CONF- 9609/33--1

# Earthquake Early Warning System Using Real-time Signal Processing

Richard R. Leach, Jr.
Farid U. Dowla

This paper was prepared for submittal to the
1996 IEEE Workshop on Neural Networks for Signal Processing
Keihanna, Kyoto, Japan
September 4-6, 1996

February 1996
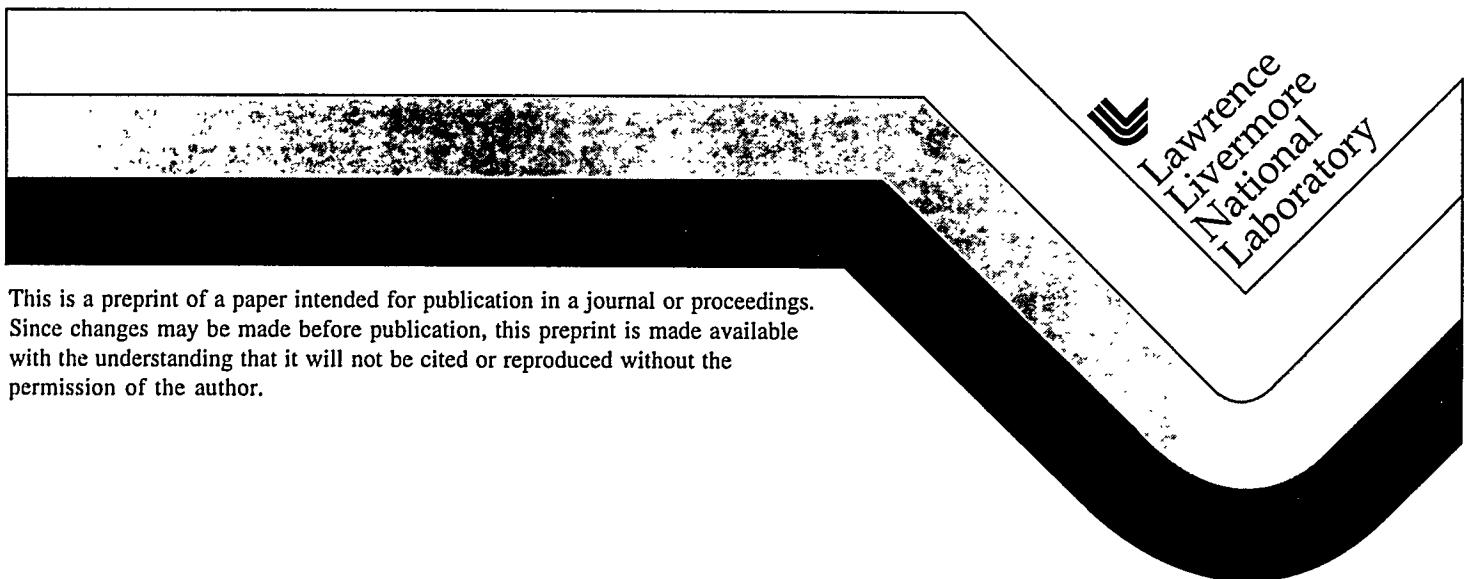
Lawrence Livermore National Laboratory

MASTER

DISCLAIMER

DISCLAIMER

Portions of this document may be illegible
in electronic image products.   Images are
produced from the best available original
document.
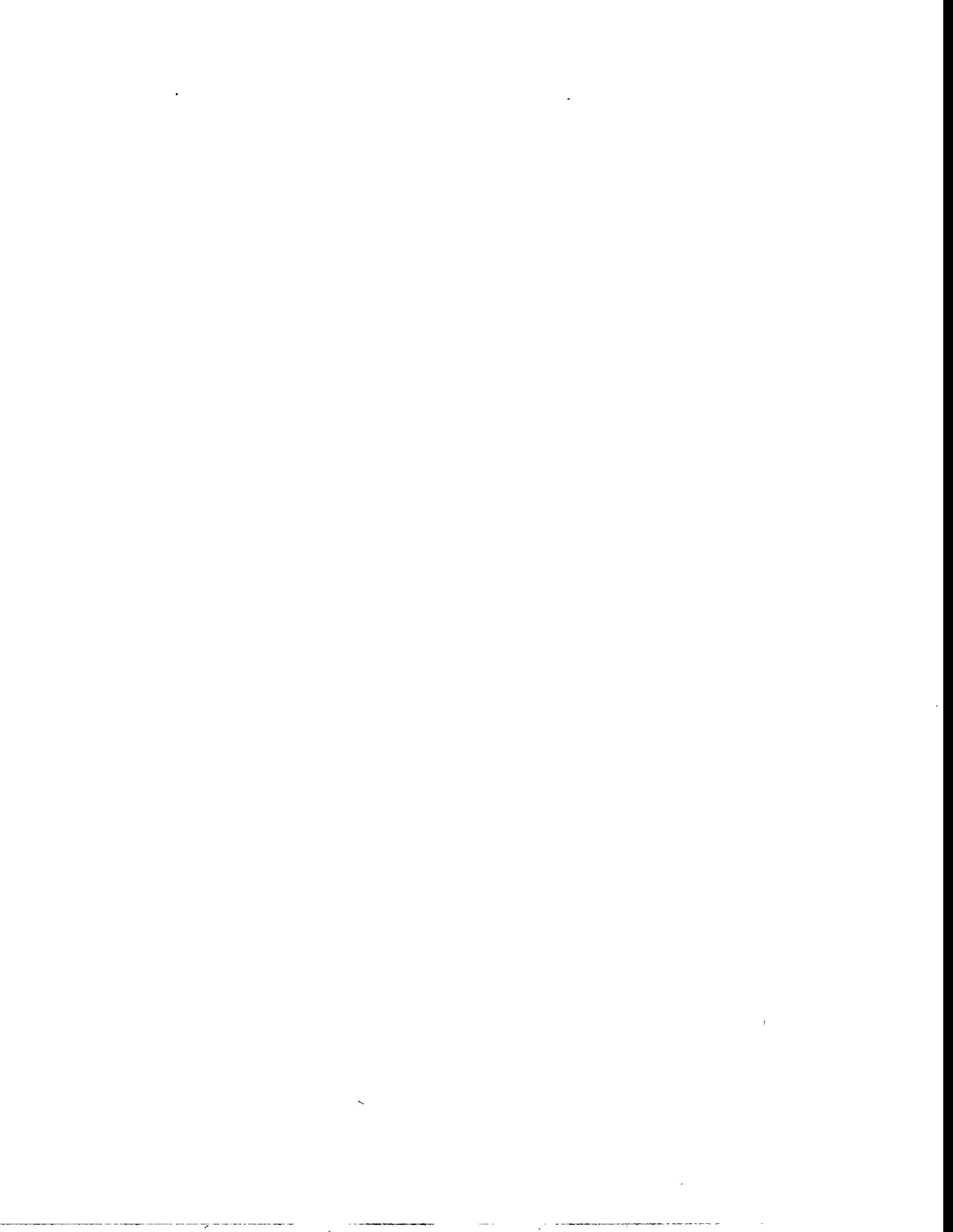
Earthquake Early Warning System using Real-time Signal Processing

Richard R. Leach Jr.
Farid U. Dowla

Lawrence Livermore National Laboratory
University of California, Livermore, CA 94550

Abstract

An earthquake warning system has been developed to provide a time series profile from which vital parameters such as the time until strong shaking begins, the intensity of the shaking, and the duration of the shaking, can be derived. Interaction of different types of ground motion and changes in the elastic properties of geological media throughout the propagation path result in a highly nonlinear function. We use neural networks to model these nonlinearities and develop learning techniques for the analysis of temporal precursors occurring in the emerging earthquake seismic signal. The warning system is designed to analyze the first-arrival from the three components of an earthquake signal and instantaneously provide a profile of impending ground motion, in as little as 0.3 sec after first ground motion is felt at the sensors. For each new data sample, at a rate of 25 samples per second, the complete profile of the earthquake is updated. The profile consists of a magnitude-related estimate as well as an estimate of the envelope of the complete earthquake signal. The envelope provides estimates of damage parameters, such as time until peak ground acceleration (PGA) and duration. The neural network based system is trained using seismogram data from more than 400 earthquakes recorded in southern California. The system has been implemented in hardware using silicon accelerometers and a standard microprocessor. The proposed warning units can be used for site-specific applications, distributed networks, or to enhance existing distributed networks. By producing accurate, and informative warnings, the system has the potential to significantly minimize the hazards of catastrophic ground motion. Detailed system design and performance issues, including error measurement in a simple warning scenario are discussed in detail.

Introduction

Providing a reliable and informative warning on the order of a few seconds to a minute before hazardous ground motion arrives could significantly reduce property damage as well as injuries and lives lost. One application of such a warning is in the implementation of automatic shutdown systems. These systems would benefit government and industrial facilities, hospitals, schools, police and fire, emergency services, and private homes. They would result in safer handling of hazardous materials, minimal interruption of power, water, or communication systems, safer transportation systems (including trains and military and civilian aircraft and spacecraft), and vehicle control at bridges and overpasses. The warning could also trigger, where appropriate, public announcements (e.g. 'duck and cover') or evacuation response messages.

Most of the research and development in earthquake warning systems has been applied to regional seismic networks. In 1989, about 50 different regional networks with a total of over 1500 seismic sensors were located within the United States (National Research Council, et. al., 1990). The contributions of these types of networks have been significant, and the promise of using these networks in a true real-time warning effort is great. They have the potential to provide the optimum warning time for impending hazardous ground motion. However, there are several obstacles associated with implementing a real-time earthquake warning system using regional networks. Much of the instrumentation in place today is inadequate because of low dynamic range, narrow bandwidth, and one component recording (National Research Council, et. al., 1991). A host of developmental challenges must be met to transmit such a large volume of network sensor information to a central computer to perform real-time processing on the data. A working system will require significant initial capital and extensive long-term maintenance

This study focused on a complimentary approach to real-time earthquake warning using a site-specific paradigm. In other words the first arriving signal from an impending earthquake is analyzed at a single site where sensors measure the three components of ground motion (east-west, north-south, up-down). The ground motion is processed at the site in real-time and a continuous estimate of hazardous ground motion, which has not yet arrived at the site, is provided. This approach to real-time earthquake warning presents several advantages, foremost being its versatility. It could be implemented at one location, for example next to a nuclear reactor, and could provide usable, reliable, warning time for incoming hazardous

ground motion. Determining the location of the earthquake focus is not of primary importance, since it is the shaking of the ground arriving at the detection site that is being analyzed, regardless of the direction or distance of the earthquake's origin. Development costs are low, since the system can be easily implemented as one complete unit on a personal computer or integrated into an imbedded system. In addition, the system can potentially replace each element in a regional network for increased warning times as well as enhanced reliability.

Often, what is termed a real-time or near real-time system provides warning after transmission delays and extensive processing of sensor data. This can shorten warning time by many seconds up to several minutes. We define a "real-time" system in three ways. First, the system must provide useful estimates of any incoming hazardous earthquake, regardless of distance or direction. Second, estimates must begin within 1 second after the first arrival of earthquake ground motion for any particular earthquake. Third, warning updates must be provided at least 10 times per second as the earthquake emerges. The neural network based system described in this paper provides estimates within 0.3 sec after the first arrival occurs and is updated every 0.04 sec.

For the purpose of this study, we identify two major areas of earthquake signals: the primary arrival and the secondary arrival. In many earthquakes these arrivals are felt as an initial jolt followed by another shock seconds later. Each of these two arrivals consist of many individual elastic waves that have traveled from the earthquake focus or source location to recombine at the recording site as a function of their respective velocities, focal distances, and propagation paths. Body waves propagate within a body of rock and appear in the first arrival. The fastest of these body waves, the primary, or P-wave, is the first elastic wave to reach the recording site. The secondary arrival contains body and surface waves such as S, Rayleigh, and Love waves. These later arriving waves often produce both horizontal and vertical ground motion whose peak velocities, peak accelerations, and duration of time over which the shaking occurs often tend to cause significant damage to structures (Bolt, 1978). One type of these later arriving waves, the regional surface wave (L_g and R_g), is composed of a family of trapped waves that travel through the earth's low velocity crust. The crust acts as a waveguide that provides little attenuation loss. The interaction of these waves as they propagate can produce considerable amplification of ground motion near the surface. This is called the *free surface effect* and occurs when upgoing and downgoing reflected waves are in phase and of considerably greater wavelength than the thickness of the

crust (Reiter, 1991). Figure 1 illustrates two different elastic wave propagation paths. A seismogram is also shown, noting primary and secondary arrivals.

Figure 1 **Elastic waves travel at different velocities and through different paths. The paths may be through propagation media with nonlinear characteristics. Primary and secondary arrivals in a recorded seismogram are also illustrated.**

Our results indicate that the first moments of the seismic activity from an emerging earthquake can be used to estimate the size and duration of later arriving waves. The P-wave segment is the most direct arrival and is rich in information about the earthquake source function and the characteristics of the elastic wave's propagation path. For example, an estimate of the source azimuth can be obtained from the P-wave segment of a three-component seismogram signal. In addition, the integral of P-wave amplitudes will often correlate to the magnitude of the earthquake. This was seen in previous work, where a neural network was used to show that the first 6 sec of seismic signal from underground nuclear explosions produces accurate estimates of the explosion magnitude (Leach, et. al., 1993). Also, since anelastic attenuation will selectively damp out higher frequencies as the wave propagates, spectral characteristics of the seismic signal can be used to produce an estimate of the source distance. Details of these and other parameters will be discussed in a later section.

Depending on the distance between the earthquake hypocenter and the seismograph, the first arriving energy could provide up to tens of seconds of warning at the detection site. The length of warning time varies as a result of the different velocities at which the elastic wave energy travels. For example, using a L_g velocity of 3 km/sec, a P_g velocity of 6 km/sec, and an hypocentral distance of 20 km from the detection site, the L_g effects will be felt about 4 sec after the onset of the earthquake. If the detection system is designed to alarm at the L_g arrival within 1 sec after the onset of the earthquake, about 3 sec of warning time is available. At a distance of 100 km, the warning time increases to about 15 sec. Using later arriving warning parameters such as the peak ground acceleration (PGA) will increase warning times. Figure 2 illustrates warning times from the primary wave arrival to the PGA for the earthquakes used in this study. Distances varied from 15 to 275 km from the earthquake source to the seismometer.

Figure 2 **Warning time varies with respect to distance. 434 earthquake recordings are plotted illustrating the warning time measured from the P-wave arrival to the peak ground acceleration (PGA) of each earthquake. For**

example, at a distance of 100 km, there is approximately 15 sec of warning time between the earthquake onset and the arrival of the PGA.

Many site-specific warning systems that are in use today wait for ground motion measurements to exceed a preset threshold. Simple thresholding almost always results in serious false alarm issues. If the threshold is set too low, a large number of false alarms rapidly destroy the credibility of the system. If the threshold is set too high, missed or underestimated alarms offer a potentially disastrous false sense of security. Glitches and noise occurring in the primary wave can also produce false alarms; to avoid this, some systems set the threshold to trigger on the later arriving S waves, but this often leaves little or no real-time warning at the detection site. Other systems measure and report some portion of the total energy of the emerging signal, taking advantage of the positive correlation between first- and second-arrival energy. This approach in itself is weak due to the nature of the earth and its non-linear propagation paths. The relationship between first and second arrival often exhibits large errors when modeled as linear. More sophisticated systems such as Japan's Urgent Earthquake Detection and Alarm System (UrEDAS), compile several estimates from measured parameters within the P-wave. The results of this work have produced good results (Bito and Nakamura, 1986), but linear methods of mapping particular measurements in the P-wave to incoming hazardous ground motion may exhibit unacceptable errors.

Artificial neural networks (ANN) offer a promising new approach that avoids many of the limitations of present systems. Neural networks are non-linear systems that use parallel processing units and have proven to be extremely powerful in the analysis of complex patterns and signals (Dowla, et. al. 1990; Hopfield 1988). Instead of using heuristic algorithms, the ANN can adoptively "learn" the nonlinear function of the paths of the primary and secondary waves from the sensor to the earthquake source. A neural network can be "trained" with either recorded or synthesized seismograms. Using this approach, we can quantitatively evaluate candidate training parameters. The neural network's use of nonlinear functions makes these systems attractive, because nonlinearities are common in the real world. ANN's are stable, they are able to generalize (Hopfield 1987), and are less sensitive to noisy and spurious input. Most importantly, once trained, neural networks are compact and operate at speeds that make them ideal for real-time applications. They can easily be implemented in a compact integrated circuit. The high-speed, miniaturized capability makes ANNs attractive for a large distributed system with many sensors as well as a low-cost single user approach, such as for a

5

typical homeowner.

In this study, we trained a neural network with >400 earthquake seismic signals such that when parameters from a first arrival segment of a previously unseen earthquake seismic signal were input, a profile and total energy estimate of the complete seismogram was produced. The neural network "learned" the nonlinear relationship between the source and path characteristics of early arriving (the P wave in particular) and later arriving elastic waves of an earthquake seismic recording. Minimally, the output of the system described here provides a time series profile from which three vital parameters can be derived: the time until strong shaking starts, the magnitude of the shaking, and the duration of the shaking. The neural-network-based algorithm can easily be implemented in digital electronic hardware for a single-station warning system or as an element in a distributed regional warning system. The theoretical basis for our approach follows.

Theoretical Basis: Linear System Model

In this section, we describe the theoretical basis of our method by presenting a grossly simplified linear system model. We assume large distances ("far field") and a homogeneous, isotropic, unbounded medium with a source acting with general time-varying strength producing a unidirectional force originating at a particular point. The earthquake source function is represented by $f(t)$. Considering the linear propagation case, we can represent the received primary arrival or P-wave signal $p(t)$ by the convolution,

$$p(t) = f(t) * h_p(t), \tag{1}$$

where $h_p(t)$ represents the propagation Green's function (impulse response) for the primary arrival or P-wave segment in the seismogram (Aki, et. al., 1980). Similarly, for the segment of the seismogram following the initial P wave, or the secondary arrival $s(t)$, we can write,

$$s(t) = f(t) * h_s(t), \tag{2}$$

where $h_s(t)$ represents the Green's function of the secondary arrival or post-P-wave propagation waveform. From equations (1) and (2), it can be shown that,

6

$$s(t) = p(t) * h_p^{-1}(t) * h_s(t) = p(t) * g(t),$$ (3)

where

$$g(t) = h_p^{-1}(t) * h_s(t),$$ (4)

and

$$h_p^{-1}(t) * h_p(t) = 1.$$ (5)

From the above, if it is possible to approximate $g(t)$ using $p(t)$ from a set of representative earthquakes, or if it is possible to associate classes of $g(t)$'s to $p(t)$'s, then, from equation (3), we can estimate $s(t)$, the secondary arrival or post-P-wave signal, from $p(t)$, the primary arrival P-wave signal. By training with temporal precursors within the primary arrival $p(t)$ for many different earthquakes, a neural network learns a general rule for equation (4) and is able to solve, within some system error for $s(t)$ in equation (3) given a previously unseen $p(t)$.

Intensity Measures from Amplitudes of the First Arrival

The selection of the temporal precursors or measures within the primary arrival is vital to successful neural network training. Measures that relate to earthquake intensity, magnitude, and distance have the potential of producing a useful forecast of later arriving, damaging waves. The identification of these parameters within the primary arrival is one the most challenging and important aspects of this research. The questions we now ask are: what primary signal measures might be useful, and how well do they correlate with their secondary signal counterparts?

Numerous parameters have been used to measure earthquake damage potential, such as PGA. PGA is given by the maximum value reached within an acceleration record $a(t)$, or:

$$PGA = max\{a(t)\}.$$ (6)

Several forms of energy measurements from seismograph recordings have also

been used. In this paper, we will refer to the energy of a signal as a general measurement E using the equation,

$$E = \int_{t_s}^{t_e} x^2(t)\, dt.$$
(7)

where E refers to the energy in the signal $x(t)$ from time t_s to t_e.

One such damage parameter is named the ARIAS intensity (Arias, 1970) and is an earthquake energy measurement of the form (7), where $x(t)$ is replaced by the acceleration record $a(t)$, and the equation is multiplied by a constant. For the earthquakes used in this study, the origin time t_0 of the earthquake is used as the reference time and is set at zero. t_s denotes the first arrival of an earthquake signal and is the time between t_0 and the first arrival of earthquake shaking at the seismometer site. t_e denotes the final arrival of an earthquake signal and is the time between t_0 and the end of the earthquake shaking at the seismometer site. Cumulative absolute velocity (CAV) is a similar parameter that uses the earthquake's acceleration recording $a(t)$, and is defined as:

$$CAV = \int_{t_s}^{t_e} |a(t)|\, dt.$$
(8)

Another damage potential measurement is the peak spectral amplitude (PSA), which is defined as the maximum value of the Fourier transform of the acceleration recording $a(\omega)$ or $max\{a(\omega)\}$. The Fourier transform of a general signal $a(t)$ is given by,

$$a(\omega) = \int_{t_s}^{t_e} a(t)\, e^{-j\omega t}\, dt.$$
(9)

The Electrical Power and Research Institute (EPRI) reviewed these and other parameters that have been used to assess the damage potential of ground motion (Reed, et. al., 1988). Using regional strong motion recordings, eleven damage measures such as the PGA and peak PSA were ranked as per their ability to measure earthquake damage potential. EPRI found that the ARIAS intensity and the CAV proved to be leading measures for earthquake strong motion hazard and intensity measurements

To explore the usefulness of some of the damage measures, we would like to

8

define global and local measures within an earthquake signal. Figure 3 illustrates several representations of one component of an earthquake signal, including the velocity recording, the velocity recording differentiated to acceleration with the primary and secondary arrival noted, and the fast Fourier transform of the primary and secondary arrival portions of the acceleration signal. Primary and secondary wave signals are denoted as group velocity windows $a_p(t)$ and $a_s(t)$. The primary signal is defined a time window whose duration spans from the first arrival of the earthquake signal (t_s) to $\Delta/3.6$, where Δ is the distance from the hypocenter to the recording station. The constant 3.6 was chosen to force $a_p(t)$ to end shortly before the L_g phase arrival. The remaining portion of the signal is the secondary signal $a_s(t)$.

Figure 3 **Several plots from an earthquake signal, including its velocity recording (top), the same signal differentiated to acceleration with the primary $a_p(t)$ and secondary $a_s(t)$ signals noted (middle), and the fast Fourier transform of the primary and secondary signals of the acceleration signal (bottom). Some potential ANN training measures are denoted with arrows. Local peak ground acceleration is shown by PGA$_p$ for primary and PGA$_s$ for secondary PGA. Local peak spectral amplitude is shown by PSA$_p$ for primary and PSA$_s$ for secondary PSA.**

Note that the damage measures previously discussed are global measures: that is, the entire record is used in their derivation. The duration of the regional seismograph records used in this work generally do not exceed 2 to 3 minutes in length. We now introduce similar measures to the ones described above, which are local in scope. They are derived from a windowed portion of the earthquake record, specifically during primary or secondary arrival. As shown in Figure 3, we denote local measures such as PGA and PSA, within the duration of the primary and secondary wave. For each earthquake used in this study, a local PGA$_p$, PSA$_p$, and CAV$_p$ were calculated for the primary wave as described earlier. Similar local calculations for PGA$_s$, PSA$_s$, and CAV$_s$ were taken for the secondary wave. Scatter plots of primary vs. secondary measures are illustrated in Figure 4.

Figure 4. **Scatter plots of local measures in the primary wave against similar local measures in the secondary wave are shown for PGA, PSA, and CAV. Correlation coefficients are also indicated. The parameter that had the strongest correlation is the cumulative absolute velocity measure.**

These plots show a positive correlation between primary and secondary arrivals; the CAV having the strongest correlation. This correlation occurs because

primary and secondary arrival share similar source, site, and path effects. These results indicate that the linear system model described earlier is somewhat viable in that there is a positive correlation between the primary and secondary measures. Though each measure in itself produces errors that make it impractical as a warning system, a nonlinear, multi-dimensional system might be designed using these data collectively to form a nonlinear mapping between primary and secondary signals. We take advantage of this result by using some of these measures as training input to the neural network.

Distance Measures from the Spectrum of First Arrival

Spectral amplitude from the first arrival alone provides a poor indicator of distance. For any given frequency, it is difficult to assess whether the amplitude is due to a small earthquake nearby or a large earthquake at a farther distance. The increasing area of the wavefront or geometric spreading contributes to wave attenuation, yet is a frequency invariant process (Kulhanek, 1990). Losses at interfaces (diffraction, scatter, mode conversion, reflection, and refraction) also contribute to the attenuation of the waves.

One solution lies in the fact that the media is not perfectly elastic; therefore, spectral wave amplitudes diminish from internal friction resulting from the inelastic property of various types of rocks and material grains. This anelastic attenuation contributes to losses in a frequency-dependent manner where the highest frequencies exhibit the greatest attenuation. Thus, anelastic attenuation can provide information by which the neural network can form a magnitude-independent distance estimate. To illustrate how this might occur, consider wave amplitude A as a function of distance travelled x and frequency in radians ω or:

$$A(x, \omega) = A_0 \exp\left[-\frac{\omega x}{2cQ}\right].$$

(10)

where c is the phase velocity, Q is a dimensionless measure of internal friction, and A_0 is the value of $A(0)$ (Aki, et. al., 1980). It is clear from equation (10) that higher frequencies will be attenuated to a greater degree than lower frequencies. By using a sampling of spectral values for each earthquake as training input, the neural network can be forced to learn this relation. For instance, the network can learn a relationship between spectral amplitude values and distance measures such as duration or location of the PGA in an emerging seismic signal.

10

To illustrate this concept, we conducted the following experiment. A group velocity window was defined from the arrival time of the earthquake signal t_s to $t_s +$ 1 for >400 earthquake seismograms. A Bartlett spectral estimate $b(\omega)$ was calculated from this signal to determine if a spectral representation of as little as one second of emerging earthquake signal can contain evidence of anelastic attenuation. High and low-frequency measurements of spectral energy bands were calculated using equation (7) by replacing $x(t)$ with $b(\omega)$. For the low band B_l, limits of integration of 0 to 1.5 Hz were used. For the high band B_h, limits of integration from 6.5 to 8.0 Hz were used. Figure 5 depicts the boundaries of B_h and B_l as well as the Bartlett spectral estimate for two example earthquakes. One earthquake is at distance 213 km and magnitude M_l of 4.2. The other is at distance 50 km and M_l of 3.0. In this example the discriminant between the two earthquake distances for the higher frequency energy relative to the lower frequency energy is due to the slight difference in magnitude as well as the effects of anelastic attenuation.

Figure 5 Bartlett spectral estimates for two earthquakes illustrating two frequency bands used to form a ratio of low to high frequency energy. Spectral falloff results somewhat from the differences in the magnitudes as well as from anelastic attenuation where the ratio of the integral of energy in the two frequency bands is relatively larger for the more distant earthquake.

Using B_h and B_l, a ratio R was formed by:

$$R = log_{10}\left(\frac{B_l}{B_h}\right).$$
(11)

In equation (11), R will be low for earthquakes at closer distances and high for earthquakes at farther distances due to amplitude attenuation as per equation (10). Also, R is a dimensionless measure in which amplitude information cancels, emphasizing that frequency characteristics alone contain distance related information.

Using equation (11), R was calculated for >400 earthquakes. The results are illustrated in Figure 6. A scatter plot of R values is plotted against the hypocentral distances of the same earthquakes. The results show a correlation between the R values and their respective distances. Although the correlation is weak, this is an important result because it offers evidence that a simple spectral ratio can provide distance information while constrained to a signal taken moments after the onset an

earthquake. Knowing the earthquake distance is a vital feature in a reliable warning system. The wide scatter figure 6 appears daunting, but consider that the input to the neural network includes more than a simple ratio. The complete sampled spectrum is used as one feature of the neural network input. In this manner, the neural network is free to use R or any function of one or more frequency bands to map distance dependent parameters, such as duration, from the inputs to the output waveform estimate. Furthermore, though anelastic attenuation's frequency-damping properties are fairly well known, the neural network is free to learn any frequency variant relation that might exist in the training inputs, such as the velocity dispersion of Love and Rayleigh waves.

Figure 6   Scatter plot of R values versus hypocentral distances for >400 earthquakes. The overall positive correlation illustrates anelastic attenuation as a function of distance. The superimposed line is a least squares fit of the data. The wide variation occurs because the range of earthquake magnitudes for the various distances were plotted simultaneously.

Choosing a Neural Network

In general, ANNs have been successfully used to develop nonlinear, nonparametric, calibration systems, as well as signal discriminators from ground motion signals (Dowla, et. al., 1990). They also have been used to do real-time classification of signals from three-component seismic sensors (Bowman, et. al., 1992) as well as class identification of seismic activity (Maurer, et. al., 1991). In addition, the results from a neural network showed that the first few seconds of a seismic signal from a nuclear explosion can be used to predict the magnitude of that explosion (Leach, et. al., 1993). Since the ultimate goal of this work is to develop a system for the detection and early warning of earthquakes, a neural-network-based system is a good choice, not only because of its successful heritage, but also because of its speed and simplicity of implementation.

We considered two types of ANNs in this study: the self-organizing Kohonen network and the backpropagation network. Because of the many different types of ANNs, an adequate discussion of the taxonomy and applications is beyond the scope of this paper. For a more complete treatment of ANNs, interested readers are referred to a journal article by Lippman (1987), and books written by Pao (1989) and Hertz et al. (1990).

Due to the sparsity of readily available training data, the backpropagation network was chosen in this application. The backpropagation network is a powerful

tool for learning nonlinear mapping from the input to the output. More importantly though, a backpropagation neural network can readily interpolate as well as extrapolate missing information. This is vital for seismic data, because very large magnitude earthquake recordings are naturally scarce. Also, a generalized training set, which is well-sampled with respect to magnitudes, locations, and geophysical parameters, is often unavailable.

The three-layer perceptron or backpropagation neural network (Lippman, 1987) was used in this study with the conjugate gradient error search method (Johansson, et. al., 1990) for training time reduction and training stability. The three-layer back-propagation network can form arbitrary complex decision regions and can separate meshed classes. The complexity and apparent overlap of features within the P-wave segments made this an attractive option.

A representation of the system is shown in Figure 7. During the training phase, the neural network is presented with the first arrival of a seismogram. The error is defined as the difference between the output of the network and the complete seismogram signal. The internal weights of the network are adjusted to reduce this error. This process is repeated until the error is lowered sufficiently, and the network has learned to generalize the nonlinear relationship between the input and output. During the testing phase, the neural network is presented with first arrival data from previously a unseen seismogram. The neural network estimates the complete seismogram by weighing the first arrival inputs with internally stored information learned during the training phase.

Figure 7. First neural networks are trained to learn the relationship between the first arrival and the complete seismogram of >400 different seismograms. Then the neural network can use the first arrival of a new seismogram to estimate the complete seismogram.

Database

The earthquakes used in this study were recorded by the Lawrence Livermore National Laboratory (LLNL) seismic observatory station at Landers, California. Landers is located at 34.39° latitude by −116.41° longitude. Using United States Geological Society (USGS) records, all earthquakes were located that occurred in the region surrounding Landers during the period from January 1, 1988 to December 31, 1992 as shown in Figure 8. A window bounded by 32.2° to 36.5° latitude to −118.8° to −113.8° longitude was defined. This window defines an area approximately 300 km from Landers.

Figure 8 Earthquakes that occurred from 1988 to 1992 surrounding Landers, CA. This study focused on earthquakes at distances from 20 to 294 km from Landers recording station.

The training data set for the neural network consisted of earthquake recordings with good (>2) signal-to-noise ratio, yet did not saturate the instrumentation. The criteria for each earthquake candidate were (1) that the radial distance from Landers to the epicenter be greater than 20 km, and (2) that the local magnitude divided by the log of the radial distance be greater that 1.5. This excluded earthquakes with poor signal to noise ratio. It also excluded earthquakes occurring at distances that were too close to provide useful warning time. Earthquake recordings that saturated the seismograph or contained dropouts or glitches were removed. This resulted in a final set of over 400 earthquakes.

System Output Design

We approached this problem with the idea of using only data from seismogram records. Though it is expected that a system that augments the input with other known earthquake parameters, such as medium or source location, would theoretically improve performance and produce a more accurate earthquake profile, we wanted to isolate what an artificial neural network can "learn" about earthquake seismograms when trained with seismic recordings only. Producing acceptable results using this data-driven approach provides a simple, reliable system that could be easily implemented.

We represent a three-component seismogram by a signal:

$$c(t) = [c_e(t), c_n(t), c_z(t)], \tag{12}$$

where t represents time, and $c_e(t)$, $c_n(t)$, and $c_z(t)$ represent east, north, and vertical channels of velocity recordings. The discrete time signal $c(n)$ is stored in computer memory and represents $c(t)$. Preprocessing the data included removing the mean, windowing the data with a Hamming cosine taper, and filtering each component with a seventh-order Butterworth bandpass filter with corner frequencies at 0.5 and 10 Hz. It has been shown that good yield estimation can be achieved in this frequency range (Leach, et. al., 1993). These filtered velocity signals $v_e(n)$, $v_n(n)$, and $v_z(n)$ were integrated to provide displacement signals $d_e(n)$, $d_n(n)$, and $d_z(n)$ and

were differentiated to provide acceleration signals $a_e(n)$, $a_n(n)$, and $a_z(n)$. For example, the filtered vertical signals from the three-component spatial motion records for one of the earthquakes are illustrated in figure 9. The time of first arrival of ground motion at Landers was marked in each record and is denoted by t_s.

Figure 9 Data example of earthquake velocity recorded at the Landers recording station showing east-west, north-south, and up-down motion. This particular event had a magnitude (m_l) of 3.2 and originated 122 km from Landers.

An envelope function was formed with the three components of the displacement signals using a Hilbert transform. In general, the envelope $e(n)$ of a signal $x(n)$ is defined by,

$$e(n) = \sqrt{x^2(n) + y^2(n)} \, , \tag{13}$$

where $y(n)$ is the Hilbert transform of $x(n)$. Equation (13) was applied to $d_e(n)$, $d_n(n)$, and $d_z(n)$ after removing the mean and taking the absolute value of each signal. A three-component incoherent beamforming of the three resulting signals represented the general size and shape or profile of the earthquake. The resulting envelope signals for several example earthquakes are illustrated in Figure 10.

Figure 10 Several preprocessed earthquake displacement envelope signals. Increasing amplitudes were chosen to illustrate the arrival of predominant shaking as a function of distance.

Envelope or profile signals as system output offer a useful, viable warning for imminent ground motion. A continuous broadband signal would be the ideal output, but a neural network would require a computationally prohibitive number of example earthquakes to train these types of signals successfully. We recognize that training the neural network with profile signals does remove the higher frequency characteristics of the earthquake signal but regardless, they are still a good choice since they will provide a viable warning and contain several important measures. First, the signals would provide a measure of energy that could be scaled to a magnitude or intensity value. sec, they would measure the time available before the onset of maximum ground shaking, for example the L_g phase arrival. Third, they would provide a measure of the duration of the large amplitude shaking. Many other designs for system outputs were considered, such as ground motion harmonics, distance, coda, and the complete point by point earthquake record.

Some of these ideas were eliminated because of limited data availability or limited neural network capabilities. Others will be used in future work.

Neural-network-training output signals were taken from envelopes of $d_e(n)$, $d_n(n)$, and $d_z(n)$ for each earthquake used in this study. The duration or window of the output signal was set at 60 sec. and sampled every 0.25 sec. This resulted in 240 training outputs, each representing a sample point in the 60-sec earthquake profile from arrival time t_s to $t_s + 60$. Two distinct output categories were derived from the output envelope: the earthquake scale and the earthquake profile or shape.

The first output category consisted of the earthquake scale which we now define. For each earthquake, a vector **O** was formed whose elements contained the 240 output envelope values. In general, the length or norm of any vector **X** containing N elements is given as,

$$\|\mathbf{X}\| = \sqrt{\sum_{n=1}^{N} (x_n)^2}, \tag{14}$$

With i representing the earthquake index, the length or norm $\|\mathbf{O}_i\|$ was calculated for all earthquakes using equation (14). The resulting norm was defined as the earthquake scale. The earthquake scales have a strong correlation to the energy and intensity of the earthquake signal.

The dynamic range of these scale values spanned many orders of magnitude. To provide a better scale distribution for neural network training, we normalize the set of scales as follows. In general, any vector **X** with n elements has both a global maximum (x_{max}) and minimum (x_{min}). A normalized set of data is formed when each element x in **X** is replaced by a normalized value x_n and is defined as,

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{15}$$

Equation (15) was applied to the log10 of each earthquake's norm to provide a normalized scale output or $\|\mathbf{O}\|_n$. The set of normalized scale values of $\|\mathbf{O}\|_n$ comprised the first category of system output.

The second category of system output is the shape or profile. Each output

vector **O** was normalized to length 1 using equation (16) to normalize the total energy of the earthquake and emphasize the shape of the envelope. A normalized earthquakes profile vector $\mathbf{O}_{ni}$ is formed by,

$$O_{ni} = \frac{O_i}{\|O_i\|} \, , \tag{16}$$

where $\|\mathbf{O}_{ni}\| = 1$.

Selection of Input Window Sizes

Neural network input parameters are derived from a primary window or segment using the three-component seismic signals $v_e(n)$, $v_n(n)$, and $v_z(n)$. Windows always begin at the earthquake arrival time t_s. Each earthquake can provide multiple training examples by defining any number of durations or window sizes Δt. For example, if a window duration is defined from t_s to $t_s + 1$ for the earthquake seismogram, then Δt equals 1 sec. The input parameters are then derived from this 1-sec segment of the seismogram. Using the same earthquake, but a different window size Δt of 3-sec, the input parameters are derived from this 3-sec segment of the seismograph. Many examples of each earthquake can be created for neural network training by choosing multiple windows in this manner. Note that for any given earthquake, although the input window durations vary, the training output is the same, again the 60-sec profile of the earthquake from t_s to $t_s + 60$.

In this study, six windows were defined, resulting in the generation of six training examples for each earthquake. The motivation of using multiple input windows was to provide a robust representation of the primary wave for each earthquake in the training set. The duration or Δt of each of the six windows was determined as follows. For any given earthquake record, the focal distance was divided by a group velocity of 3.5 km/sec. The result is an approximation of the secondary arrival time for that earthquake. The arrival time t_s is subtracted from this secondary arrival time and defines the duration between the start of the earthquake shaking and secondary arrival. This duration is used as the window length for the first window. The second window was one-half the duration of the first, the third window's duration was one-half the second, and so on for all six windows. If the window duration calculated in this manner was less than 0.3 sec, the window duration was set to 0.3 sec. The purpose of this approach is to train the neural

17

network with 6 windows or perspectives of the primary wave from each earthquake. All 6 windows began at t_s and ended before the arrival of the secondary waves. The 0.3 sec constraint was used because, for very near earthquakes, the Δt reduces to a few data points and is not useful for neural network training.

Preparing a typical training input started with cutting out a segment of the seismogram for a specific Δt. The cut segment was used to calculate the local measures that were input to the neural network. Details of how the local measures are derived will be discussed in the following section. We note here that the value of Δt is also used as an input to the neural network. This provides the network with the window duration as well as the parameters derived from that window. By using multiple windows for each earthquake, the network trains with six perspectives of each earthquake. It is generally known that neural networks have good interpolative capabilities; therefore, window durations that can be presented to the trained network are unrestricted, since any window duration will fall between the training example's window durations. This design also allows the network to learn emergent properties or patterns of an earthquake signal as a function of a growing window. It also offers a system that can produce an updated output after every data point of an arriving earthquake since each new data point simply represents a new window duration. This property is vital to the implementation of a real-time system. Table 1 shows the window durations used for several of the earthquake examples used in this study.

Table 1: Window durations for each of six windows used for training input to the neural network. Five earthquake examples are listed.

	Magnitude (m₁)	Distance (km)	Training window duration Δt					
	Magnitude (m_l)	Distance (km)	1 (sec)	2 (sec)	3 (sec)	4 (sec)	5 (sec)	6 (sec)
Earthquake 46	2.8	41.0	4.5	2.3	1.1	0.6	0.3	0.3
Earthquake 144	3.2	50.8	5.7	2.8	1.4	.7	0.4	0.3
Earthquake 260	4.6	121.1	14.0	7.0	3.5	1.8	.9	.44
Earthquake 393	2.7	226.1	27.4	13.7	6.9	3.4	1.7	0.9
Earthquake 431	3.5	269.5	37.9	18.9	9.5	4.7	2.3	1.1

System Input Design

Although each input or set of inputs in itself may offer a marginal estimate of impending ground motion, collectively the neural network will form a mapping of

input to output using a nonlinear combination of all the input. The network is constrained to learn or weigh the value of each input based on the example training set. Inputs can contribute as much or as little as needed to train the system to an successful performance level.

We used a total of 30 neural network input values. These inputs consisted of a time series profile, its local CAV and PGA values, the duration or window size of the input, a measurement of its backazimuth and angle of incidence, and a spectral representation. A schematic of the input to the neural network is shown in figure 11.

Figure 11 Schematic of neural network training input. Inputs consist of the duration in sec after t_s or Δt, a profile of the earthquake, local CAV and PGA values, a measure of backazimuth and angle of incidence, and a spectral representation. All inputs are local measures taken during t_s to $t_s + \Delta t$ in the seismogram.

The first ten input are formed as described earlier for the output profile, but in this case the input seismogram time series signal is used to form an input envelope. Again an envelope is formed using equation (13) from the displacement signals $d_e(n)$, $d_n(n)$, and $d_z(n)$, but the window duration used is from t_s to $t_s + \Delta t$. The first ten inputs to the neural network are the linearly sampled normalized input envelope. These input provide the general shape of the emerging ground motion signal. These shapes or particular patterns within them can be compared to the shape of the output time series envelope.

The next three inputs are the local CAV_p and PGA_p and the value of Δt. CAV_p and PGA_p are calculated as per equations (6) and (8), using local signals. Local, in this case, again refers to the segment of the seismogram signal from t_s to $t_s + \Delta t$. The motivation for using these parameters as inputs is that they offer a strong correlation between primary and secondary wave energy as seen in figure 4. The value of Δt provides a reference for the window duration or portion of the primary wave under consideration.

The next two inputs are estimates of the backazimuth and the angle of incidence calculated from the three-component polarization of the local displacement signals. These input were chosen to incorporate depth and angle measurements to provide focal distance information as input.

Finally, the last fifteen inputs consist of a sampling of a spectral transformation of the local seismic displacement signals. A Bartlett estimator was

19

used to form a spectral estimate of the displacement signals. This estimate was sampled into fifteen neural network inputs in the following manner. Previously described filtering put the maximum frequency f_{max} at 12.5 Hz and the minimum frequency f_{min} at 0.5 Hz For frequency input values i, where i ranges from 1 to 15 and $N = 15$, f_i is determined by:

$$f_i = f_{min} \left[(f_{min}/f_{min})^{(i-1)/(N-1)} \right],$$ (17)

The purpose of the above sampling procedure is to emphasize the lower frequencies that provide better overall signal-to-noise ratios for the earthquake spectral inputs.

Experimental Method

The task of scale estimation was separated from the task of profile estimation by training two different neural networks, one for each task, in two phases. The input remained the same for both networks and consisted of the 30 parameters described in the previous section. The output for each phase was taken from one of two categories: the set of scale values as calculated by equation (15) for the first phase, and the set of profiles as calculated by equation (16) for the second phase. Figure 12 illustrates the architectures of the neural networks. A 3-layer conjugate gradient backpropagation neural network was used in both phases with 30 input values. Training was performed with backpropagation learning for a multilayer feed-forward neural network, using an algorithm developed at Lawrence Livermore National Laboratory (Johansson et al., 1992). The neural network trained to estimate earthquake scale used 5 hidden nodes and 1 output. node. The neural network trained to estimate earthquake profile used 7 hidden nodes and 240 output nodes.

Figure 12 Neural network architectures used in this study: (a) neural network used to estimate scale of the earthquake; (b) neural network used to estimate earthquake profile. In both cases, a 3-layer conjugate gradient backpropagation network was used.

In the first phase, a neural network is trained to estimate earthquake scale using the architecture shown in figure 12a. Fifty randomly chosen earthquakes were removed from the original 434 earthquakes in the data base. These 50 earthquakes were not part of the training process and were set aside for neural network performance measurements. The remaining 384 earthquakes were used to train the

neural network. The training input for each training example consisted of the 30 previously described parameters. Each of the 384 earthquakes provided six training examples by using six different window sizes as described earlier. The total number of training examples was 2304. The training output for each earthquake was the normalized scale input or $\|O\|_n$, which was calculated using equations (14) and (15). Once trained, previously unseen examples were fed through the network, and neural network scale estimates were produced. The results are discussed in the next section.

For the second phase, a separate neural network was trained to estimate earthquake profile using the architecture shown in figure 12b. This network was trained with the 2304 examples described in the first phase, having removed the same 50 randomly chosen earthquakes. The training output for each earthquake in the second phase was the normalized earthquake profile vector O_{ni}, which was calculated using equation (16). The training input for both phases was identical. Again, once trained, new examples were fed through the network, but in this phase, neural network profile estimates were produced.

Results and Discussion

The earthquakes used in this study were, in general, weak motion, regional events with distances ranging from 20 to 294 km. Earthquake magnitudes varied from 2.5 to 4.6 m_l. We are currently increasing this range to include strong motion earthquakes. Earthquake focal distance measures tended to dominate the scale values due to the relatively limited range of magnitudes. We now discuss the results of the first phase, where a neural network was trained to estimate earthquake scale.

The input derived from the earthquake onset to 1 sec after the onset ($\Delta t = 1$) in the 50 previously unseen earthquakes were fed-forward through the trained neural network to produce scale estimates. We emphasize that the scale neural network can produce estimates at any time after the earthquake arrival, regardless of earthquake distance or magnitude. However, a fixed value of 1 sec was selected for two purposes. First, it illustrates that the network can perform well even when forced to interpolate using window sizes different from training set window sizes. secondly, a reasonably accurate estimate after 1 sec will offer useful warning time in general, as seen in figure 2.

Results are illustrated in the scatter plots in figure 13. Local PGA and CAV were calculated using the first sec of the 384 earthquake training examples. A

regression analysis was done using these measures as well as the true scale of each earthquake. Figure 13a shows the true scale values for the 50 removed earthquakes plotted against the regression-produced PGA estimates. Figure 13b shows the true scale values again plotted against the regression-produced CAV estimates. Figure 13c shows the true scale values plotted against the neural network scale estimates with $\Delta t = 1$. After 1 sec of shaking, there is a strong correlation between the neural network scale estimates and the true scale values of the earthquake signals.

Figure 13 Scatter plots showing the scale of the output envelope against (a) linear regression fit of the local PGA value (b), a linear regression fit of the local CAV measurement, and (c) the neural network estimate of the envelope scale. The lines were plotted to represent a perfect correlation or correlation coefficient of 1.

It is clear from figure 13 that the network can produce good scale estimates. The PGA and CAV generated estimates of figures 13a and 13b each in itself produces a scale estimate with a positive correlation to true scale. However, the neural network is able to improve the correlation. This illustrates how neural networks use multiple input features to improve performance. The network is using the PGA and CAV measures as input during training and, therefore, has the option of using these parameters to form its estimate. The network results are also consistent across the range of scale values. The regression estimates are less accurate at lower scale values. This illustrates the tolerance of the network to noisy input, as the lower scale values correspond to lower magnitudes, and large focal distances correspond to poor signal to noise.

Examination of the most serious outlier in figure 13c revealed an earthquake with magnitude m_l of 3.6 and a focal distance of 154 km. The signal revealed an apparent foreshock, which confused the neural network, producing an underestimation of the scale value. Note that the PGA and CAV regression estimates had the same problem with this event. As the window size was increased, the neural-network scale estimate improved significantly. The cause of this error stems from the selection of earthquake arrival times by visually marking the records. A more sophisticated algorithm, which ignores small foreshocks would help reduce this error. We are currently developing detection algorithms to accomplish this. Another potential solution is to form a more robust training example set that contains many occurrences of earthquakes with and without small foreshock activity.

In a second experiment, neural network estimates of scale values as a function of Δt were obtained to examine the emergent properties of the scale estimator as well as the accuracy at widely different scale values. In larger earthquakes, the faulting may continue for many seconds. The system has to be able to continually update its estimate during the evolution of the earthquake source. For example, there is not much difference between magnitude estimates for moderate and large earthquakes in the initial (0.5 sec) portion of the signal (Anderson, et. al., 1995). To examine the emergent capabilities of the system, each earthquake was fed into the trained neural network scale estimator 15 times, using values of Δt ranging from 0.2 to 10 sec. The results obtained from 4 arbitrary earthquakes are listed in

Table 2: Neural network scale estimates using Δt values ranging from 0.2 to 10 sec for four previously unseen earthquakes

	Earthquake 213		Earthquake 265		Earthquake 245		Earthquake 429	
	Distance (km)	87	Distance (km)	122	Distance (km)	116	Distance (km)	258
	Magnitude (m_l)	4.2	Magnitude (m_l)	3.5	Magnitude (m_l)	2.6	Magnitude (m_l)	3.0
	Scale (normalized)	.80	Scale (normalized)	.53	Scale (normalized)	.33	Scale (normalized)	.16
Δt (sec)	Neural network scale estimates (normalized)							
0.2	0.61		0.53		0.51		0.13	
0.3	0.72		0.57		0.40		0.18	
0.5	0.81		0.52		0.32		0.20	
0.7	0.87		0.50		0.37		0.22	
0.9	0.86		0.56		0.36		0.26	
1.1	0.85		0.64		0.41		0.28	
1.3	0.84		0.60		0.42		0.26	
1.5	0.83		0.59		0.39		0.23	
1.7	0.81		0.63		0.42		0.19	
2.0	0.83		0.63		0.51		0.21	
3.0	0.80		0.66		0.47		0.20	
4.0	0.76		0.64		0.44		0.22	
6.0	0.81		0.62		0.40		0.19	
8.0	0.79		0.59		0.37		0.17	
10.0	0.79		0.59		0.39		0.15	

Table 2. Overall, the network scale values were excellent as the earthquake emerged for 10 sec. A warning system that incorporated a running average of these values would perform quite well as a scale estimator in this case. This can be seen more readily in figure 14 which shows a plot of the four sets of earthquake estimates over 10 sec as well as the true scale for each earthquake.

Figure 14 Neural network scale estimates using Δt values ranging from 0.2 to 10 sec for four previously unseen earthquakes. The superimposed line plots indicate the target or true scale value for each earthquake.

The scale estimates in figure 14 appear somewhat unstable for the first sec or two of the emerging earthquake. This is caused partly by inadequate window sizes, which resulted in poor earthquake signal spectral representation being input to the

neural network. Also, especially as focal distance increases, the portion or percentage of the input window relative to the complete primary wave is quite small. The network is forced to make a decision on a very short 'look' at the primary wave. It is clear that as the earthquake emerges, the scale estimates tend to converge to reasonable values. Note that the network has an overall tendency to overestimate, which is a desirable property in a warning system, given that the overestimation is minimal.

In the second phase of the study, we trained a second neural network to estimate earthquake profile. Again the input derived from the earthquake onset to one sec after the onset were fed-forward through the trained neural network. Results for the same set of 50 previously unseen earthquakes with Δt equal to 1 produced 50 profile estimates. The correlation coefficient C_c between the true earthquake profile and the neural network profile estimate was calculated. To find C_c, the earthquake and neural network estimate profiles are represented by vectors **e** and **n**, respectively. After removing the mean, the correlation coefficient is then given by,

$$C_c = \frac{\mathbf{e} \cdot \mathbf{n}}{(\|\mathbf{e}\|)\,(\|\mathbf{n}\|)}. \tag{18}$$

The correlation coefficient will provide an indication of how well the shape of the two profiles match. For example, a C_c value of 1 indicates a perfect match. The C_c for all 50 pairs of earthquake and neural network signals is plotted in Figure 15a.

Figure 15 The correlation coefficient C_c between the true earthquake profile and the neural network profile estimate is shown in (a) for 50 previously unseen earthquakes. Likewise, the correlation coefficient C_c between the true earthquake profile and the average earthquake signal is shown in (b) for the same 50 earthquakes. Values were calculated using only the first sec of earthquake signal.

One might conjecture that the neural network is doing little more than producing the average shape of the 384earthquakes in the training set. To address this concern, consider that for any given training set of earthquake signals, the average signal is represented by,

$$a[n] = \frac{1}{N} \sum_{i=1}^{N} \sum_{n=1}^{M} p_i[n] \cdot \qquad (19)$$

where N is the total number of earthquakes used in the average, p_i is the discretized profile signal for earthquake i, and M represents the number of elements or profile samples. The vector **a** was formed using $a[n]$ as in equation (19) with N equal to 384, and M equal to 240. Figure 15b, we plot the C_c for each earthquake signal, but use the vector **a** to replace **n** in equation (18). The result clearly illustrates that the neural network has learned to predict the overall shape of the emerging earthquake signal better than a simple average of the training set. The neural network profile estimate provided a profile estimate that was better than using the average earthquake profile in 88% of the 50 earthquakes tested.

To examine the emergent properties of the profile estimator, neural network estimates of earthquake profile as a function of Δt were examined. The results of one of the earthquakes are shown in figure 16, which contains earthquake neural network profile estimates for Δt values of 0.2, 2.0, 4.0, 6.0, and 8.0 sec. The earthquake shown in figure 16 has a magnitude m_l of 3.2 and a focal distance of 46 km. It is clear in this example that the profile estimates can improve as Δt increases.

Figure 16 Neural network profile estimates using Δt values of 0.2, 2.0, 4.0, 6.0, and 8.0 sec to illustrate emerging profile estimates. The target or true profile for this earthquake is at the bottom of (a) and in black in both (a) and (b); (b) shows the signals in (a) superimposed. This earthquake's focal distance was 46 km and had a magnitude m_l of 3.2. Note the duration of approximately 12 sec from the onset of shaking to the arrival of the most severe ground motion. Also note that for this example, the profile estimates are quite accurate and provide greater than 10 sec of warning time.

Error Measurements in a Simple Warning Scenario

It is difficult to establish a general criteria for the measurement of hazard. Safety is relative to many variables such as population density, the age of structures, location of hazardous materials, etc. However, to demonstrate the utility of the neural network output estimate performance, a simple warning scenario is now presented.

We begin by accepting the accuracy of the earthquake scale estimates as shown in figure 13c. From the normalized range of scale values 0.0 to 1.0 we define an arbitrary warning scale level of 0.58. All feed-forward input for both scale and envelope estimates in this scenario were prepared using the first 1.0 sec of each

earthquake tested. Fifty previously unseen, randomly picked earthquakes were fed-forward through the trained scale estimate neural network. Ten of the resulting fifty neural network scale estimates exceeded 0.58 and were classified as hazardous. Figure 17 illustrates the 50 earthquake scale values, warning level, and the distribution of hazardous and nonhazardous events. The ten hazardous earthquakes were passed through the profile neural network which produced an envelope estimate. From these ten envelope estimates, the system error was measured for three vital parameters: the time within the secondary signal at which strong shaking begins, the time where hazardous shaking ends, and the duration of the shaking.

Figure 17 Warning level set at the arbitrary level of 0.58 divides a set of 50 earthquakes into hazardous and non-hazardous groups. The ten hazardous events had distances ranging from 32 to 122 km with magnitudes ranging from 2.9 to 4.2 m_l.

Each discreet earthquake output envelope signal O given by equation (16) was modeled by an exponentially decaying signal O_m where each element n of the model replaces the elements in O by,

$$o_m[n] = o[n_{max}] \cdot e^{-\left(\frac{n}{\tau}\right)} . \tag{20}$$

where τ is determined from the envelope signal by,

$$\tau = \frac{1}{N - n_{max}} \cdot \sum_{n = n_{max}}^{N} \frac{-n}{\log_e \cdot o_n} . \tag{21}$$

and n_{max} represents the time of the local maximum peak value of the secondary envelope signal or the start of the strong shaking. n_{max} denotes the time after the start of the earthquake when the secondary arrival signal O exceeds the product of a constant k and the maximum value of O. For the output envelope target signals, k was 0.707. Figure 18a illustrates an earthquake envelope signal with the exponential model O_m superimposed.

The neural network envelope estimate signal Ô is modeled in the same manner, where equations (20) and (21) are used to produce $\hat{O}_m$. In the case of neural network estimates, the value of 1.0 is used for the constant k in determining n_{max}. This is because the neural network output signal is less noisy than the target earthquake output envelope signal. The neural network envelope estimate of the

true earthquake shown in Figure 18a is plotted in Figure 18b.

Figure 18 Envelope signals with exponential models superimposed showing the hazard begin (t_{hb}), end (t_{he}), and duration time (Δt_h) for (a) an earthquake signal classed as hazardous by a neural network scale estimate and (b) the profile neural network estimate of the earthquake signal in (a) formed 1 sec after the start of the earthquake.

The first of the three hazard parameters measured is the time corresponding to the location of n_{max} which we define as the hazard begin time or t_{hb}. The second hazard parameter is defined as the time corresponding to when the exponential model drops to below the arbitrary value of 25% of its maximum peak. This time is defined as the hazard end time or t_{he}. A third hazard parameter Δt_h is the duration of the hazard which is simply the time between t_{hb} and t_{he}. An example of these three hazard parameters is shown in Figure 18a and 18b for both a true earthquake as well as the neural network estimate profile of the same earthquake. Using these parameters, system performance was measured for hazard start, stop, and duration time. Table 3 lists the system performance and error for estimation of Δt_h, t_{hb}, and t_{he} in both the true earthquake signals as well as the neural network estimates for the ten earthquakes which fell above the hazard line in Figure 17.

Table 3: Profile neural network hazard estimates and error values for beginning, end, and duration of heavy shaking for ten earthquakes. Earthquakes were chosen using scale neural network estimate values above an arbitrary threshold. Shaded column is plotted in Figure 18 and shows the network's estimate of 5.25 sec of the hazard start and 15.0 sec of hazard end. The real earthquake corresponding to this estimate had a hazard start time of 5.25 sec and a hazard end time of 15.5 sec. All neural network estimates were made 1 sec after the first arrival of each earthquake.

Earthquake			11	76	116	163	179	184	213	217	265	266
Distance (km)			32	45	49	53	60	61	87	88	122	122
Magnitude (m_l)			2.9	3.1	3.5	3.5	4.0	3.3	4.2	3.8	3.5	3.6
Scale (normalized)			0.69	0.50	0.62	0.67	0.72	0.64	0.85	0.72	0.60	0.62
Hazard measurements	Start	t_{hb} (sec)	3.75	4.75	5.0	5.5	6.75	7.5	9.5	9.25	16.3	16.5
		$\hat{t}_{hb}$ (sec)	5.25	5.25	5.25	5.25	5.75	5.5	5.25	5.25	16.0	16.0
		ε_{hb} (sec)	-1.5	-0.5	-0.25	0.25	1.0	2.0	4.25	4.0	0.25	0.5
	Stop	t_{he} (sec)	13.8	14.5	14.5	15.5	17.5	18.0	19.8	19.3	25.3	25.3
		$\hat{t}_{he}$ (sec)	15.3	15.3	15.3	15.0	16.5	16.0	15.3	15.3	25.5	25.3
		ε_{he} (sec)	-1.5	-0.75	-0.75	0.5	1.0	2.0	4.5	4.0	-0.25	0.0
	Duration	Δt_d (sec)	10.0	9.75	9.5	10.0	10.8	10.5	10.3	10.0	9.0	8.75
		$\hat{\Delta t}_d$ (sec)	10.0	10.0	10.0	9.75	10.8	10.5	10.0	10.0	9.5	9.25
		ε_d (sec)	0.0	0.25	0.5	0.25	0.0	0.0	0.3	0.0	0.5	0.5

The results of Table 3 show that the neural network was able to estimate the hazard start, stop, and duration time with surprising accuracy. In most cases, the start error was small or positive, which is desirable as large negative start time indicates warning after the hazard has begun. The duration errors were quite low, which may be due to the simplicity of the exponential model used as well as the distribution of the data base with respect to distance and magnitude. In the worst cases, the start and stop times were each about four seconds early. This apparent shift, as well as a somewhat lack of error increase as a function of distance, indicate that the distance related input to the neural network is weak and should be improved. In future work, we plan to add to the training input, a distance estimate which would presumably be available from existing regional seismic networks. This could improve the performance in a symbiotic manner using distributed warning networks in conjunction with a trained sensor.

29

Conclusion and Proposals For Future Studies

A neural network has been successfully trained with >400 recorded earthquakes to produce a scale and profile estimate of the earthquake signal in 0.3 sec and to continually update the estimate as the earthquake emerges. The system has been implemented in hardware using silicon accelerometers and a standard pc. It is clear from the results that the future of "smart sensors" holds great promise. However, there are many research issues that need to be addressed.

One issue is measuring the portability of the system, or the system error of the trained neural networks when moved to a new location. We plan to study whether a trained neural network will perform equally well at a new site (though trained at a different site). It may be possible to add a small set of representative earthquakes from the new site, retrain, and regain system performance. It is also a subject of future work to design the neural network to continually learn and upgrade the internal weights as a background process, allowing it to become more site-specific as a function of time.

Another issue is the expansion of the earthquake training set to include large motion earthquakes. This area holds great promise, as the availability of good quality strong motion events is increasing at a rapid rate via modern seismic stations and data availability on the internet. We have done some preliminary work in this area and will be expanding our data base to include strong motion earthquakes. The improved signal to noise of strong motion recordings improves the input and the neural network sees fewer inconstancies. For example, one of our neural networks estimated the magnitude of the Loma Prieta earthquake to be 6.8 m_l using less than 0.5 sec of the earthquake signal.

An important contribution of this work is to illustrate how a backpropagation neural network can be used to estimate continuous, real-time parameters. This contrasts to more common uses for neural networks such as a discrimination or classification task, such as pattern recognition of an explosion as opposed to an earthquake. In future work, we plan to test to see how well our system can identify and classify of non-earthquake signals, as well as multiple events

Acknowledgments

REFERENCES

Aki, K., and P. G. Richards (1980). *Quantitative Seismology, Volume I*, W. H. Freeman and Company, New York, NY.

Anderson, J. G., and Q. Chen (1995). Beginnings of Earthquakes in the Mexican Subduction Zone on Strong-Motion Accelerograms. *Bull. Seism. Soc. Am.* **85**, 1107–1115.

Arias, A. (1970). A Measure of Earthquake Intensity, *Seismic Design For Nuclear Power Plants*. The MIT Press, Cambridge, MA.

Bito, Y., and Y. Nakamura (1986). Urgent Earthquake Detection and Alarm System, *Civil Engineering in Japan*, 103-116.

Bolt, B. A. (1978). *Earthquakes, A Primer.* W. H. Freeman and Company, San Francisco, CA.

Bowman, B. C., and F. U. Dowla (1992). Real-time classification of signals from three-component seismic sensors using neural nets. *IEEE Transactions on Instrumentation and Measurement*; Lawrence Livermore National Laboratory report UCRL-JC-109410.

Draper, N. and H. Smith (1966). *Applied Regression Analysis*, 2nd ed., Wiley, New York, NY.

Dowla, F. U., S. R. Taylor, and R. W. Anderson. (1990). Seismic discrimination with artificial neural networks: preliminary results with regional spectral data. *Bull.*

Seism. Soc. Am. **80**, 1346–1373.

Hertz J., K. Anders, and R. G. Palmer (1990). *Introduction to the Theory of Neural Computation.* Lecture notes Vol. 1., Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley, Menlo Park, CA.

Hopfield, J. J., and J. Denker (1987). Large automatic learning rule extraction and generalization. *Complex Systems* **1**, 877-922.

Hopfield, J. J. (1988). Artificial neural networks, *IEEE Circuits and Devices*, **4**, 3-10.

Johansson, E. M., F. U. Dowla, and D. M. Goodman (1992). Backpropagation learning for multi-layer feed-forward neural networks using the conjugate gradient method, *International Journal of Neural Systems* **2**(4).

Kulhanek, O. (1990). *Anatomy of Seismograms*, Seismological Section, University of Uppsala, Sweden, Elseveir.

Leach, R. R., F. U. Dowla, and E. S. Vergino (1993). Yield estimation using bandpass-filtered seismograms: preliminary results using neural networks with $m_b(P_n)$, short-time, long-time, and coda energy measurements. *Bull. Seism. Soc. Am.* **83**, 488–508.

Lippmann, R. P. (1987). An introduction to Computing with Neural Nets, *J. IEEE ASSP*, April 4–22.

Maurer, W. J., F. U. Dowla, and S. P. Jarpe (1991) Seismic event classification using self-organizing neural networks. Lawrence Livermore National Laboratory report UCRL-JC-108630.

National Research Council, Panel on Regional Networks, Committee on Seismology, Board on Earth Sciences and Resources, and Commission on Geosciences, Environment, and Resources, (1991). Assessing the Nation's Earthquakes, *Real-Time Earthquake Monitoring,* National Academy Press, Washington, DC.

National Research Council, Panel on Regional Networks, Committee on

Seismology, Board on Earth Sciences and Resources, and Commission on Geosciences, Environment, and Resources, (1990). Assessing the Nation's Earthquakes, *The Health and Future of Regional Seismograph Networks*, National Academy Press, Washington, DC.

Pao, Y. (1989). *Adaptive Pattern Recognition and Neural Networks*. Addison-Wesley, Menlo Park, CA.

Reed, J. W., N. Anderson, N. C. Chokshi, R. P. Kennedy, W. J. Metevia, D. K. Ostrom, J. D. Stevenson, (1988). *A Criterion for Determining Exceedance of the Operating Basis Earthquake*, Electric Power Research Institute (EPRI) research project 2848-16. EPRI NP-5930.

Reiter, L. (1991). *Earthquake Hazard Analysis*. Columbia University Press, New York, NY.

Stanley, W. D. (1982). *Electronic Communications Systems*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Whitman, R. V. (1970). Basic Concepts and Important Problems. *Seismic Design For Nuclear Power Plants*. The MIT Press, Cambridge, MA.
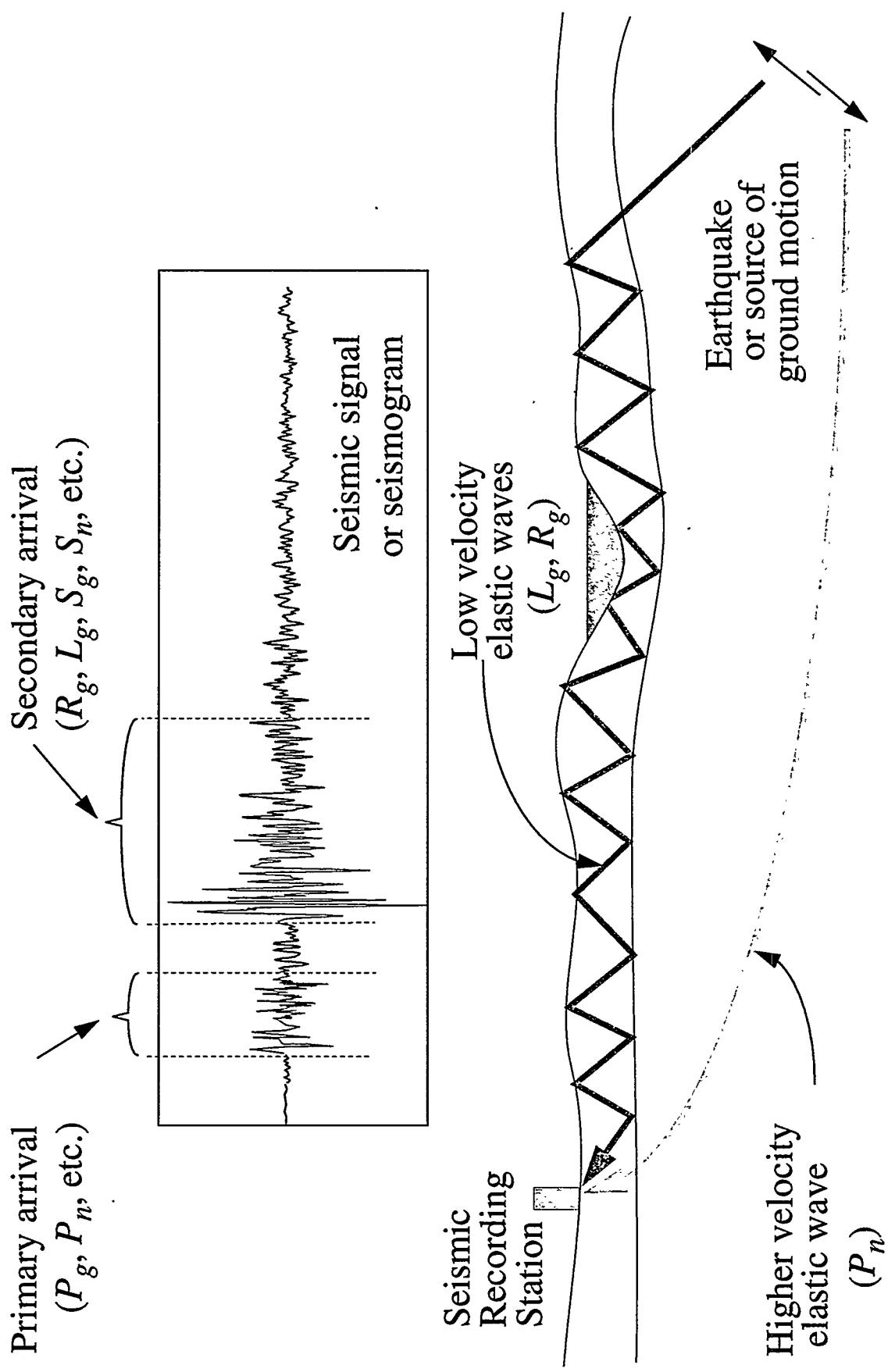
Primary arrival
(P_g, P_n, etc.)

Secondary arrival
(R_g, L_g, S_g, S_n, etc.)

Seismic signal
or seismogram

Low velocity
elastic waves
(L_g, R_g)

Earthquake
or source of
ground motion

Seismic
Recording
Station

Higher velocity
elastic wave
(P_n)

Figure 1 Elastic waves travel at different velocities and through different paths. The paths are often through highly non-linear propagation media. Primary and secondary arrivals in a recorded seismogram are also illustrated.
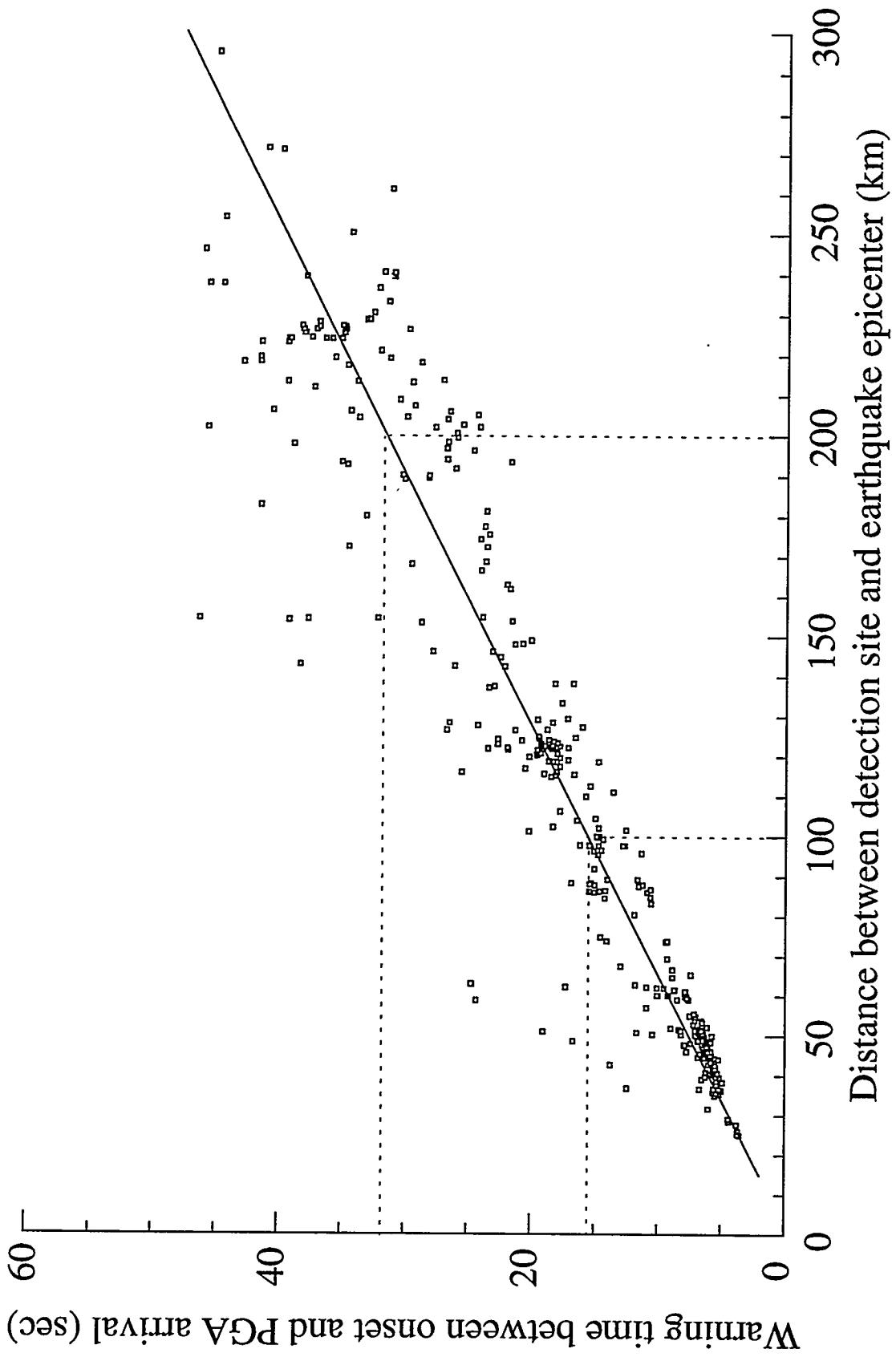
Figure 2 Warning time varies with respect to distance. 434 earthquake recordings are plotted illustrating the warning time measured from the P-wave arrival to the peak ground acceleration (PGA) of each earthquake. For example, at a distance of 100 km, there is approximately 15 sec of warning time between the earthquake onset and the arrival of the PGA.

Figure 3 Several plots from an earthquake signal, including its velocity recording (top), the same signal differentiated to acceleration with the primary $a_p(t)$ and secondary $a_s(t)$ signals noted (middle), and the fast Fourier transform of the primary and secondary signals of the acceleration signal (bottom). Some potential ANN training measures are denoted with arrows. Local peak ground acceleration is shown by PGA_p for primary and PGA_s for secondary PGA. Local peak spectral amplitude is shown by PSA_p for primary and PSA_s for secondary PGA.

Figure 4 Scatter plots of local measures in the primary wave against similar local measures in the secondary wave are shown for PGA, PSA, and CAV. Correlation coefficients are also indicated. The superimposed lines are the least squares fit of the data. The parameter That had the strongest correlation is the cumulative absolute velocity measure.
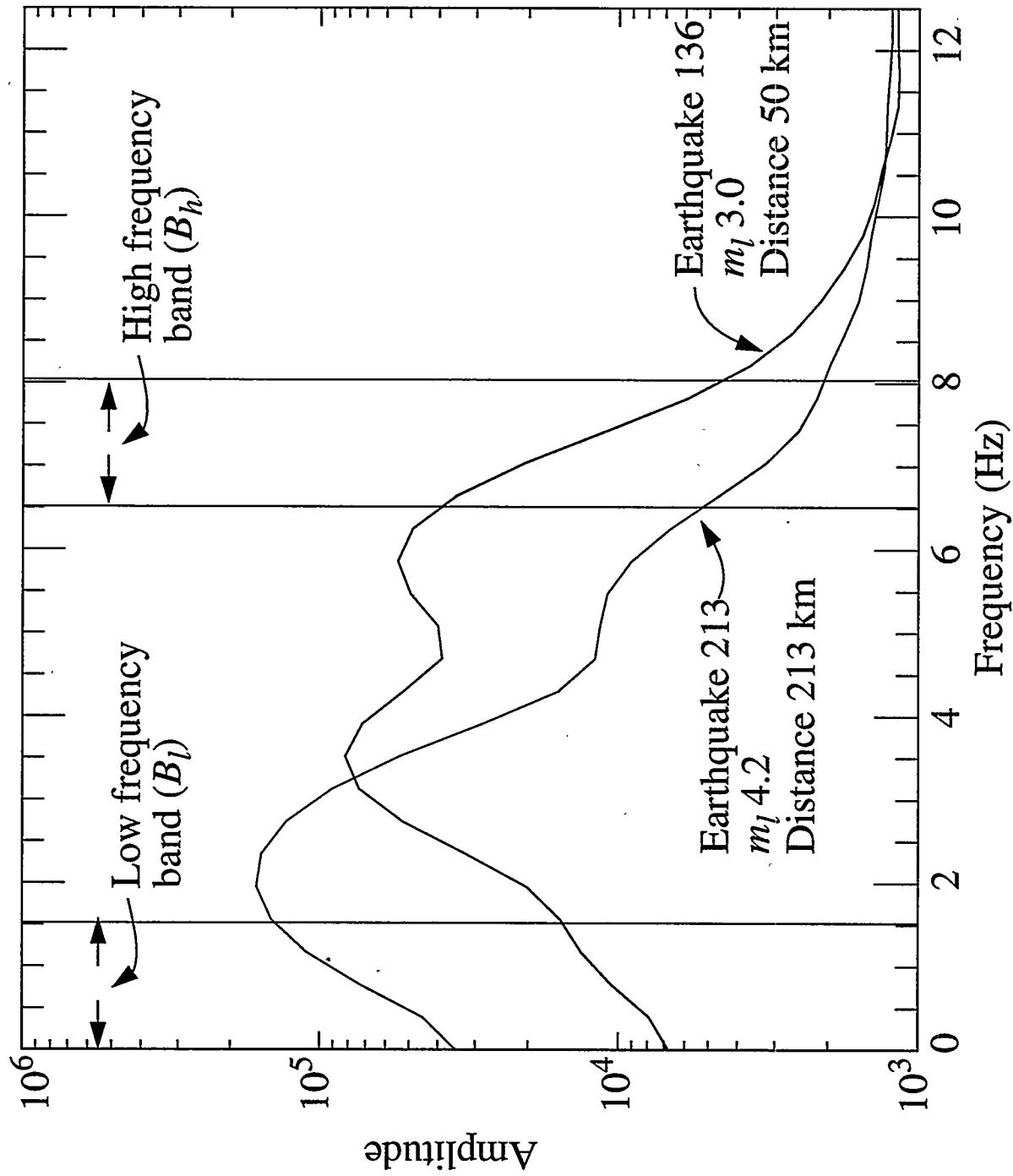
Figure 5 Bartlett spectral estimates for two earthquakes illustrating two frequency bands used to form a ratio of low to high frequency energy. Note the evidence of anelastic attenuation where the ratio of the integral of energy in the two frequency bands is relatively larger for the more distant earthquake.
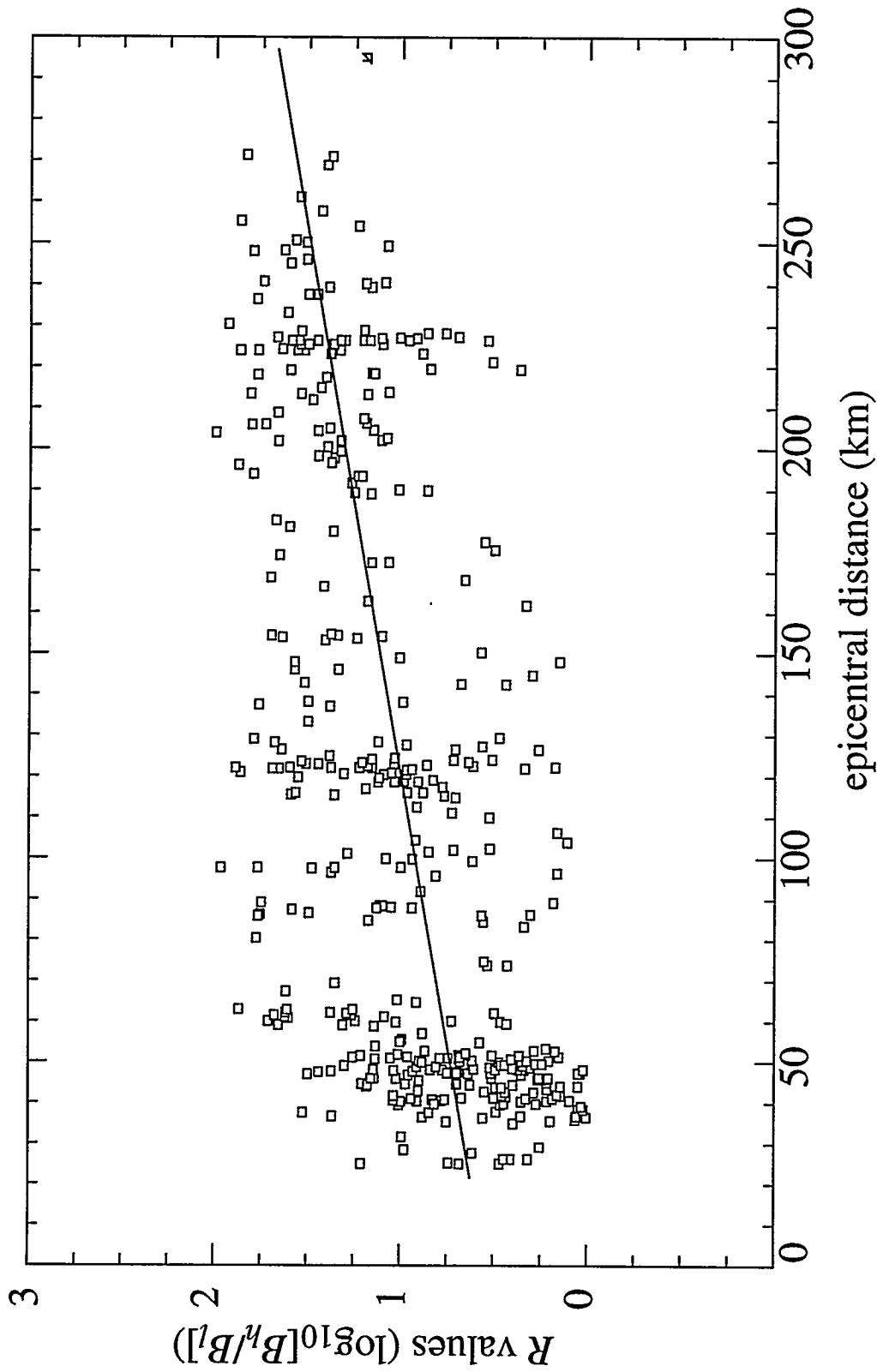
Figure 6 Scatter plot of *R* values versus hypocentral distances for >400 earthquakes. The overall positive correlation illustrates anelastic attenuation as a function of distance. The superimposed line is a least squares fit of the data. The wide variation occurs because the range of earthquake magnitudes for the various distances were plotted simultaneously.
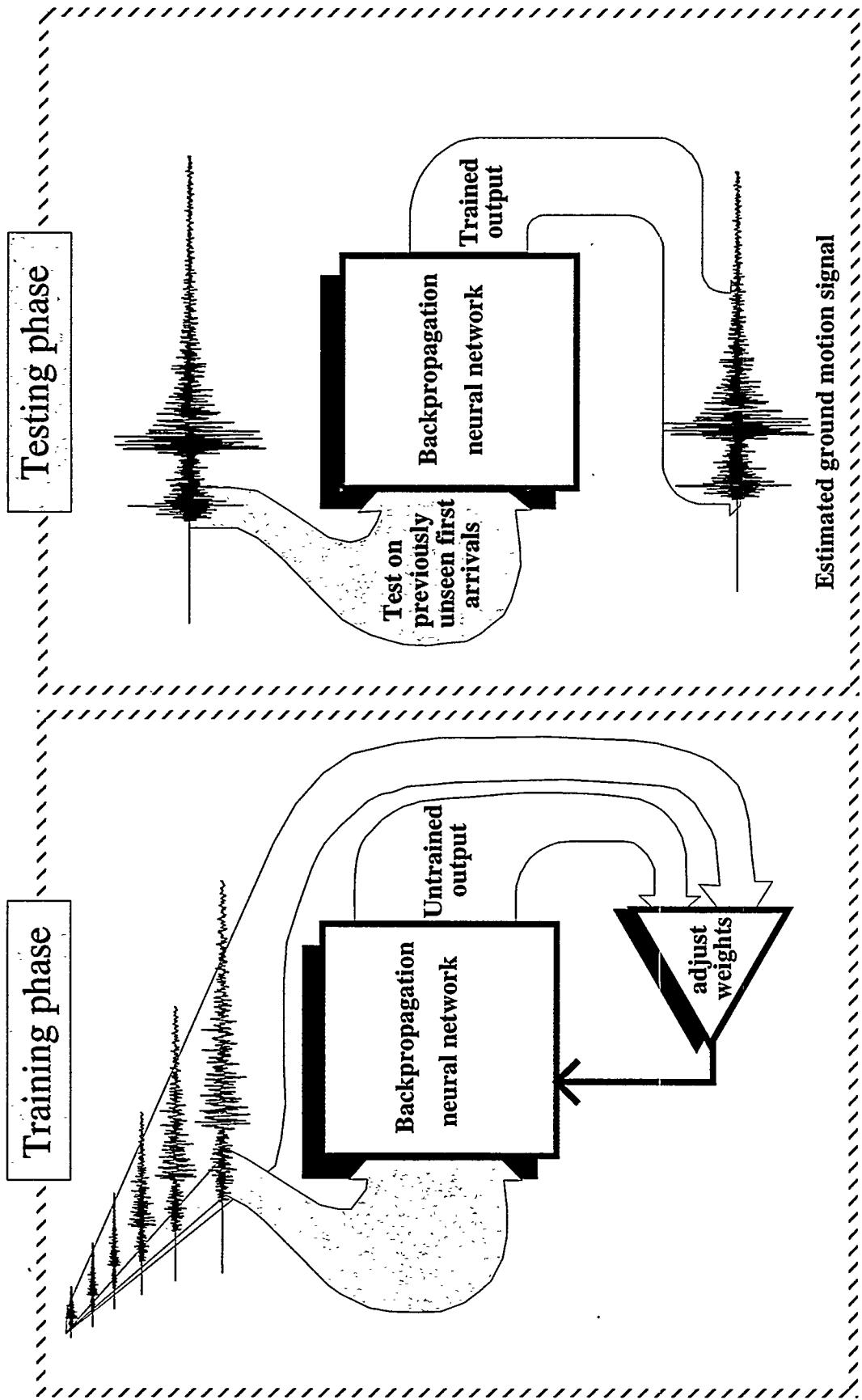
Figure 7 First, neural networks are trained to learn the relationship between the first arrival and the total seismogram of >400 different seismograms. Then the neural network can use the first arrival of a new seismogram to estimate the complete seismogram.
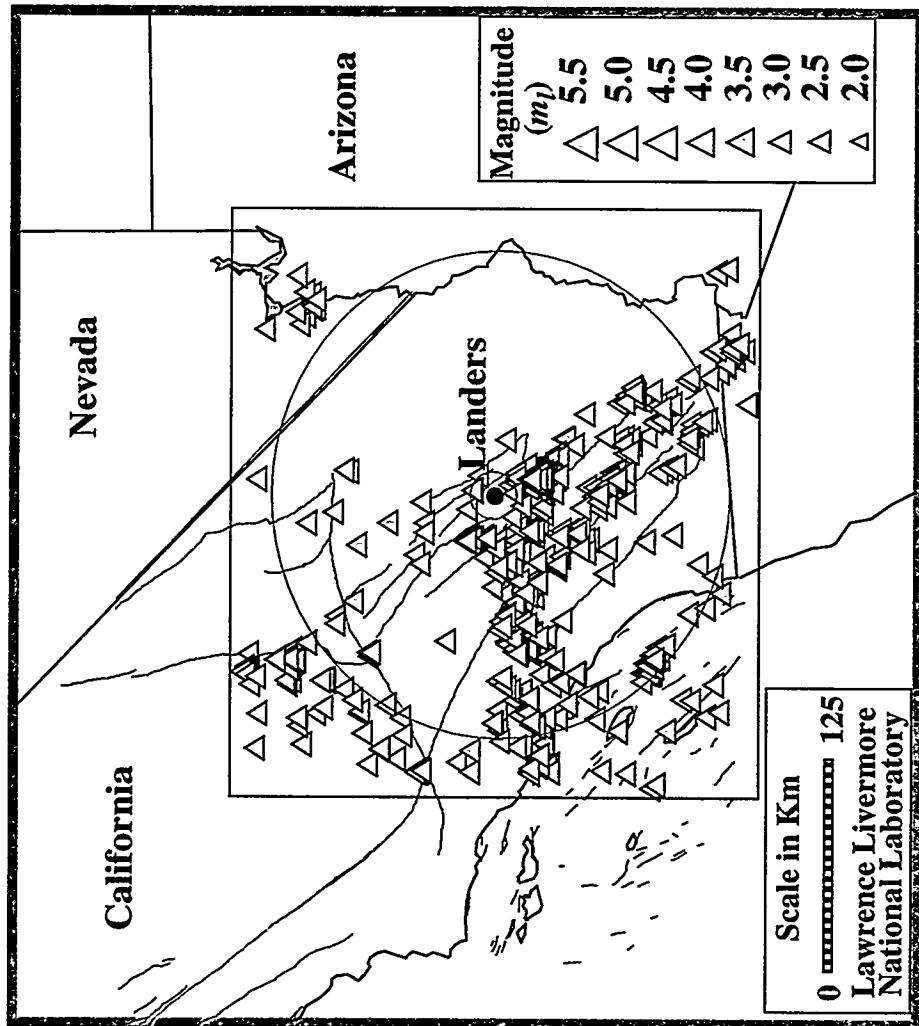
Figure 8 Earthquakes that occurred from 1988 to 1992 surrounding Landers, CA. This study focused on epicentral distances ranging from 20 to 294 km to Landers recording station.
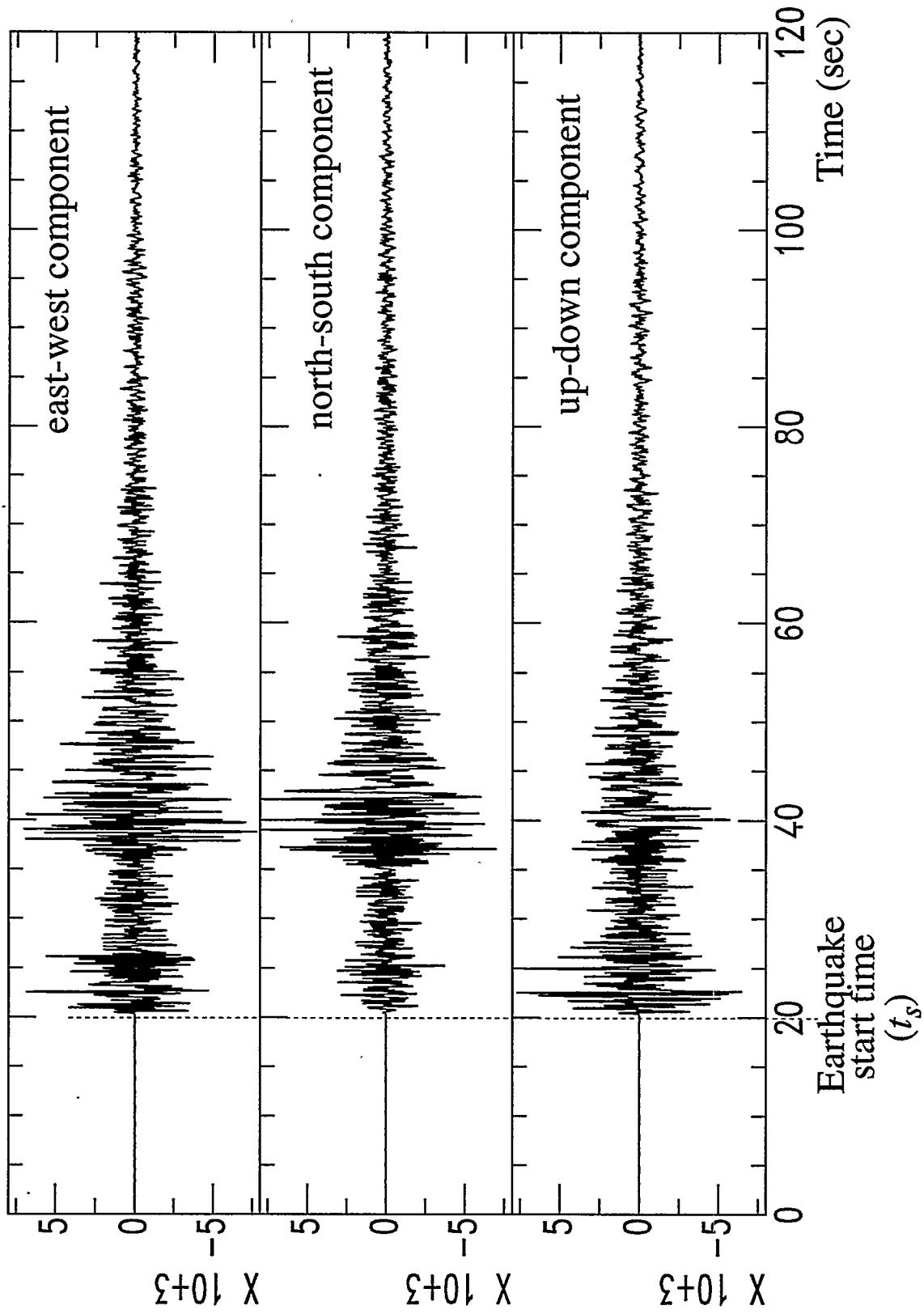
Figure 9 Data example of earthquake velocity recorded at the Landers recording station showing east-west, north-south, and up-down motion. This particular event had a magnitude (m_l) of 3.6 and originated approximately 122 km from Landers.

Earthquake 46
m_l 2.8
Distance 41 km

Earthquake 144
m_l 3.2
Distance 51 km

Earthquake 260
m_l 4.6
Distance 121 km

t_s

time (sec)
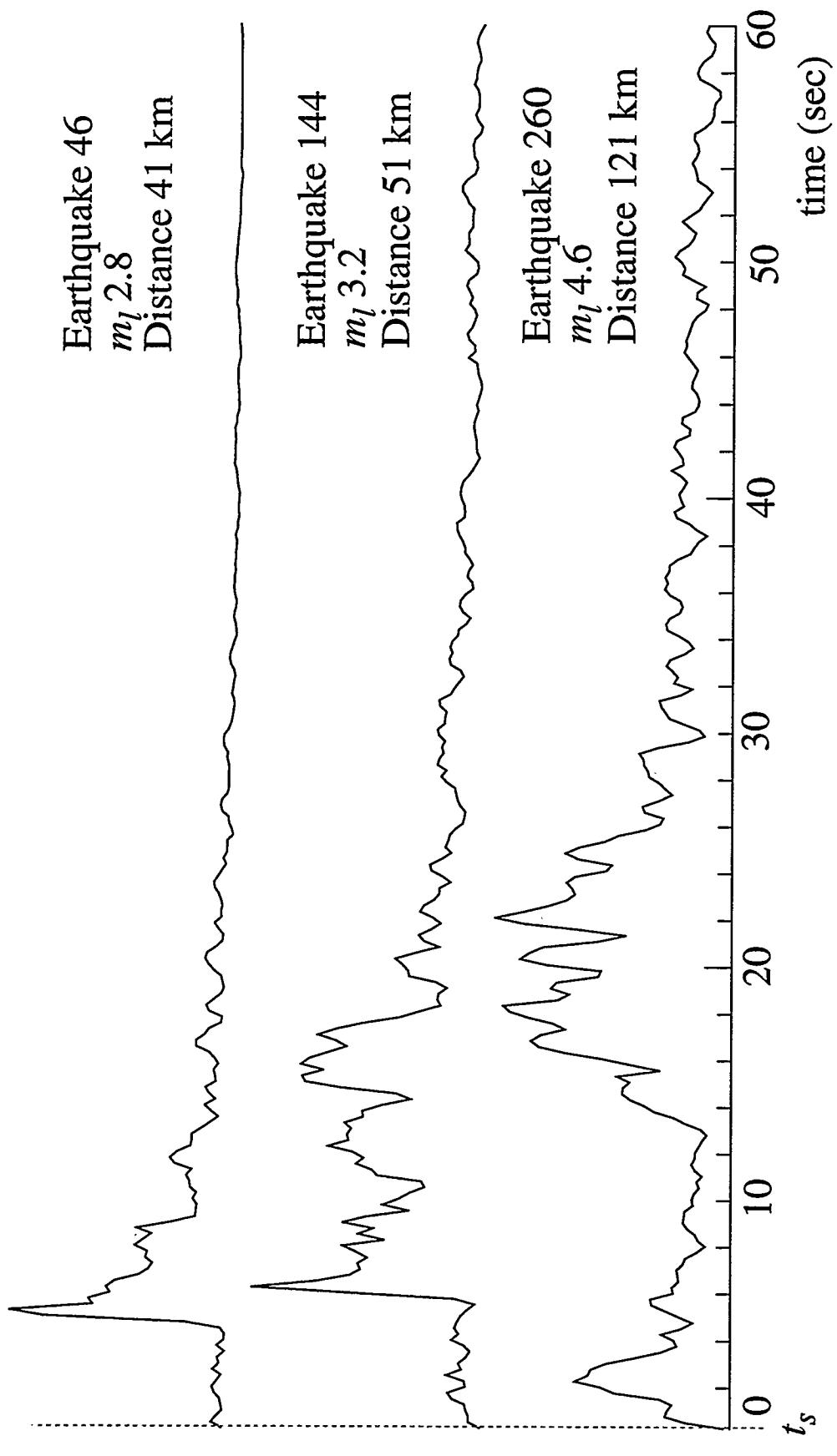
Figure 10 Several preprocessed earthquake displacement envelope signals. The signals were chosen to illustrate the arrival of predominant secondary arrival shaking as a function of distance
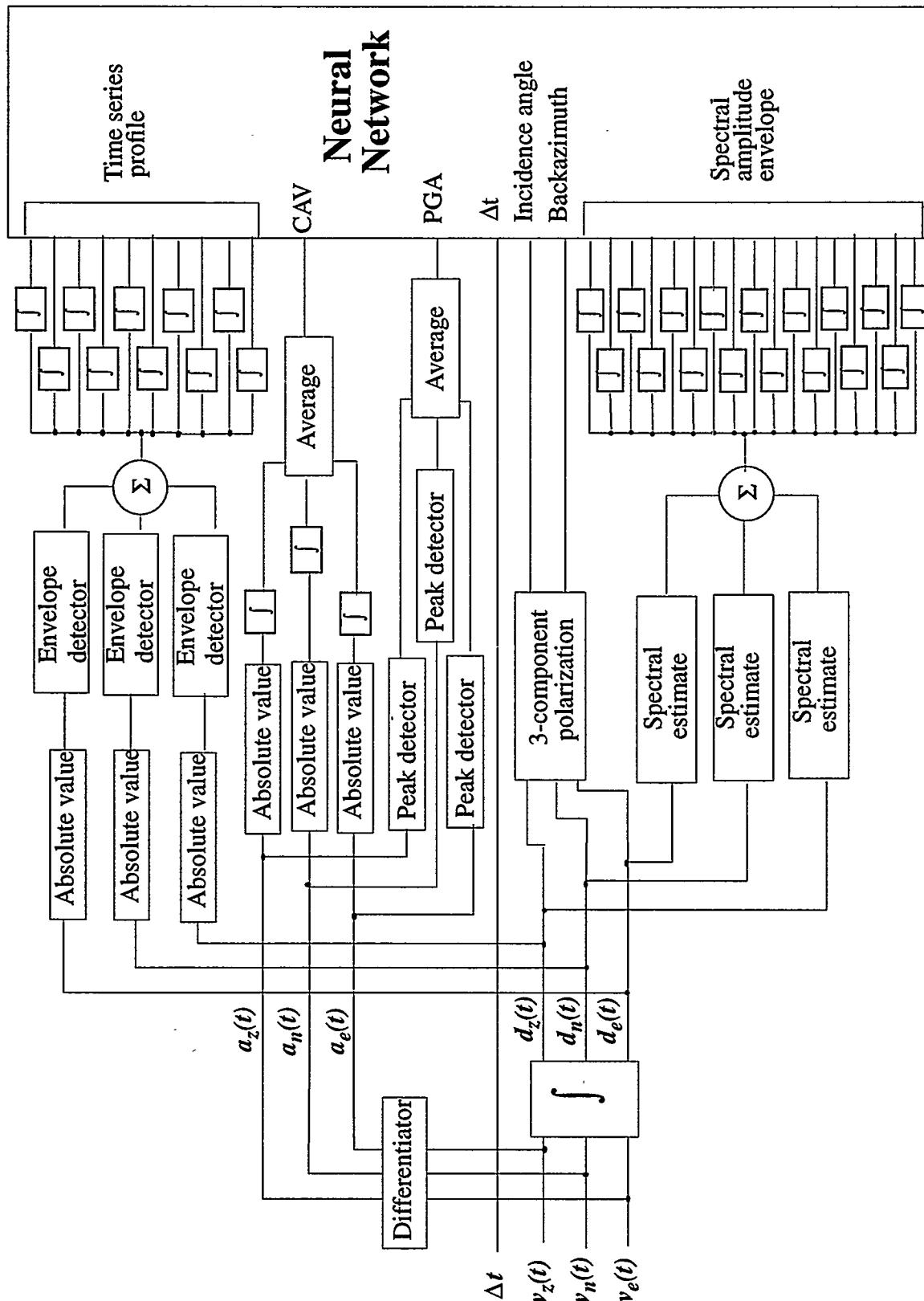
Figure 11 Schematic of neural network training input. Inputs consist of the duration in sec after t_s or Δt, a profile of the earthquake, local CAV and PGA values, a measure of backazimuth and angle of incidence, and a spectral representation. All inputs are local measures taken during t_s to $t_s + \Delta t$ in the seismogram.
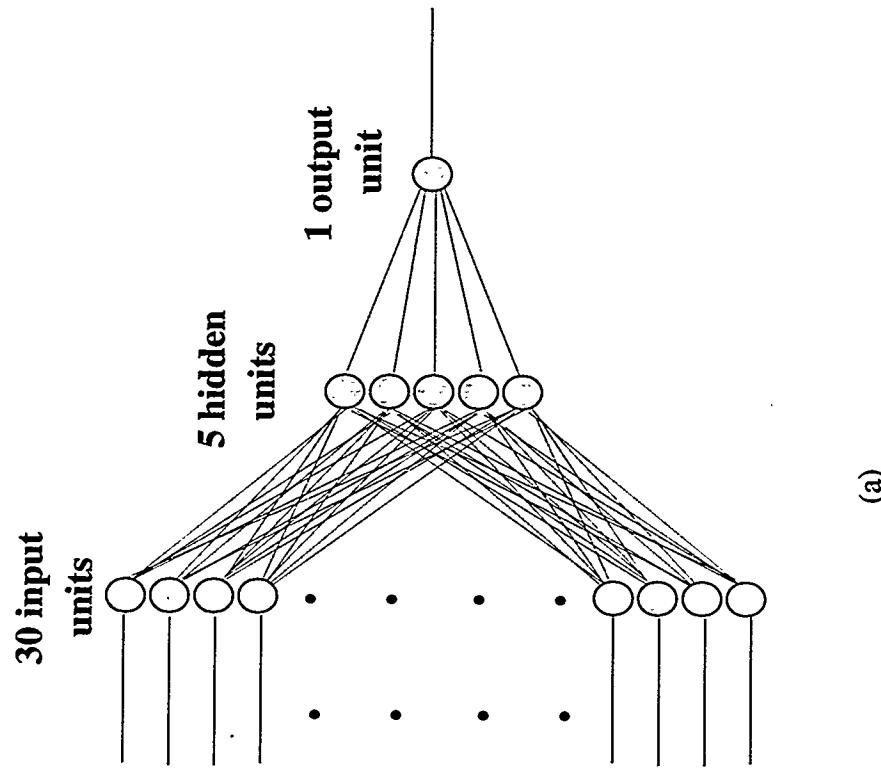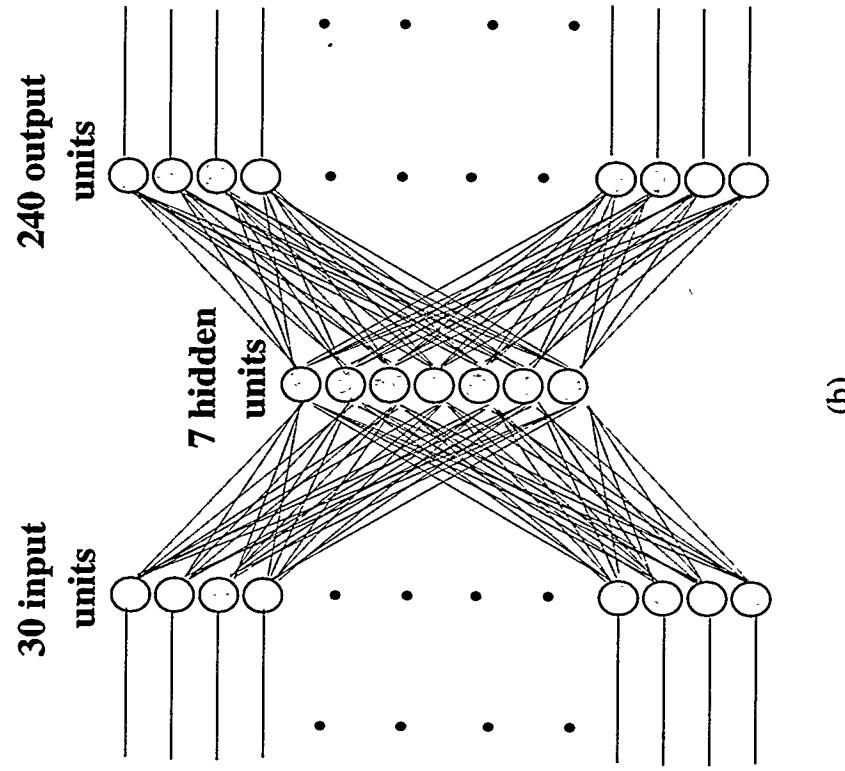
Figure 12 Neural network architectures used in this study: (a) neural network used to estimate scale of the earthquake; (b) neural network used to estimate earthquake profile. In both cases, a 3-layer conjugate gradient backpropagation network was used.

Figure 13 Scatter plots showing the scale of the output envelope against (a) linear regression fit of the local PGA value (b), a linear regression fit of the local CAV measurement, and (c) the neural network estimate of the envelope scale. The lines were plotted to represent a perfect correlation or correlation coefficient of 1.

Figure 14 Neural network scale estimates using Δt values ranging from 0.2 to 10 sec for four previously unseen earthquakes. The superimposed line plots indicate the target or true scale value for each earthquake.

(a)

Correlation coefficient (C_c)

Distance (km)
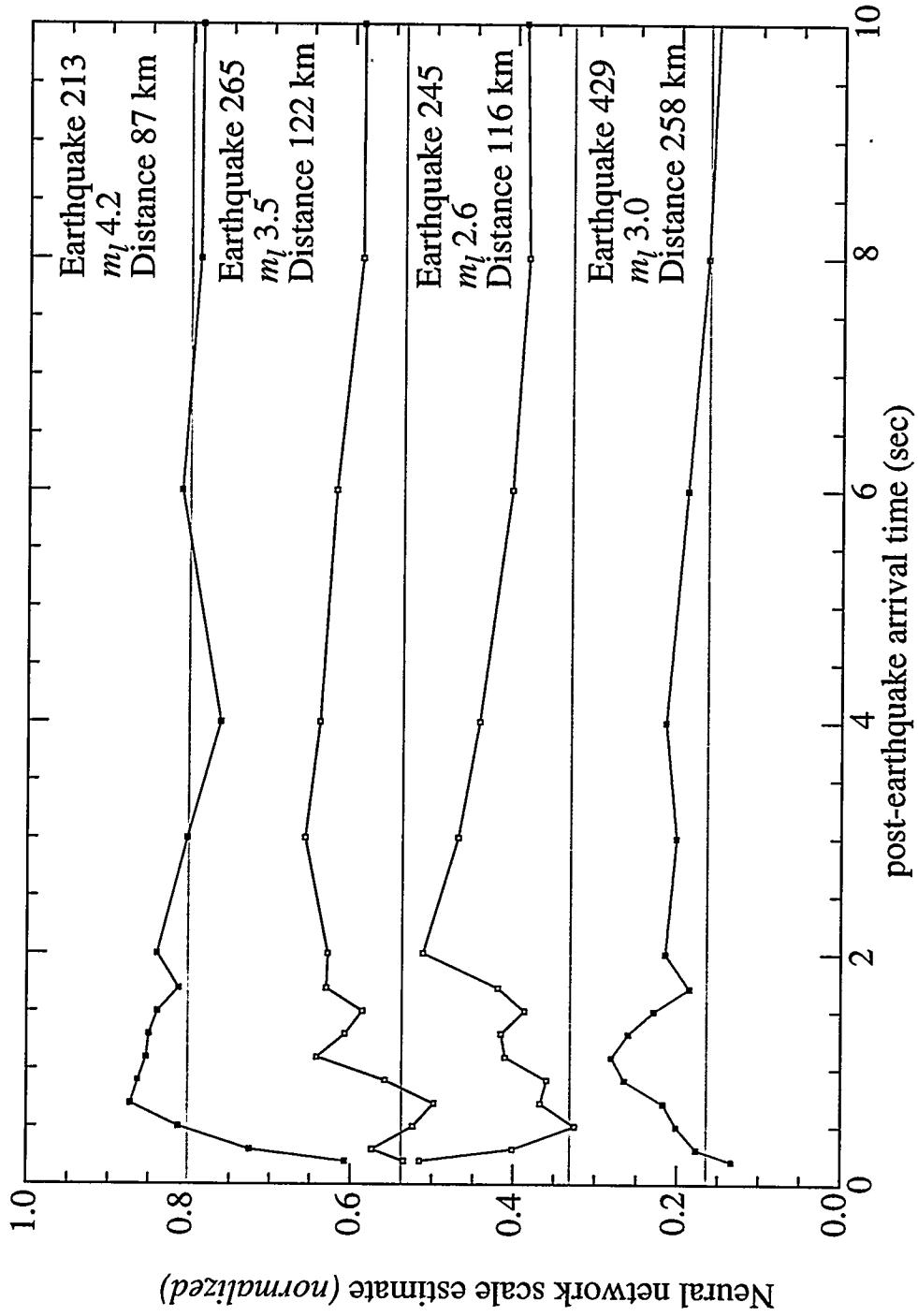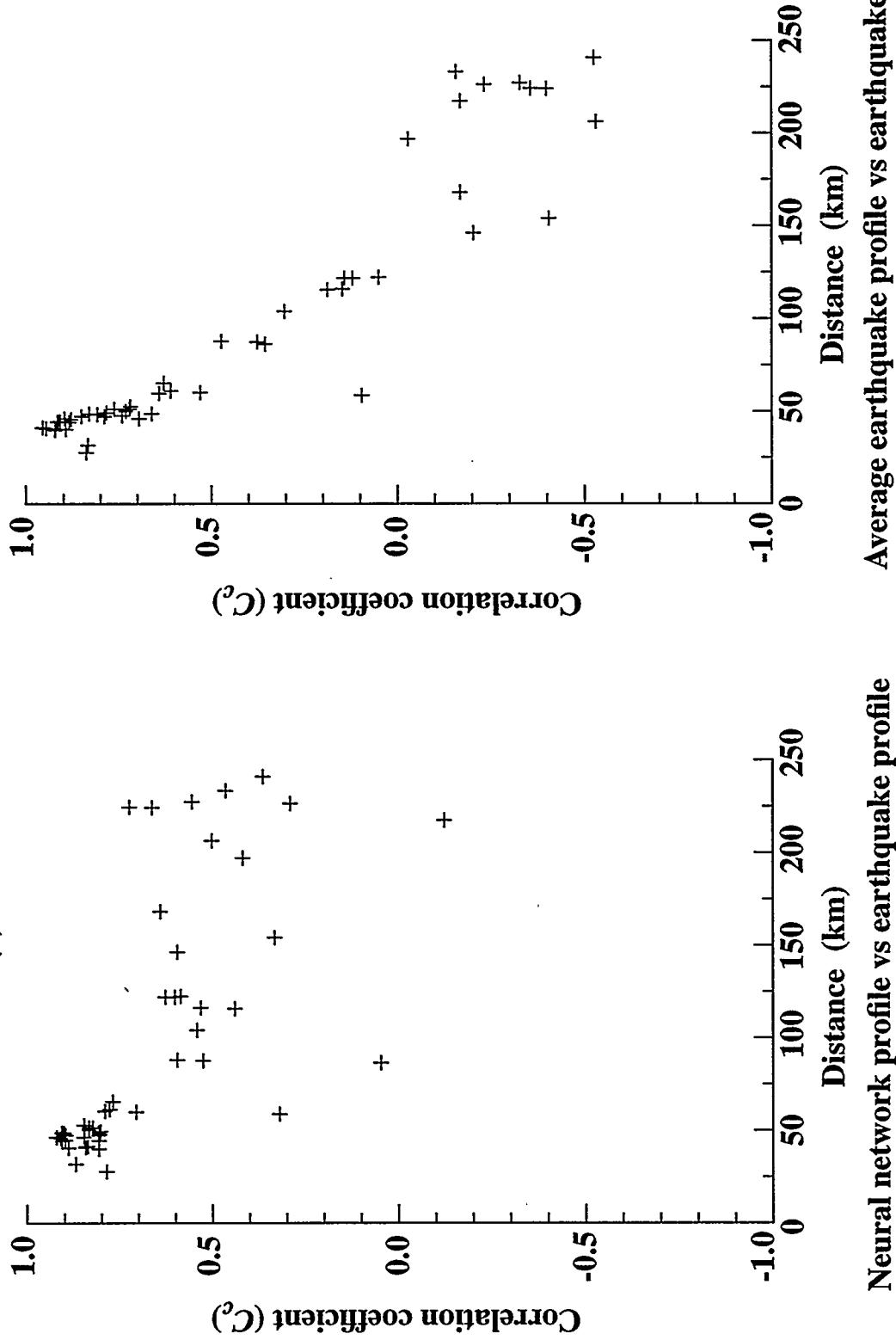
Neural network profile vs earthquake profile

(b)

Correlation coefficient (C_c)

Distance (km)

Average earthquake profile vs earthquake profile

Figure 15 The correlation coefficient C_c between the true earthquake profile and the neural network profile estimate is shown in (a) for 50 previously unseen earthquakes. Likewise, the correlation coefficient C_c between the true earthquake profile and the average earthquake profile is shown in (b) for the same 50 earthquakes. Values were calculated using only the first sec of earthquake signal.

Figure 16 Neural network profile estimates using Δt values of 0.2, 2.0, 4.0, 6.0, and 8.0 sec to illustrate emerging profile estimates. The target or true profile for this earthquake is at the bottom of (a) and in black in both (a) and (b); (b) shows the signals in (a) superimposed. This earthquake's focal distance was 46 km and had a magnitude m_l of 3.2. Note the duration of approximately 12 sec from the onset of shaking to the arrival of the most severe ground motion. Also note that for this example, the profile estimates are quite accurate and provide greater than 10 sec of warning time.

Figure 17 Warning level set at the arbitrary level of 0.58 divides a set of 50 earthquakes into hazardous and non-hazardous groups. The ten hazardous events had distances ranging from 32 to 122 km with magnitudes ranging from 2.9 to 4.2 m_l.

Figure 18 Envelope signals with exponential models superimposed showing the hazard begin (t_{hb}) end (t_{he}) and duration time (Δt_h) for (a) an earthquake signal classed as hazardous by a neural network scale estimate and (b) the profile neural network estimate of the earthquake signal in (a) formed 1 sec after the start of the earthquake.