Hengchang Dai and Colin MacBeth:

Arrival Type Identification in Local Earthquake Data using an Artificial Neural Network

Tiré à part des
Cahiers du Centre Européen de Géodynamique et de Séismologie.
Volume 12 - 1996

ARRIVAL TYPE IDENTIFICATION IN LOCAL EARTHQUAKE DATA USING AN ARTIFICIAL NEURAL NETWORK

Hengchang Dai[1,2] and Colin MacBeth[1]

[1] *British Geological Survey Murchison House, West Mains Road. Edinburgh EH9 3LA Scotland, UK.*
[2] *Department of Geology and Geophysics, University of Edinburgh. Edinburgh EH9 3JW, Scotland. UK.*

ABSTRACT

A preliminary study is performed to test the ability of an artificial neural network (ANN) to identify seismic arrival types from local earthquake data after they are picked. This is achieved using the degree of polarization (DOP) for a segment of three-component time series, as the ANN input. The ANN was designed to classify arrivals into three groups: P-arrivals, S-arrivals and noise, corresponding to the maximum output of the output nodes. The ANN was trained with nine groups of segments of *P*- and s-arrivals and noise. 327 pre-triggered recordings from a station in a local earthquake network are firstly processed by an ANN picker for all possible *P*- and S-arrivals and measured their onset times. Segments of the DOP are selected according to these onset and then fed into the trained ANN. Compared with manual analysis, the trained ANN can correctly identify 84% P-arrivals and 63% S-arrivals. Its performance has inherent limitations due to the complexity of DOP patterns which cannot be improved by simply adding new training datasets. The example is shown that the ANN trained with data from one station fails to deal with seismic data from another station as the DOP patterns are station-dependent. This limitation shows that selecting the input information is critical. The ANN has potential as a tool to identify the arrivals type automatically but needs to be associated with other information.

Keywords: artificial neural network, seismic arrival identification, degree of polarization.

1. INTRODUCTION

The most important procedure of analysing earthquake events is the estimation of the arrival times of the primary (P) and secondary (*S*) waves, as these measurements form the basis of subsequent analysis schemes employing processing for event location, event identification, source mechanism analysis and spectral analysis. These tasks are often performed by the trained analyst who manually picks arrival times according to his individual experience, involving an intensive amount of pattern recognition. With the increase in the number of digital seismic networks being established worldwide, there is a pressing need to provide an automatic alterative, which is more reliable, robust, objective and less time-consuming.

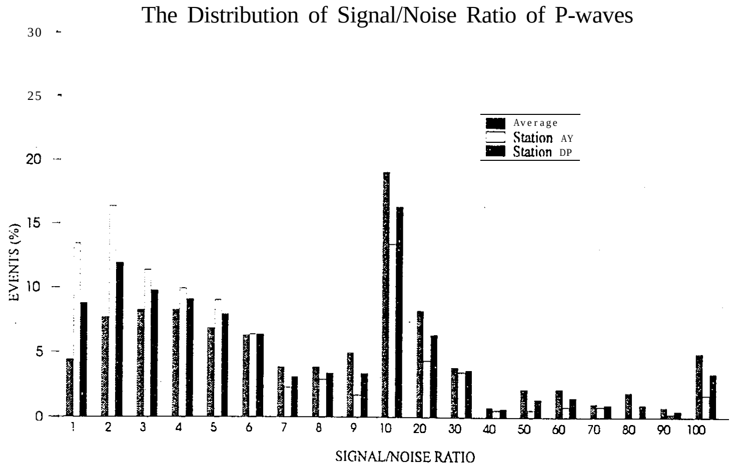
The estimation of arrival times includes two steps: 1) arrival picking which is a reliable and accurate estimation of the onset time of a definite seismic arrival; 2) arrival identification which classifies individual arrivals into categories relating to their polarization, their amplitude and the nature of their propagation. A great deal of effort, stretching back several decades. has been devoted to the automation of arrival picking (Allen, 1982; Bache *et al*, 1990; Bear and Kradolfer, 1987; Chiaruttini, Roberto and Saitta, 1989; Chiaruttini and Salemi, 1993, Dai and MacBeth, 1995; Houliston, Waugh and Langhlin, 1984; Jowsig, 1990, 1995; Jowsig and Schulte-Theis, 1993; Kracke. 1993; Klumpen and Jowsig, 1993. Pisarenko, Kushnir and Savin, 1987; Takanami and Kitagawa, 1988, 1993). Identifying arrival types is more difficult than picking their onset time and still is an unresolved problem. For a seismic network, arrival identification based on horizontal velocity from an f-k filter can provide a major simplification of the interpretation task (Mykkeltveit and Bungum, 1984; Bache *et al*, 1990; Kvaerna and Ringdal, 1992). Der, Baumgardt and Shumway (1993) have investigated the feasibility of adaptive; automatic recognition of regional arrivals by a wavefield extrapolation scheme for data from a mini-array. However for single station data, there are few methods which can be used to pick special type arrivals. Roberts, Christoffersson and Cassidy (1989), based on the auto- and cross-correlations of the three orthogonal components within a short time window, detect the arrival of a *P-wave* or a linearly polarized S-wave. Cichowicz (1993) developed a S-phase picker which depends on a well defined pulse of the first-arrival P-wave. In this paper, we will introduce an artificial neural network (ANN) approach to the identification problem.

2. THE LOCAL EARTHQUAKE DATA

In this work, real earthquake data are used to design and test an ANN approach. The data are local earthquake events recorded at station DP which is located near the centre of the TDP3 seismic network and station AY which is on the edge of the TDP3 seismic network (Lovell, 1989) between April 1984 and December 1984. Several hundred local earthquakes are recorded on three-component seismometers at a $10ms$ sampling interval. These recordings are not continuous and are triggered by a digital system (Evans et *al.* 1987). All are local, with depths from $2km$ to $14km$ and epicentral distances less than $30km$ from the stations, and most are closer to station DP than to station AY. For these local events, we identified predominant *Pg* and *Sg* waves in the seismogram records. Most events have magnitudes (M_L) between -0.3 and 1 .0, and possess a wide distribution of signal-to-noise ratio (SNR) which are shown in Fig. 1 for the complete dataset. All SNRs lie between 1 and 200, with station DP being of higher fidelity than station AY. Not all of these recording can be used due to following reasons. Some events were not earthquakes and some small earthquakes, recorded on Station DP or AY, were not confirmed by network data. These recordings are discarded by comparing with network data. In some cases, the seismometers did not function properly and either one or two components were inactive or possessed high amplitude noise so that some of the three component sets were incomplete; and some recordings have excessive noise preceding the events or ringing throughout the record which produce many false alarms. Here we manually selected the recordings in which the earthquake event has been confirmed by the seismic network data. We must be aware that our statistics will appear more successful than if this procedure had been applied to all the data irrespective of quality. In total, 327 recordings in station DP and 282 recordings in station AY were selected respectively for further processing. We can visually pick 333 P-arrivals and 317 S-arrivals at DP and 283 P-arrivals and 261 S-arrivals at AY. All these recordings were processed by an ANN arrival picker (Dai and MacBeth, 1995) to measure the onset times of all possible *P*- and S-arrivals. Compared with the manual analysis, the ANN picks 326 (97%) P-waves and 286 (92%) S-waves at station DP and 242 (87%) P-waves and 235 (90%) S-waves at station AY.

3. THE DEGREE OF POLARIZATION

The identification of different arrival types is accomplished using a combination of the degree of polarization (DOP) of arrival and the vector modulus of its three-component motion. In



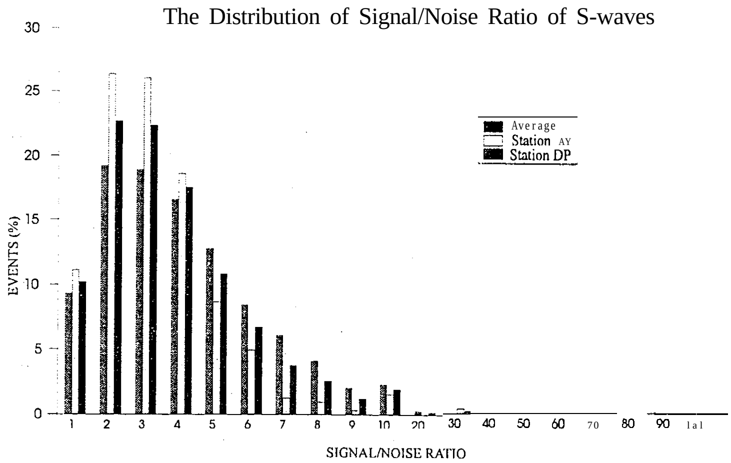


Figure 1 The distribution of signal/noise ration (SNR) of P-waves and S-waves in the dataset. SNR is defined as the ration between the maximum of modulus of the three component recordings before and after the onset time. Notation for SNR represents 1 for [1,2), 2 for [2,3), ... 10 for [10,20), ... and 100 for [100, 200).

each single component, the seismic signals are strongly dependent on the source position and ray direction, which may otherwise give rise to a misleading interpretation. We must separate this dependency from the seismic recordings. In this paper, we input the DOP which is independent of the source position. The DOP is calculated from the covariance matrix of 3-C recordings which is a useful measure of the polarization of seismic signal (Samson, 1977; Cichowicz, Green and Brink, 1988; Cichowicz, 1993). The covariance matrix is defined as:

$$C = \begin{vmatrix} COV(X,X) & COV(X,Y) & COV(X,Z) \\ COV(Y,X) & COV(Y,Y) & COV(Y,Z) \\ COV(Z,X) & COV(Z,Y) & COV(Z,Z) \end{vmatrix} ;$$

where the covariance is measured for N samples:

$$COV(X,Y) = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})$$

where $\bar{x}$ and $\bar{y}$ are the average values of X and Y. N is determined from the signal predominant frequency (Cichowicz, 1993), which is ten samples in this case.

The diagonalization of the covariance matrix gives the principal axis of this matrix. The direction of polarization is measured by considering the eigenvector of the largest principal axis. This direction is parallel to the propagation direction for a P-wave and is perpendicular to the propagation direction for a S-wave in an isotropy medium. It is difficult to use this direction, related to the source position, as a decision parameter for arrival identification. Some parameters which are independent of the source location should be defined to extract the polarization properties. Samson (1977) defines the degree of polarization as:

$$F(t) = \frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}{2(\lambda_1 + \lambda_2 + \lambda_3)^2} = \frac{3\,trS^2 - (trS)^2}{2\,(trS)^2}$$

where the λ_1, λ_2 and λ_3 are eigenvalues of the covariance matrix of a moving window of width N samples; tr**S**, defined as $\lambda_1+\lambda_2+\lambda_3$, is the trace of **C** and tr**S**^{2}, is defined as $\lambda_1^2+\lambda_2^2+\lambda_3^2$. This equation shows that the function can be calculated without having to diagonalize the covariance matrix. As these are independent of the coordinate system, they also are independent of the source location and depend only on the polarization state. According to this definition, if only one eigenvalue is non-zero, then F=1, and the signal is linearly polarized; if all of the eigenvalues are equal. then F=0, and the signal can be considered as completely unpolarized or circularly polarized. Thus $F(t)$ enables us to study the evolution of the degree of polarization (DOP).

For the data used, most P-arrivals have high values of F(t) and most S-arrivals have medium or low values of F(t). Fig. 2 shows an example. The patterns of polarization are too complex to find

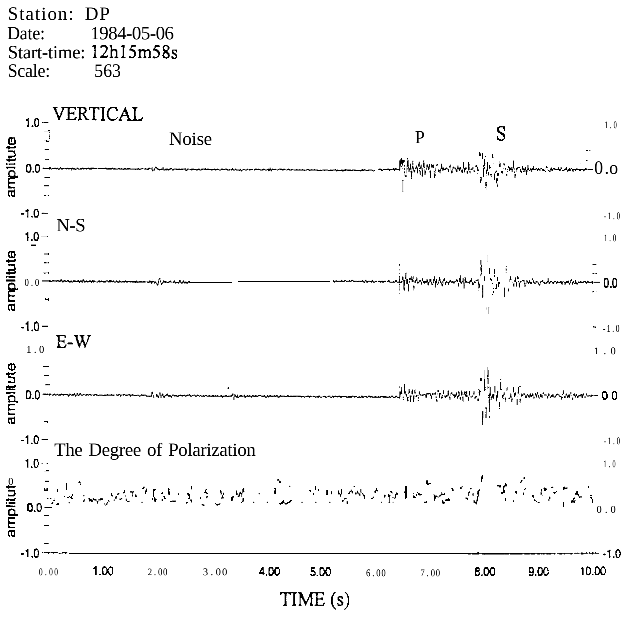


Figure 2. The degree of polarization and three-component seismograms. Three vertical lines indicate the arrivals onset times of noise, a P-arrival and a S-arrival. The degree of polarization has a low value for the noise, a high value for the P-arrival and a middle value for the S-arrival.

threshold values to distinguish them. Most of the F(t) patterns of P-arrivals usually differ from those of S-arrivals. There are also some noise bursts whose F(t) patterns are similar to those from the seismic arrivals. Spikes manifest as special patterns in which F(t) is very high, near unity, with about ten point length and can be easily discarded by a conventional program. Comparing the data from Stations DP and AY, the F(t) patterns are different even for the same earthquake and arrival.

To calculate the degree of polarization, all three components must have the same frequency bandwidth, the same scale and the same noise level. If one of the three-components has a different property, the DOP is highly biased. In this case, even manual identification using only the DOP cannot be used. This particular definition of DOP does not take into account the signal intensity. Different arrival types not only have different polarization characteristics. but also have different amplitude characteristics. To take into account both polarization and amplitude information, we define a modified function of the DOP:

$$MF(t) = F(t) \times M(t)$$

where M(t) is the smoothed relative function of modulus M(t), defined as $[x(t)^2+y(t)^2+z(t)^2]^{1/2}$, of 3-C recording in a window which is also independent of the source position. The normalization factor is taken from the window between the onset point and following ten points, in which the maximum is defined as unity. Note that MF(t) and F(t) may have slightly different patterns. MF(t) is now presented to the neural network in segments selected from a window in which the centre is the onset-time of arrival.

4. IDENTIFYING ARRIVAL TYPES USING AN ANN

4.1 ANN structure

The ANN used in this study is a nonlinear, multilayer, feed-forward and back-propagation of error (Rumelhart, Hinton, and Williams, 1986). This is the most popular type of ANN in use today as it is well understood. It also incorporates a back-propagation learning algorithm, or *Delta Rule* which is usually used to train this type of ANN -- a good mathematical summary is given by Pao (1988).

This ANN has three layers, the input having 60 nodes, giving a MF(t) segment with a fixed $590ms$ (60 samples) length which is chosen to include several complete cycles of a wave. There are three nodes in the output layer to flag the result: the output is (1 ,0,0) for noise; (0,1,0) for a P-arrival; and (0,0,1) for a S-arrival in training. The number of hidden nodes depends on various factors such as input nodes, output nodes, system error, pattern error, and training samples. There is no fixed generic relationship between the number and these factors for this type of ANN. However, we do know that in ANN learning, generalization is increased and memory is reduced by limiting the number of hidden nodes (Dowla, Taylor and Anderson, 1990). Too few hidden nodes will lead to a long learning or no convergence. In this case we chose ten hidden nodes after a process of trial and error with different training runs.

4.2 Training procedure

As each segment of MF(t) is fed into the trained ANN, the output will be three values: o_1, o_2, and o_3. If the segment is the same as the training segment, the output will be perfectly (1,0,0) for noise; (0.1 ,0) for the P-wave; and (0,0,1) for the S-wave. For non-training segments, the output (o_1, o_2, and o_3) is the measurement of similarity between the new segment and training segment. To identify segment types, we simply seek the maximum of the three outputs (o,, o_2, 0,). If o_1 is the maximum, this segment belongs to the noise category; if o_2 is the maximum, it belongs to the P-wave; and if o_3 is the maximum, it belongs to the S-wave. This method is also applied to some segments which are far different from training segnrnents and have low outputs.

In this procedure, only a small number of recordings from station D? are used to train the ANN and the remainder to test its performance. The ANN performance depends on the training datasets, if we use incorrect or inconsistent data to train the ANN, we cannot expect it to give a correct answer for new data. P- and S-arrivals with similar MF(t) patterns should be avoided. At the beginning of training, we select only three MF(t) segments for noise, P- and S-arrivals. Using manual analysis results, we select another three segments, to combine with the former training segments, to train again. This procedure is repeated until the performance of the trained ANN cannot be improved by increasing the training dataset or we are satisfied by its performance. Fig. 3 shows all training segments of MF(t) used in this particular study.

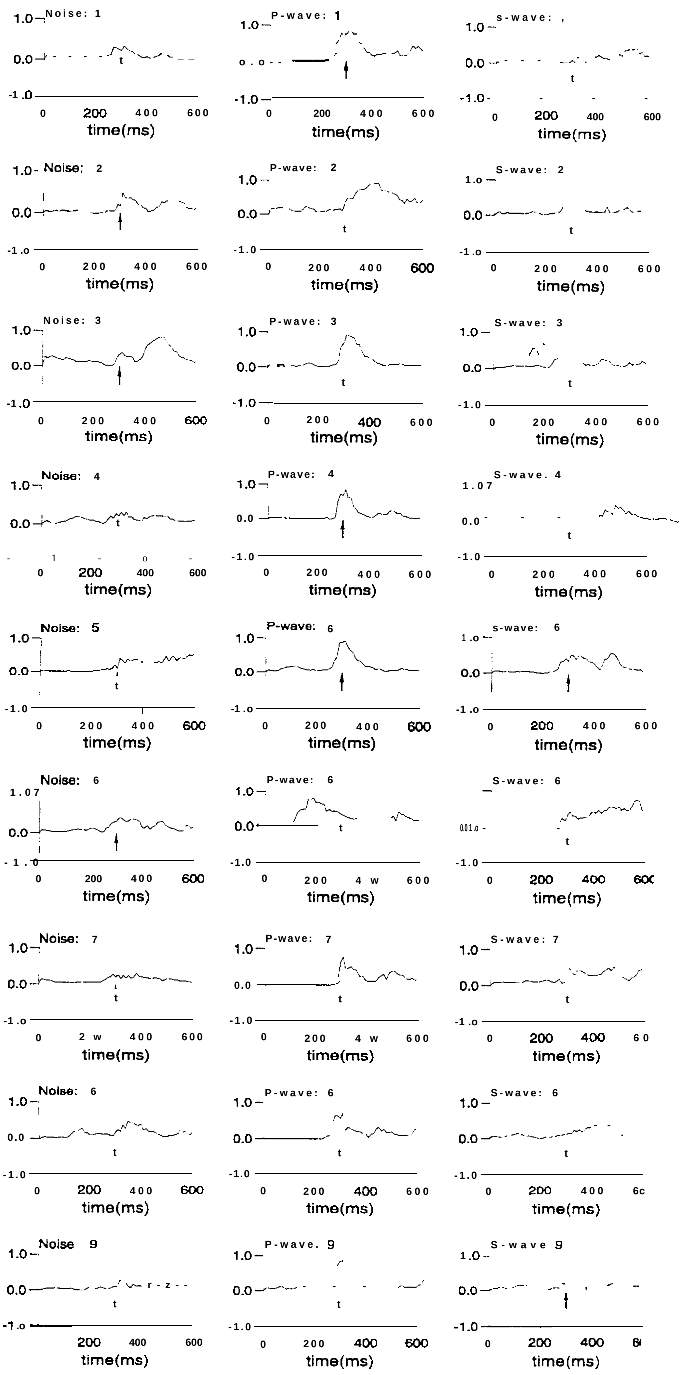


Figure 3. Nine group of MF(t) segment of noise, P-waves ar S-waves. The arrows indicate the adjusted onset times

4.3 Testing procedure

Here we focus on the data from station DP. Unlike our earlier ANN arrival picking, the ANN is not used as a filter to deal with an entire seismic trace, but arrival segments are input which are picked previously. The results are time-sensitive and a small shift in segment can greatly effect the output. Because the onset-times of picked arrivals have errors, we must adjust the onset time to ensure the performance of the trained ANN is not affected by the onset time error. For each MF(t) segment, we set the first local maximum after the onset-time on the centre of the segment. To test the trained ANN performance, we input all the pre-picked arrival segments into this neural network.

Table 1: The performance of the ANN trained with 9 groups of training segments. The ANN has 60 input nodes.

	P-arrivals (326)	S-arrivals (286)	Noise (146)
P-arrivals	84% (274)	20% (58)	10% (16)
S-arrivals	11% (35)	63% (180)	42% (62)
Noise	5% (17)	17% (49)	47% (69)

4.4 The performance of the trained ANN

For a three-layer ANN with 60, 10, 3 nodes, the final training was taken by using 27 training segments (nine noise, nine P-arrivals and nine S-arrivals). Training takes 1422 iterations, about 1.5 minutes on a VAX 4000. Table 1 shows the performance of the final trained ANN, with Fig. 4 displaying an example of correct identification. However, if we use this to deal with Station AY, it only identifies 43% of P-arrivals and 41% of S-arrivals since many P-arrivals have low values of F(t) and most of S-arrivals have high values of F(t) which is contrary to the training. Another ANN is needed to specifically process data from station AY.

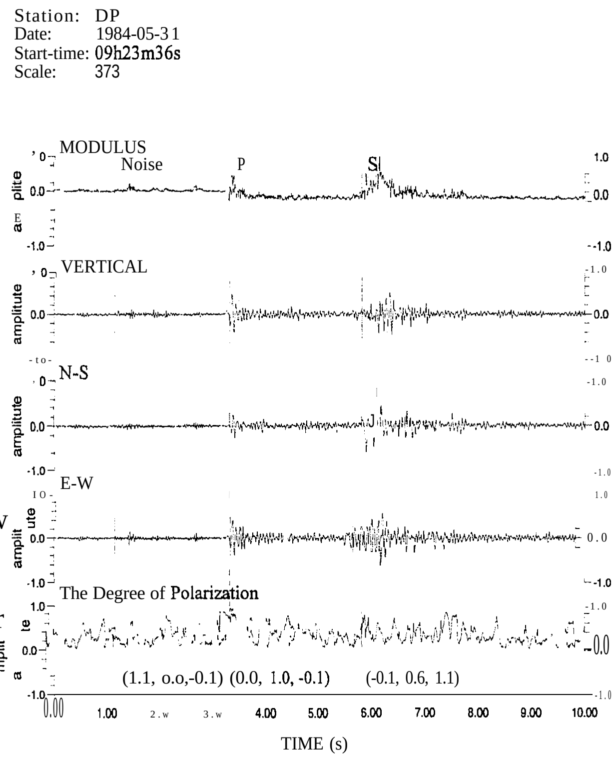


Figure 4. Three component seismograms, the vector modulus, and the degree of polarization of a local earthquake. Three vertical lines indicate the arrival onsets of noise, a P-arrivals and a S-arrival. The ANN correctly identifies them with its output (1.1, 0.0, -0.1), (0.0, 1.0, -0.1) and -0.1, 0.6, 1.1) respectively.

Table 2 shows a comparison of three trained ANNs with different datasets. As the datasets increase, the performance for identifying P-arrivals improves, but the performance for identifying S-arrivals and noise becomes worse. This is due to the complexity of the MF(t) patterns. MF(t) patterns of *P*-waves are more typically alike, and MF(t) patterns of *S*-waves are quite different. In addition, some P-arrivals, S-arrivals and noise have similar MF(t) patterns. If we use such a P-arrival pattern to train the ANN, the trained ANN will generate all them as P-arrivals regardless of what they are. It seems that using the MF(t) alone is not enough to solve this problem. Other properties, for example, the direction of polarization, or some other methods, such as an expert system, may be needed.

Table 2: The comparison of the performances of three ANNs trained with different training dataset Only correct identification oercentage are shown. The best performance is with 9 training groups

	P-arrivals	S-arrivals	Noise
8 training groups	82%	58%	51%
9 training groups	84%	63%	47%
10 training groups	87%	59%	38%

Table 3: The comparison of the performance of the ANN with different input nodes trained with nine groups of segments. Only correct identification percentages are shown in this table. The best performance is from the ANN with 60 input nodes.

	P-arrivals	S-arrivals	Noise
50 input nodes	89%	45%	47%
60 input nodes	84%	63%	47%
70 input nodes	88%	44%	48%

We also investigated the sensitivity to the input segment length as this decides the ANN structure. Various input nodes were tested, between 50 and 70 nodes, retaining the same hidden nodes and output nodes. The training procedure is the same: beginning with one group of training segments and increasing to nine groups. The training segments are different for these three ANNs due to their different performance at every training stage (Table 3). On balance, the ANN with 60 input nodes has the best performance. From this a rough guideline is suggested: the segments should include several complete cycles of a wave. This reflects the general observation that network architecture must be specifically tailored to individual applications. Further optimization is required to adapt to particular event types.

5. CONCLUSIONS

An ANN is used as a tool to identify P- and S-arrivals from local earthquake data, using the polarization state of three-component records. Our results demonstrate that an ANN trained using a small subset of the data can identify? most P-arrivals (84%) and S-arrivals (63%) simultaneously. This high performance, combined with the advantage of not requiring programs to construct special variables and parameters with complicated mathematics: suggest that the ANN is a natural choice for such applications. The method is adaptive, and training sets can be altered to enhance particular features of different datasets. Adding new training datasets and retraining an ANN is easy and quick, and can improve its performance. However it also appears to have a limitation due to the inter-station complexity of the DOP.

Although the training time can be long, especially as the ANN architecture becomes large, once trained the ANN is sufficiently quick to operate in most real-time applications. However. the ANN cannot be viewed as all encompassing, as the performance still depends upon the training set and its ability to predict cannot lie too far outside its experience. The exact boundaries of this behaviour have not yet been completely explored. Another limitation is in finding an optimal architecture for a particular application.

This work forms part of an ongoing programme of research to develop a fully automatic system for earthquake analysis. It is ultimately hoped to integrate other ANN units into a processing flow for record editing and event classification.

ACKNOWLEDGEMENTS

This research was sponsored by the Global Seismology Research Group (GSRG) at the British Geological Survey (BGS), Edinburgh. and is published with the approval of the Director of BGS. We thank Dr. Chris Browitt, Programmes Director of BGS, Mr. Terry Turbitt, Programmes Manager of GSRG, Dr. David Booth and Mr. John Love11 for supplying the earthquake data. Thanks are also extended to the staff and students of GSRG for their support and encouragement with this work.

REFERENCES

Allen, R. V., 1978, Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America*, **68**. 1521 - 1532.

Bache, T. C., Bratt, S. R., Wang, J, Fung, R. M., Kobryn, C., and Given. J. W.. 1990, The intelligent monitoring system, *Bulletin of the Seismological Society of America*, **80**,1833-1851.

Bear, M. and Kradolfer, U., 1987, An automatic phase picker for local and teieseismic event. *Bulletin of the Seismological Society of America. 77*, 1437- *1445.*

Chiaruttini, C., Roberto, V., and Saitta, F.. 1989, Artificial intelligence techniques in seismic signal interpretation, *Geophysical Journal International,* **98**, *223-232.*

Chiaruttini, C. and Salemi, G., 1993, Artificial intelligence techniques in the analysis of digital seismograms, *Computer and Geosciences*, **19**, 149- 156

Cichowicz, A., Green, R., and Brink. A., 1988, Code polarization properties of high-frequency microseismic events, *Bulletin of the Seismological Society of America, 78*, 1297- 13 18.

Cichowicz, A., 1993, An automatic S-phase picker, *Bulletin of the Seismological Society of America, 83*, 180-1 *89.*

Dai, H., and MacBeth, C., 1995, Automatic picking of seismic arrivals in local earthquake data using an artificial neural network. *Geophysical Journal International.* **120,** 758-**774.**

Der, Z., Baumgardt, D., and Shumway, R.. 1993. The nature of particle motion in regional seismograms and its utilization for phase identification. *Geophysical Journal International,* **115,** 1012-1024.

Dowla. F. U., Taylor. S. R., and Anderson. R. W., 1590, Seismic discrimination with artificial neural networks: preliminary results with regional spectral data, *Bulletin of the*

Seismological Society of America, 80, 19 10- 1933.

Evans, S. A., Beamish, D., Crampin, S., and Ucer, S. B., 1587. The turkish Dilatancy Project (TDP3): multidisciplinary studies of a potential earthquake source region, *Geophys. J. R. astr. Soc., 91, 265-286*

Houliston, D. J., Waugh, G., and Laughlin, J., 1984, Automatic real-time event detection for seismic networks, *Computer and Geosciences,* **10, 43** l-436.

Joswig, M., 1990. Pattern recognition for earthquake detection, *Bulletin of the Seismological Society of America, 80,* 170- 186

Joswig, M., 1995, Automated classification of local earthquake data in the BUG small array. *Geophysical Journal International, 120, 262-286.*

Joswig, M., and Schulte-Theis, H., 1993, Master-event correlations of weak local earthquake by dynamic waveform match, *Geophysical Journal International, 113, 562-574.*

Klumpen, E., and Joswig, M., 1993, Automated reevaluation of local earthquake data by application of generic polarization patterns for P- and S-onset, *Computer and Geosciences,* **19, 223-23** 1.

Kracke, D., 1993, A three-component event detector based on waveform analysis, *Computers and Geosciences,* **19,** 117-122

Kvaerna, T. and Ringdal, F., 1992, Integrated array and there-component processing using a seismic microarray, *Bulletin of the Seismological Society of America, 82, 870-882.*

Lovell, J., 1989, Source parameters of a microearthquake swarm in Turkey, *Master oj Philosophy thesis, University of Edinburgh. Scotland, UK.*

Mykkeltveit, S. and Bungum, H., 1984, Processing of regional seismic events using data from small aperture arrays, *Bulletin of the Seismological Society of America,* **74**, 23 13-2333.

Pao, Y. H., 1988, *Adaptive pattern recognition and neural networks,* Addison-Wesley Publishing Company, Inc, New York.

Pisarenko, V. F., Kushnir, A. F., and Savin, I. V., 1987, Statistical adaptive algorithms foı estimation of onset moments of seismic phases, *Physics of Earth and Planetary Interiors, 47, 4-10.*

Roberts, R. G., Christoffersson, A., Cassidy, F., 1989, Real-time event detection. phase identification and source location estimation using single station three-componen seismic data, *Geophysical Journal International, 97, 47* l-480

Rumelhart, D. E., Hinton, G. E., and Williams, R. J., 1986, Learning internal representatioı by backpropagating errors, *Nature, 323,* 533-536.

Samson. J., 1977, Matrix and Stokes vector representations of detectors for polarized waveforms: theory, with some applications to teleseismic waves, *Geophysics J. R. astr Soc.* **, 51, 583-603.**

Takanami, T., and Kitagawa, G.. 1988, A new efficient procedure for the estimation of onse times of seismic waves, *Journal of Physics of Earth, 36,* 267-290.

Takanami, T., and Kitagawa. G., 1993. Multivariate time-series model to estimate the arriva times of S-waves: *Computers and Geosciences.* **19.** 295-301.