Bulletin of the Seismological Society of America, Vol. 80, No. 1, pp. 150–169, February 1990

DETERMINATION OF TELESEISMIC RELATIVE PHASE ARRIVAL TIMES USING MULTI-CHANNEL CROSS-CORRELATION AND LEAST SQUARES

BY J. C. VANDECAR AND R. S. CROSSON

ABSTRACT

We have developed a semi-automated method of determining accurate relative phase arrival times and uncertainty estimates for teleseisms recorded on regional networks. Our analysis begins by obtaining preliminary arrival times with a single-trace phase-picking algorithm. For each possible pair of traces we then perform a search for the maximum of their cross-correlation function in order to obtain relative delay times. Depending on event magnitude, the best results are obtained by using correlation windows containing 2 to 4 sec of the initial energy pulse of the phase. The cross-correlation derived delay times are then used to generate an overdetermined system of linear equations whose solution is an optimized set of relative arrival times. We solve for these times using least squares. Cycle skipping is eliminated through the automatic re-evaluation of cross-correlation functions which yield high equation residuals. Quantitative estimates of timing uncertainty are obtained from the variance of equation residuals associated with each trace.

Using data from the Washington Regional Seismograph Network, we have found that for reasonably high-quality events, the rms uncertainty in arrival time estimates is on the order of the sample interval (0.01 sec). Reproducibility of delay anomalies is excellent for events from the same geographic locations despite significant differences in waveform character. In connection with a study of the deep structure of the Cascadia Subduction Zone, we include a preliminary examination of the variation of residual patterns with azimuth.

INTRODUCTION

Teleseismic relative arrival times provide the primary source of data for the modeling of velocity structure to depths of several hundred kilometers (Aki *et al.*, 1977; Humphreys *et al.*, 1984; Nolet, 1987). In addition to insufficient data coverage and model approximations, three problems that degrade resolution and accuracy in such models are (1) errors in arrival time measurement, (2) the lack of quantitative uncertainty estimates for use in weighting observations during inversion, and (3) the rejection of possibly useful data perceived to have too low of a signal-to-noise ratio. The last problem is especially critical in studies which have poor data coverage from certain azimuths, since often the sparse number of events from these regions have low magnitudes and shallow hypocenters resulting in recorded waveforms of relatively low quality. When large velocity perturbations are present in the region to be modeled and nonlinear methods are required, problems (1) and (2) may even act to preclude solution convergence (Nakanishi and Yamaguchi, 1986). The need to improve in these three areas together with the constraints imposed by the large volume of data regularly acquired by networks of up to a hundred or more stations makes it apparent that there is a demand for an efficient, effective, and automated method of relative arrival time estimation. While single-trace phase-picking algorithms (e.g., Allen, 1978; Crosson and Hesser, 1983) are objective and have proven to be useful for the analysis of local and regional data, they are not normally

suitable for use with teleseismic phases due primarily to the relatively emergent character of teleseismic waveforms.

The similarity of teleseismic waveforms leads naturally to the use of correlation methods (e.g., Bungum and Husebye, 1971) and even manual methods using first breaks, peaks, troughs, or zero crossings make qualitative use of signal correlation (e.g., Iyer *et al.*, 1981; Achauer *et al.*, 1986). In this paper, we present a novel method for determining relative arrival times which provides both high-quality relative arrivals and dependable uncertainty estimates. We use a multi-channel cross-correlation procedure with adjustment by least squares which allows us to make quantitative estimates of the uncertainty in relative arrival times as well as to make use of data that may be unusable with visual methods. Previous studies using large aperture arrays such as the Swedish Seismograph Network (Husebye and Jansson, 1966), the Norwegian Seismic Array (NORSAR) (Jansson and Husebye, 1968; Bungum *et al.*, 1971), and the Large Aperture Seismic Array of Montana (LASA) (Archambeau *et al.*, 1965; Mack, 1969) have shown that teleseismic waveforms remain suitably coherent for cross-correlation analysis. These studies found that the waveform coherency between stations as a function of station separation decreases somewhat to distances of 10 to 20 km, beyond which there is virtually no change. When stations are separated by over 10 to 20 km (as is the case with many regional networks) the waveform distortion induced by multi-pathing appears as uncorrelated noise. We have confirmed this result for the Washington Regional Seismograph Network (WRSN, Fig. 1) and later in this paper demonstrate the lack of dependence on station separation of both waveform coherency and rms timing uncertainty.

Clearly when considering only a single station and therefore each waveform as a sample from a multiple number of events, waveform distortion cannot be thought of as a purely random process. This is due to the fact that at least some waveform distortion will undoubtedly arise from multi-pathing (multiple reverberations in the upper crust) and therefore a contribution to waveform distortion will vary predictably with respect to the azimuth and distance of each source. Likewise, given two stations, the cross-correlation function between their traces will not be distorted in a purely random manner when considering multiple events from the similar azimuth and distance. However, when considering a single event recorded on a multiple number of stations (separated by sufficient distances), the distortions in the cross-correlation functions, and therefore shifts induced in the positions of cross-correlation maxima, do manifest themselves as nearly random uncorrelated noise. This is due to the incoherence of multi-pathing between stations (as opposed to between events) and is the statistical property which we seek to exploit.

A simple and effective method of obtaining the time lag between two traces containing similar waveforms is to locate the offset of the maximum of their cross-correlation function. One way to apply this approach to more than two traces is to choose a single channel with a characteristic waveform and then cross-correlate it with remaining channels to find their relative time shifts. The problem with this method is that waveform distortions from multi-pathing and even random noise (since the time series used are necessarily of finite length) will introduce first-order distortions into each cross-correlation function and therefore first order errors into the relative arrival time estimates. Even the relative magnitudes of these errors are difficult, if not impossible, to reliably estimate and are a function of which trace was chosen as the standard to cross-correlate against. As an alternative to selecting a single representative trace, Bungum and Husebye (1971) developed a procedure

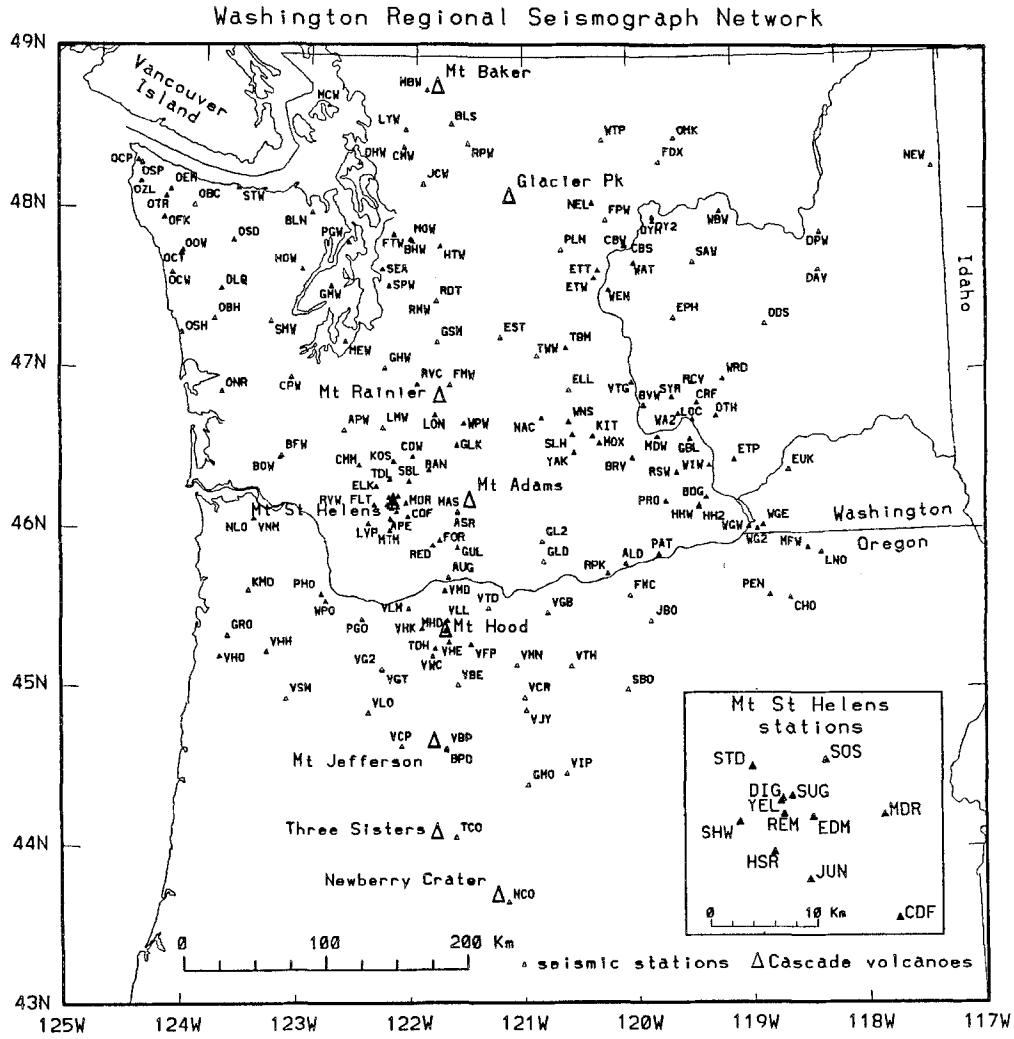


FIG. 1. Map view of seismic stations in the Washington Regional Seismograph Network (WRSN) showing stations which are either currently operating or have collected digital teleseismic data in the past. Inset shows relative location of stations in the Mount St. Helens area.

which cross-correlates each trace against a summed beam trace. They then generate a new beam with the resultant relative delay times and iterate on this process. Although superior to simply choosing a single trace to correlate against, their method is still influenced in a first-order manner by the nonrandom distortions present in individual cross-correlation functions and does not converge in all cases. Neither of these procedures are capable of detecting cycle skipping—the selection of an incorrect maximum of a cross-correlation function—nor of producing quantitative uncertainty estimates.

Our goal was to improve on these methods in order to provide a high-quality set of data for imaging the deep structure of the Cascadia Subduction Zone (CSZ). To do this we have developed a straightforward, semi-automated procedure which uses the cross-correlations between all possible pairs of traces and estimates both accurate delays and quantitative timing uncertainties. We assume that errors in cross-correlation derived delay times, when taken between many different stations,

are primarily random in nature and therefore an improved estimate of waveform relative arrival times can be made through an overdetermined system of equations and least squares. Our procedure can be broken into three major parts: (1) estimation of the relative delay time between each pair of traces via an efficient search of their cross-correlation function, (2) minimizing the inconsistencies between all delay time estimates in a least-squares sense in order to eliminate cycle skipping and invert for a single waveform arrival time on each trace, and (3) final interactive optimization of the solution to eliminate signals of unacceptably low quality. In this paper we describe each of these steps in detail. Using examples from the WRSN, we explore the optimization of correlation and least-squares weighting parameters and analyze the validity of our uncertainty estimation procedure. Finally, some examples of the application of the method are presented.

Cross-Correlations

Prior to any multi-channel processing we remove both the high- and very low-frequency energy present in the data by applying a zero-phase, bandpass Chebyshev filter with corner frequencies of 0.5 and 5.0 Hz. Although the precise filter specifications are not important, we find that bandpass filtering increases the reliability of our signal quality and timing uncertainty estimates while also making waveforms more recognizable during visual inspection. We have found by experiment that filtering has little effect on our optimized solution indicating that, for our purposes, high-frequency energy may be considered incoherent noise. Figure 2 shows an example of P-wave arrivals for data so filtered, recorded on 85 WRSN stations. This event, originating near the east coast of Kamchatka, will be used as an example throughout the analysis portions of this paper and referred to as event 1 (see Table 1 for event parameters). It is of relatively low quality compared to many teleseisms recorded by the WRSN.

For each trace we normally calculate a preliminary arrival time using a single-trace phase-picking algorithm developed at the University of Washington (Crosson and Hesser, 1983; Hesser, 1982). This method, which was designed primarily for regional and local earthquake phase-picking, estimates approximate arrival times. With event 1, for purposes of comparison, we have used visual picks of the first minimum which are indicated on Figure 2 by single vertical lines. Any picking procedure that locates the approximate onset of the major energy packet of the phase may be used. We then take advantage of the similarity that waveforms exhibit between stations through the use of their cross-correlation functions. In discrete form, the truncated estimate of the cross-correlation function between the ith and jth traces is

$$\phi_{ij}(\tau) = \frac{\delta t}{T} \sum_{k=1}^{T/\delta t} x_i(t_i^p + t_o + k\delta t + \tau) x_j(t_j^p + t_o + k\delta t) \tag{1}$$

where x_i = digital data from ith trace; t_i^p = ith trace's preliminary arrival time estimate; τ = lag time relative to preliminary arrival time estimates; T = length of correlation window (sec); t_o = time between preliminary arrival time estimate and when correlation window begins; and δt = sample interval (≈ 0.01 sec for WRSN).

Two traces to be cross-correlated, with the parameters T and t_o depicted, are shown at the top of Figure 3. Since on a routine basis we are only interested in the position of the maximum and its magnitude, the entire cross-correlation function need not be calculated. Instead we perform an efficient search for the maximum's

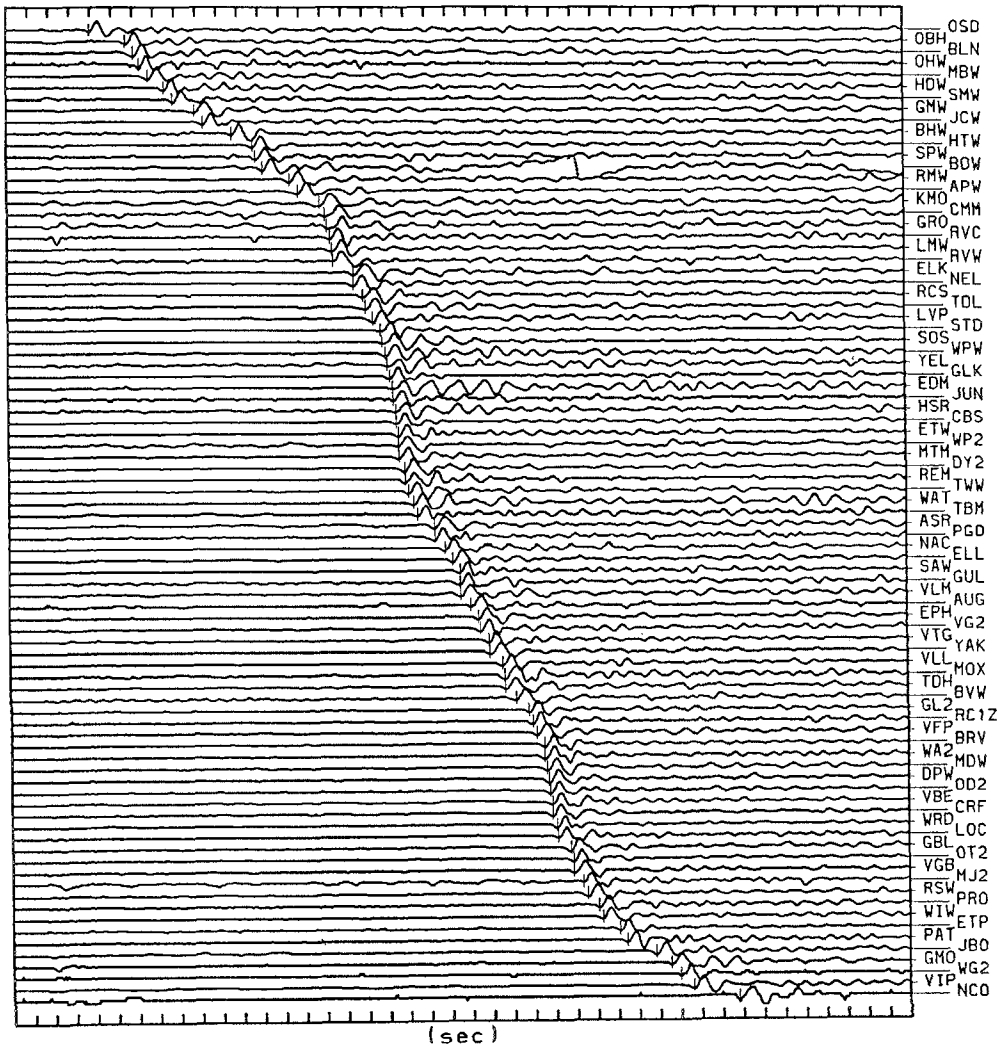


FIG. 2. Bandpass-filtered seismic traces recorded on 85 WRSN stations illustrating the direct P phase waveform of event 1 (see Table 1 for event parameters). Traces are ordered by arrival time with vertical line markers delineating preliminary arrival time estimates made with visual picks of first minimum. Each trace is individually scaled with respect to the signal energy around the initial pick.

TABLE 1

EARTHQUAKE PARAMETERS FOR EXAMPLE EVENTS

Event Number	Date mm/dd/yy	Time (UTC)	Location (PDE)		Depth (km)	Magnitude		Azimuth (w.r.t WRSN)	Δ
			Latitude	Longitude		m_b	M_{ss}		
1	07/18/89	10:41	53.4N	160.4E	24	5.4	5.0	311°	51°
2	06/25/89	11:15	32.9N	39.6W	10	5.2	4.6	70°	63°
3	05/29/89	22:22	23.8S	70.3W	43	5.5	4.9	134°	81°
4	05/28/89	09:46	16.4S	173.4W	51	5.7		232°	76°

position, $\tau_{ij}^{\max}$, through the use of coarse estimates of the function. These estimates are performed by ranging lag times in steps of m sample intervals and multiplying only every mth datum. We proceed through two levels of coarseness ($m = 10$ and 5) before performing the full resolution cross-correlation ($m = 1$). In each case, the

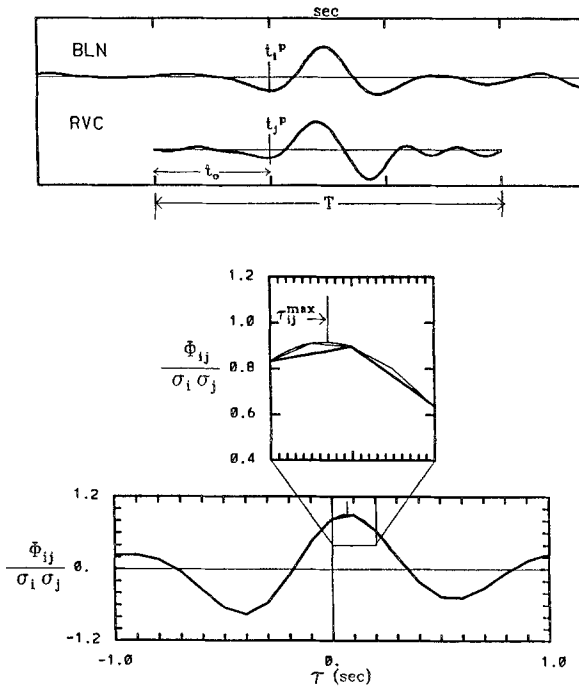


FIG. 3. Estimate of the (normalized) cross-correlation function between trace data from stations BLN and RVC for event 1. At the top of the figure are the portions of trace data used for the cross-correlation. Below are the coarse estimate of the function and a detail of the higher resolution portions.

next finer portion of the function is calculated within only $\pm m$ sample intervals of the current estimate. For instance with $T = 3$ sec and lag time, τ, ranging ± 1 sec this method results in a reduction in computation time to $\frac{1}{16}$ of that for computing the same range at full resolution, thereby making multi-channel cross-correlation feasible for use in routine analysis. The bottom of Figure 3 illustrates the cross-correlation function of the two traces above, showing both the coarse estimates of the function and a detail of the higher resolution portions. We are able to ignore negative extrema of the cross-correlation functions by having corrected the polarity of each trace before processing. This eliminates the possibility of large negative side-lobes being erroneously identified as the desired cross-correlation maximum.

Combining these lag times with the preliminary arrival time estimates defines the *cross-correlation derived relative delay time* between the ith and jth traces:

$$\Delta t_{ij} = t_i^P - t_j^P - \tau_{ij}^{\max} \tag{2}$$

From the maximum magnitude of the cross-correlation function, $\phi_{ij}(\tau_{ij}^{\max})$, an estimate of the *cross-correlation coefficient* between the ith and jth traces is given by

$$r_{ij} = \frac{\phi_{ij}(\tau_{ij}^{\max})}{\sigma_i \sigma_j} \tag{3}$$

where ϕ_{ij} is defined in equation (1) and σ_i^2 represents the sample variance of the ith trace data computed over the appropriate correlation window (when $\tau = \tau_{ij}^{\max}$).

Since r is bounded by unity and therefore not normally distributed we introduce Fisher's transform (e.g., Snedecor, 1956):

$$z = \frac{1}{2} \ln\left[\frac{(1 + r)}{(1 - r)}\right]. \tag{4}$$

This new variable, z, is distributed almost normally, thereby allowing for the calculation of appropriate means, variances, and confidence intervals which are then inverse transformed for use in analysis. All cross-correlation coefficient statistics which we refer to below have been calculated in this manner.

We initially quantify the relative signal quality of a given trace by using the sample mean and standard deviation of all cross-correlation coefficients associated with that trace. These statistics are plotted at the bottom of Figure 4 for the filtered traces of event 1. For data of moderate quality the sample means range from 0.80 to 0.95, corroborating our assertion from visual inspection that the majority of

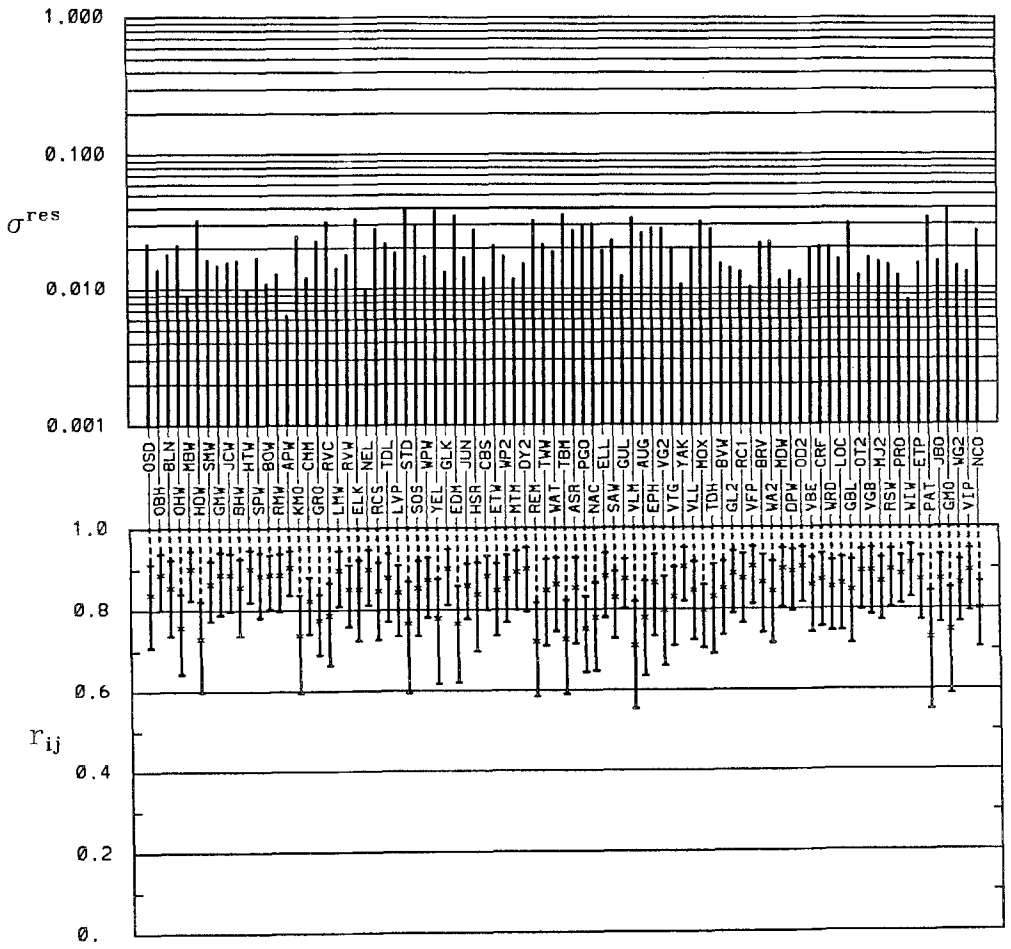


FIG. 4. Statistics associated with each trace of the filtered version of event 1. Above are plotted the rms timing errors associated with each trace (see equation (8)) and below are the mean and standard deviation ($\pm 1\sigma$) of the maximum cross-correlation coefficients associated with each trace (equation (3)). The standard deviations are not symmetric about the mean since the cross-correlation coefficient statistics were calculated from Fisher-transformed values (equation (4)) and then inverse transformed for analysis.

waveforms are highly coherent across the WRSN. The cross-correlation statistics provide useful keys in assessing a trace's relative quality and for spotting any large errors that may have been made during the preliminary picking process. For example, if a poor initial pick causes some of the cross-correlation functions associated with the trace to not include the true maximum while others do, the sample variance will be abnormally high. If a waveform is badly distorted or completely missed by the initial picking, that trace will produce consistently low cross-correlation coefficients usually with a small variance.

Least Squares and Uncertainty Estimation

Since waveforms are never completely coherent from station to station and there is always noise present, the cross-correlation derived relative delay times are not perfectly consistent (i.e., $\Delta t_{12} + \Delta t_{23} \neq \Delta t_{13}$). For n stations this inconsistency allows us to generate a system of $n(n-1)/2$ overdetermined equations given by

$$t_i - t_j = \Delta t_{ij} \qquad i = 1, 2, \ldots, n-1; \quad j = i+1, i+2, \ldots, n \tag{5a}$$

to which we add the constraint equation

$$\sum_{i=1}^{n} t_i = 0 \tag{5b}$$

to force the arbitrary mean of the optimized relative arrival times, t_i, to zero. This system may be expressed symbolically as $\mathbf{At} = \Delta\mathbf{t}$ where the n length solution vector $\mathbf{t}$ represents the optimized relative arrival time of the waveform on each trace (with zero mean), $\Delta\mathbf{t}$ is the $n(n-1)/2 + 1$ length data vector of cross-correlation derived relative delay times, and $\mathbf{A}$ is a sparse $n(n-1)/2 + 1 \times n$ coefficient matrix. For the simple case of only 5 traces this system would take the form

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{bmatrix} = \begin{bmatrix} \Delta t_{12} \\ \Delta t_{13} \\ \Delta t_{14} \\ \Delta t_{15} \\ \Delta t_{23} \\ \Delta t_{24} \\ \Delta t_{25} \\ \Delta t_{34} \\ \Delta t_{35} \\ \Delta t_{45} \\ 0 \end{bmatrix}$$

For the more typical case of 90 or more stations the number of equations grows to over 4000.

Adding the zero-mean constraint makes the system nonsingular and allows us to analytically calculate the least-squares solution $\mathbf{t}^{\text{est}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\Delta\mathbf{t}$ (e.g., Menke, 1984). Since $\mathbf{A}^T\mathbf{A} = n\mathbf{I}$ (where $\mathbf{I}$ is the identity matrix), and therefore $(\mathbf{A}^T\mathbf{A})^{-1} = (1/n)\mathbf{I}$, it follows that $\mathbf{t}^{\text{est}} = (1/n)\mathbf{A}^T\Delta\mathbf{t}$, or explicitly the simple sum

$$t_i^{\text{est}} = \frac{1}{n}\left(-\sum_{j=1}^{i-1} \Delta t_{ji} + \sum_{j=i+1}^{n} \Delta t_{ij}\right). \tag{6}$$

If we wish to weight the equations to reflect varying quality of observations, then we must instead perform a numerical inversion of the system

$$\mathbf{t}_{\mathbf{w}}^{\text{est}} = (\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}\Delta\mathbf{t}$$

(where $\mathbf{W}$ represents an $n(n-1)/2+1 \times n(n-1)/2+1$ diagonal weighting matrix). Due to the sparseness of $\mathbf{A}$ and its regular form it is possible to calculate $\mathbf{A}^T\mathbf{W}\mathbf{A}$ without constructing its component matrices, thereby limiting the maximum storage requirement to a size $n \times n$ matrix.

We have tested two different weighting schemes. The first involves weighting each equation by its associated cross-correlation coefficient and the second via a weighted function of equation residuals from the previously determined unweighted estimate,

$$\text{res}_{ij} = \Delta t_{ij} - (t_i - t_j) \tag{7}$$

These weighting schemes have little effect on the solution itself except when large outliers are present, in which case they make the inversion more robust by down-weighting those equations. The weighting does, however, make the uncertainty estimates more realistic since outliers have a strong effect on variance estimation. Since, for the unweighted case, the least-squares solution to (5) is simply a normalized sum of cross-correlation derived relative delay times given by (6), the standard deviation of the residuals associated with the ith trace,

$$\sigma_i^{\text{res}} = \sqrt{\frac{1}{n-2}\left[\sum_{j=1}^{i-1} \text{res}_{ji}^2 + \sum_{j=i+1}^{n} \text{res}_{ij}^2\right]}, \tag{8}$$

represents an estimate of its rms timing uncertainty. We have found that these residuals are nearly normally distributed with a tendency to have slightly heavy tails. Their standard deviation therefore provides a reliable, if somewhat pessimistic, estimate of rms timing uncertainty. At the top of Figure 4 we plot rms timing uncertainty for the filtered traces of event 1 associated with the unweighted least-squares inversion.

Timing uncertainty, σ_i^{res}, normally has a strong inverse correlation with mean cross-correlation coefficient, as one would expect, but under certain circumstances this correlation breaks down. The breakdown occurs when relatively high-frequency noise (4 to 10 Hz) is present in an otherwise undistorted waveform. The result is low cross-correlation coefficients but highly consistent and accurate delay times producing low timing uncertainties. Therefore the least-squares derived timing uncertainty, rather than correlation coefficient, gives the most reliable estimate of arrival-time accuracy.

We have found that very large equation residuals ($\text{res}_{ij} > 0.5$ sec) are usually the result of cycle skipping occurring during the cross-correlation analysis. To alleviate this problem our procedure automatically computes a new delay time, Δt_{ij}, for each cross-correlation function associated with one of these large equation residuals. In order to choose the correct maximum, we now search only the region of the cross-correlation function near the time lag predicted by the least-squares solution ($\tau = t_i^{\text{est}} - t_j^{\text{est}}$). This method has proven to be effective when the number

of cases of cycle-skipping is less than 10 to 20% of the total number of cross-correlations with a given trace. In the rare case when cycle-skipping is more prevalent, the initial least-squares solution is likely to be thrown off too far to be useful, and manual correction by an analyst is required.

The first manual intervention with our algorithm normally occurs on a computer workstation after the first round of least-squares optimized picks have been made. Adjustments may be made in several ways. The analyst may delete dead or highly distorted traces and re-run only the least-squares portion of the procedure. This is accomplished in several seconds. If excessive cycle-skipping has occurred or an improper phase has been chosen, the analyst may move one or more traces into alignment within visual accuracy and then re-run the full cross-correlation and least-squares procedures with a short range of lag times (e.g., $-0.25 \leqq \tau \leqq 0.25$ sec). Figure 5 is a plot showing details both of unfiltered and filtered traces of event 1 aligned on the left by preliminary arrival time estimates, t_i^p, and on the right by their respective least-squares optimized arrival times, t_i^{est}.

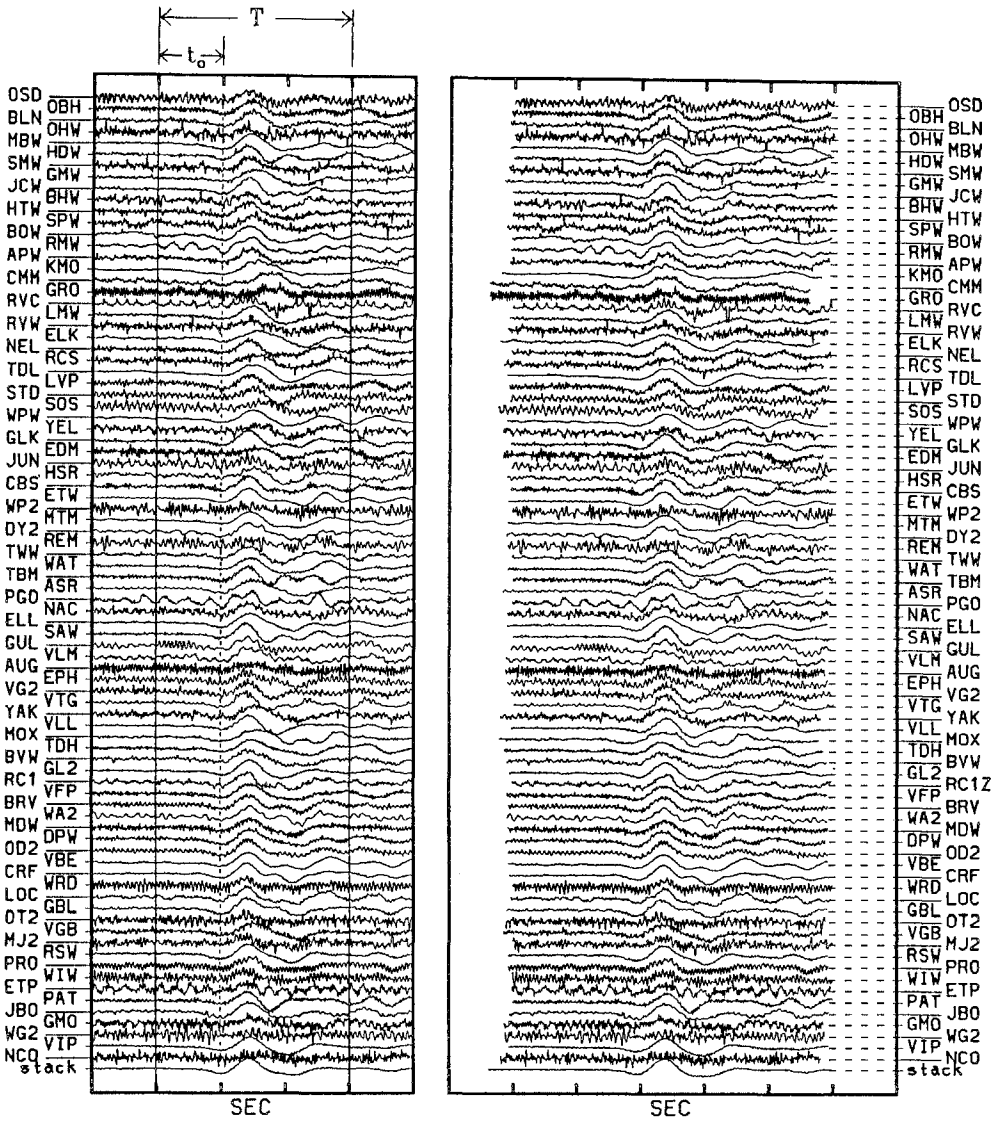
When examining the various extrema in the waveforms of event 1 it can be seen that any one of these features varies substantially from the optimized alignment in Figure 5. For the highest quality data (e.g., from large, deep events) all the methods that we tested yielded similar results. For slightly lower quality events, if the visually obtained arrival times from several features are averaged, the result once again approximates that of the least-squares optimized relative times. For yet lower quality data, however, such as in event 1, the major features are distorted to such an extent that even averaged times may be severely biased for many of the traces.

Selection of Parameters and Waveform Coherency

The adjustable parameters in our cross-correlation analysis are the cross-correlation window length (T), the lag time range ($\Delta\tau$), and the position of the correlation window starting time (t_o). As long as the true maximum is included in the estimated cross-correlation function, lag time range has no effect on our results; thus this parameter need only be chosen large enough to guarantee inclusion of that maximum. The global statistic that we use to evaluate the quality of the relative arrival times determined with a given set of parameters is the mean cross-correlation coefficient. We find that the most consistent and computationally efficient results are obtained with window lengths of 2 to 4 sec beginning at the onset of the first major energy packet of the phase. Figure 6 illustrates the three-dimensional representation of mean cross-correlation coefficient as a function of window length, T, and window starting position, t_o. Plotted beneath for reference is the stacked trace for this event, computed by

$$x_{\text{sum}}(k\delta t) = \sum_{j=1}^{n} x_j(t_{\text{ref}}^p - t_{\text{ref}}^{\text{est}} + t_j^{\text{est}} + k\delta t) \tag{9}$$

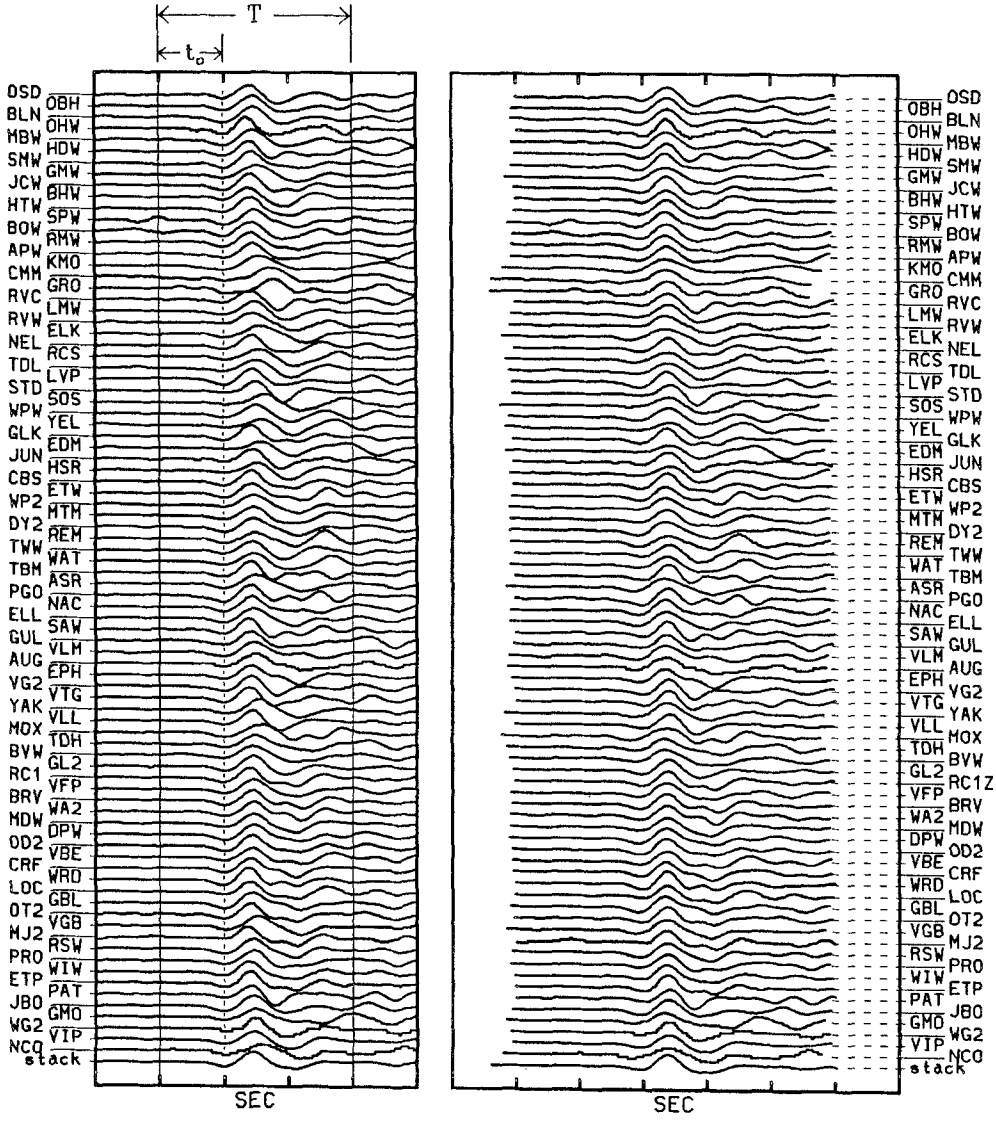
where the subscript *ref* indicates times associated with an arbitrary reference station. The stacked trace illustrates how the coefficient means are influenced by the signal energy at a given position in the waveform. Bands of high coefficient means that trend diagonally across the plot represent major packets of coherent waveform energy being reached with shorter window lengths as the window position moves down the trace. These correlation coefficient means are all calculated with respect



(a)

FIG. 5. Portions of waveforms used for cross-correlation analysis for (a) the unfiltered version of event 1, aligned by preliminary arrival-time estimates on the left and least-squares optimized relative arrival times on the right obtained with unfiltered data, and (b) the bandpass-filtered data. Each trace is individually scaled by the signal energy within the window of data shown.

to the optimized alignment at $t_o = -1.0$ sec and $T = 3.0$ sec (indicated by an * on the plot) so that each point represents the amount of waveform energy coherent with that alignment. It can readily be seen from both the short window length means ($T = 0.5$ sec) and the stacked trace that there is a considerable amount of coherent energy out 2 to 2.5 sec beyond the initial correlation window starting position ($t_o = -1.0$ sec). For larger magnitude events, this length of coherency extends later into the traces as one would expect, since larger events tend to have longer source time functions. It is therefore the source-time function, or equivalently



(b)

FIG. 5. (*Continued*).

magnitude of the earthquake, which controls the optimum window length for any particular event. Extending the correlation windows beyond this region has little effect on our solution but does increase the computer time necessary to process each event. Using small windows ($T < 2$ sec) generates solutions which tend to be unstable and highly sensitive to parameter changes. This is due in part to the edge effect arising from the truncation inherent in the windowing process. The effect can be partially alleviated with the use of tapering, but only at the expense of discarding a sizable portion of the information contained in the (already small) window.

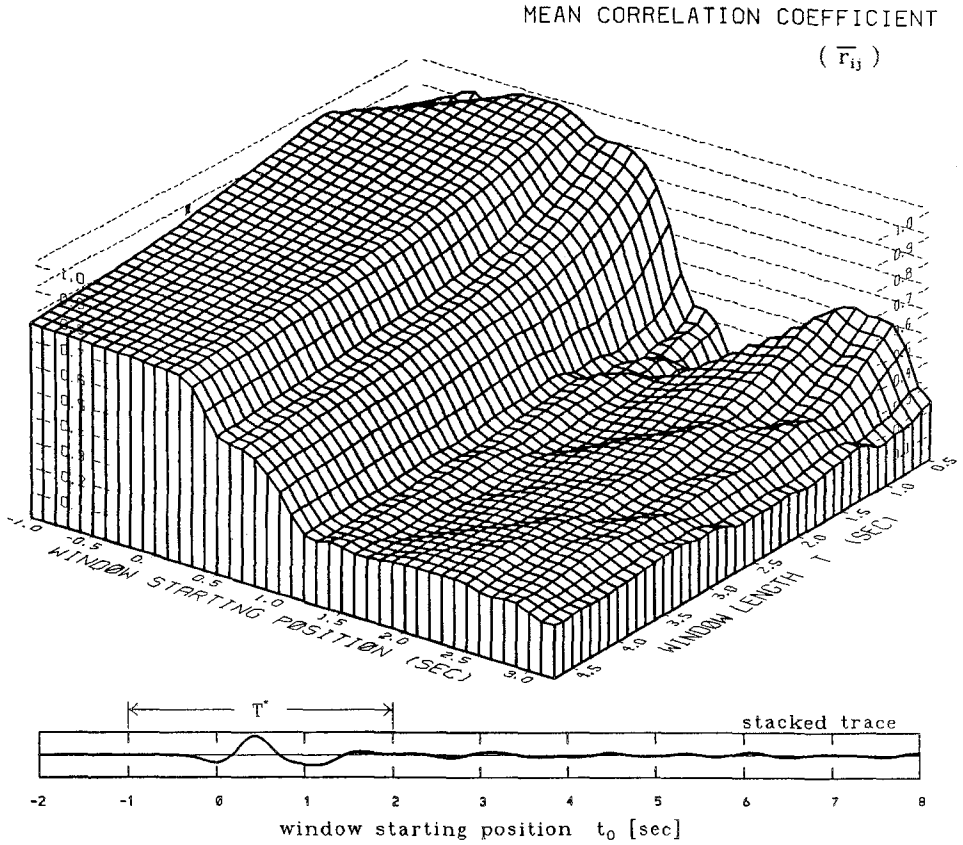
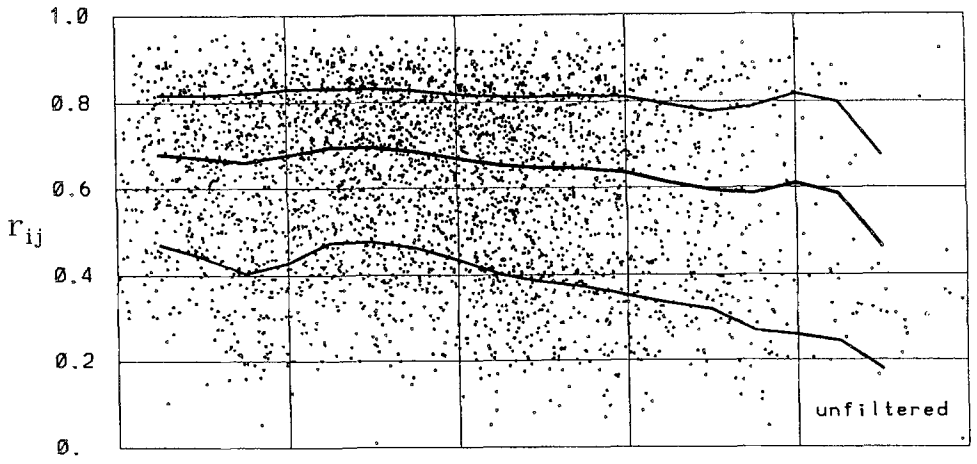
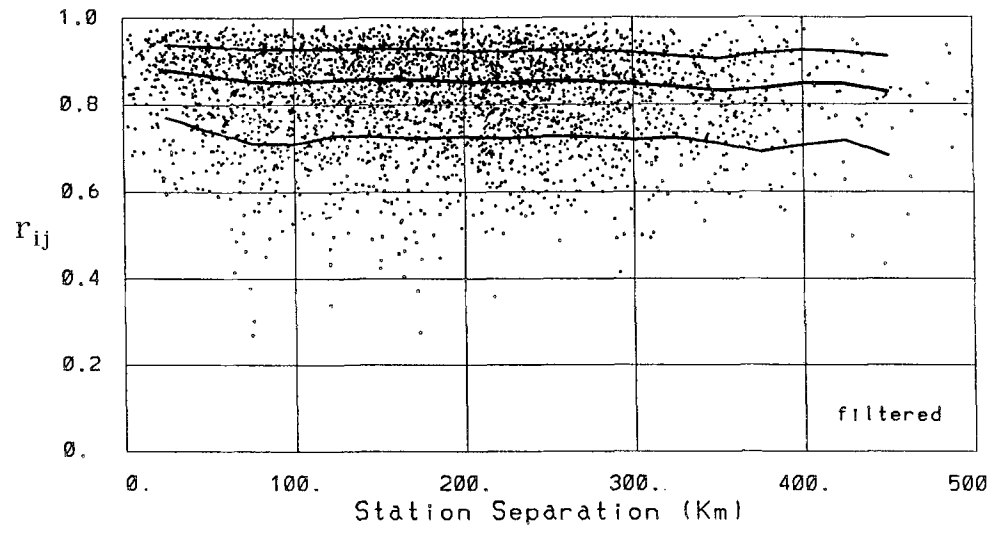


FIG. 6. Three-dimensional perspective of mean cross-correlation coefficient as a function of window position, t_o, and window length, T, for event 1 (see text). Plotted below for reference is the stacked trace for event 1 (equation (9)). T^* represents the window length used to obtain the least-squares optimized trace alignment. The corresponding point on the diagram is noted with an *. The rest of the diagram represents the amount of energy coherent with the alignment obtained with these parameters (T = 3.0 sec, t_o = −1.0 sec) with various window lengths and positions.

We have found that there is little if any correlation between station separation and waveform coherence for the WRSN. To quantify this result we have examined the coherence, as measured by the maximum cross-correlation coefficient, as a function of station separation. The correlation coefficients, r_{ij} as defined in equation (3) are plotted in Figure 7 as a function of station separation for both the unfiltered and filtered traces of event 1. Also shown in Figure 7 is a running mean and one standard deviation from that mean ($\pm\sigma$), once again calculated from Fisher-transformed coefficients (equation (4)) and then inverse transformed. Likewise, Figure 8 illustrates the dependence of solution residuals on station separation. If there were coherence in multi-pathing on scale lengths spanned by the network, then one would expect to see both a definite decline in waveform coherence and increase in residual magnitude with station separation. We have found that these dependencies are small (if at all existent) and change from event to event showing no general trend. This outcome is consistent with the findings of Archambeau *et al.* (1965) and also those of Mack (1969) who concluded that there was coherent multi-



(a)



(b)

FIG. 7. Maximum correlation coefficients as a function of station separation for (a) unfiltered traces and (b) filtered traces of event 1. The heavy line represents a running mean (sample window = 50 km) calculated from the Fisher-distributed coefficients (equation (4)) with lighter lines representing one standard deviation from that mean ($\pm 1\sigma$).

pathing within LASA sub-arrays (0.5–7 km scale lengths) but that between the sub-arrays (10 to 200 km span) multiples were uncoherent so that "phased summation tends to cancel them."

APPLICATIONS

As with studies relying on visually chosen arrival times (e.g., Michaelson and Weaver, 1986; Rasmussen and Humphreys, 1988), we have found that the delay residual patterns over the WRSN vary greatly with azimuth and at places angular distance. Figure 9 shows relative arrival times across WRSN for event 1. The overall

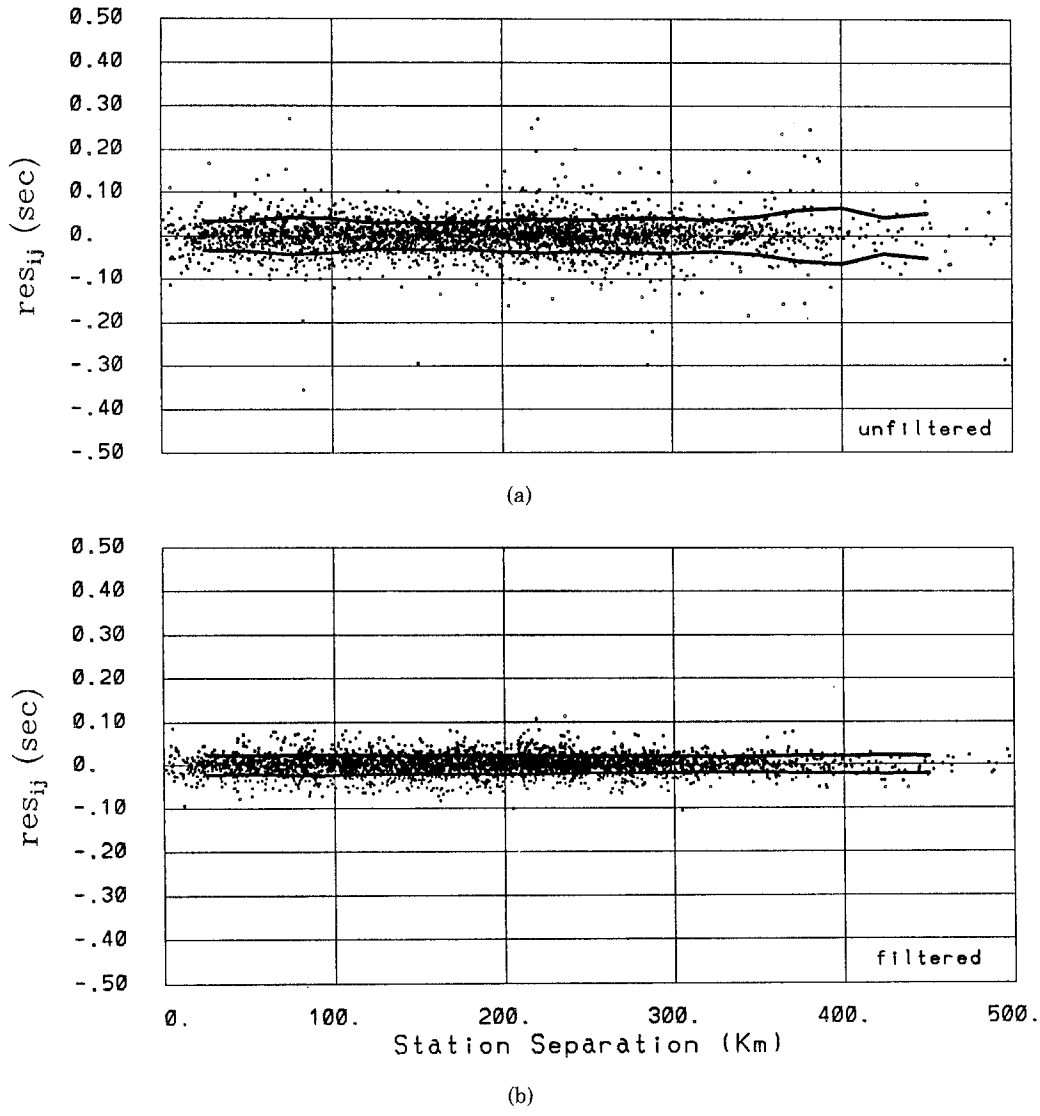


FIG. 8. Residuals to least-squares solution, res_{ij} (equation (7)), as a function of station separation for (a) unfiltered traces, and (b) filtered traces of event 1. The heavy lines represent the running standard deviation of this data (sample window = 50 km). Means are in general indistinguishable from zero.

pattern is one of a plane wave crossing the array but distortions can be seen even in these raw arrival times. To obtain the relative arrival residual patterns, we subtract expected times calculated from PDE locations and the PREM earth model (Dziewonski and Anderson, 1981). We calculate the expected arrival time at each station, remove the mean of those times, and then subtract them from our zero-mean optimized relative arrival times. Figure 10a is a map of these residuals contoured for event 1. The relatively small magnitude of many features of this pattern (0.2 to 0.6 sec) points out the need for highly accurate relative arrival times in order to resolve this signal. The signal dramatically changes with both azimuth and distance as is illustrated in Figures 10a through d. The degree of reproducibility of

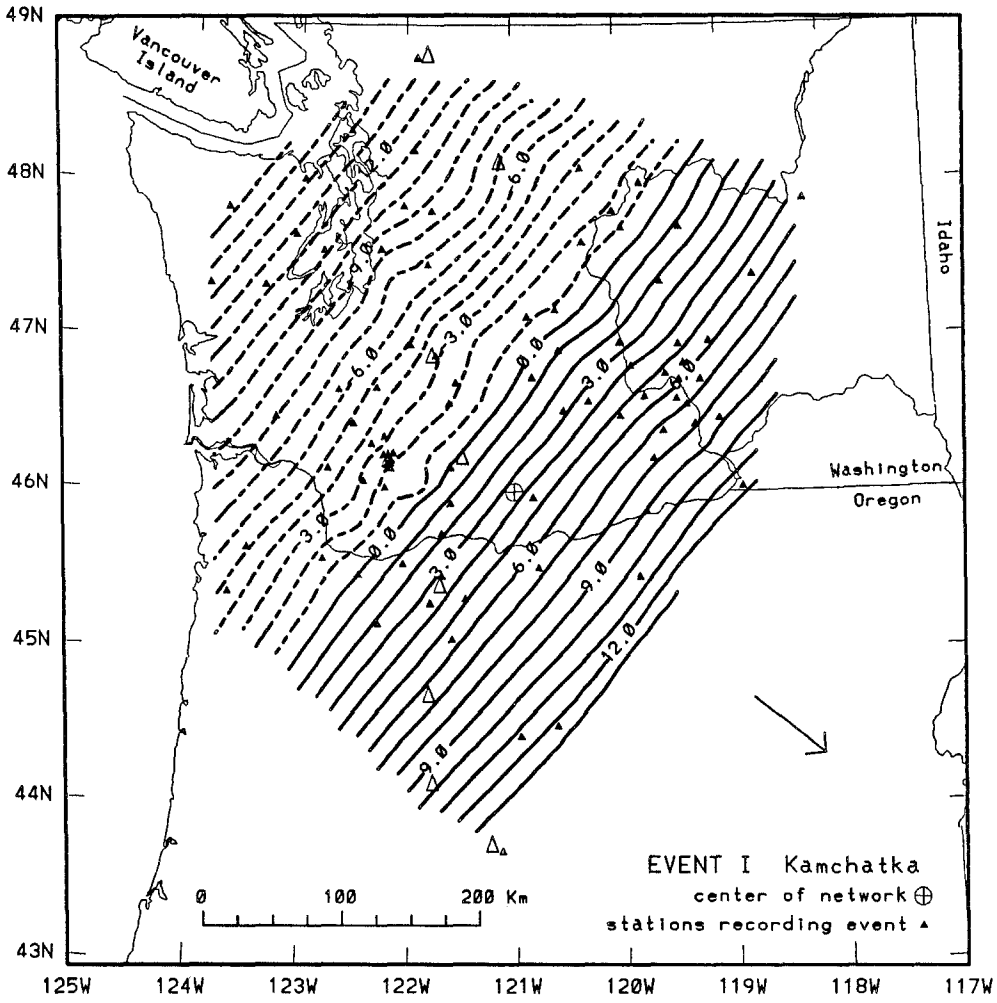


FIG. 9. Contour map of P phase relative arrival times for event 1. Contour interval is 1.0 sec. Arrow represents azimuthal direction with length proportional to epicentral distance from center of network. Azimuth to event 1 is 311° and angular distance, Δ, is 51°.

these residual patterns is very high, even for low-amplitude features, with statistical correlations between events from similar regions commonly ranging from 0.96 to 0.99. To achieve this level of reproducibility with visual picks requires fairly high-quality data and often the averaging of several waveform features; a relatively time-consuming process.

The large variation of residual pattern with azimuth indicates that deep structures play an important role in their generation. Some features such as the late arrivals over the north Puget Sound lowlands and the early arrivals over the North Cascades are relatively constant and depend only moderately on azimuth, possibly reflecting fairly shallow structure. On the other hand, features such as the band of large magnitude early arrivals from eastern azimuths (events 2 and 3) are certainly not present in data from the west, most likely reflecting deep slab structure (Rasmussen and Humphreys, 1988).

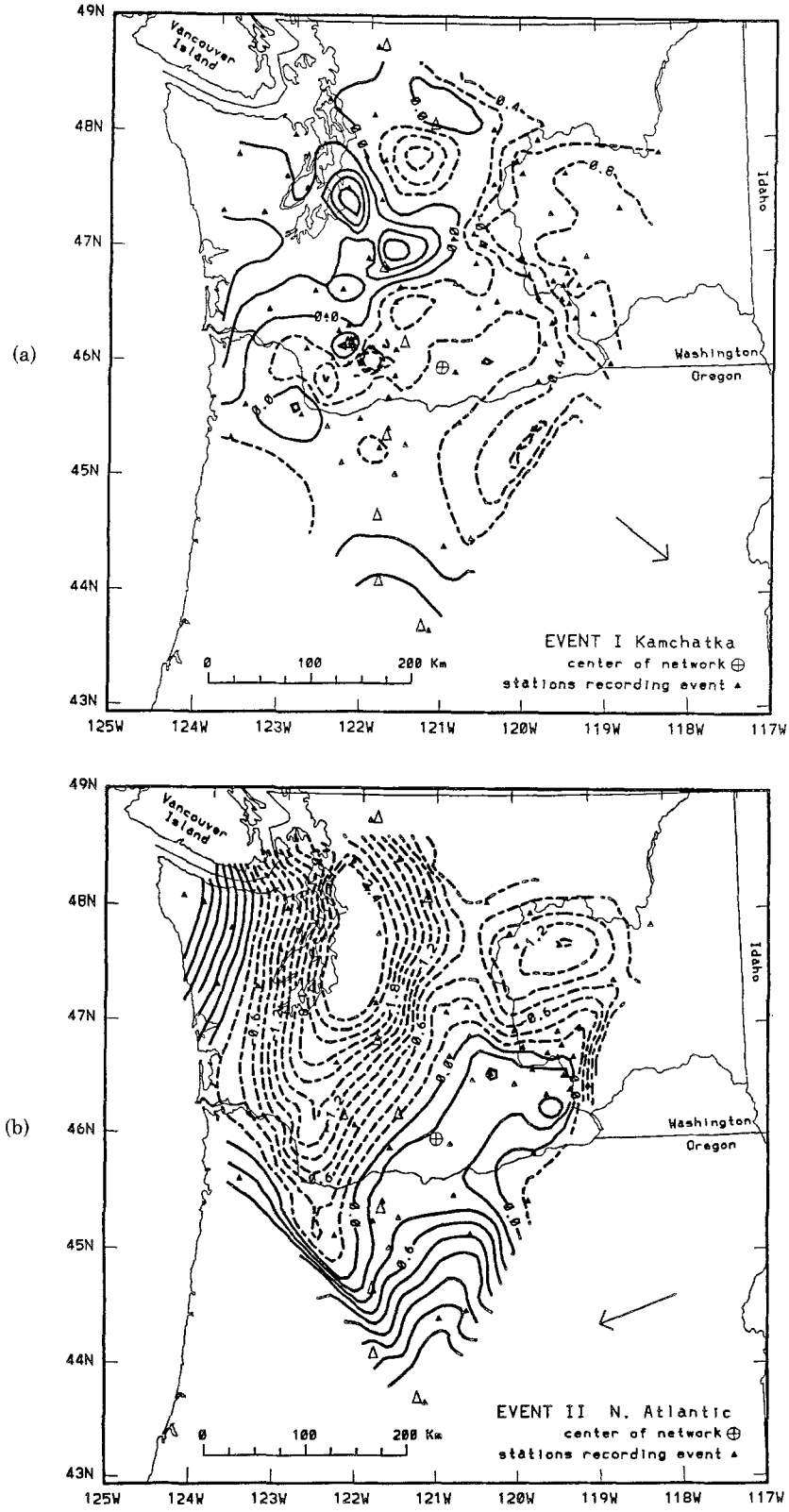
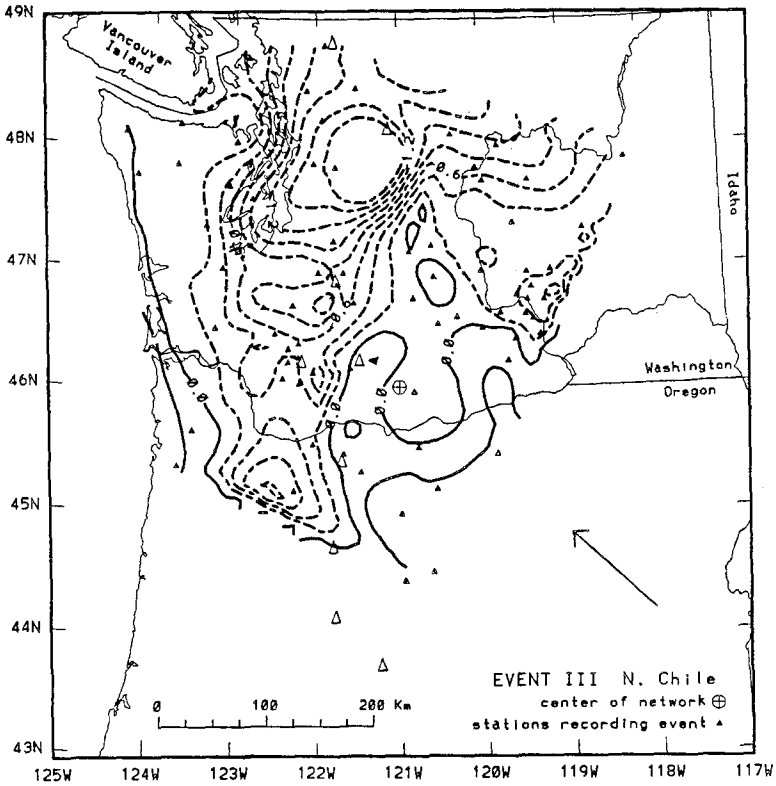


FIG. 10. Contour maps of P phase relative arrival-time residuals for (a) event 1, azimuth = 311°, Δ = 51°, (b) event 2, azimuth = 70°, Δ = 63°, (c) event 3, azimuth = 134°, Δ = 81°, and (d) event 4, azimuth = 232°, Δ = 76°. Residuals are the zero-mean difference between the least-squares optimized relative arrival times and those predicted from PDE locations and the Preliminary Reference Earth Model (PREM) (Dziewonski and Anderson, 1981). Contour interval is 0.2 sec with dashed lines representing negative contours. See Table 1 for additional event parameters.

(c)



(d)

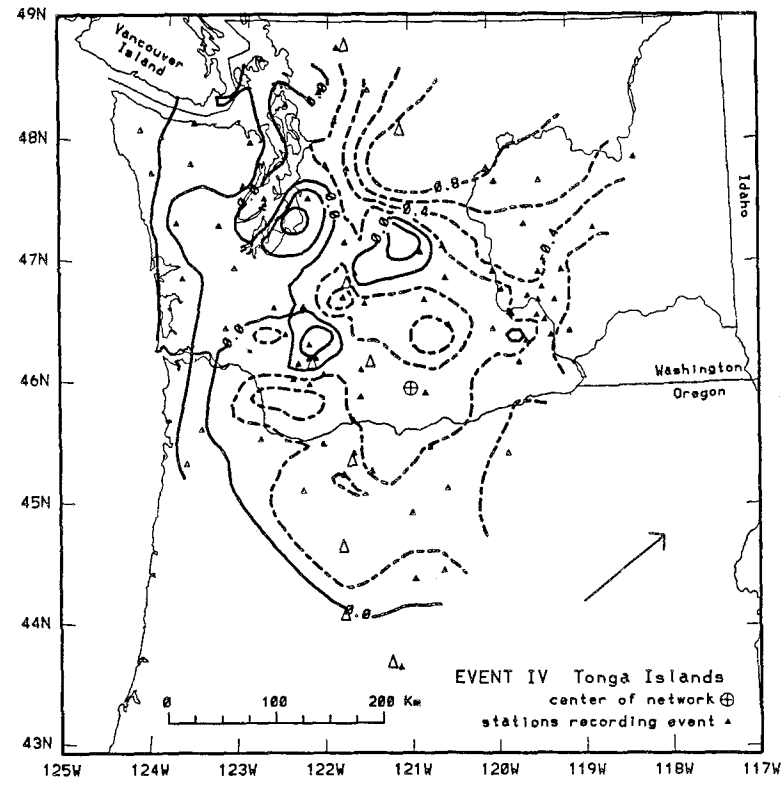


FIG. 10. (*Continued*).

Conclusions

We have developed a procedure which illustrates the potential for simultaneously obtaining highly accurate teleseismic arrival times over regional networks and reliable, quantitative estimates of their associated rms timing uncertainties. The method uses least squares to take advantage of the nearly normal distribution of errors in the cross-correlation derived lag times between all possible pairs of stations. We have found that these optimized relative arrival times remain virtually unaffected by either high-frequency noise or moderate amounts of waveform distortion (low-frequency noise). This robustness suggests that these waveform distortions influence cross-correlation functions in a random manner over scale lengths of tens to hundreds of kilometers. To illustrate this characteristic, we have tested the dependence of waveform coherence and residuals to the least-squares solution on station separation and found no correlation for the Washington Regional Seismograph Network. This finding corroborates studies done with other networks.

We would like to reiterate here that the arrival times that we measure with our procedure are those of the major phase arriving in each time window. If in fact an earlier phase is present, it will not necessarily represent the measured arrival time unless it is sufficiently separated from the major phase and of sufficient amplitude. This is a problem with any method of arrival-time estimation for teleseismic data and should caution users of travel-time data to make sure that both the amplitude of the earliest phase is indeed sufficiently large and its arrival sufficiently separated from other phases to have been picked.

While the precision of manual or single-trace automated picking allows the same gross features in the data to be observed, the increased accuracy of cross-correlation derived arrival times allows subtler details to be resolved. This increased accuracy along with the addition of quantitative uncertainty estimates may in fact be crucial to the convergence of tomographic inversions when nonlinear iterative methods are required. As with previous studies, the anomaly patterns which we obtained for the Pacific Northwest are strongly dependent on azimuth, indicating that lateral heterogeneities extend to considerable depth. These anomalies range from a few tenths of a second up to several seconds depending on azimuth and distance of the source. The increased accuracy and quantitative error estimates which we have obtained with this method should allow for greater resolution and decreased uncertainty in subsequent tomographic inversions using teleseismic delay times.

Acknowledgments

The authors would like to thank David Boore and an anonymous reviewer for helpful comments which improved the quality of this manuscript. This research was supported by USGS contract 14-08-0001-G-1390.

References

Achauer, U., L. Green, J. R. Evans, and H. M. Iyer (1986). Nature of the magma chamber underlying the Mono Craters Area, eastern California, as determined from teleseismic travel time residuals, *J. Geophys. Res.* **91,** 13,873–13,891.

Aki, K., A. Christoffersson, and E. S. Husebye (1977). Determination of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res.* **82,** 277–296.

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bull. Seism. Soc. Am.* **68,** 1521–1532.

Archambeau, C. B., J. C. Bradford, P. W. Broome, W. C. Dean, E. A. Flinn, and R. L. Sax (1965). Data processing techniques for the detection and interpretation of teleseismic signals, *Proc. IEEE* **53**(12), 1860–1884.

Bungum, H., E. S. Husebye, and F. Ringdahl (1971). The NORSAR array and preliminary results of data analysis, *Geophys. J. Roy. Astr. Soc.* **25,** 115–126.

Bungum, H. and E. S. Husebye (1971). Errors in time delay measurements, *Pure Appl. Geophys.* **91**(VIII), 56–70.

Crosson, R. S. and D. H. Hesser (1983). An algorithm for automated phase picking of digital seismograms from a regional network, *EOS* **64,** 775.

Dziewonski, A. M. and D. L. Anderson (1981). Preliminary reference earth model, *Phys. Earth Plan. Interiors* **25,** 297–356.

Hesser, D. H. (1982). Linear predictive processing and pattern recognition for automated classification of local seismic events, *M.S. Thesis*, Univ. of Wash., Seattle, 240 pp.

Humphreys, E. D., R. W. Clayton, and B. H. Hager (1984). A tomographic image of mantle structure beneath southern California, *Geophys. Res. Lett.* **11,** 625–627.

Husebye, E. S. and B. Jansson (1966). Application of array data processing techniques to the Swedish Seismograph Network, *Pure Appl. Geophys.* **63,** 82–104.

Iyer, H. M., J. R. Evans, G. Zandt, R. M. Stewart, J. M. Coakley, and J. N. Roloff (1981). A deep low-velocity body under the Yellowstone caldera, Wyoming: delineation using teleseismic *P*-wave residuals and tectonic interpretation, *Geol. Soc. Am. Bull.* **92**(Part I), 792–798.

Jansson, B. and E. S. Husebye (1968). Application of array data processing techniques to a network of ordinary seismograph stations, *Pure Appl. Geophys.* **69**, 80–99.

Mack, H. (1969). Nature of short-period *P*-wave variations at LASA, *J. Geophys. Res.* **74**, 3161–3170.

Menke, W. (1984). *Geophysical Data Analysis*: *Discrete Inverse Theory*, Academic Press, Orlando, Florida, 260 pp.

Michaelson, C. A. and C. S. Weaver (1986). Upper mantle structure from teleseismic *P* wave arrivals in Washington and Northern Oregon, *J. Geophys. Res.* **91**, 2077–2094.

Nakanishi, I. and K. Yamaguchi (1986). A numerical experiment on nonlinear image reconstruction from first arrival times for two-dimensional island arc structure, *J. Phys. Earth* **34**, 195–201.

Nolet, G. (1987). Seismic wave propagation and seismic tomography, in *Seismic Tomography*, G. Nolet (Editor), Reidel, Dordrecht, The Netherlands, 1–23.

Rasmussen, J. R. and E. D. Humphreys (1988). Tomographic image of the Juan de Fuca plate beneath Washington and western Oregon using teleseismic *P*-wave travel times, *Geophys. Res. Lett.* **15**(12), 1417–1420.

Snedecor, G. W. (1956). *Statistical Methods*, Iowa State Press, Ames, Iowa, 173–180.

GEOPHYSICS PROGRAM, AK-50
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195

Manuscript received 17 August 1988