Discrimination of teleseismic events in Central Asia with a local network of short period stations

Timo Tiira and Matti Tarvainen
Institute of Seismology, University of Helsinki, Finland

Abstract
The difficult problem of distinguishing underground nuclear explosions from earthquakes at teleseismic distances was approached using short period seismic data from 6 stations in South and Central Finland. The events were nuclear tests mostly from the Semipalatinsk and Lop Nor test sites and earthquakes from adjacent areas. The magnitude range of the events was from 4.1 to 6.6. The features of the two classes of events were examined by computing spectral ratio, third moment of frequency (TMF) and complexity from *P*-wave signals. The spectral discrimination parameters were extracted from spectra computed in 5 different ways in order to obtain all possible information even from weak events. The standard FFT spectra were computed from raw data, after noise adaption and data adaption, from correlograms and using combinations of adaption and correlation methods. This was done to employ not only the spectral differences of the events but also the temporal variation of energy and lack of it as a function of frequency. The optimum frequency windows for spectral ratio and TMF were defined using stacked spectra of about 10 events from both classes. No single discriminant could classify all the events. Their performance varied significantly for different stations, but on average the spectral discriminants had slightly higher discrimination capability than complexity. The distributions of all discriminants were studied and a group separation function was formed using an optimum set of discriminants. Instead of discriminant values their relative positions in the corresponding distributions of nuclear tests and earthquakes were used as inputs to the function. A weight for each discriminant was derived from the amount of overlap in the distributions of earthquakes and nuclear tests. All 75 events in the data set were correctly classified with the method. The testing was performed with a jack-knife method to create an independent test data base.

Key words *teleseismic discrimination – nuclear explosions – short-period data – Central Asia*

1. Introduction

For nuclear test ban verification, reliable and efficient identification methods are essential. In recent years international seismic centers have located approximately 15 000 events annually. Discriminating possible nuclear tests from this number of events is a difficult task that has attracted the attention of many researchers during the last few decades. Several methods have been developed for seismic discrimination and a large amount of literature exists concerning the classification of seismic events (Evernden 1977; Evernden and Kohler 1979; Husebye and Mykkeltveit 1981; Pomeroy *et al.*, 1982; Blandford 1982; Evernden *et al.*, 1986; Kennett 1993; Wüster 1993). The main focus of classification today is on relatively weak events at regional distances. However all areas are not within regional distances from closest seismic station. There still exist vast areas with insufficient station coverage for reliable regional seismic discrimination. Also, it is not guaranteed that data from all stations are always accessible. Consequently, there exists a need for seismic discrimination of teleseismic events, which is the scope of our work. Such research has not been popular in recent years. Only few studies have been published this area in this decade (Taylor and Marshall 1991; Tsvang *et al.*, 1993).

The first tools for seismic discrimination were event location and depth. Shallow events close to known nuclear test sites could be considered as suspicious events requiring further analysis. Perhaps the most reliable discrimination parameter for teleseismic events has been the ratio of body wave magnitude and surface wave magnitude, but this has the disadvantage that teleseismic surface waves are usually recorded only from large events. Other potential parameters are complexity and spectral parameters like spectral ratio and third moment of frequency (TMF). These are the three basic discrimination parameters used in this study. Tsvang *et al.* (1993) used also complexity and spectral variants, but in a different manner. They used the logarithm of normalized power spectra taken from both initial *P* and its coda as spectral discriminant parameters. They also used peak spectral frequencies and ratio of *P*/*P*-coda spectral maximums, which resemble complexity. Our choice of parameters is complexity, spectral ratio and third moment of frequency as they are traditionally defined (Dahlman and Israelsson, 1977). The spectral parameters were computed from data processed in several different ways.

Instead of one array or single station or several stations and a voting method we are using a net of stations together to define a single discrimination factor, which is a weighted linear combination of a set of parameters from different stations. By using this method it is possible to compute the discrimination factor even with a limited number of stations. A shortage of data does not significantly limit the validity of the method. Reliability decreases with fewer stations, but it is possible to define the discrimination factor using only one station.

2. Event discrimination – physical basis

Earthquakes can be assumed to involve sudden movements along fault planes, beginning at a certain point under tension and continuing as a rupturing process towards the borders of the tension field. The distant displacement due to any kinematic model of an earthquake is expected to have a spectrum with constant values at lower frequencies and proportional to some negative power at higher frequencies (Haskell, 1967). In contrast, explosions are viewed as point sources with isotropic moment tensors.

Discrimination between explosions and earthquakes can be effective and reliable even at teleseismic distances, if the events considered are of large magnitude ($m_b >$ 4.5). The far-field spectrum is characterized by a low-frequency level, a corner frequency and the power of a high-frequency asymptote. Savage (1972) calculated the corner frequency f_c for *P*-waves for explosions assuming bilateral faulting with *P*-wave velocity α and final rupture length L

$$f_c = 1.2\frac{\alpha}{L} \tag{2.1}$$

Also, he found the corner frequency to be a geometric mean of two corner frequencies associated with the finite elastic radius and the rise time. Consequently, explosions having small L values have higher f_c, and consequently higher TMF and spectral ratio values than earthquakes. According to Dahlman and Israelson (1977) the corner frequency can be used as a discriminant reliably only on larger events. Considering the expected small dynamic radius of an explosion and its very short rise time, the radiation emitted by explosions should have more high frequency content than earthquakes (*e.g.* fig. 2). In the cases shown here, peak spectral amplitude of nuclear explosions occurs at higher frequencies than earthquakes. If explosion spectra in general showed this systematic difference, it would be easy to construct a pure «explosion detector», but the signal wavetrains from different parts of the world vary due to geological differences in the source area and along the wave path. Also, the structure near the stations may cause azimuth dependent variations in the frequency content of an incoming signal.

Taylor and Denny (1991) examined Shagan river explosions at regional distances and noticed that they radiated more high-frequency energy than earthquakes in the same area. In contrast, they had found previously that Nevada test site explosions had more low-frequency and less high-frequency content than earthquakes in the Western United States. These observations may be due to differences in the dynamic responses of the near-source geology. Near-field data may have contributions of secondary sources, such as spall or collapse in explosion chimneys, that are of less impact at teleseismic distances than at regional distances.

The solutions of simple time functions of pressure such as a step or decaying pulse are straightforward. Some attempts have been made to model the exact source time function for nuclear explosions with the aid of near-field empirical measurements. The displacement pulse caused by an explosion, can be expressed as a Heaviside step function with an exponential fall-off. The generalized explosion source time function $\Psi(t)$ proposed by Haskell (1967) has the form

$$\frac{\Psi(t)}{\Psi(\infty)} = 1-e^{-kt}\left[1+kt+\frac{(kt)^2}{2}+\frac{(kt)^4}{6}-B(kt)^2\right] \tag{2.2}$$

where t is time, $\Psi(\infty)$ is an asymptotic value for large t and k and B are constants depending on the yield and medium respectively.

3. The data set

Since the late 1970s digital data recordings from nuclear tests and other events have been archived from 6 short-period seismograph stations in Finland (KEF, SUF, KAF, PVF, PKK, NUR). Recently, the number of digitally recording stations has increased and the data archiving has expanded first to store all detected events and then to continuous data from several stations. We used in this study only the 6 stations mentioned earlier since they offer the largest data base of digital nuclear test recordings by the Finnish stations. The sampling rate of the stations is 20 Hz, which is sufficient when we are dealing with teleseismic data. More detailed description of the stations is done by Teikari and Suvilinna (1992).

We concentrated on the well known and frequently used nuclear test site near Semipalatinsk. Nuclear events have been selected from the SIPRI yearbook (Ferm 1992). Earthquakes in adjacent areas have been obtained from the ISC epicenter lists, PDE lists and the analysis results of the Institute of Seismology in Helsinki. In addition to the Semipalatinsk events, we added 2 nuclear tests conducted at the Chinese test site at Lop Nor, another from Central Siberia and several earthquakes close to Lop Nor. The events used in this study are shown in fig. 1. The number of events that could be used was limited by their size. We tried to include the smallest possible events. Weak events can hardly be separated from the noise at some stations and their signals disappear into the noise within 15 s, even after filtering. Such events were however used in order to find true limits of teleseismic discrimination with this method, instead of using an easy data set. Some of the large events had to be excluded because dynamic range of the stations was insufficient and the data were clipped. The m_b

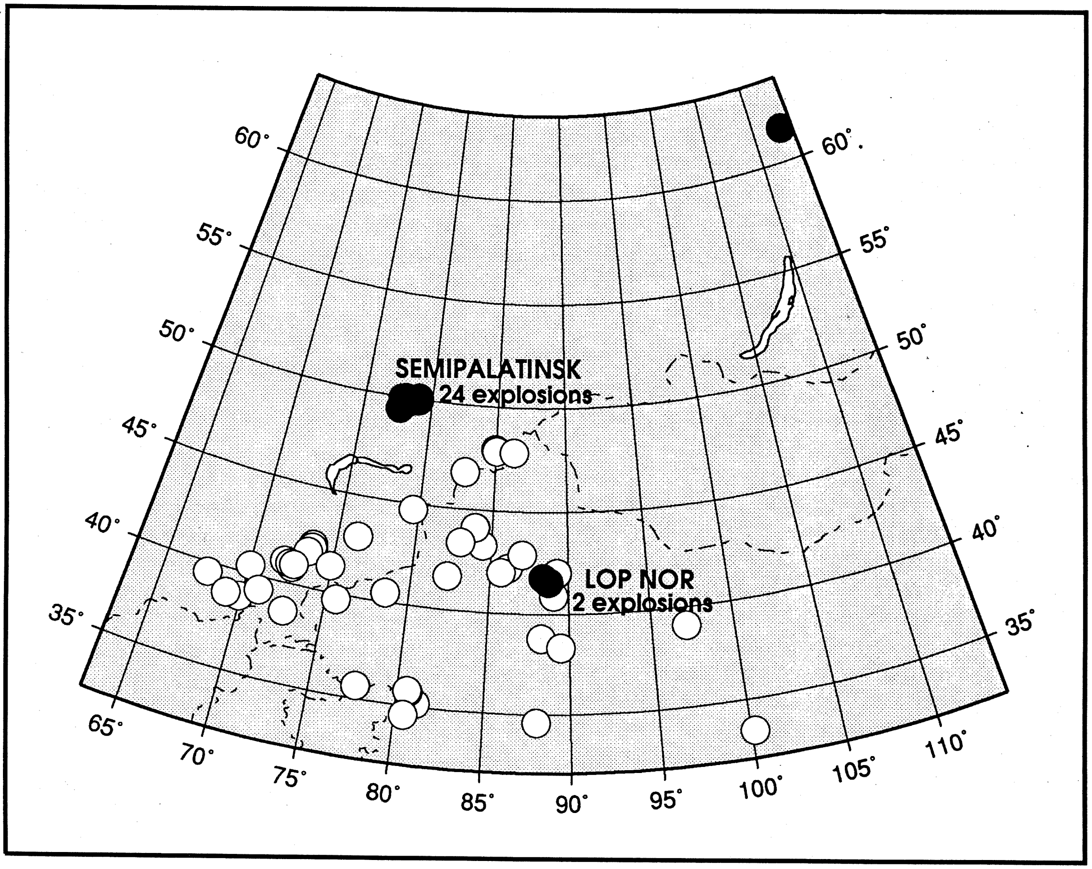

Fig. 1. Geographical distribution of the events used in the study. Closed circles denote nuclear tests and open circles earthquakes.

magnitudes of the events were between 4.1-6.6, and the median was 4.9. The distance to the Semipalatinsk test site from network center is approximately 31 and to Lop Nor 42°. The nuclear explosions and the earthquakes used are listed in table I.

4. Data analysis

The problem tackled is more demanding than discrimination at local and regional distances. With small events at teleseismic distances the initial P is often the only phase that is available for analysis. Attenuation has reduced the frequency band of the signal, especially at high frequencies, and the amount of spectral information is limited compared to that at shorter distances. Many of the events were rather small considering that we are using a net of single short-period stations and the signal to noise ratios (SNR) can be very low. In order to maximize information we processed the data in several different ways before computing the discriminants. The main reason for doing this was noise reduction but we also hoped to gain benefits in other aspects, as discussed later. The complexity values were computed from both raw and

Table I. Events used in the study. The numbering is same as in fig. 4. The EQ denotes earthquakes and NU confirmed nuclear explosions, respectively. The origin of focal parameters is denoted in the last column. The focal parameters are taken from ISC and PDE bulletins, preliminary analysis of the Institute of Seismology, University of Helsinki (HEL) and from the SIPRI yearbook for one event.

Date and time				EQ/NU	Latitude	Longitude	Depth	Magnitude	Reporting authority
1992179	June	27	13:21:20.9	EQ	35.14	81.13	33	5.0	PDE
1992096	April	5	07:47:47.6	EQ	35.70	80.66	18	5.5	PDE
1989265	Sept.	22	02:25:50.9	EQ	31.58	102.43	14	6.1	PDE
1992232	Aug.	19	02:04:37.4	EQ	42.14	73.57	27	6.6	PDE
1992258	Sept.	14	13:18:28.6	EQ	33.40	100.19	-	5.3	HEL
1990307	Nov.	3	17:25:13.8	EQ	40.88	89.07	22	5.1	PDE
1992136	May	15	08:08:02.9	EQ	41.02	72.43	50	5.7	PDE
1989263	Sept.	20	16:22:00.9	EQ	39.09	97.11	33	4.9	PDE
1993093	April	3	05:02:52.0	EQ	39.73	75.66	24	4.7	PDE
1992113	April	22	19:59:38.1	EQ	39.51	67.67	33	4.6	PDE
1992137	May	16	09:19:59.7	EQ	40.82	72.59	33	4.6	PDE
1990215	Aug.	3	09:15:06.1	EQ	47.96	84.96	33	6.0	PDE
1989297	Oct.	24	13:38:28.9	EQ	41.64	82.29	33	4.8	PDE
1990165	June	14	14:18:10.6	EQ	47.89	85.05	37	5.2	PDE
1991006	Jan.	6	15:46:38.6	EQ	38.77	88.31	22	4.9	PDE
1993054	Febr.	23	15:20:31.1	EQ	47.85	86.29	33	4.4	PDE
1992359	Dec.	24	17:08:48.1	EQ	35.76	80.61	33	4.8	PDE
1992328	Nov.	23	23:11:06.7	EQ	38.62	72.64	41	5.6	PDE
1992357	Dec.	22	16:42:37.2	EQ	34.57	88.05	33	5.1	PDE
1992233	Aug.	20	02:46:52.6	EQ	41.97	73.62	17	4.7	PDE
1992232	Aug.	19	04:06:20.9	EQ	42.00	73.55	19	5.0	PDE
1990021	Jan.	21	07:53:31.9	EQ	41.53	88.73	33	4.6	PDE
1993033	Febr.	2	16:05:14.1	EQ	42.22	86.13	33	5.7	PDE
1993098	April	8	03:49:33.2	EQ	35.65	77.65	42	5.0	PDE
1990297	Oct.	24	23:46:57.6	EQ	44.12	83.88	22	5.3	PDE
1992311	Nov.	6	07:21:57.8	EQ	41.05	72.51	40	5.1	PDE
1988085	March	25	02:07:55.8	EQ	44.71	79.60	33	4.5	PDE
1992233	Aug.	20	01:28:02.5	EQ	41.75	73.36	33	4.6	PDE
1992313	Nov.	8	20:50:13.0	EQ	38.78	69.86	64	5.3	PDE
1992332	Nov.	27	16:09:09.1	EQ	41.98	89.28	14	5.3	PDE
1992204	July	22	20:56:41.7	EQ	42.99	76.28	33	-	PDE
1988146	May	25	18:21:58.0	EQ	42.01	85.69	22	5.2	PDE
1993048	Febr.	17	02:00:25.8	EQ	38.32	89.48	15	5.1	PDE
1992338	Dec.	3	22:10:25.5	EQ	40.40	70.09	-	-	HEL
1993104	April	14	08:31:09.7	EQ	42.90	87.04	33	4.4	PDE
1992118	April	27	00:31:54.7	EQ	39.32	70.94	33	4.1	PDE
1993076	March	17	10:15:03.8	EQ	41.06	72.05	21	4.8	PDE
1993021	Jan.	21	21:05:46.3	EQ	38.80	69.06	63	4.5	PDE

Table I. (*continued*) Events used in the study. The numbering is same as in fig. 4. The EQ denotes earthquakes and NU confirmed nuclear explosions, respectively. The origin of focal parameters is denoted in the last column. The focal parameters are taken from ISC and PDE bulletins, preliminary analysis of the Institute of Seismology, University of Helsinki (HEL) and from the SIPRI yearbook for one event.

Date and time				EQ/NU	Latitude	Longitude	Depth	Magnitude	Reporting authority
1992232	Aug.	19	06:57:15.2	EQ	41.29	74.90	-	-	HEL
1992338	Dec.	3	22:08:48.9	EQ	43.28	84.41	10	4.6	PDE
1992142	May	21	07:07:00.9	EQ	41.00	72.41	33	4.6	PDE
1992142	May	21	14:28:19.7	EQ	40.96	72.67	20	5.0	PDE
1990297	Oct.	24	23:38:15.1	EQ	44.12	83.86	20	5.2	PDE
1991006	Jan.	6	13:07:34.9	EQ	43.37	82.95	13	4.7	PDE
1992177	June	25	09:41:49.8	EQ	40.47	78.55	33	4.7	PDE
1993116	April	26	16:24:07.2	EQ	34.54	80.51	33	4.5	PDE
1992302	Oct.	28	21:15:52.0	EQ	46.79	82.90	-	-	HEL
1988320	Nov.	15	16:56:46.2	EQ	42.02	89.29	33	5.0	PDE
1984328	Nov.	23	03:55:05.1	NU	49.90	78.11	0	4.7	ISC
1980361	Dec.	26	04:07:07.4	NU	49.98	78.01	0	4.5	ISC
1990228	Aug.	16	04:59:57.6	NU	41.56	88.77	0	6.2	PDE
1981088	March	29	04:03:50.1	NU	49.98	79.02	0	5.6	ISC
1982359	Dec.	25	04:23:05.5	NU	49.83	78.12	0	4.8	ISC
1980181	June	29	02:32:57.8	NU	49.91	78.86	0	5.7	ISC
1981356	Dec.	22	04:31:02.6	NU	49.84	78.21	0	5.1	ISC
1992269	Sept.	25	07:59:59.9	NU	41.76	88.39	0	5.0	PDE
1989277	Oct.	4	11:29:57.7	NU	49.75	78.01	0	4.6	PDE
1981147	May	27	03:58:12.2	NU	49.94	79.01	0	5.5	ISC
1989245	Sept.	2	04:16:57.3	NU	50.04	79.02	0	5.0	PDE
1989048	Febr.	17	04:01:06.9	NU	49.87	78.08	0	5.0	PDE
1988166	June	14	02:27:06.4	NU	50.04	79.00	0	5.0	PDE
1988037	Febr.	6	04:19:11.1	NU	49.80	78.06	0	4.8	PDE
1988317	Nov.	12	03:30:03.7	NU	50.08	78.99	0	5.3	PDE
1988292	Oct.	18	03:40:06.6	NU	49.87	78.08	0	4.9	PDE
1980116	April	25	03:56:57.4	NU	49.92	78.81	0	5.5	ISC
1980213	July	31	03:32:58.0	NU	49.81	78.14	0	5.3	ISC
1988113	April	22	09:30:06.9	NU	49.82	78.12	0	4.9	PDE
1980095	April	4	05:32:57.4	NU	49.97	77.78	0	5.1	SIPRI
1981198	July	17	02:37:15.7	NU	49.79	78.17	0	5.2	ISC
1987126	May	6	04:02:05.8	NU	49.80	78.11	0	5.6	ISC
1980101	April	10	04:06:58.0	NU	49.82	78.08	0	5.0	ISC
1988328	Nov.	23	03:57:06.7	NU	49.82	78.07	0	5.3	PDE
1981226	Aug.	14	02:27:12.9	NU	49.75	78.07	0	5.0	ISC
1982283	Oct.	10	04:59:56.9	NU	61.53	112.86	0	5.3	ISC
1979047	Febr.	16	04:03:58.2	NU	49.97	77.74	0	5.4	ISC

filtered data. We used 4th order Butterworth band-pass filter with cut off frequencies at 1 and 7 Hz. In addition to the conventional way, the spectral discriminants: spectral ratio and TMF were computed from spectra modified with noise and data adaption, from spectra of correlograms and from spectra of a combination of noise adaption and correlation method. We considered it necessary to try to get all the possible information from the signal since some of the events were very weak. The results provided information on the capabilities of various methods and a cross-check of their performance with data from different stations. The total number of discriminants for each station was 12.

Noise and data adaption were used to improve the spectral contents of the recorded signals. Noise adaption was done by subtracting the noise spectrum calculated just prior to the P-onset from the spectrum of the event. The time window for computing both spectra was 16 s. Data-adaption was carried out by using a set of overlapping 4 s long data windows starting at the P-onset. The last window started 16 s after the onset. The step between windows was just one sample. The data-adaption was done by computing the differences of the spectra of consecutive data windows and averaging the differences. This can be expressed as

$$G(f) = \frac{1}{L} \sum_{l=0}^{L-1} G_l(f) - G_{l+1}(f) \tag{4.1}$$

where L is the number of overlapping spectral windows, G_l denotes the lth signal spectrum, and G_{l+1} is the adapting spectrum. Note the subtraction order: the spectrum of the new data window is subtracted from the previous spectrum. The data-adaption enhances frequencies which had more variation and a faster decay rate in the signal compared to those frequencies which had steady energy levels. It brought into the discrimination process the fact that the coda of a nuclear explosion decays faster than that of a similar sized earthquake from the same area. Both adaption methods suppress the effects of steady noise from the discriminants.

The third method used to compute modified spectra for forming spectral discriminants was the use of correlogram functions. Sixteen seconds of data after the P-wave onset were divided into 4 equal length non-overlapping windows. The autocorrelograms of all windows were computed separately. The mutual crosscorrelograms of the autocorrelograms were determined and the spectrum was computed from their arithmetic mean using Fourier transform. The autocorrelograms were formed according to standard the methods applied for digital recordings. The autocorrelation function $R^A(\tau)$ of a sequence y is expressed by

$$R^A(\tau) = \frac{1}{T} \int_0^T y(t)y(t + \tau)\, dt \tag{4.2}$$

where T is length of the section y, t denotes time and τ time lag. The mutual cross correlograms $R^C(\tau)$ in turn were formed from these autocorrelograms as

$$R^C(\tau) = \frac{1}{T} \int_0^T R^A_m(t) R^A_n(t + \tau)\, dt \tag{4.3}$$

and their average $\Re(\tau)$ is expressed by

$$\Re(\tau) = \frac{\sum_{m=1}^{M-1} \left[\sum_{n=m+1}^{M} R^C(\tau) \right]}{\sum_{h=1}^{M} M - h} \tag{4.4}$$

where the subscripts m and n denote the corresponding autocorrelograms and M is the number of autocorrelograms. This method enhanced those frequencies, that were dominant throughout the signal. We also used the correlation method in conjunction with the noise adaption. In this case the spectra were scaled so that the high frequency ends were at the same level, because correlograms have a normalized amplitude.

The frequency windows for spectral ratio and TMF were defined separately for each station. From 10 to 7 nuclear tests and earthquakes were selected at each station. Their spectra were computed in several ways as explained above. The spectra were normalized and summed together to get a general view of frequency content of earthquakes and nuclear tests obtained with each type of spectral computation. The summed spectra were used to select the frequency windows for spectral discriminants. These spectra for station KEF are shown in fig. 2a-e. The most profitable windows were selected automatically by testing all possible frequency ranges. The results varied considerably from station to station. The number of events was limited to 10 or less to prevent the discriminants from being too dependent on the data set. The averages of the upper and lower frequency limits of TMF and spectral ratio were computed for each station to determine their overall performance at different frequencies. The station KAF seemed to have the strongest discrimination capability at low frequencies, and the highest upper limits for frequency windows were usually at stations PVF and NUR.

5. The discrimination method

The discrimination parameters form a set A of K input vectors a_i each with 72 possible discriminants b_j for further analysis

$$A = \{a_1,...,a_K\}; \quad a_i = (b_1,...,b_k) \tag{5.1}$$

The discriminants are listed and numbered in table II. Several statistical methods exist for classification and discrimination analysis. Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) are well-known methods and also Tsvang *et al.* (1993) used them successfully in seismic discrimination. FACT (Fast Algorithm for Classification Trees) by Loh and Vanichsetakul (1988) uses LDA and other standard statistical techniques to build a classification tree of the data set. Yet another statistical method worth considering is MARS (Multivariate Adaptive Regression Splines) by Friedman (1991). Also artificial neural networks have been used successfully for seismic discrimination *e.g.* Dowla *et al.* (1990) and Dysart and Pulli (1990). The type of neural network which was used in these and some other works in seismology is the multilayer perceptron with a back-propagation learning method. We chose to use a weighted linear combination of discrimination parameters, which in fact is very close to a one layer perceptron with only one neuron in the output layer.

The learning session of neural networks was replaced by computing a set W of weights w_j. using the distributions of earthquakes and nuclear tests for each discriminant

$$W = \{w_1,...,w_k\} \tag{5.2}$$

The set A was divided into 2 subsets A^E for earthquakes and A^N for nuclear tests. The weight of each discrimination parameter was computed using percential overlap O_j in distributions of earthquakes A^E and nuclear tests A^N for each discriminant b_j

$$\begin{aligned} O_j \geq 0, \quad & w_j = 1 - \frac{O_j}{100} \\ O_j < 0, \quad & w_j = 0 \end{aligned} \tag{5.3}$$

When the distributions had an opposite orientation compared to that physically expected, the value of O_j was given a negative sign. The weights w_j correspond to the discrimination capability of each individual discriminant b_j. The input vectors y_i^N of the discrimination function $DF^N(y^N)$ are transformed from the original input vectors a_i.

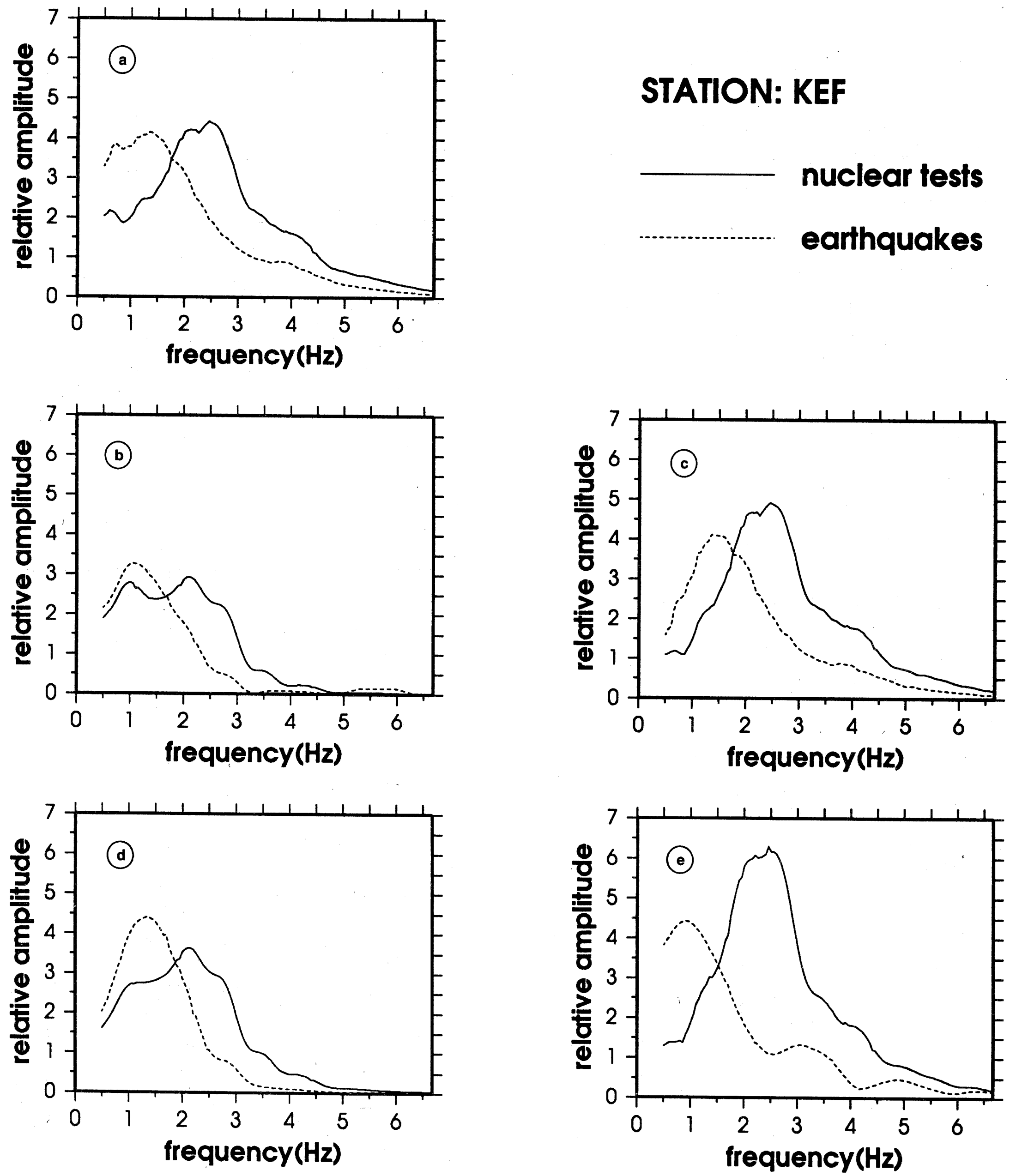

Fig. 2a-e. Summed spectra from station KEF for each type of spectra computation: a) spectra of raw data; b) spectra of correlograms; c) spectra of correlograms after noise cancellation; d) spectra of raw data after noise cancellation (noise adaption); e) spectra after data adaption. The spectra were computed from 8 earthquakes and 9 nuclear tests with nearly equal magnitude distributions.

$$y_i^N = T(a_i); \quad y_i^N = (y_1^N, ..., y_K^N);$$
$$0 \le y_{ij}^N \le 1 \qquad (5.4)$$

where y_{ij}^N is the relative position of the corresponding discriminant value b_j in the distribution of nuclear tests. A value of 0 denotes that b_j is outside the distribution pointing to earthquakes and $y_{ij}^N = 1$, that b_j lies on the other side of the distribution showing an extremely high probability of the event being a nuclear test. The discrimination was done using a linear function

$$DF^N(y^N) = \frac{\sum_{j=0}^{k} w_j x_j^N}{\sum_{c=0}^{C} k_c^C} \qquad (5.5)$$

C denotes the number of stations and k_c^C the number of discriminants extracted from the station. By using x_j^N instead of the actual value of the discriminant b_j, we created a group classification function $DF^N(y^N)$ for nuclear explosions. The function value relates to the probability with which the test event belongs to the set A^N. The maximum value 1 indicates an event which exceeds all events in the training data base with every discriminant in giving discrimination values pointing to a nuclear test. In general a classification problem can be solved using either categorical or continuous valued parameters, or a combination of both. In this method each parameter is partly categorical and partly continuous valued. The value of an individual parameter is either 0 indicating an earthquake, or a continuous valued number in the range 0.0-1.0 suggesting a nuclear test.

A similar group classification function $DF^E(y^E)$ was formed for earthquakes and the final classification was done using the group separation function $DF(y^N, y^E)$

$$DF(y^N, y^E) = DF^N(y^N) - DF^E(y^E) \qquad (5.6)$$

The $DF(y^N, y^E)$ defines a hyperplane which separates the earthquake and nuclear test populations. Neither of the group classification functions could classify all the events but the group separation function $DF(y^N, y^E)$ separated the two groups totally as will be shown in the next section.

6. Results and discussion

When the discrimination capabilities of individual parameters were studied it was found that none of them could separate the two groups of events completely. In fig. 3a are shown complexity values as a function of TMF for recordings at the seismograph station SUF, and in fig. 3b complexity and spectral ratio from station KEF. The discriminants show some capability in discrimination but the events are not clearly divided into two groups. This situation is typical. No single discriminant could separate all the events of different type. The results of individual discriminants are comparable to those Taylor and Marshall (1991) presented for spectral ratios. When all the discriminants from all stations were examined a large variation in discrimination capability was found. None of the discriminants performed well at all stations. The weight values for each discriminant for all stations and their means and deviations are found in table II. The spectral discriminants were slightly more capable than complexity. Between TMF and spectral ratio there was significance difference. Differences between stations were more significant. The Central Finland stations KEF, SUF and KAF were clearly superior compared to PVF, PKK and NUR in Southern Finland. This can be explained by the higher noise levels at these stations due to their location closer to the Baltic Sea, and to environmental noise (Tarvainen, 1985).

The validation of any discrimination work should be done with a separate test data set. A system with a large number of parameters could memorize the training data set and testing with the same data set

Table II. List of discriminants and their numbers as used in this study. Discriminants weights are defined in eq. 5.2. The mean value and standard deviation are given for each discriminant and station.

	KEF	SUF	KAF	PVF	PKK	NUR	Mean	SD
1	0.71227	0.71316	0.77152	0.52725	0.70683	0.61103	0.67368	0.08842
2	0.61862	0.63492	0.84062	0.38905	0.41896	0.58185	0.58067	0.16421
3	0.79200	0.73958	0.64289	0.52681	0.58024	0.57912	0.64344	0.10301
4	0.65897	0.63339	0.77777	0.40609	0.45767	0.55522	0.58152	0.13712
5	0.64558	0.64654	0.59226	0.51129	0.43349	0.41593	0.54085	0.10275
6	0.57450	0.49402	0.55297	0.40023	0.33051	0.55468	0.48449	0.09861
7	0.74974	0.80091	0.84520	0.41889	0.47252	0.57174	0.64317	0.17977
8	0.79049	0.81209	0.87187	0.49509	0.65430	0.61949	0.70722	0.14180
9	0.53028	0.56304	0.71211	0.33352	0.39434	0.46173	0.49917	0.13433
10	0.65679	0.78060	0.81346	0.51739	0.61832	0.60046	0.66450	0.11278
11	0.56469	0.65685	0.74026	0.35209	0.37787	0.56079	0.54209	0.15256
12	0.65008	0.63297	0.71545	0.37316	0.42144	0.42008	0.53553	0.14676
Mean	0.64194	0.68831	0.71804	0.42871	0.48375	0.51641		
SD	0.08696	0.08115	0.12691	0.06740	0.13091	0.06803		

1: TMF; 2: TMF from correlograms; 3: TMF from noise adaptive spectra; 4: TMF from correlogram with noise adaption; 5: TMF from data adaptive spectra; 6: complexity; 7: complexity from filtered data; 8: spectral ratio; 9: spectral ratio from correlograms; 10: spectral ratio from noise adaptive spectra; 11: spectral ratio correlogram with noise adaption; 12: spectral ratio data adaptive spectra.

would not correctly represent the capabilities of the system. The data set was not large enough to divide into two parts so that one part could be used for training and the other for testing. A popular way to create an independent test data set with a small data base is the so-called jack-knife or leave-K-out method. One of the events is left out for testing and the discrimination functions are formed using all the other events. After the distributions and discrimination functions were formed the system was tested with the excluded event. This was done for all events in turn and the outputs were combined to get a set of test results.

With linear discrimination functions the groups have to be separable with one hyperplane. Finding the optimum subset for parameters is essential when none of the individual discriminants give good classification results. Using as few discriminants as possible is usually the most rewarding approach in selecting the input parameters. However, when there is significant variation in the performance of the discriminants with different events, a rather large set of parameters gives the smallest number of misclassified events. Employing a small set of the best parameters would result in the largest variation in results between the two groups but the variation in the groups would also be large due to events which have non-uniform input vectors. Applying too few discriminants could provide a good overall separation of the groups, but also some misclassified events due to separate outlying discriminant values.

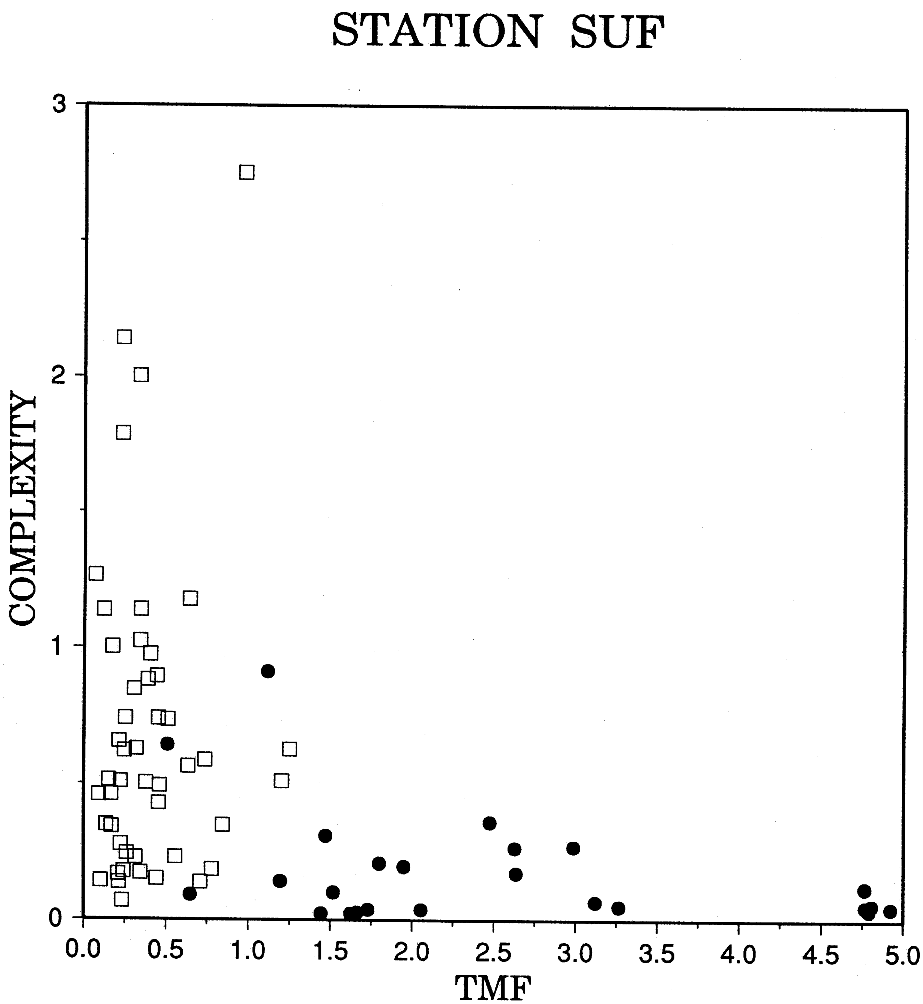

Fig. 3a. The distribution of complexity values from filtered data as a function of third moment of frequency (discriminants 7 and 1) from station SUF. Black squares denote nuclear tests and open squares earthquakes.

The optimum number of parameters was selected empirically. The number of parameters was limited using the weight value w_j of each discriminant. Several threshold values were tested and a minimum in separation between the most difficult events was gained with $w_j = 0.46$, when number of accepted parameters was 56. The threshold value limits the maximum amount of overlap in the distributions of the two classes to 27%. Input parameters with more overlap were rejected. All the discarded parameters were discriminants from stations PVF, PKK and NUR from Southern Finland (PVF: 2, 4, 6, 7, 9, 11, 12; PKK: 2, 4, 5, 6, 9, 11, 12; NUR: 5, 12). The results computed with the leave-K-out method are shown in fig. 4. All events are classified, al-

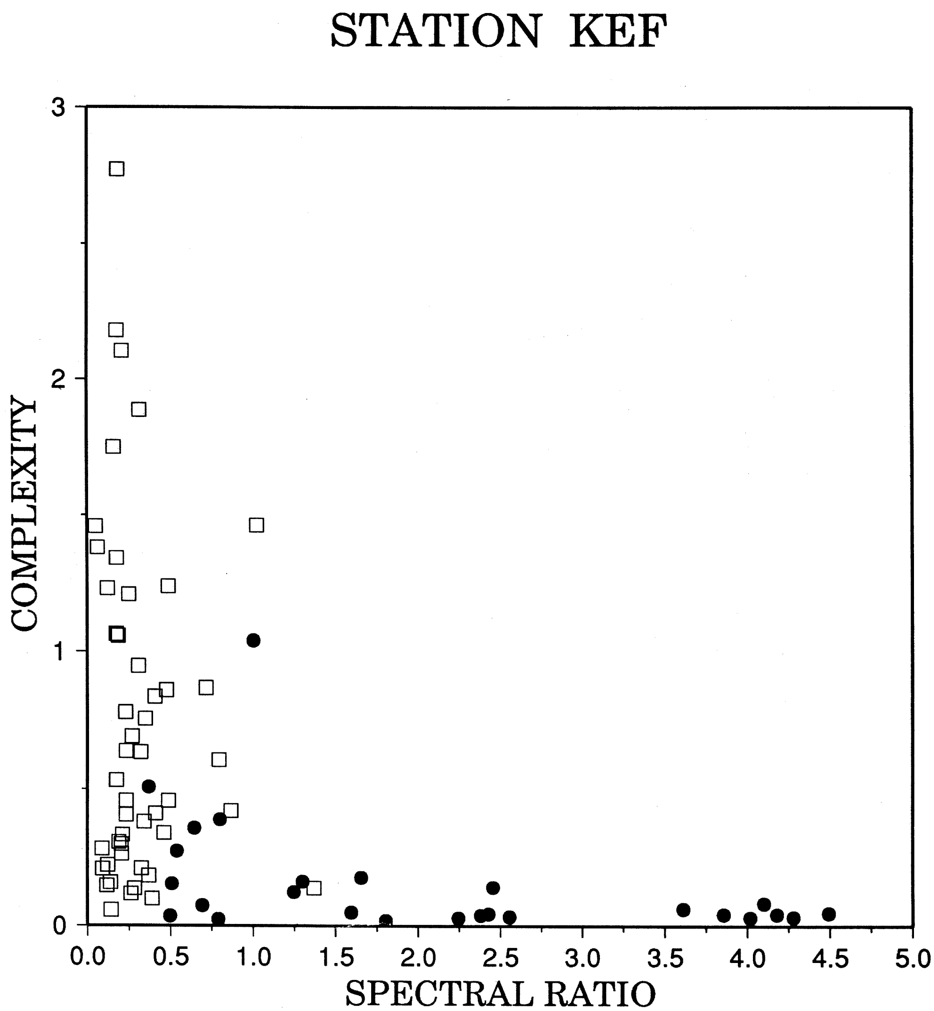

Fig. 3b. Complexity and spectral ratio (discriminants 6 and 8) from station KEF. Black squares denote nuclear tests and open squares earthquakes.

though 3 nuclear tests lie rather close to the earthquakes. The 2 nuclear tests closest to misclassification had the smallest SNRs of the whole data set. The third one was one of the 2 events from Lop Nor test site and also one of the two explosions with $m_b >$ 6.0. The other event from Lop Nor and the so-called peaceful nuclear explosion from Siberia were clearly discriminated, showing that the method is valid for a larger region than just the Semipalatinsk test site where most of the explosions were located. In fig. 5 the results are plotted as a function of magnitude. As could be expected, the discrimination capability declines at lower magnitudes. This phenomenon was stronger for explosions because their coda decays faster than that of earthquakes, and so the usable window was only 16 s long including the coda. Tests made with pure noise samples produced discriminant values characteristic of earthquakes for all discriminants. The few largest events are less well classified because the number of large events was small in the data set and the discrimination function was formed mainly by utilizing features of the smaller events. Large magnitude seemed to have a more significant negative influence on the classification capability than unusual geographical location.

Using the relative position of discriminants in the distribution instead of the value of the discriminant makes the method less dependent on the shape of the distribution. Asymmetric distributions with long tails as were present in this study are not harmful if they truly describe the behavior of the discriminant. Also, prescaling of the input parameters was not necessary. Of the three basic discriminants, TMF had the most symmetric distribution and complexity always had a heavy tail at large values. The distributions of spectral ratio had varying shapes for different computational methods.

Both nuclear tests and earthquakes must be selected from the same area in order to obtain reliable results. If events of only one class were added from other areas, the signal differences would arise not only from nature of the source but also from the ef-

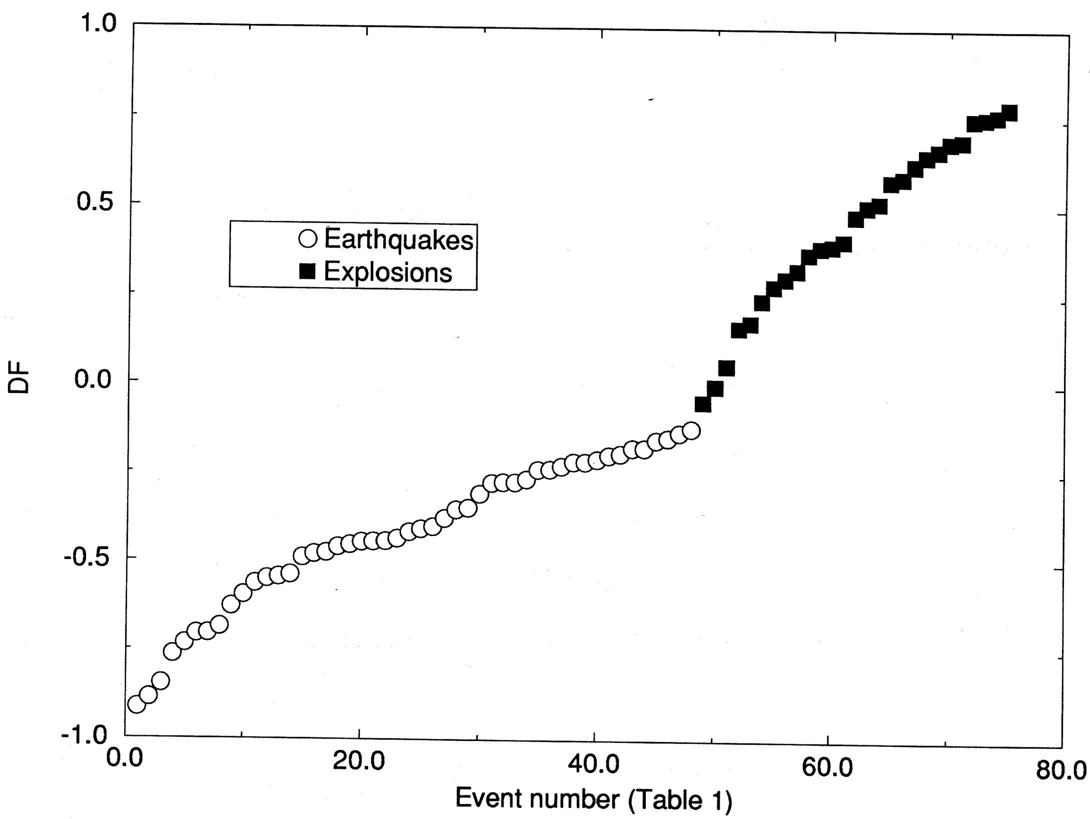

Fig. 4. Classification results computed using the jack-knife method. The events appear in the same order as in table I. Black squares denote nuclear tests and open squares earthquakes.

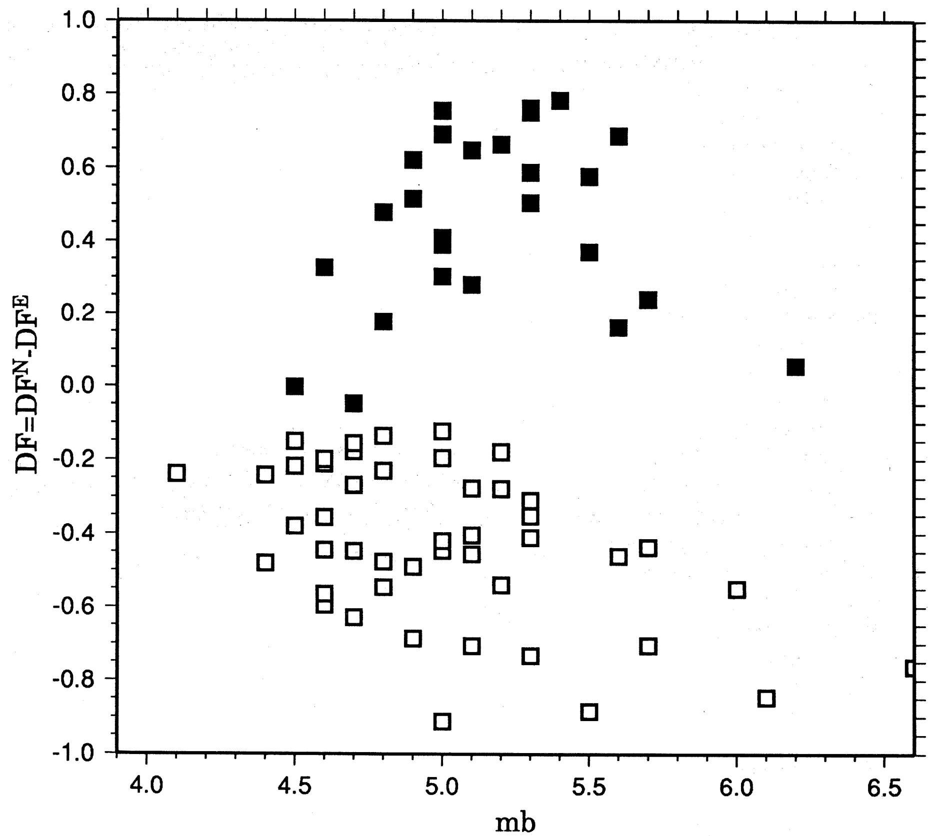

Fig. 5. Discrimination as a function of magnitude. Black squares denote nuclear tests, and open squares earthquakes.

fects of different source area and propagation path. Consequently, the discrimination function is applicable only to the area covered by the original data set.

7. Conclusions

The discrimination method presented in this study was capable of correctly classifying nuclear tests from the Semipalatinsk and Lop Nor test sites, and earthquakes from adjacent areas. One peaceful nuclear explosion from Central Asia was also correctly classified.

The total number of the events used was 75 (27 nuclear explosions and 48 earthquakes). The three basic discriminants that were employed were complexity, spectral ratio and third moment of frequency. The complexity was computed from both filtered and unfiltered data and spectral discriminants were computed from spectra computed in 5 different ways producing 12 possible discriminants for each of the 6 stations from Central and Southern Finland. The input vectors to the discrimination function consisted of the relative positions of discriminant values in the earthquake and nuclear test distributions of corresponding discriminants in the training data base. Using relative positions instead of the original values provided better independence on the distribution shapes.

A linear discrimination function was computed using a combination of discriminants from all stations. The results are similar to those Tsvang *et al.* (1993) obtained with events from the Semipalatinsk area using NORESS data even though we had events from a wider area and the single station data we used provides lower SNR ratios than beams from an array. The reason for this is probably the combined use of discriminants from several stations. The geological structure near the receivers causes different features to be recorded with varying quality from station to station. Consequently, the relative classification capability of the discriminants varied at different stations. Part of the variation is also due to different noise conditions at the stations. In this study we employed the best information provided by each station. Using selected discriminants from several stations together yields obvious benefits in seismic discrimination. The number of discriminants in the selected optimal set was relatively large. We predict that with better single discriminant classification capability, or with a more powerful discrimination method that can utilize hidden relations between discriminants, a smaller set would be sufficient for optimal results.

Part of the success in classification can be attributed to the large set of discriminants from which the optimal subset was selected. Even small differences in the behavior of nearly similar discriminants can give much needed extra information for the difficult process of separating events at teleseismic distances. The use of several types of discriminants may have even more significance when a search is made for more complex decision regions than used in this study.

Acknowledgements

The authors wish to express their thanks to Prof. Eystein S. Husebye, University of Bergen, Norway for his invaluable advice during this work. The comments of Dr. Luosto, Director of the Institute of Seismology, University of Helsinki, concerning this work are also highly appreciated. This work was financially supported by the Finnish Cultural Foundation (Maili Aution rahasto).

REFERENCES

BLANDFORD, R.R. (1982): Seismic event discrimination, *Bull. Seismol. Soc. Am.*, **72**, S69-S88.

DAHLMAN, O. and H. ISRAELSSON (1977): *Monitoring Underground Nuclear Explosions* (Elsevier, Amsterdam) pp. 440.

DOWLA, F.U., S.R. TAYLOR and R.W. ANDERSON (1990): Seismic discrimination with artificial neural networks: preliminary results with regional spectral data, *Bull. Seismol. Soc. Am.*, **80**, 1346-1373.

DYSART, P.S. and J.J. PULLI (1990): Regional seismic event classification at the NORESS array: seismological measurements and the use of trained neural networks, *Bull. Seismol. Soc. Am.*, **80**, 1910-1933.

EVERNDEN, J.F. (1977): Spectral characteristics of the *P* wave codas of Eurasian earthquakes and explosions, *Bull. Seismol. Soc. Am.*, **67**, 1153-1171.

EVERNDEN, J.F. and W.M. KOHLER (1979): Further study of spectral composition of *P* coda of earthquakes and explosions, *Bull. Seismol. Soc. Am..*, **69**, 483-511.

EVERNDEN, J.F., C.B. ARCHAMBEAU and E. CRANSWICK (1986). An evaluation of seismic decoupling and underground nuclear test monitoring using high-frequency data, *Rev. Geophys.*, **24**, 143-215.

FERM, K. (1992): Nuclear explosions 1945-1991, in *SIPRI Yearbook 1992: World Armaments and Disarmament* (Oxford University Press, New York), 107-119.

FRIEDMAN, J.H. (1991): Multivariate adaptive regression spline, *Ann. Statist.*, **19**, 1-141.

HASKELL, N.A. (1967): Analytic approximation for the elastic radiation from a contained underground explosion, *J. Geophys. Res.*, **72**, 2583-2587.

HUSEBYE, E.S. and S. MYKKELTVEIT (Editors) (1981): *Identification of Seismic Sources-Earthquakes and Underground Explosions*, NATO ASI series (D. Reidel Publishing Co., Dordrecht, The Netherlands) pp. 876.

KENNETT, B.L.N. (1993): The distance dependence of regional phase discriminants, *Bull. Seismol. Soc. Am.*, **83**, 1155-1166.

LOH, W. and N. VANICHSETAKUL (1988): Tree-structured classification via generalized discriminant analysis, *J. Am. Statist. Assoc.*, **83**, 715-728.

POMEROY, P.W., W.J. BEST and T. V. MCEVILLY (1982): Test ban treaty verification with regional data – a review, *Bull. Seismol. Soc. Am.*, **72**, S89-S129.

SAVAGE, J.C. (1972): Relation of corner frequency to fault dimensions, *J. Geophys. Res.*, **77**, 577-592.

TARVAINEN, M. (1985): *Results of the Noise Studies in Finland 1981-1984*, Institute of Seismology, University of Helsinki, report **T-30**.

TAYLOR, S.R. and M.D. DENNY (1991): An analysis of spectral differences between Nevada test site and Shagan river nuclear explosions, *J. Geophys. Res.*, **96**, 6237-6245.

TAYLOR, S.R. and P.D. MARSHALL (1991): Spectral discrimination between Soviet explosions and earthquakes using short-period array data, *Geophys. J. Int.*, **106**, 265-273.

TEIKARI, P. and I. SUVILINNA (1992): *Finnish Seismic Stations 1991*, Institute of Seismology, University of Helsinki, report **T-53**.

TSVANG, S.L., V.I. PINSKY and E.S. HUSEBYE (1993): Enhanced seismic source discrimination using NORESS recordings from Eurasian events, *J. Geophys. Int.*, **112**, 1-14.

WÜSTER, J. (1993): Discrimination of chemical explosions and earthquakes in Central Europe – a case study, *Bull. Seismol. Soc. Am.*, **83**, 1184-1212.