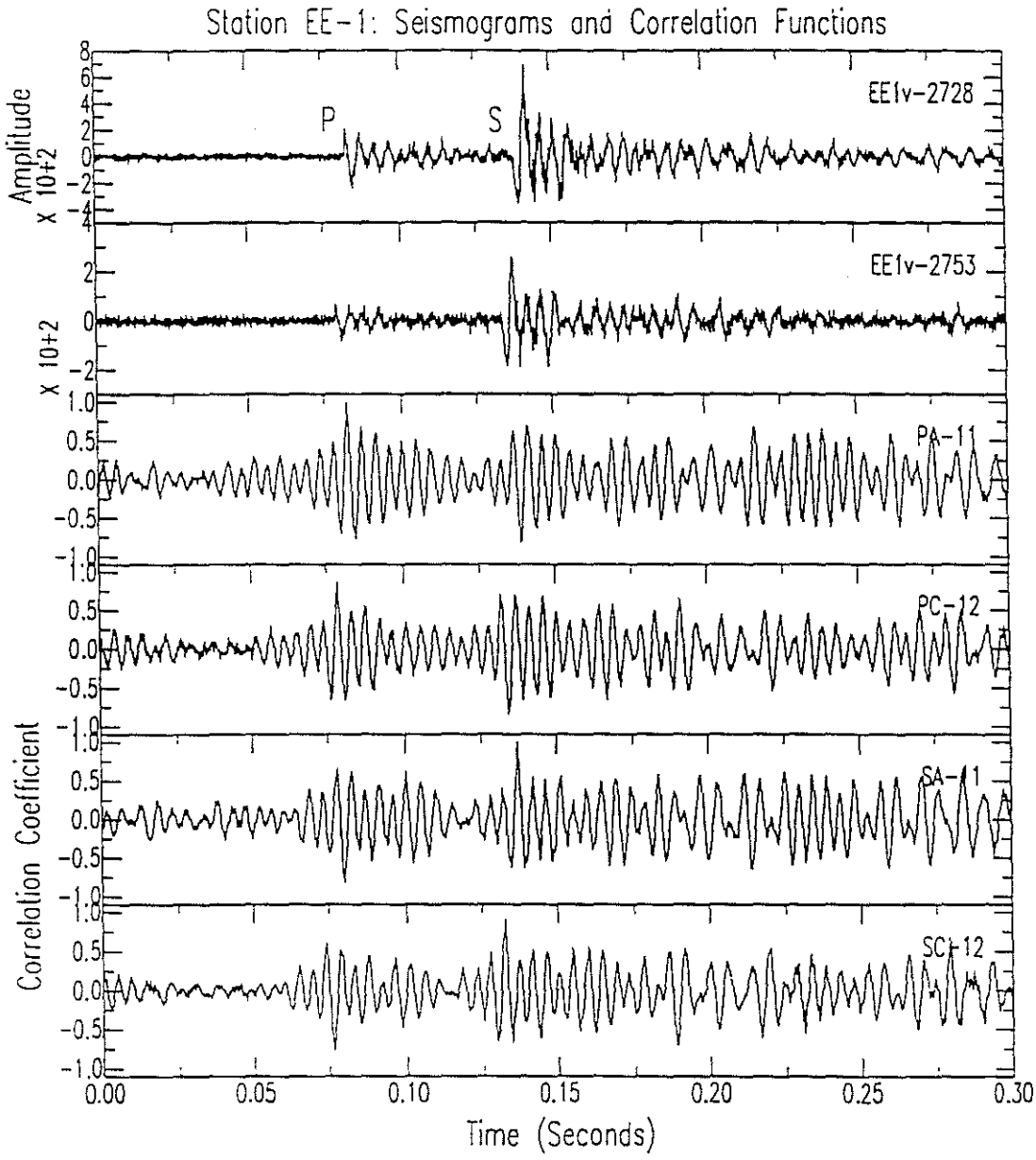

Figure 5: Example of differential arrival time determination from cross-correlation functions. The top two traces are the observed seismograms for events 2728 and 2753. The middle two traces, respectively, are the cross-correlation of each seismogram with the P wave segment extracted from the event 2728 seismogram. The bottom two traces are the cross-correlations of each seismogram with the S wave segment for Event 2728. The P and S wave segments used were each 30 ms long.

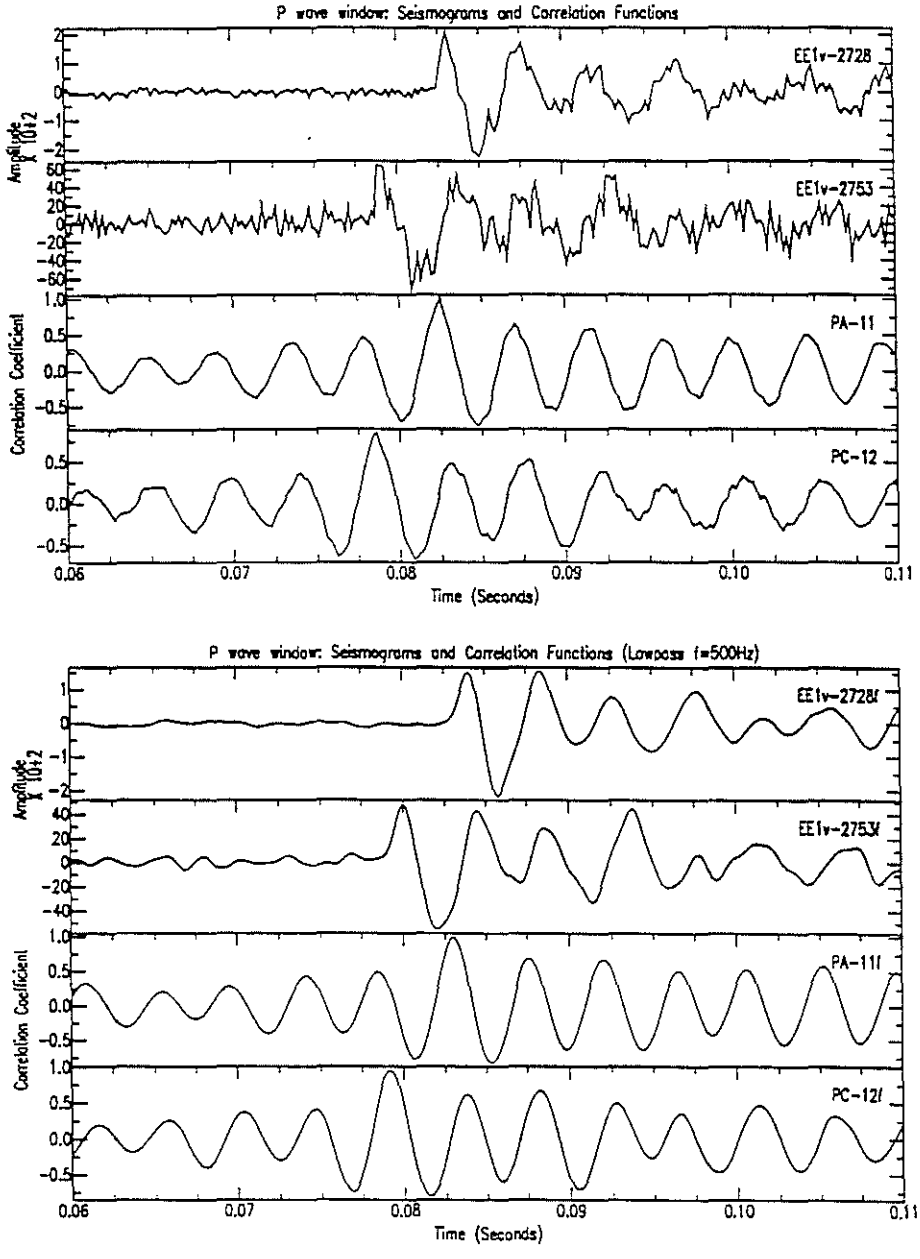

Figure 6: The effect of lowpass filtering on differential P wave arrival time estimate. The top frame shows the P wave portion of the seismograms (traces 1 and 2) and correlation functions (traces 3 and 4) from the example in Figure 5 (i.e., traces 1–4 from Figure 5 on an expanded time scale.) The bottom frame shows the same seismograms after lowpass filtering (500 Hz cutoff) and the correlation functions obtained from the filtered seismograms. The correlation functions from filtered data (bottom frame, traces 3 and 4) yield the same differential P wave time as the correlation functions from unfiltered data (top frame, traces 3 and 4).

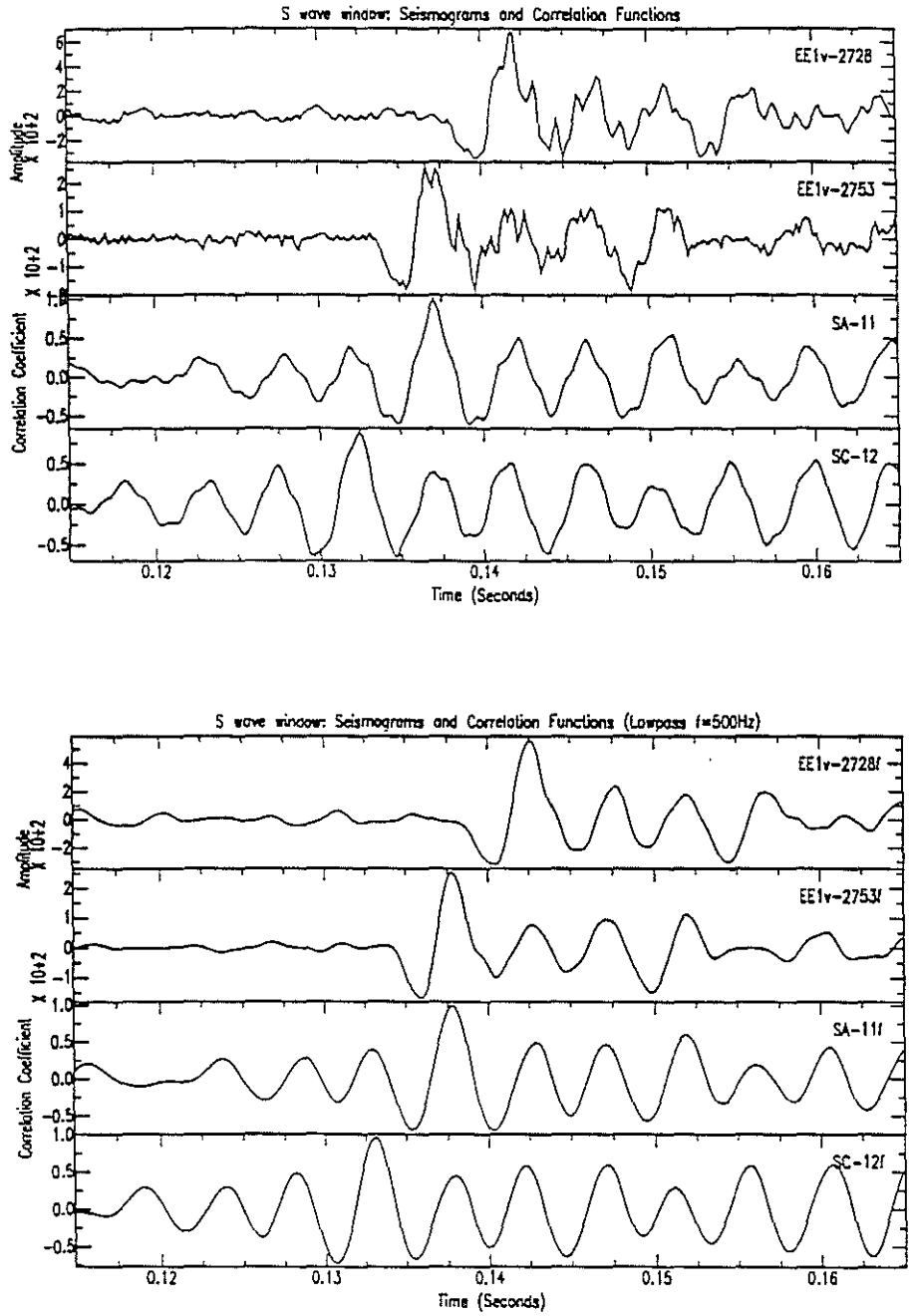

Figure 7: The effect of lowpass filtering on differential S wave arrival time estimate. The results shown are analogous to those in Figure 6 but are obtained with the S wave time window. The traces in the top frame are the same as Figure 5, traces 1, 2, 5 and 6, on an expanded time scale, while the bottom frame are the corresponding results using lowpass filtered data. The results with and without filtering yield the same differential S wave time.

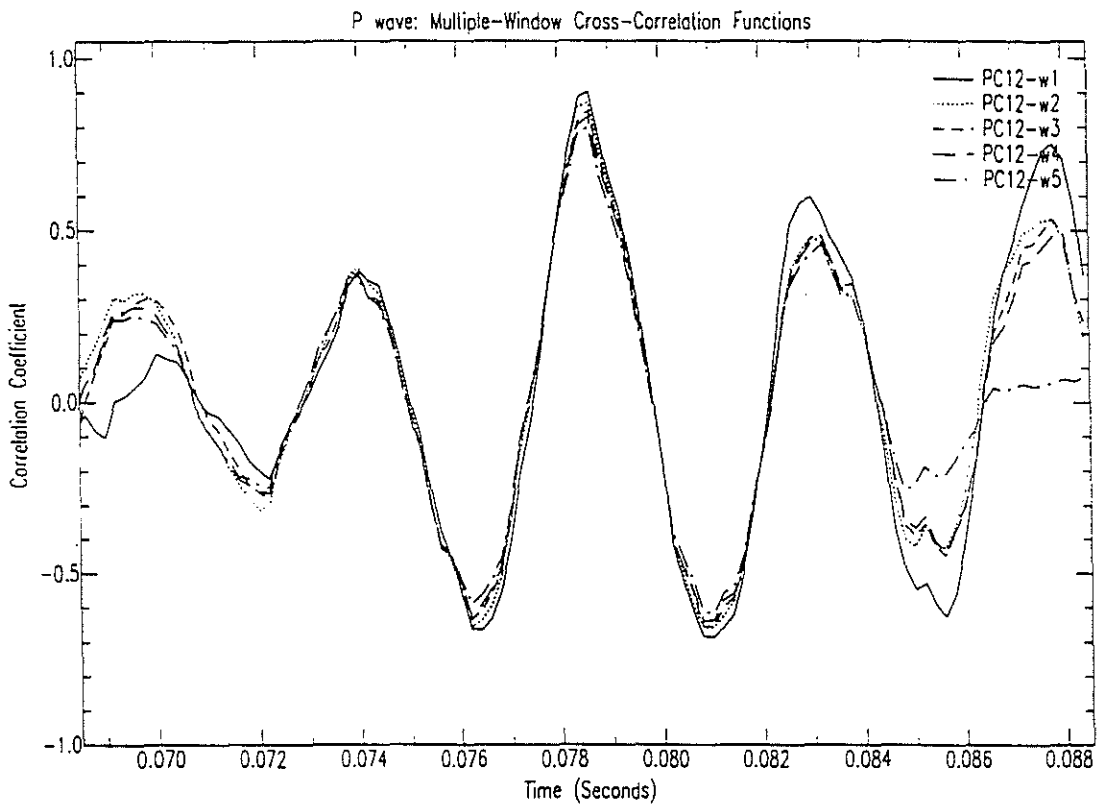

Figure 8: P wave cross-correlation function (Station EE-1, Events 2728 and 2753) derived with time windows of length 10, 20, 30, 40 and 50 ms. The time of peak correlation does not depend on the window length.

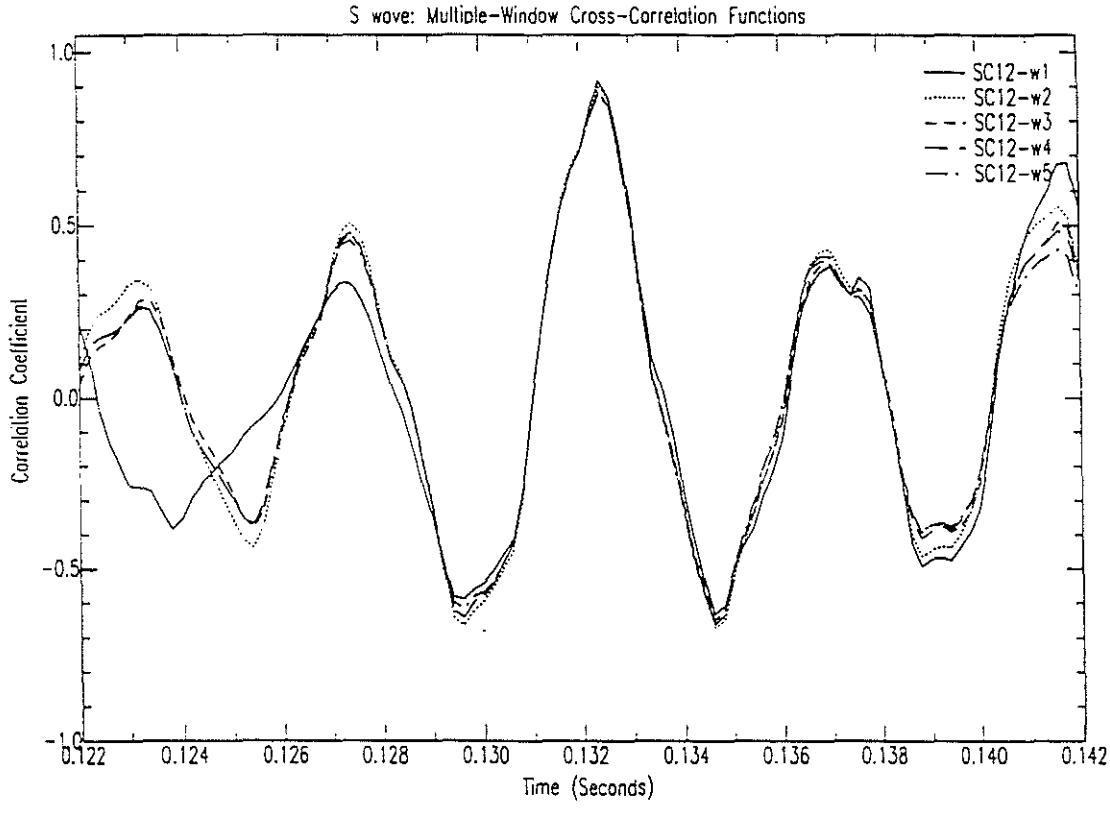

Figure 9: S wave cross-correlation function (Station EE-1, Events 2728 and 2753) derived with time windows of length 10, 20, 30, 40 and 50 ms. As in Figure 8, the time of peak correlation does not depend on the window length.

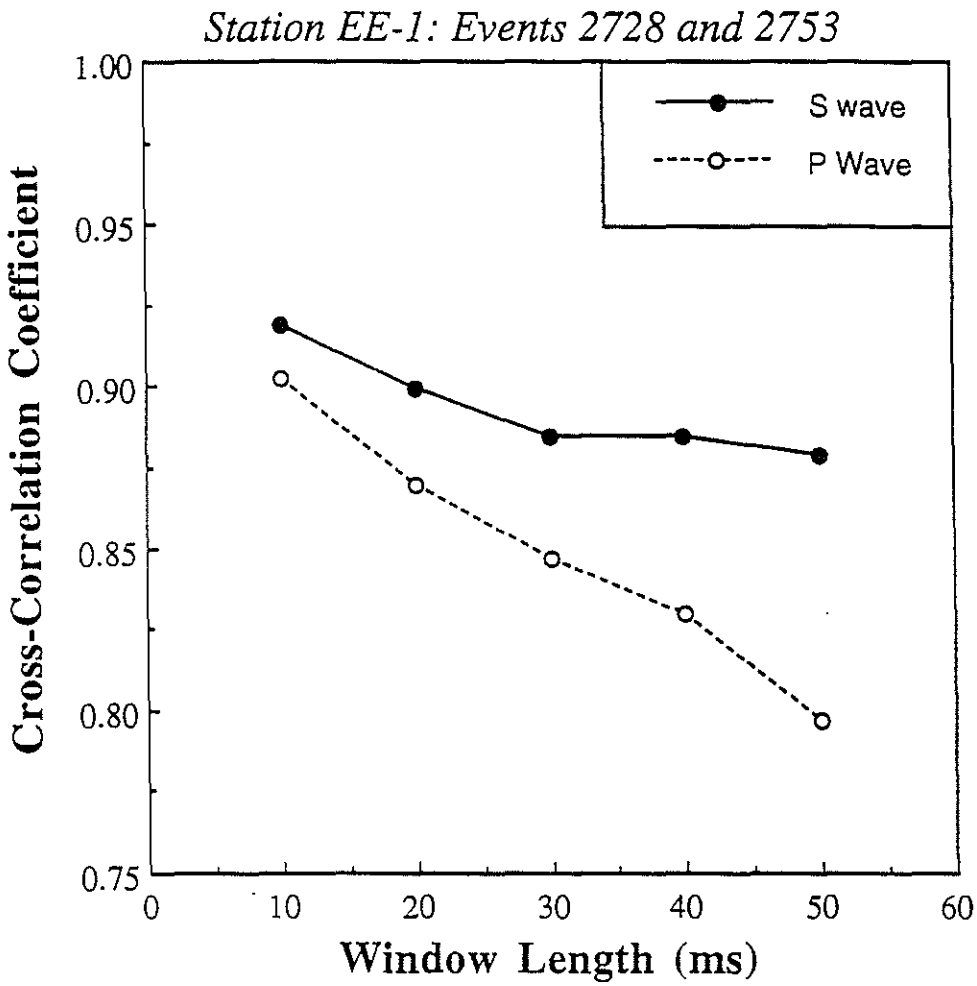

Figure 10: P wave (dashed line) and S wave (solid line) correlation coefficient between events 2728 and 2753 at station EE-1 as function of window length. The correlation coefficient is defined as the maximum of the cross-correlation function.

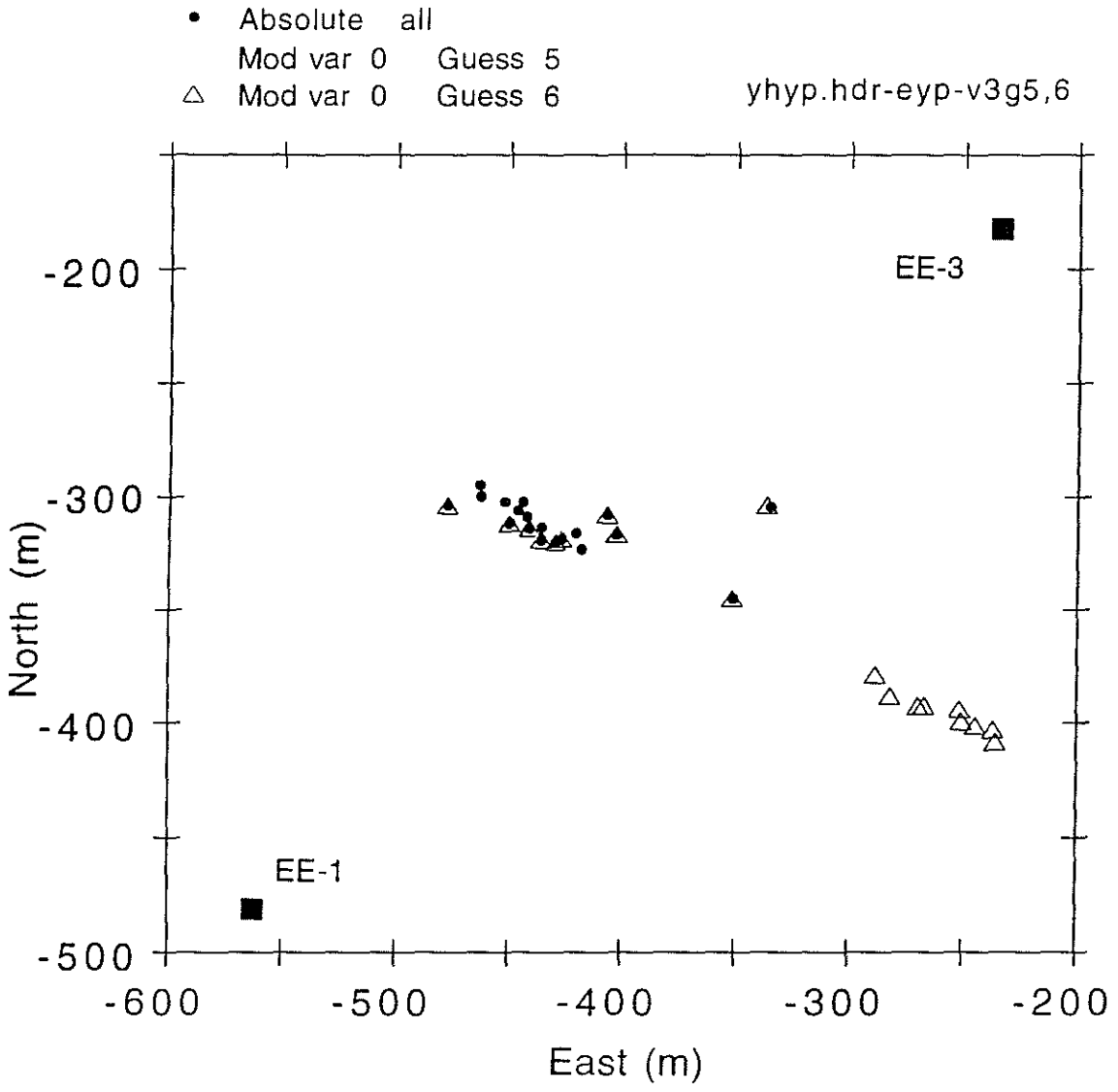

Figure 11: Map view of the hypocenters of 19 microearthquakes at the Los Alamos HDR injection site, determined from absolute arrival times at four downhole stations. Two solutions are shown (filled circles vs. open triangles) corresponding to different initial guesses of the event hypocenters.

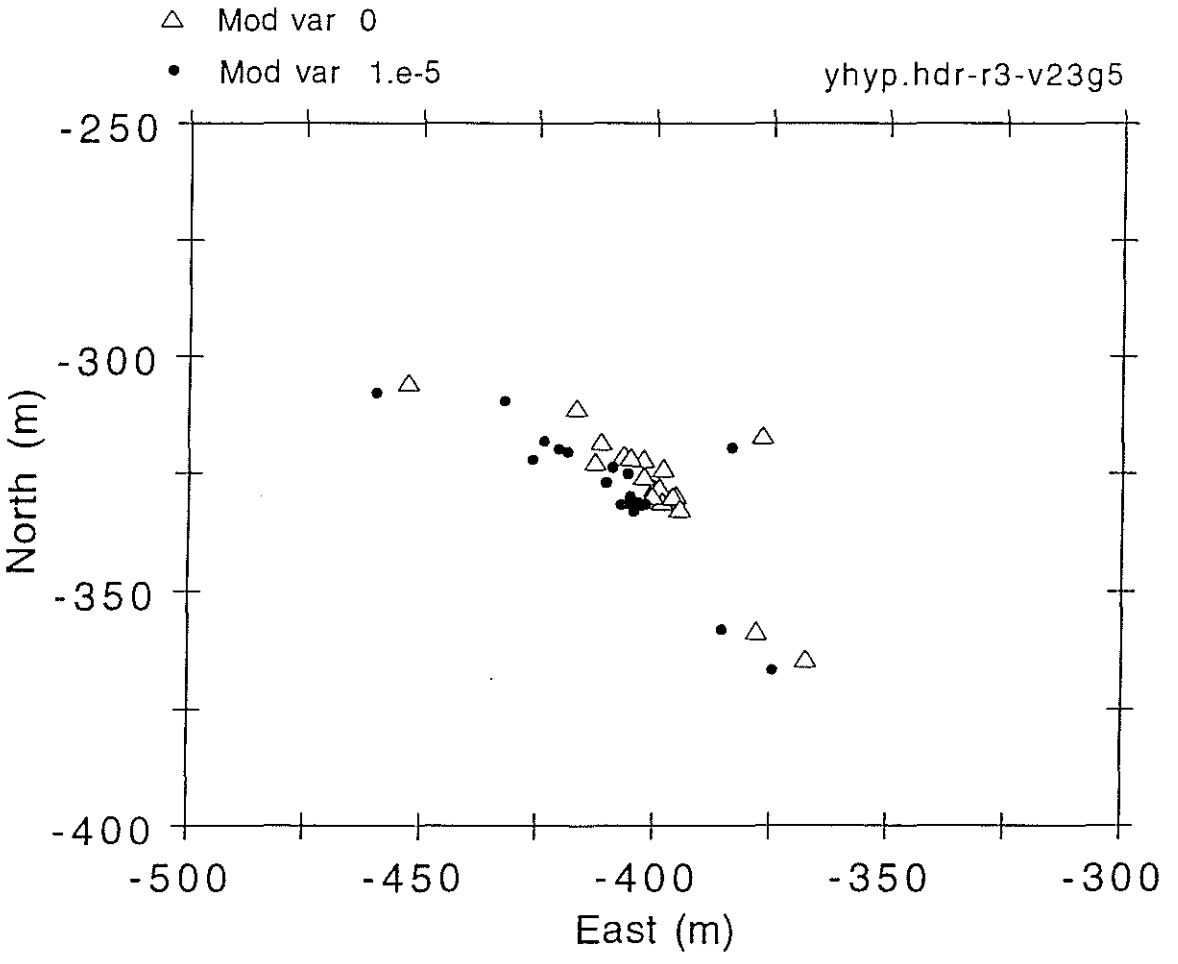

Figure 12: Map view of the microearthquake hypocenters determined from the data set of differential and absolute arrival times. Two solutions are compared: one with the modeling error variance parameter $\mu = 10^{-5}$ (filled circles) and one with $\mu = 0$ (open triangles).

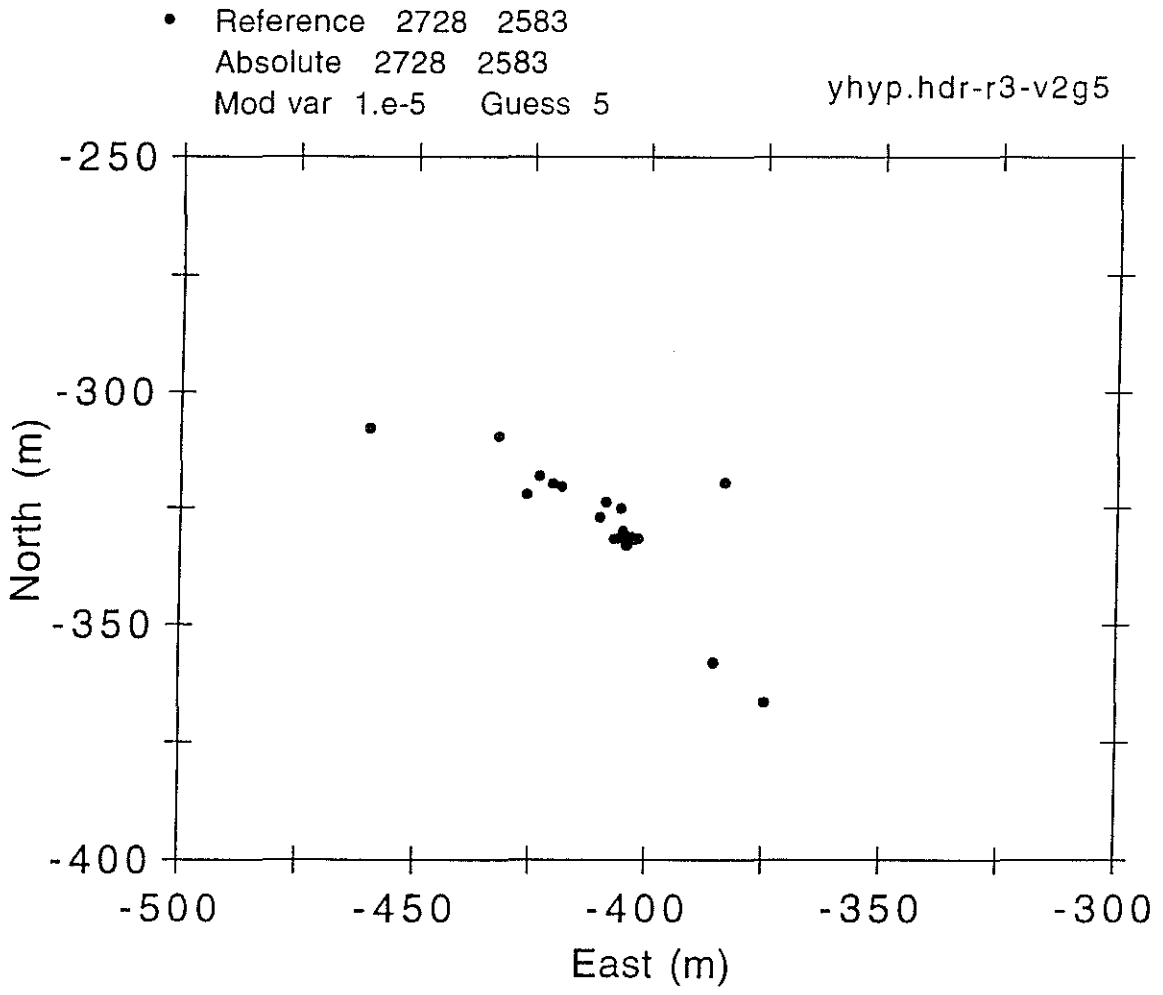

Figure 13: Map view of the microearthquake hypocenters determined from the data set of differential and absolute arrival times ($\mu = 10^{-5}$).

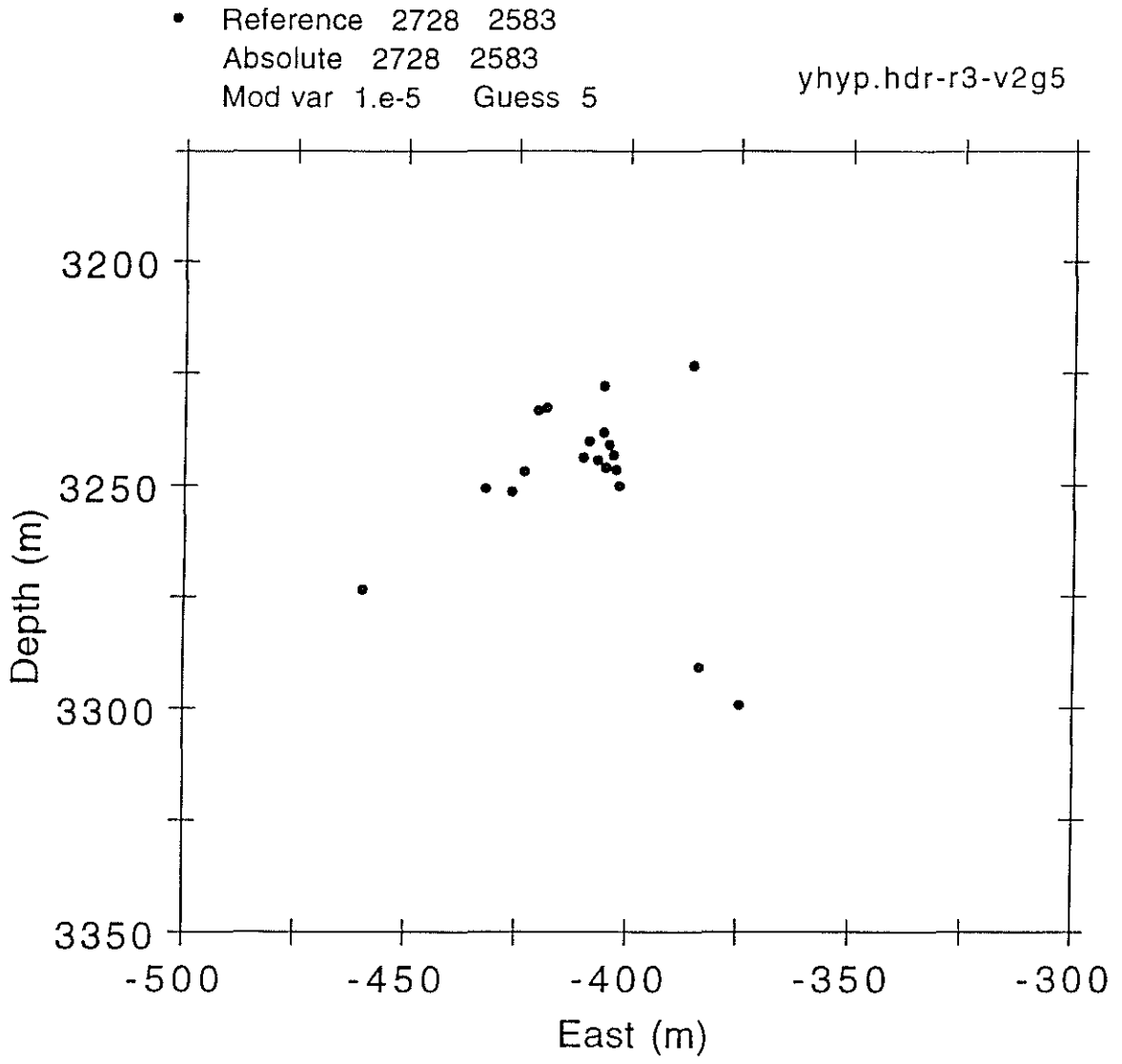

Figure 14: East-west vertical section of the microearthquake hypocenters determined from the data set of differential and absolute arrival times ($\mu = 10^{-5}$).

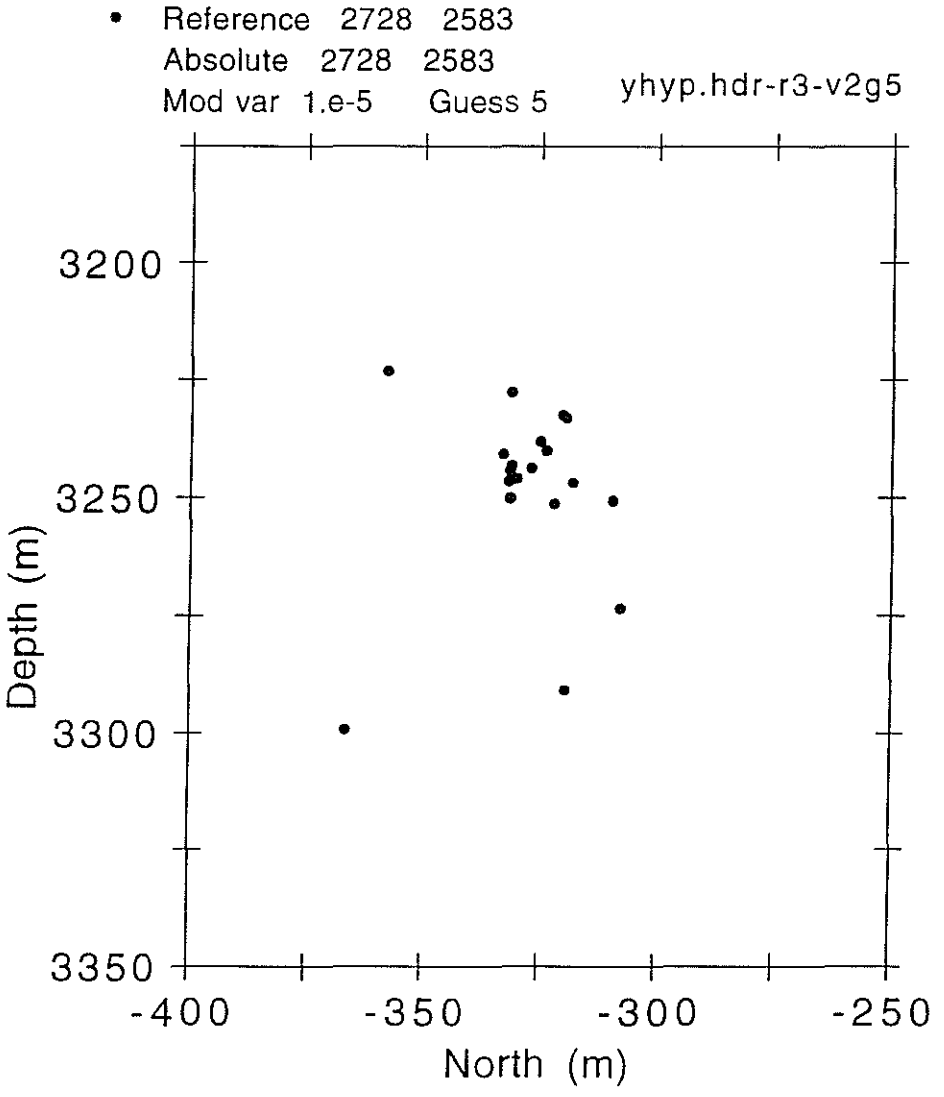

Figure 15: North-south vertical section of the microearthquake hypocenters determined from the data set of differential and absolute arrival times ($\mu = 10^{-5}$).

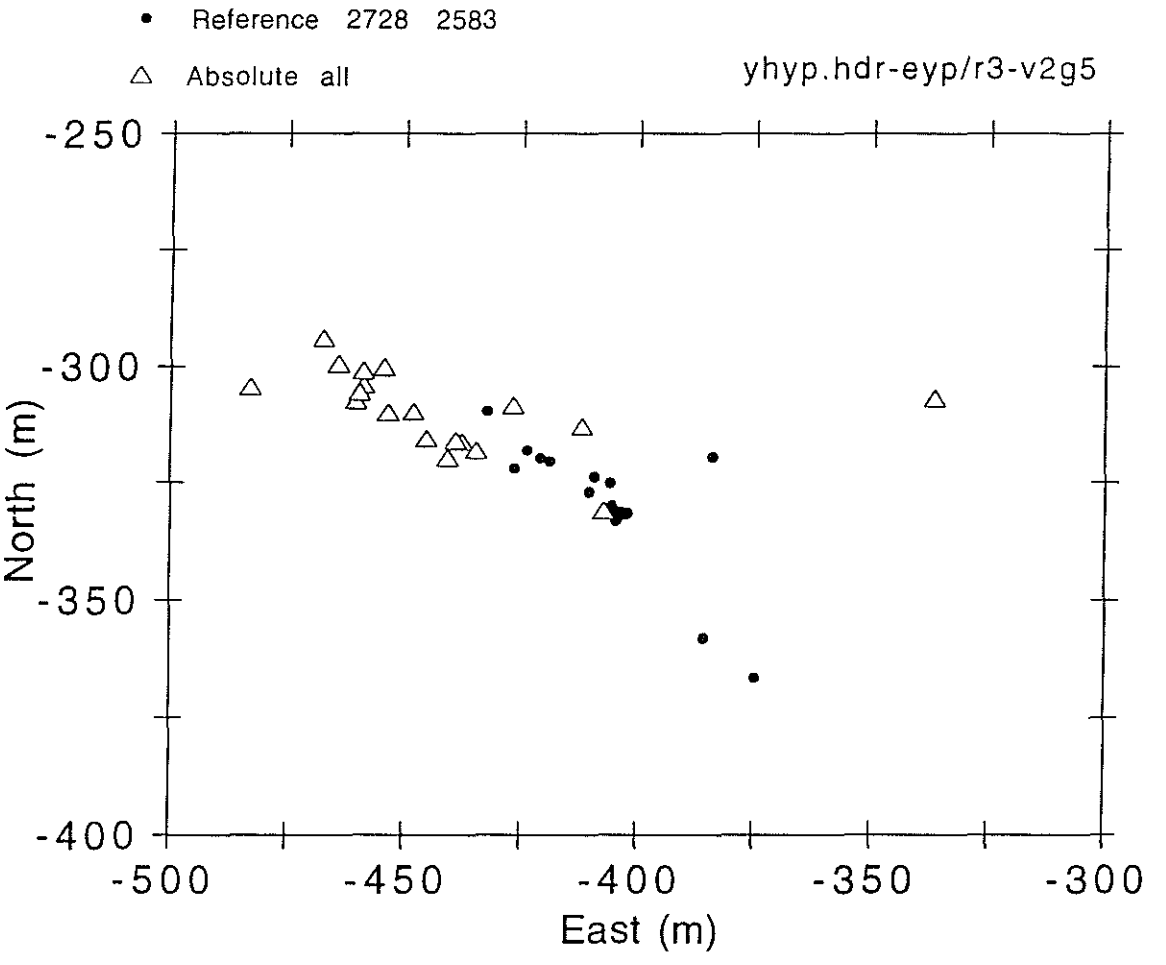

Figure 16: Comparison of hypocenter solutions obtained with absolute arrival times only (open triangles) and with the combined data set of differential and absolute arrival times (filled circles) ($\mu = 10^{-5}$).