

The application of back-propagation neural network to automatic picking seismic arrivals from single-component recordings

Hengchang Dai¹ and Colin MacBeth

British Geological Survey, Edinburgh, Scotland

Abstract. An automatic approach is developed to pick *P* and *S* arrivals from single component (1-C) recordings of local earthquake data. In this approach a back propagation neural network (BPNN) accepts a normalized segment (window of 40 samples) of absolute amplitudes from the 1-C recordings as its input pattern, calculating two output values between 0 and 1. The outputs (0,1) or (1,0) correspond to the presence of an arrival or background noise within a moving window. The two outputs form a time series. The *P* and *S* arrivals are then retrieved from this series by using a threshold and a local maximum rule. The BPNN is trained by only 10 pairs of *P* arrivals and background noise segments from the vertical component (V-C) recordings. It can also successfully pick seismic arrivals from the horizontal components (E-W and N-S). Its performance is different for each of the three components due to strong effects of ray path and source position on the seismic waveforms. For the data from two stations of TDP3 seismic network, the success rates are 93%, 89%, and 83% for *P* arrivals and 75%, 91%, and 87% for *S* arrivals from the V-C, E-W, and N-S recordings, respectively. The accuracy of the onset times picked from each individual 1-C recording is similar. Adding a constraint on the error to be 10 ms (one sample increment), 66%, 59% and 63% of the *P* arrivals and 53%, 61%, and 58% of the *S* arrivals are picked from the V-C, E-W and N-S recordings respectively. Its performance is lower than a similar three-component picking approach but higher than other 1-C picking methods.

Introduction

The task of estimating arrival times for *P* and *S* waves found in recordings of earthquake events forms an important foundation for schemes employing automatic processing for event location, event identification, source mechanism analysis, and spectral analysis. In traditional picking methods the trained analyst visually checks the seismograms and picks out *P* and *S* arrivals according to his individual experience. This task is time consuming and subjective. With the increase in the number of digital seismic networks being established worldwide, there is a pressing need to provide a more reliable and robust alternative, which is less time consuming and perhaps more objective. In order to achieve this, *Dai and MacBeth [1995]* developed an approach which used a back-propagation neural network (BPNN) to pick seismic arrivals from the vector modulus of three-component (3-C) recordings of local earthquake data. This BPNN is trained by presenting different *P* arrivals and segments of background noise. After training is accomplished, it can indicate arrivals from a variety of new seismograms. A BPNN trained with nine pairs of *P* arrival and background noise segments can pick 94% of *P* arrivals and 86% of *S* arrivals of two stations of a local seismic network. Unfortunately, not all seismic stations have 3-C seismometers or can provide consistently high quality 3-C recordings due to the

failure of one or more components. There is therefore a practical need to pick seismic arrivals from single-component (1-C) recordings.

There is no shortage of techniques in the literature to pick seismic arrivals from the 1-C recordings [*Allen, 1978; Anderson, 1978; Bache et al., 1990; Baer and Kradolfer, 1987; Chiaruttini et al., 1989; Chiaruttini and Salemi, 1993; Houlston et al., 1984; Joswig, 1990, 1995; Joswig and Schulte-Theis, 1993; Pisarenko et al., 1987; Takanami and Kitagawa, 1988, 1993*]. However, most of them are conventional methods which are not adaptive, working well only under certain conditions and usually picking one type of arrival. There are also approaches using a BPNN to pick first arrivals from seismic prospecting or earthquake data [*McCormack et al., 1993; Murat and Rudman, 1992; Wang and Teng, 1995*]. These approaches take a window from a seismic trace and calculate a certain number of associated values such as mean amplitude or spectral properties. These values are presented to a neural network as its input pattern and the network has to decide whether or not it is a first arrival. However, these values may lose information and involve too much computing time. Using the full waveforms as the network input might be a better choice. As our previous approach [*Dai and MacBeth, 1995*] has shown a good performance in picking seismic arrivals from 3-C recordings, we shall adapt this BPNN approach to pick seismic arrivals from either vertical or horizontal recordings. This is achieved by utilizing the absolute values of 1-C recording as the BPNN input. A BPNN is trained with small number training samples and is then used as a filter action on the entire seismogram by sliding a window along the seismograms. The output of the BPNN yields a time series which is interrogated for a decision regarding the seismic arrivals. Its performance depends among others on the input characteristic. As the 1-C recordings

¹Also at Department of Geology and Geophysics, University of Edinburgh, Edinburgh, Scotland.

Copyright 1997 by the American Geophysical Union.

Paper number 97JB00625.
0148-0227/97/97JB-00625\$09.00

have different overall characteristics from the 3-C recordings, the original approach might not give rise to an equally good performance in picking arrivals from I-C recordings. It is necessary to alter the training data set and the BPNN structure, so that a good performance is obtained.

Approach

Back Propagation Neural Network

The BPNN used in this approach is a three-layer feed-forward neural network trained by the generalized delta rule [Rumelhart *et al.*, 1986; Pao, 1989]. It is made up of sets of nodes arranged in layers, an input layer, a hidden layer, and an output layer. Each node, the basic processing unit, has a nonlinear activation function. The outputs of the nodes in one layer are transmitted to nodes in another layer through links called "weights". These weights are real numbers, which are applied as simple multiplicative scalars, and effectively amplify or attenuate the signals. With the exception of the nodes in the input layer, the net input to each node is the sum of the weighted outputs of nodes in the previous layer. Each node is then activated in accordance with the summed input using its activation function and a threshold for the function. In the input layer, the net inputs to each node are the components of the input pattern. The signals feed through the network only in a forward direction.

This network is trained by using the generalized delta rule. This method attempts to find the most suitable solution (numerical values of weights and thresholds) for a global minimum in the mismatch between the desired output pattern and its actual value for all of the training examples. The degree of mismatch for each input-output pair is quantified by solving for unknown parameters between the hidden and output layer and then by propagating the mismatch backwards through the network to adjust the parameters between the input layer and hidden layer. In this learning procedure, the first pattern is presented as input to a randomly initialized network, and these weights and thresholds are then adjusted in all the links. Other patterns are then presented in succession, and the weights and thresholds adjusted from the previously determined values. This process continues until all patterns in the training set are exhausted (an iteration). The final solution is generally accepted to be independent of the order in which the example patterns are presented. A final check can be performed by looking at the pattern error, which is defined as the square of the mismatch between desired and actual output for each pattern, and the system error, which is defined as the average of all of these pattern errors, to determine whether the final network solution satisfies all the patterns presented to it within a certain threshold error. The set of weights and thresholds in the network are now specifically tailored to "remember" each input and output pattern and can consequently be used to recognize or generate new patterns given an unknown input.

Characteristics of Seismic Recordings

In contrast with 3-C recordings, 1-C recordings only represent behavior of the particle motion of a seismic signal in orthogonal directions. Despite the fact that complete polarization and propagation information cannot be obtained from 1-C recordings which are strongly dependent on the source position and ray direction, the arrival of a seismic signal may still be observed on each separate 1-C recording by changes in the amplitude, frequency, and polarization characteristics which might be obtained by taking different characteristic functions. Several basic

characteristic functions of I-C recordings can be used for this purpose, such as, the absolute value function [Anderson, 1978], the square function [Swindell and Snell, 1977], Allen's function [Men, 1978], the envelope function [Kanasewich, 1981], and the modified differential function [Stewart, 1977]. Here we choose the absolute value function as the input signal of the BPNN because it has the highest fidelity and processing speed and is most objective amongst these functions. We do not use the I-C recording itself because the first motion of an arrival has two directions (up and down) and is source dependent. Figure 1 schematically shows the method with this modification. The amplitude of 1-C recordings is also strongly dependent on the magnitude and epicentral distance of an earthquake. In order to utilize a small training data set to cover all the recordings with different amplitudes, each segment of absolute values is individually normalized before it is fed into the BPNN. Usually, the vertical component recordings are used for picking arrivals in many other methods, so they are used in our first attempt.

BPNN Structure and Training Data Set

We know that for each problem there exists an optimum BPNN, but finding it might be difficult. Generally, there is no fixed rule to select the structure, and it must be determined by a process of trial and error. We finally select a three-layer BPNN with 40 nodes in its input layer, 10 nodes in its hidden layer, and two nodes in its output layer. The input of this BPNN is a normalized segment of absolute values of the V-C recordings of local earthquake data with length 390 ms (40 samples). Two output nodes (o_1 and o_2) of the BPNN flag the input segment with (0,1) for P arrivals and (1,0) for background noise. It seems that one output node can also flag the two states. However, if there are more than two output states, one node cannot properly be used. For this kind of pattern recognition, the output node should be equal to the desired output states [McCormack, 1991; Haykin, 1994]. In order to be consistent with this, most authors prefer two nodes for the problem with two desired output states [Dowla *et al.*, 1990; Gorman and Sejnowski, 1988; McCormack *et al.*, 1993; Penn *et al.*, 1993]. We also use two nodes in this approach.

The training data set is reselected from the V-C recordings of local seismic data because the onset times of P arrivals in V-C recordings are, sometimes, different from those in the vector modulus. The number of training patterns required is determined

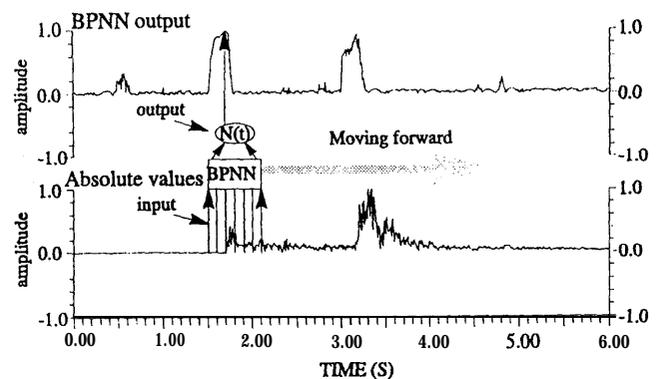


Figure 1. Schematic diagram showing the method of using a BPNN to pick seismic arrivals. The input of this BPNN is the absolute values of the 1-C recordings. A trained BPNN is treated as a filter moving across the entire 1-C recording by a sliding window. The output of BPNN yields a time series which enhances the changes in the absolute values to indicate the arrival onset.

by the nature of the problem and the anticipated performance of a trained BPNN. Although some rules give guidance on this [Baum and Haussler, 1989; Faussett, 1994], experience still plays an important role in selecting the suitable number. Too small a number may result in a poorly trained BPNN, and too large a number may result in the learning procedure becoming too long. It seems better to train a BPNN with a small training data set first, and then to improve its performance by subsequently adding more training data sets. In order to balance the performance of a trained BPNN for P arrivals and background noise, the same number of P arrivals and background noise segments are required to train the BPNN. Selecting the training data set is simple. After training with a selected set of training samples, the network is tested with other seismograms. If the test proves unsuccessful, we include this in the overall training set. This procedure is repeated until the success performance is acceptable. The training sample should contain the general characteristics of P arrivals. Finally, by a process of trial and error, 10 pairs of P arrival and background noise segments (Figure 2) from station DP of the TDP3 seismic network [Lovell, 1989] are used to train this BPNN.

Rules for Picking Arrivals

After training, the BPNN is applied to the entire 1-C recording by sliding a small window, of the size of the input layer, along

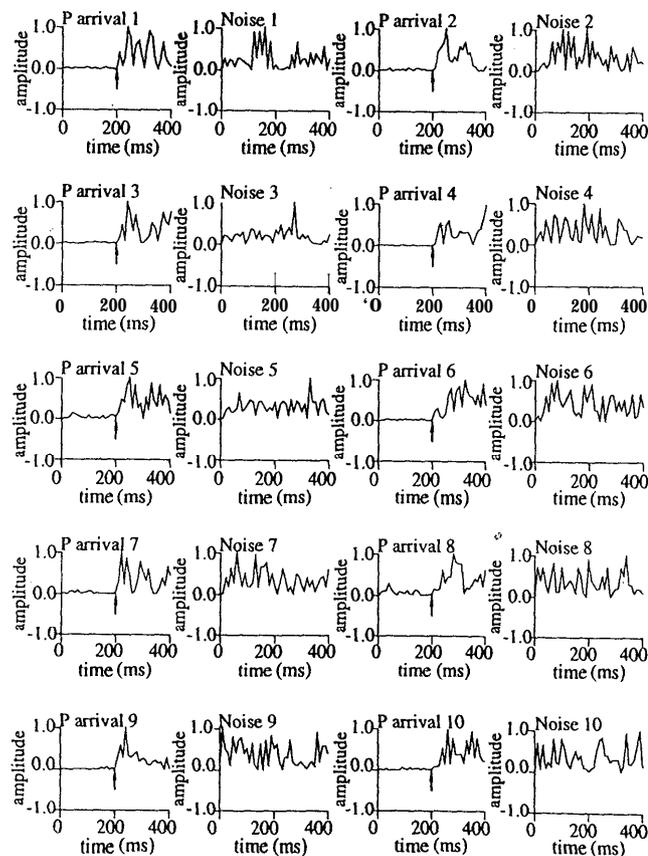


Figure 2. Ten pairs of P arrivals and background noise segments of absolute values of V-C recordings used for training a BPNN. Noise segments are extracted prior to the P arrivals in the same seismograms. Arrows on P arrival segments, all at the 21st sample, indicate arrival onset used to train the BPNN. These segments are individually normalized before being fed into the BPNN.

the absolute trace values. The resulting outputs are converted into a time series $N(r)$:

$$N(t) = \frac{1}{2}[(1 - o_1(t))^2 + o_2(t)^2] \quad (1)$$

which is interrogated for a decision regarding arrivals. This function exaggerates the difference between the desired output and the background noise. A peak in $N(r)$ corresponds to characteristic change in amplitude, and its maximum indicates the onset of the arrivals. If the change is similar to a training P arrival, the peak value should be close to 1. Usually, the peak value indicates the similarity between a change and a training P arrival. The rules used to detect and pick arrivals from this time series are the same as in our previous approach [Dai and MacBeth, 1995]: (1) detect an arrival using a simple threshold rule on $N(t)$; and (2) pick the arrival onset time using the local maximum of $N(r)$. The local maximum is in the window, beginning when the $N(t)$ exceeds the threshold, with a length of the input segments. Note that the onset time is relatively insensitive to the picking threshold.

Postprocessing for Discarding Noise Bursts and Spikes

In the $N(t)$ curve, some peaks correspond to noise bursts with low amplitude and signal-to-noise ratio (SNR), and some to spikes which are typically one or two samples of anomalously large amplitude compared with the background signal. They may be discarded by a BPNN trained with samples including such noise bursts and spikes. However, it will involve a large network structure and a long training procedure. Because discarding the noise bursts and spikes is a simple task, we prefer to discard them by using conventional algorithms described below. Note that this postprocessing can only reduce false alarms, it cannot improve the BPNN's picking capability.

Discarding small noise burst. The features of a small noise burst are low amplitude and low SNR. Two criteria, mean amplitude criterion and mean SNR criterion, may therefore be used to discard this noise burst. The mean amplitude of a segment is calculated from the onset time to the end of the sliding window with the length of input segment.

$$\text{mean amplitude} = \frac{1}{N} \sum_{i=1}^N m_i \quad (2)$$

where m_i is the amplitude value, and N is the window length. A threshold obtained from the background signals is then applied to the mean amplitude. If the mean amplitude is below the threshold, it is a small noise burst. For the entire data set used, the mean amplitude threshold is fixed to be 10.

The mean SNR criterion is defined as the ratio between the mean amplitudes after and before the onset time.

$$\text{mean SNR} = \frac{\text{mean amplitude}_{\text{onset}}}{\text{mean amplitude}_{\text{before onset}}} \quad (3)$$

If the mean SNR is below a threshold, it is a small noise burst. For the entire data set used, the threshold is fixed to be 1.7.

Discarding spikes. The feature of spikes is that only one or two samples have much large values than other samples. So we define a spike amplitude ratio criterion to discard the spike. Three steps are used to define this criterion. First, in the sliding window in which an arrival is picked, the peaks of amplitude are chosen as p_i ($i < \text{window length}$). Second, the mean amplitude of peaks

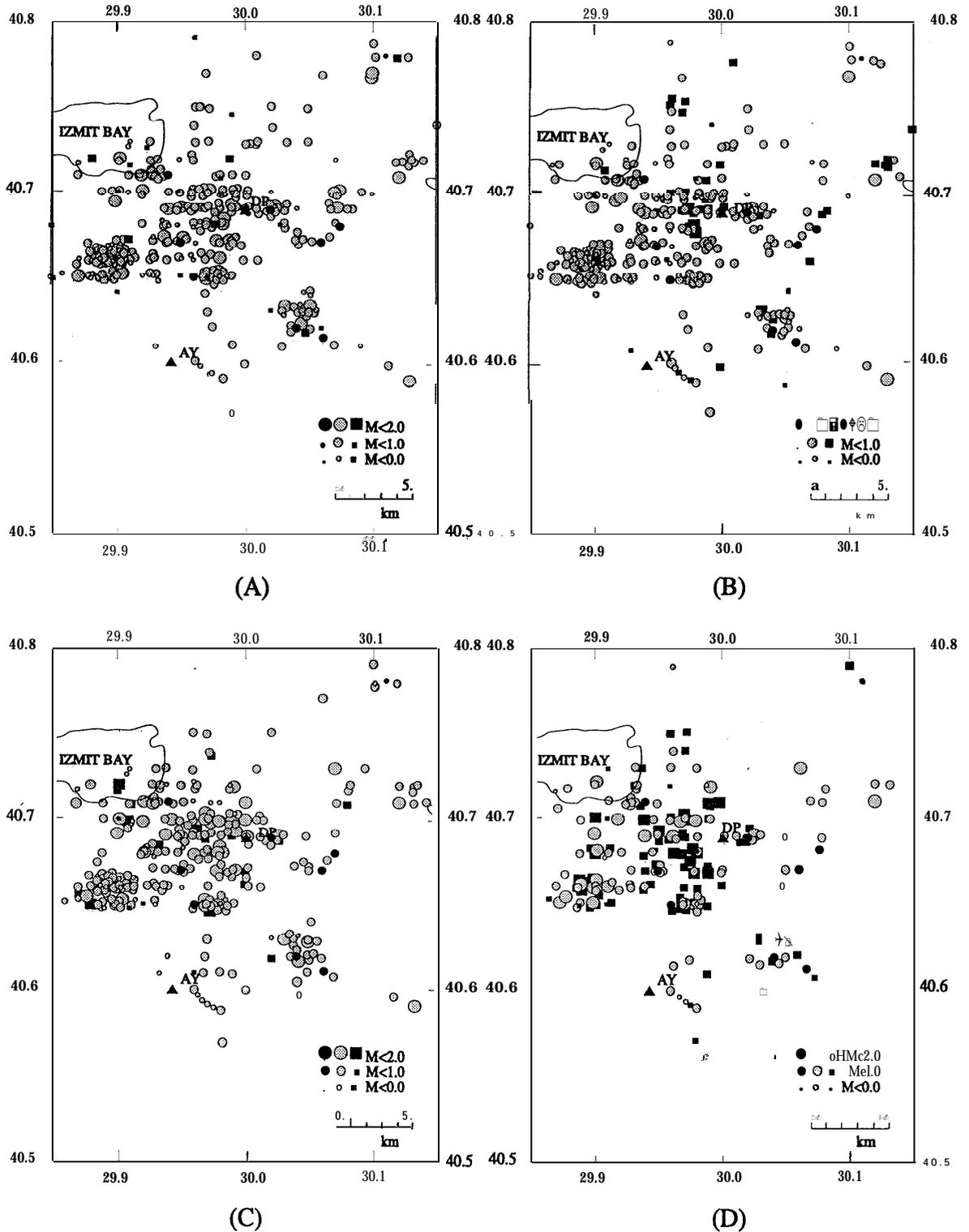


Figure 3. The map of events recorded on vertical component recordings. (a) shows the P arrivals from station DP, (b) the S arrivals from station DP, (c) the P arrivals from station AY, and (d) the S arrivals from station AY. Each symbol represents an event whose (P or S) arrival was manually picked. The grey circles represent arrivals picked by the trained BPNN, and dark grey squares represent arrivals missed by this BPNN. Black circles represent the training P arrivals from station DP. Two solid triangles represent stations DP and AY.

is calculated, excepting the largest two peaks. Third, the ratio defined as:

$$\text{spike amplitude ratio} = \frac{\text{mean peak amplitude}}{\text{maximum of peaks}} \quad (4)$$

If this ratio is below a given threshold, it is a spike. For the entire data set used, the threshold value is fixed to be 0.1.

The values of these criteria are easy to choose as they are not sensitive. These criteria can be used to filter out most small noise

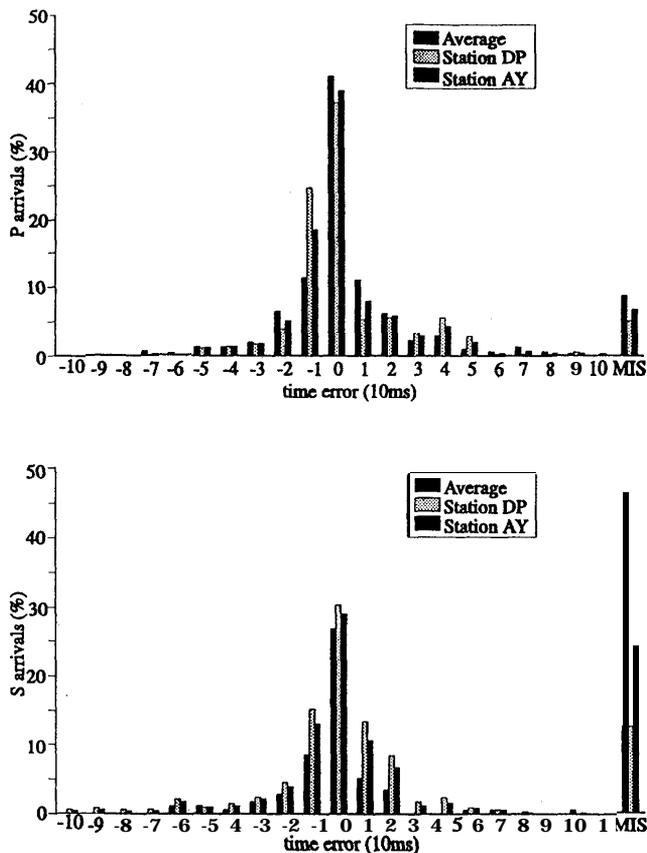


Figure 4. Statistical comparison of (top) P and (bottom) S arrivals picked by the BPNN with manual picks from V-C recordings. A negative value of time error indicates a later pick by trained BPNN approach than that given by visual analysis, and the positive value indicates an earlier pick. MIS refers to all those unpicked arrivals, defined as picks with error larger than 10 sample increments (100 ms). The success rate of trained BPNN relative to manual reference picks is quoted as a percentage.

bursts and spikes, however, other kinds of noise burst which are more similar to seismic arrivals are still left for further identification. We confine the discussion of such discrimination to a secondary stage of the analysis scheme, that is, the arrival identification.

Testing on a Complete Data Set

After optimizing the BPNN's structure and training data set, the trained BPNN is applied to a data set including 762 recordings (371 from station DP and 391 from station AY) of TDP3 seismic network [Lovell, 1989] which was used in our previous approach [Dai and MacBeth, 1995]. Due to the ray path, some P arrivals and some S arrivals are low amplitude or absent from 1-C recordings, especially from smaller events, and some arrivals have different onset times on different 1-C recordings due to phase change. We must manually pick the P and S arrival onset times from three 1-C recordings and use them as references to compare with the BPNN's picking results. The trained BPNN will be used to deal with three 1-C recordings, respectively.

Testing Results

The testing results do not show any obvious relationship between the picking ability of this trained BPNN and event

positions. Figure 3, as an example, displays the picking results from vertical recordings related to event positions. Although the P arrivals used to train this BPNN have high SNR, this trained BPNN can also pick P and S arrivals with low SNR.

Figures 4, 5, and 6 show the statistics of the time accuracy of the picked onset time compared with manual picks. As picked onset times are insensitive to the picking threshold, only the results with threshold 0.6 are shown. These figures show that stations DP and AY have different performances. Detecting and picking results from station AY are poorer than those from station DP. This is largely a consequence of the data quality and different characteristics. The recordings on station DP usually have stronger energy and higher SNR than those on station AY. The arrivals on station DP are much clearer than on station AY, so the arrival times picked by the BPNN from station DP are more accurate than those from station AY. The source orientation has a larger effect on the picking of S arrivals than that on the picking of P arrivals. For example, S arrivals of the V-C recordings in station AY are masked by the P wave coda for some small events, so it is difficult to detect and pick these S arrivals, even by visual analysis. Figure 7a shows such an example in which the S arrival in V-C recordings has a smooth first motion and is consequently not picked by the BPNN, but on the horizontal components it is clearer. This S arrival is picked from both N-S and E-W components (Figure 7b and 7c). In contrast with the S arrival, the P arrival is missed from the two horizontal components. However, both P and S arrivals are picked from the modulus (Figure 8). It is difficult to say from which component the method has the best performance. For example, for P arrivals the best performance on station DP is from the V-C recordings, but on station AY the best

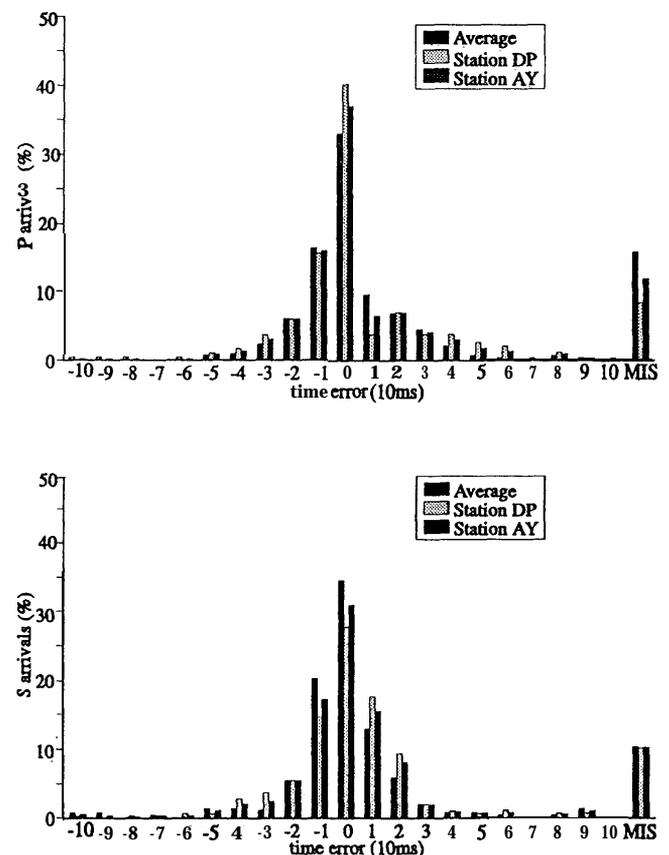


Figure 5. Statistical comparison of (top) P and (bottom) S arrivals picked by the BPNN with manual picks from E-W recordings.

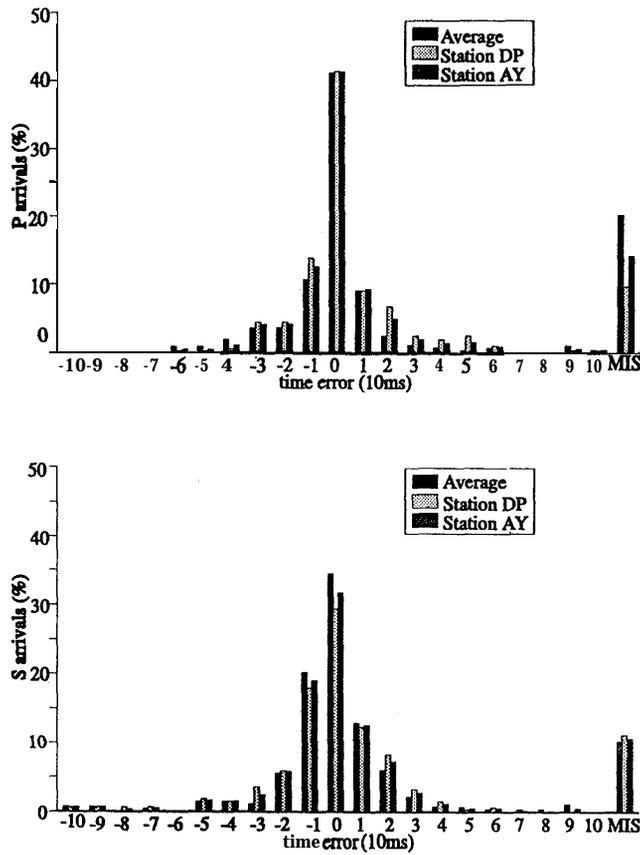


Figure 6. Statistical comparison of (top) *P* and (bottom) *S* arrivals picked by the BPNN with manual picks from N-S recordings.

one is from the E-W component; for *S* arrivals, on station DP the best performance is still from the V-C recordings, but on station AY the best one is from the N-S component.

Three onset times obtained from three 1-C recordings of the same arrival might be different due to the data quality and SNR which are dependent on the source position and many other factors. For the event shown on Figure 7, the *P* onset time picked from V-C recordings is 40 ms (four samples) later than that from modulus, and both the *S* onset times from the E-W and the N-S component are 10 ms (one sample) later than the modulus results. In comparison with the manual picks, the onset times picked from the modulus are more accurate.

Overall Performance From Three 1-C Recordings

Tables 1 and 2 summarize the picking results from three 1-C recordings. Note that this BPNN is trained only with *P* arrivals and background noise from Station DP. However, it can be applied to different stations for picking other kinds of arrivals, although the picking ability varies as the data quality is changed. This shows that seismic arrivals have some general characteristics. This BPNN trained with *P* arrivals has learned these general characteristics and can use them to pick other kind of arrivals, even from other stations. However, if an arrival has specific characteristics which are quite different from this general character, the BPNN will not be able to pick it.

This BPNN approach is capable to suppress the noise bursts. Table 3 shows the recording numbers in which noise bursts are picked by the BPNN. Amongst these recordings, most of them have only one noise burst detected in an entire recording. These

noise bursts are similar to the seismic arrivals, so even in manual analysis, it is also difficult to discard them by only using the noise burst segments themselves. An analyst must compare the noise burst with other picks, even with the data from other stations to discard them.

The above results suggest that a trained BPNN can be applied to all three components individually. It is then possible to combine and compare results from each component, in which some arrivals may be absent, to obtain both *P* and *S* arrival onset times, such as the case shown on Figures 7 and 8. Even if one of the three components fails, the other two can give useful results. The arrival

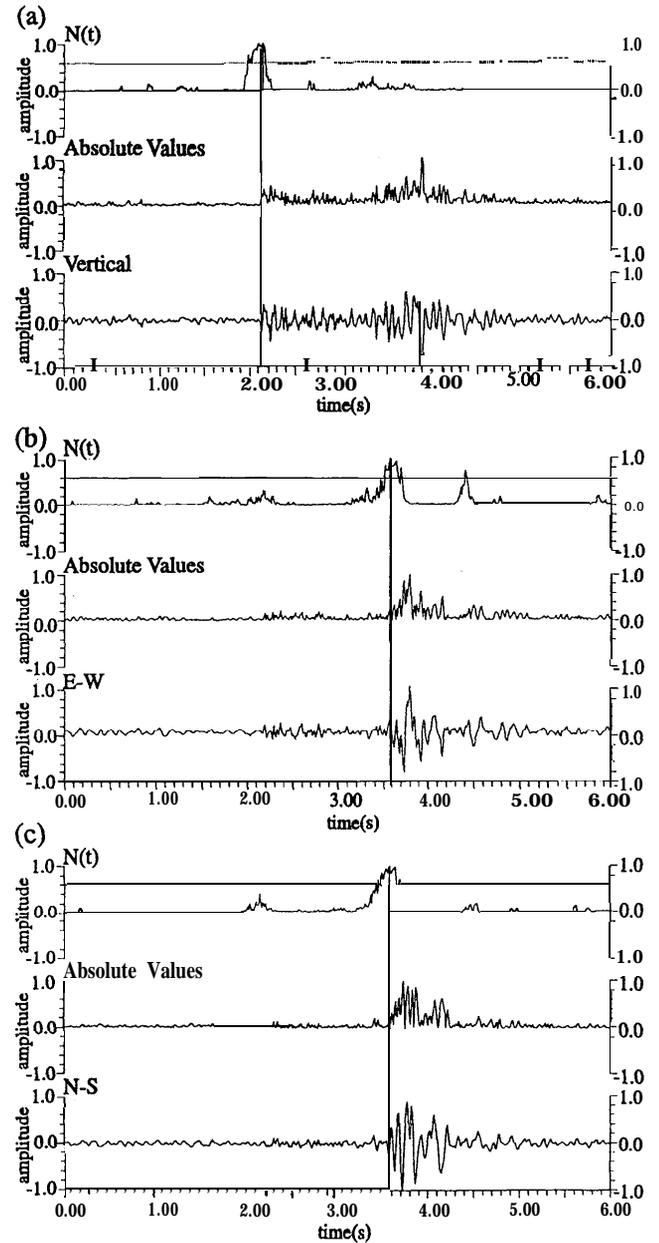


Figure 7. Examples of the BPNN output $N(t)$, with the three 1-C recordings and their absolute values of a local earthquake. Vertical lines are automatically drawn by this approach. Dashed line is the picking threshold (0.6) applied to $N(t)$. The BPNN only picks the *P* arrival from Vertical recording and the *S* arrival from E-W and N-S recordings. Station AY; May 3, 1984; 0452:38UT; Scale 641 in Figure 7a, 1086 in Figure 7b and 1334 in Figure 7c.

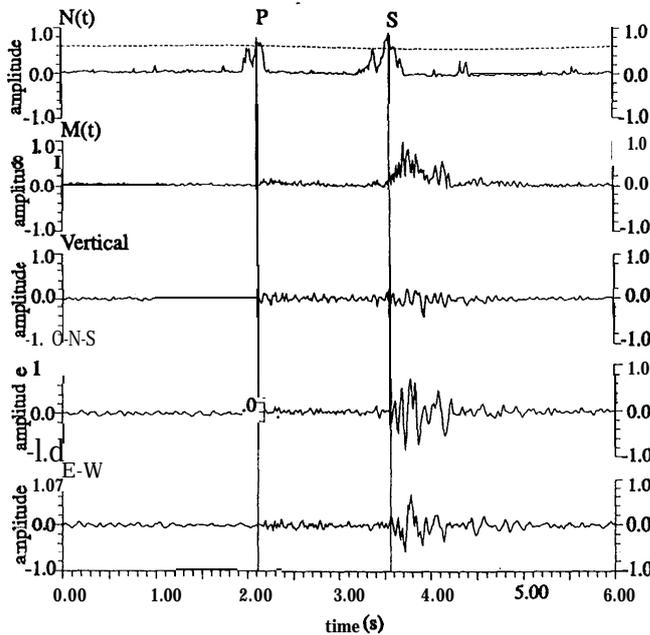


Figure 8. An example of the BPNN output $N(t)$, with the three component recordings and its modulus of the same earthquake in Figure 7. The method can pick both the P and S arrivals. Station AY; May 3, 1984; 0452:38UT; Scale 1650.

times picked by the BPNN on three 1-C recordings are usually different, which is thought to be due to the effect of ray path or variations in the inhomogeneous upper crust.

Comparison With Other 1-C Methods

As discussed in the introduction, there are many picking algorithms available. However, only a few of them are in common use. Table 4 gives a comparison of their principles and general performance against our method. Because articles tend to describe principles and show a few examples, these cannot be directly or wholly compared with our result which is applied to a specific data set of local events, so that this table may not be truly representative of the optimal forms of each technique. However, it does appear that the high success percentage and the small estimation error for both P wave and S wave analysis is potentially encouraging for future work. We believe that an additional strength of the neural network is that it can deal with raw data once it has been trained appropriately. This contrasts with many other techniques, which rely upon preprocessing steps to generate control parameters. The network presented here is relatively quick to train and has been shown to be adaptive to various types of waves.

Comparison With 3-C Method

Tables 1 and 2 also show the picking results from our previous 3-C method. Compared with the results from the 1-C method, the 3-C method has quite obviously a better performance, being more accurate. Visually checking the seismograms shows that the arrivals onset times picked from the 3-C modulus are closer to the manual picks than those from 1-C recordings. The seismic arrivals on 1-C recordings are sometimes not clear or missing, especially for the S arrivals. For example, only 175 S arrivals are manually picked from the V-C recordings of station AY, half of those (305) picked from the modulus of 3-C recordings. But even of the 175 S arrivals, some have no clear first motion, or very low SNR

and amplitude. In this case, the BPNN loses its ability to detect them. This is due to the fact that the BPNN is trained by manual picking results and uses this experience to deal with new data, so that it cannot go beyond the manual picking results.

Comparing the noise bursts picked from 1-C recordings and 3-C recordings, the 3-C method is found to be superior to the 1-C method and can efficiently suppress the noise output (Table 3). This is because the seismic arrivals in three 1-C recordings are synchronous, and the noise is not synchronous. The modulus of 3-C can sum up the changes of an arrival which are projected on the three 1-C recordings, but the noise in the three 1-C recordings may counteract each other. Using 3-C recordings also requires a small BPNN so that its speed in training and processing is quicker than that of 1-C recordings.

Summary

A BPNN is used as a tool to pick P and S arrivals from three 1-C recordings. The input of this BPNN is the absolute value of the 1-C seismic recordings. This BPNN is trained by a small subset of the data (10 P arrival and background noise segments) from the V-C recording of station DP and can successfully detect and pick seismic arrivals not only from V-C recordings but also from the other two horizontal components and other stations. The performance is different for each of the three 1-C recordings due to strong effects of ray directions and source positions. The effect is smaller on P arrivals than on S arrivals. The accuracy of the onset times picked from each individual 1-C recordings is similar.

Comparing with other 1-C methods, this approach is flexible and can pick both P and S arrivals simultaneously. It has a better

Table 1. Summary of Picking Performances for Arrivals From Different Component Recordings

N(t)	Threshold	DP		AY		Overall	
		0.6	0.5	0.6	0.5	0.6	0.5
<i>P Arrivals Picked by the BPNN</i>							
V-C	Manual	350	350	298	298	648	648
	BPNN	328	330	275	275	603	605
	Percent	93.7	94.3	92.2	90.2	93.1	93.4
E-W	Manual	356	356	295	295	651	651
	BPNN	327	330	255	257	582	587
	Percent	91.9	92.7	86.4	87.1	89.4	90.2
N-S	Manual	356	356	265	265	621	621
	BPNN	321	327	195	201	516	528
	Percent	90.2	91.8	76.2	75.8	83.1	85.0
3-c	Manual	356	356	300	300	656	656
	BPNN	345	352	274	277	619	629
	Percent	96.9	98.8	91.3	92.3	94.3	95.8
<i>S Arrivals Picked by the BPNN</i>							
V-C	Manual	334	334	175	175	509	509
	BPNN	289	306	93	115	382	421
	Percent	86.5	91.6	53.1	65.2	75.0	82.7
E-W	Manual	342	342	281	281	623	623
	BPNN	308	315	258	265	566	580
	Percent	90.1	92.1	91.8	94.3	90.9	93.1
N-S	Manual	341	341	282	282	623	623
	BPNN	302	307	241	254	543	561
	Percent	88.6	90.0	85.4	90.1	87.2	90.0
3-c	Manual	342	342	285	285	627	627
	BPNN	302	316	240	259	542	575
	Percent	88.3	92.4	84.2	90.9	86.4	91.7

Table 2. Summary of Picking Rate With Onset Time Error \leq 10 ms (One Sample Increment)

	DP	AY	Overall
<i>Picking Rate for P arrivals</i>			
V-C	67.1	65.1	66.2
E-W	58.9	60.0	59.2
N-S	63.8	55.8	60.3
3-C	80.0	68.1	74.5
<i>Picking Rate for S arrivals</i>			
V-C	58.9	40.6	52.7
E-W	54.6	69.0	61.2
N-S	59.5	55.7	57.7
3-C	68.1	57.2	63.2

Values in percent.

The picking threshold is 0.6.

Table 3. Number of Recodings Contained Noise Burst Picked by BPNN

	DP (371*)	AY (391*)	Overall (762*)	Overall Percent (%)
V-C	73	21	94	12.3
E-W	104	54	158	29.7
N-S	95	44	139	18.2
3-C	74	24	98	12.9

* Total recording number.

Table 4. Comparison of Selected 1-C Picking Methods

Reference	Method	Wave Type	Picking Result	Time Error
Allen [1978]	STA/LTA*	P	60-80%	\leq 0.05 s
Baer and Kradolfer [1987]	modified STA/LTA	P	local: 65.9% region: 79.5% teleevent: 90%	\leq 1 sample \leq 3 sample
Joswig and Shulte-Theis [1993]	master-event correlation	P	80% for weak events	\leq 1 sample
Dai and MacBeth [this paper]	artificial neural network	P and S	90% for P 83% for S 63% for P 59% for S	\leq 100 ms \leq 100 ms \leq 10 ms \leq 10 ms

* STA/LTA: Short term average/Long term average.

performance than others. In contrast with our previous 3-C picking, the performance of 1-C picking is lower, with the associated disadvantage of a larger BPNN structure. It depends on the ray directions and the accuracy of arrival onset times is reduced. But it has an obvious advantage of flexibility over 3-C picking as it can be used on 1-C recordings when 3-C recordings are not available.

The work in this paper and our previous paper [Dai and MacBeth, 1995] demonstrates the adaptive nature of the BPNN. Both the 1-C approach and the 3-C approach employ the same computer programs. Only the BPNN structure and the training data set are changed. Further, without changing the programs, this BPNN approach might also be used to deal with regional and teleseismic earthquake data.

Acknowledgments. This research is sponsored by Global Seismology Research Group (GSRG), British Geological Survey (BGS) of the Natural Environment Research Council (NERC), and is published with the approval of the Director of the BGS (NERC). We thank Chris Browitt, David Booth, and John Lovell of BGS for supplying the earthquake data. Thanks are extended to the staff and student of GSRG for their support and encouragement with this work.

References

- Allen, R V., Automatic earthquake recognition and timing from single trace, *Bull. Seismol. Soc. Am.*, **68**, 1521-1532, 1978.
- Anderson, K. R., Automatic processing of local earthquake data, Ph.D. thesis, Mass. Inst. of Technol., Cambridge, 1978
- Bache, T. C., S. R Bratt, J. Wang, R M. Fung, C. Kobryn and J. W. Given, The intelligent monitoring system, *Bull. Seismof. Soc. Am.*, **80**, 1833-1851, 1990.
- Baer, M. and U. Kradolfer, Automatic phase picker for local and teleseismic event, *Bull. Seismol. Soc. Am.*, **77**, 1437-1445, 1987.
- Baum, E. B. and D. Haussler, What size net give valid generalization, *Neural Comput.*, **1**, 151-160, 1989.
- Chiaruttini, C. and G. Salemi, Artificial intelligence techniques in the analysis of digital seismograms, *Comput. Geosci.*, **19**, 149-156, 1993.
- Chiaruttini, C., V. Roberto and F. Saitta, Artificial intelligence techniques in seismic signal interpretation, *Geophys. J. Int.*, **98**, 223-232, 1989.
- Dai, H. and C. MacBeth, Automatic picking of seismic arrivals in local earthquake data using an artificial neural network, *Geophys. J. Int.*, **120**, 758-774, 1995.
- Dowla, F., S. R Taylor, and R W. Anderson, Seismic discrimination with artificial neural networks: Preliminary results with regional spectral data, *Bull. Seismoi. Soc. Am.*, **80**, 1346-1373, 1990.
- Fausett, L., *Fundamentals of Neural Networks*, Prentice-Hall, Englewood Cliffs, N. J., 1994.
- Gorman, R P. and T. J. Sejnowski, Analysis of hidden units in a layered network trained to classify sonar targets, *Neural Networks*, **1**, 75-89, 1988.
- Haykin, S., *Neural Networks, A Comprehensive Foundation*, Macmillan College Publishing, New York, 1994.
- Houliston, D. J., G. Waugh, and J. Laughlin, Automatic real-time event detection for seismic network, *Comput. Geosci.*, **10**, 431-436, 1984.
- Joswig, M., Pattern recognition for earthquake detection, *Bull. Seismof. Soc. Am.*, **80**, 170-186, 1990.
- Joswig, M., Automated classification of local earthquake date in the BUG small array, *Geophys. J. Int.*, **120**, 262-286, 1995.
- Joswig, M. and H. Schulte-Theis, Master-event correlations of weak local earthquake by dynamic waveform match, *Geophys. J. Int.*, **113**, 562-574, 1993.
- Kanasewich, E. R., *Time Sequence Analysis in Geophysics*, Univ. of Alberta Press, Edmonton, 1981.
- Lovell, J., Source parameters of microearthquake swarm in Turkey, Master of Philosophy thesis, Univ. of Edinburgh, Edinburgh, Scotland, 1989.
- McCormack, M. D., Neural computing in geophysics, *Geophys. Leading Edge Explor.*, **10**, 11-15, 1991.
- McCormack, M. D., D. E. Zauha and D. W. Dushek, First-break refraction event picking and seismic data trace editing using neural networks, *Geophysics*, **58**, 67-78, 1993.
- Murat, M. and A. Rudman, Automated first arrival picking: A neural network approach, *Geophys. Prospect.*, **40**, 587-604, 1992.
- Pao, Y. H., *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley, Reading, Mass., 1989.
- Penn, B. S., A. J. Gordon and R F. Wendland, Using neural networks to locate edges and linear features in satellite images, *Comput. Geosci.*, **19**, 1545-1565, 1993.
- Pisarenko, V. F., A. F. Kushnir and I. V. Savin, Statistical adaptive algorithms for estimation of onset moments of seismic phases, *Phys. Earth Planet. Inter.*, **47**, 4-10, 1987.
- Rumelhart, D. E., G. E. Hinton and R J. Williams, Learning internal representation by backpropagating errors, *Nature*, **323**, 533-536, 1986.

- Stewart, S. W., Real-time detection and location of local seismic events in central California, *Bull. Seismol. Soc. Am.*, 67, 433-452, 1977.
- Swindell, W. H. and N. S. **Snell, Station processor** automatic signal detection **system**, Phase 1: Final **report**: Station processor software development, *Tan Instrum. Rep. No. Alex (01)-FR-77-01*, Tex. **Instrum. Inc.**, Dallas, Tex., 1977.
- Takanami, T. and G. Kitagawa, A new efficient procedure for the estimation of onset times of seismic waves, *J. Phys. Earth*, 36, 267-290, 1988.
- Takanami, T. and G. Kitagawa, Multivariate time-series model to estimate the arrival times of S-waves, *Comput. Geosci.*, 19, 295-301, 1993.
- Wang, J. and T. L. Teng, Artificial neural network-based seismic detector, *Bull. Seismol. Soc. Am.*, 85, 308-319, 1995.
-
- H. Dai and C. **MacBeth**, British Geological Survey, Murchison House, West Mains Road, Edinburgh EH9 3LA, Scotland. (e-mail: **e_hcd@va.nmh.ac.uk**).

(Received July 3, 1996; revised February 20, 1997; accepted February 25, 1997.)