*Royal Norwegian Council
for Scientific and Industrial Research
(NTNF)*

NORSAR

NORSAR Scientific Report No. 1-91/92

Semiannual Technical Summary

1 April — 30 September 1991

Kjeller, November 1991

7 Summary of Technical Reports / Papers Published

7.1 Statistical optimization of seismic holography algorithms for array data processing

In this paper we describe the application of statistical parameter estimation theory to the problem of locating weak seismic radiation sources in the lithosphere. The radiation may have two origins. The first is scattering generated by an earthquake or other type of source. These scattered seismic waves are to be extracted from the background primary waves. The other origin is a weak seismic emission in the medium. The problem of mapping of a weak seismic radiation source's spatial distribution on the basis of seismic array data processing is called seismic holography.

Theory

The seismic field $\xi(t,\vec{\rho})$ measured on the surface may be given as the sum of the "signal" field, produced by the source, located in the medium at point $\vec{r}$, and the residual field $\eta(t,\vec{\rho})$:

$$\xi(t,\vec{\rho}) = \int_{-\infty}^{+\infty} \mu s(t-\tau) G(\tau,\vec{r},\vec{\rho}) d\tau + \eta(t,\vec{\rho}) \tag{1}$$

where $\vec{\rho}$ is the point on the surface, $\mu s(t)$ is a scalar waveform generated by the source at point $\vec{r}$, $G(t,\vec{r},\vec{\rho})$ is the Green function of the medium, and μ is a scaling factor. According to eq. (1), the multidimensional time series $\vec{x}_t = (\xi(t/f_s,\vec{\rho}_1),\ldots,\xi(t/f_s,\vec{\rho}_m))$ ($\vec{x}_t$ is a column vector) recorded by one-component array sensors has the following structure

$$\vec{x}_t = \sum_{\tau=0}^{\infty} \mu s_{t-\tau} \vec{h}_\tau(\vec{r}) + \vec{\eta}_t = \mu s_t * \vec{h}_t(\vec{r}) + \vec{\eta}_t, \qquad (t \in \overline{1,N}) \tag{2}$$

where $s_t = s(t/f_s)$, $\vec{h}_\tau(\vec{r}) = G(\tau,\vec{r},\vec{\rho}_i), i \in \overline{1,m}$ is a column vector of the medium's impulse response for the seismic wave propagating from the source at point $\vec{r}$ to the sensors at points $\vec{\rho}_i$; $\vec{\eta}_t$ is the "noise" time series, generated by other sources located away from point $\vec{r}$; f_s is the sampling rate and * denotes convolution. In the case of scattering, η_t contains strong components generated by the primary wave, and may therefore influence the estimation of scattered wave power.

In most cases there is no a priori information about the source waveform and the "noise" field features. So from a statistical point of view it is reasonable to assume the scalar process s_t and the vector process $\vec{\eta}_t$ to be realizations of Gaussian stationary time series with zero mean and power spectral densities $\phi(\lambda)$ and $\mathbf{F}(\lambda)$, respectively. Under the additional assumption that s_t and $\vec{\eta}_t$ are statistically independent, the matrix spectral density of the time series $\vec{x}_t$ becomes:

$$\mathbf{F}_x(\lambda,\theta) = \theta\phi(\lambda)\vec{h}_r(\lambda)\vec{h}_r^*(\lambda) + \mathbf{F}_\eta(\lambda), \quad \lambda \in [0, 2\pi]$$
$$\theta = \mu^2 \tag{3}$$
$$\int_0^{2\pi} \phi(\lambda)\,d\lambda = 1$$

where

$$\vec{h}_r(\lambda) = \sum_{t=0}^{\infty} \vec{h}_t(\vec{r})e^{i\lambda t}$$

is the vector response of the propagation paths from point r to the array sensors; θ is the power of the seismic source at point $\vec{r}$ and $\lambda = 2\pi f/f_s$ is a normalized frequency.

For simple reference models, $\vec{h}_r(\lambda)$ may be computed by solving the "direct" seismic problem. With reference model of the medium, we mean the model which determines the main features of the seismic wave propagation without accounting for the inhomogeneities which are to be detected by the analysis of the scattered waves. This means that in equation (2), we ignore the effect of secondary scattering of the seismic waves caused by the inhomogeneities.

Equations (2) and (3) allow consideration of the problem of locating seismic scatterers as the statistical problem of source power θ estimation, which is successively solved for each point of the "scanned" medium area (Troitsky *et al*, 1980). Under the assumption that the power spectra $\phi(\lambda)$ and $\mathbf{F}(\lambda)$ are known, statistically optimal algorithms may be proposed for this estimation problem. But in practice these assumptions are never fulfilled and these algorithms must be modified to make them efficient also when we do not have full information about statistical features of the observations.

Statistically optimal algorithms for estimation of the seismic source power θ may be developed using the maximum likelihood method (in its asymptotical modification). Using the Bartlett formula (Bartlett, 1951) for inversion of the matrix in equation (3), we obtain:

$$\mathbf{F}_x^{-1} = \mathbf{F}_\eta^{-1} - \mathbf{F}_\eta^{-1}\vec{h}\vec{h}^*\mathbf{F}_\eta^{-1}\frac{\theta\phi}{1+\vec{h}^*\mathbf{F}^{-1}\vec{h}} \tag{4}$$

For the maximum likelihood estimate $\hat{\theta}_N$ of the power of a seismic source, located at point $\vec{r}$, Kushnir (1989) shows that the following equation holds:

$$\sum_{j=1}^{N} \frac{\phi_j V_j^2}{1+\hat{\theta}_N\phi_j V_j}\left[\left(\frac{|Z_j|^2}{V_j^2} - \frac{1}{V_j}\right) - \hat{\theta}_N\phi_j\right] = 0 \tag{5}$$

where $\phi_j = \phi(\lambda_j)$, $Z_j = \vec{h}_r^{\dagger}(\lambda_j)\mathbf{F}_{\eta}^{-1}(\lambda_j)\vec{X}_j$, $V_j = V(\lambda_j)$, $V(\lambda) = \vec{h}_r^{\dagger}(\lambda)\mathbf{F}_{\eta}^{-1}(\lambda)\vec{h}_r(\lambda)$, $\vec{X}_j = (1/(\sqrt{N}))\sum_{j=1}^{N}\vec{x}_t e^{i\lambda_j t}$. $\lambda_j = 2\pi j/N$, N is the number of observations, and $\dagger$ denotes the Hermitian conjugation.

Equation (5) must in general be solved by numerical methods, but in some practically important cases the solution may be written in "closed" form. Namely, if the source power is much less than the "noise" power, then from (5) the following approximate solution may be obtained (Kushnir, 1989):

$$\tilde{\theta}_N = \frac{\sum_{j=1}^{N}\phi_j V_j\left(\frac{|Z_j|^2}{V_j^2} - \frac{1}{V_j}\right)}{\sum_{j=1}^{N}\phi_j^2 V_j^2} \tag{6}$$

This is a consistent and asymptotically normal estimate, but it has larger variance than the exact solution of equation (5).

Estimates based on solving equation (5) do not provide answers to the problem of mapping of scattered seismic radiation sources formulated in the introduction, because such estimates assume that the power spectrum $\mathbf{F}_{\eta}(\lambda)$ of the "noise" and source waveform power spectrum $\phi(\lambda)$ are known. It can be shown that the information about $\mathbf{F}_{\eta}(\lambda)$ is very important because $\mathbf{F}_{\eta}(\lambda)$ determines the statistical properties of all power θ estimators. When $\mathbf{F}_{\eta}(\lambda)$ is unknown, it is reasonable to use the adaptive approach. In that case, $\mathbf{F}_{\eta}(\lambda)$ in equations (5) and (6) is substituted by $\tilde{\mathbf{F}}_{\eta}(\lambda)$ derived from the observations themselves. But in the first of the problems outlined in the introduction, it is impossible to consistently estimate $\mathbf{F}_{\eta}(\lambda)$ from observations. Often the same is true for the second problem. Substituting for $\mathbf{F}_{\eta}(\lambda)$ in the expressions (5) and (6), any consistent estimate $\mathbf{F}_x(\lambda)$ of the matrix power spectrum of the observations $\vec{x}_t$ yields, values $\hat{\theta}_N$ and $\tilde{\theta}_N$ approaching zero in probability. In other words, adaptation of the estimates in expressions (5) and (6) is impossible.

Nevertheless, the solution of the problem under the assumption of full *a priori* information about the observations' statistical features is useful. It allows us to find the important statistic, which is contained in all statistically optimal algorithms for array data processing. This statistic is the output of the optimal group filter (OGF):

$$y_t = \vec{r}_t * \vec{x}_t \qquad Y_j = \vec{R}(\lambda_j)\vec{x}_j \qquad t, j \in \overline{1, N}$$

where

$$\vec{R}(\lambda_j) = \frac{\vec{Z}_j}{V_j} = \frac{\vec{h}^{\dagger}(\lambda_j)\mathbf{F}_{\eta}^{-1}(\lambda_j)}{\vec{h}^{\dagger}(\lambda_j)\mathbf{F}_{\eta}^{-1}(\lambda_j)\vec{h}(\lambda_j)} \tag{7}$$

$$\vec{r}_t = (1/(2\pi))\int_0^{2\pi}\vec{R}(\lambda)e^{i\lambda t}d\lambda$$

If the input time series $\vec{x}_t$ fits the observational model (2), (3), then it is easy to show that the filter $\vec{R}(\lambda)$ minimizes the mean square value of the output noise component provided the signal component is undistorted. The remarkable feature of the filter $\vec{R}(\lambda)$ is that replacement in the expression (7) of the unknown spectral matrix $\mathbf{F}_{\eta}(\lambda)$ of the noise by the spectral matrix $\mathbf{F}_x(\lambda)$ of the observations, does not change its output: it keeps the signal undistorted and minimizes the noise component power. The new filter

$$\vec{K}(\lambda) = \frac{\vec{h}^{\dagger}(\lambda)\mathbf{F}_x^{-1}(\lambda)}{\vec{h}^{\dagger}(\lambda)\mathbf{F}_x^{-1}(\lambda)\vec{h}(\lambda)} \tag{8}$$

is equivalent to the filter given by equation (7). This can be shown using Bartlett's formula (4). Due to this property, each statistic, depending only on the output of OGF, has the adaptive feature. This means that in this statistic we may substitute instead of the unknown noise matrix power spectrum $\mathbf{F}_{\eta}(\lambda)$ any consistent estimate $\tilde{\mathbf{F}}_X(\lambda)$ of the matrix power spectrum of the observations $\vec{x}_t$, thus providing the adaptive (AOGF) filter $\hat{K}(\lambda)$.

Let us now consider some practically interesting modifications of OGF $\vec{R}(\lambda)$. In the case when the noise field mainly consists of coherent primary waves, generated by a seismic source, the matrix spectrum $\mathbf{F}_{\eta}(\lambda)$ may be written in the form:

$$\mathbf{F}_{\eta}(\lambda) = \varepsilon\mathbf{F}_g(\lambda) + \psi(\lambda)\vec{q}(\lambda)\,\vec{q}^{\dagger}(\lambda) \tag{9}$$

where $\mathbf{F}_g(\lambda)$ is the matrix power spectrum of the "diffuse noise" component, ε is a scaling factor, $\psi(\lambda)$ is the scalar power spectrum of the primary waveform, and $\vec{q}(\lambda) = (q_e(\lambda), 1 \in \overline{1,m})$ is the frequency response of the medium for the primary wave paths from the source to the array sensors. Using the Bartlett formula (4) it is easy to show that for the matrix spectrum (9), OGF $\vec{R}(\lambda)$ has the form:

$$\vec{R}(\lambda) = \vec{R}_o(\lambda) + \vec{o}_{\lambda}(\varepsilon), \qquad \vec{R}_o(\lambda) = \frac{\vec{h}^{\dagger}(\lambda)\mathbf{B}(\lambda)}{\vec{h}^{\dagger}(\lambda)\mathbf{B}(\lambda)\vec{h}(\lambda)} \tag{10}$$

where

$$\mathbf{B}(\lambda) = \left(I - \frac{\vec{q}(\lambda)\,\vec{q}^{\dagger}(\lambda)}{|\vec{q}(\lambda)^2|}\right), \qquad \|\vec{o}_{\lambda}(\varepsilon)\| \to 0 \quad \text{if } \varepsilon \to 0$$

I is the identity matrix.

$\|\vec{f}(\lambda)\|$ is the norm of the vector function $\vec{f}$. So if the "diffuse" components of the noise and signal fields are weak and hence the matrices $\mathbf{F}_\eta^{-1}(\lambda)$ and $\mathbf{F}_x^{-1}(\lambda)$ are close to singular, it is reasonable to use the stable form of OGF with frequency response $\vec{R}_o(\lambda)$.

In the opposite particular case, when the noise field is purely "diffuse", i.e., it may be modeled as white noise, the matrix power spectrum $\mathbf{F}_\eta(\lambda) = \sigma^2(\lambda)\mathbf{I}$ and

$$\vec{R}(\lambda) = \frac{\vec{h}^{\dagger}(\lambda)}{|\vec{h}(\lambda)|^2} = \vec{\Gamma}(\lambda) \tag{11}$$

The undistorted group filter $\vec{\Gamma}(\lambda)$ is the mathematical expression for the conventional algorithm of seismic holography (Troitsky, et al, 1980), comprising the "focusing of rays" radiated from a given point $\vec{r}$ of the scanned medium area. It compensates for the phase and amplitude distortions of the array records caused by the wave propagation from the source to the array sensors.

Let us consider the dispersion $\delta_N^2 = \frac{1}{N}\sum_{t=1}^{N} y_t^2$ of the group filter $\vec{R}_o(\lambda)$ output $y_t, t \in \overline{1,N}$. In case the input $\vec{x}_t$ has a matrix power spectrum described by the formulas (3) and (9), the property of the matrix $\mathbf{B}(\lambda)$: $\mathbf{B}(\lambda)\vec{q}(\lambda) \equiv 0$ yields:

$$P - \lim_{N \to \infty} \delta_N^2 = \theta + \beta_\varepsilon \tag{12}$$

where $\beta_\varepsilon \to 0$ if $\varepsilon \to 0$. So if the "diffuse" noise (caused, particularly, by scatterers located away from point $\vec{r}$) is small enough, the statistic δ_N^2 provides a consistent estimate of the scattered wave source power, thus suppressing the array noise generated by the primary wave. The OGF $K(\lambda)$ (8) has the same important feature. And, finally, it is clear that the relation (12) is valid for the adaptive filter $\vec{K}(\lambda)$.

Thus, in the case of coherent noise, the problem of consistent estimation of the seismic wave power radiated from point $\vec{r}$ can be solved rather effectively by using OGF $\vec{R}_0(\lambda)$ or AOGF $\hat{K}(\lambda)$.

At the same time, the conventional holographic group filter $\vec{\Gamma}(\lambda)$ cannot suppress the coherent noise and does not provide a consistent estimator of the seismic wave power radiated from point $\vec{r}$.This is so because the dispersion of the holographic filter $\vec{\Gamma}(\lambda)$ output (when the input $\vec{x}_t$ matrix power spectrum is defined by (3) and (9)) does not tend to the seismic wave power radiated from the point $\vec{r}$ when the "diffuse" noise component decreases.

Results of simulation

The benefits of using the optimal statistical algorithms for mapping of scattered wave sources have been analyzed by computer modeling. The records of a Hindu Kush earthquake on 22 February 1972 with magnitude 5.6 recorded by NORSAR have been used in this modeling. To check the algorithm's capability to process data recorded at arrays with a small number of sensors and a small aperture, the six sensors of the central NORSAR subarray 01A have been chosen. The location of these sensors are shown in Fig. 7.1.1, and a record of the Hindu Kush earthquake is shown in Fig. 7.1.2.

The numerical experiments consisted of processing data from the six sensors of the central NORSAR subarray 01A. The data were composed as a sum of primary and scattered waves. As primary waves we have tested two cases: 1) observed array records of the Hindu Kush earthquake; 2) A plane P-wave with an apparent velocity and azimuth corresponding to that of the Hindu Kush earthquake. The modelling of this plane wave was done to check how the deviation from a plane wavefront affects the quality of the seismic holography algorithms. The scattered wave was simulated as a P-wave radiated from a point source located in a homogeneous medium 4 km beneath the NORSAR central subarray. The source waveform s_t of the simulated scattered wave was synthesized as a sample function of the stationary Gaussian random process with power spectrum estimated using the Hindu Kush earthquake P-wave record.

In Figs. 7.1.3-7.1.5 the results from processing four different data sets are depicted. The array data sets were the following:

a. The simulated plane wave from the Hindu Kush direction

b. The sum of a) and the simulated scattered wave from the point source beneath NORSAR

c. The original Hindu Kush earthquake records

d. The sum of c) and the simulated scattered wave from the point source beneath NORSAR.

The power of the scattered wave was equal to that of the primary wave.

In each of the four experiments the array signals were processed by the following algorithms: a) the conventional algorithm of seismic holography based on the filter $\vec{\Gamma}(\lambda)$ (11); b) the statistical algorithm based on the adaptive group optimal filter $\vec{K}(\lambda)$ (8) with the power spectral density inverse matrix $\mathbf{F}_x^{-1}(\lambda)$ estimated by the multidimensional autoregressive modeling of the observations $\vec{x}_t$; c) the statistical algorithm based on the undistorted group filter $\vec{R}_o(\lambda)$ (10), in which the vector $\vec{q}(\lambda)$ was chosen in correspondence with the plane wave with the azimuth and apparent velocity of the Hindu Kush earthquake P-wave.

The data processing procedure consisted of scanning, with steps of 0.8 km an 8 x 8 km plane parallel to the surface containing the scattered wave source in the central point. For each point scanned, the medium frequency response vector $\vec{h}_r(\lambda)$ was computed under the assumption of medium homogeneity. By this, the adjustment of the group filters for the extraction of the scattered wave radiated from the given point was provided. Then the filter output signals and their dispersion estimates were computed. As was pointed out, the dispersion of the outputs of all filters under investigation may be used as the estimates of the scattered wave power radiated from the given point. These estimates are biased because of primary wave power "leakage", and these biases have different values for the different filters.

The number of samples of the input vector time series $\vec{x}_t$ was as a rule equal to 128 in these experiments, and corresponded to the Hindu Kush earthquake P-wave record length. The matrix power spectrum $\mathbf{F}_x(\lambda)$ of the observations $\vec{x}_t$ used for the undistorting group filter adaptation was estimated by multivariate AR-modeling of order 5.

In Figs. 7.1.3-7.1.5 the scattered radiation power maps are depicted. They are calculated as the results of the numerical experiments described, with the conventional holographic filter (11) (Fig. 7.1.3), with the adaptive optimal group filter (7) (Fig. 7.1.4), and with the "spatial rejecting" filter (10) which uses the a priori information concerning the propagation direction of the primary wave (Fig. 7.1.5).

In these figures, the maps with label (a) are the results for the single simulated primary wave processing, with label (b) the results for the sum of the simulated primary wave and the simulated scattered wave processing, with label (c) the results for the Hindu Kush earthquake P-wave processing, and with label (d), the results for the sum of the Hindu Kush earthquake P-wave and simulated scattered wave processing.

Comparison of the maps in Fig. 7.1.3 shows that the conventional algorithm of seismic holography applied to processing of data from arrays with a small aperture and small number of sensors does not provide detection of the scattered wave. This may be explained by the effect of the primary wave power "leakage", practically evenly strong for all points of the scanned area. At the same time, as it is seen from Figs. 7.1.4 and 7.1.5, the statistical seismic holography algorithms provide strong peaks in the power maps, the maxima of which coincide with the point of the scattered wave source location. The explanation may be that the mentioned algorithms are capable of suppressing "leakage" by the mutual compensation of primary wave components at the different array sensors. The adaptive and "a priori" algorithms demonstrate the same ability for the suppression of the primary wave "leakage" and location of the scattered wave source, while in the case of detection of the scattered wave on the background of the Hindu Kush earthquake, the adaptive optimal group filter (Fig. 7.1.4d) has appeared more effective than the "a priori" filter (Fig. 7.1.5d). The advantage of the optimal group filter over an "a priori" filter in this experiment may be explained by the deviations of the Hindu Kush earthquake signal from a plane wave. This deviation significantly reduces the performance of the filter $\vec{R}_o(\lambda)$ (10). At the same time, for the adaptive filter $\hat{K}(\lambda)$ no assumptions are needed about the primary wave front shape, which makes it more flexible and effective.

A.F. Kushnir, E.A. Gulko, MITPAN, Moscow

References

Bartlett, M.S. (1951): An inverse matrix adjustment arising in discriminant analysis, *Ann. Math. Stat.*, Vol 22, No 1.

Kushnir, A.F. (1989): Parametric methods of geophysical time series statistical analysis. Dr. of Science Dissertation Thesis, Inst. of Physics of the Earth, USSR Academy of Sciences, Moscow.

Troitsky, P.N., E.S. Husebye and A.V. Nikolaev (1980): Nature, vol. 5, no. 284, 618-623.

Voievodin, V.V., and U.D. Kuznetsov (1984): Matrices and computational problems, M: Nauka, 318 pp (in Russian).

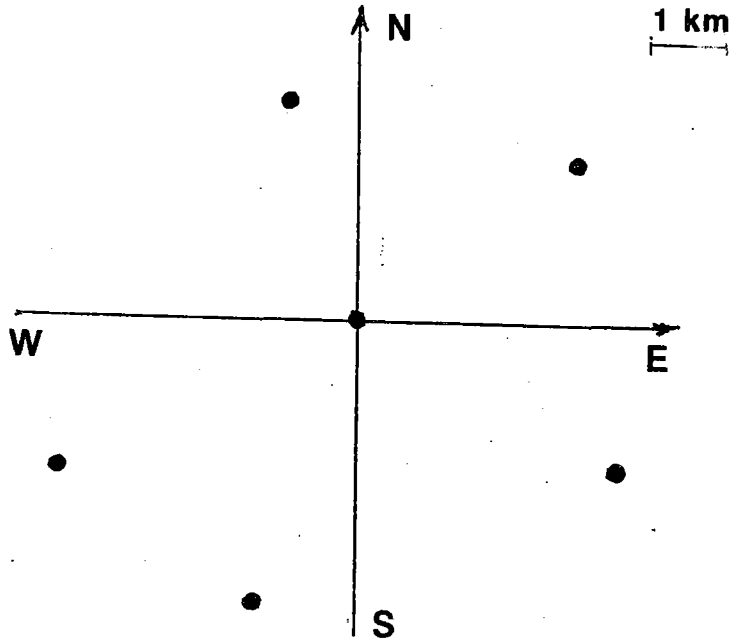

Fig. 7.1.1. NORSAR central subarray (01A) geometry.

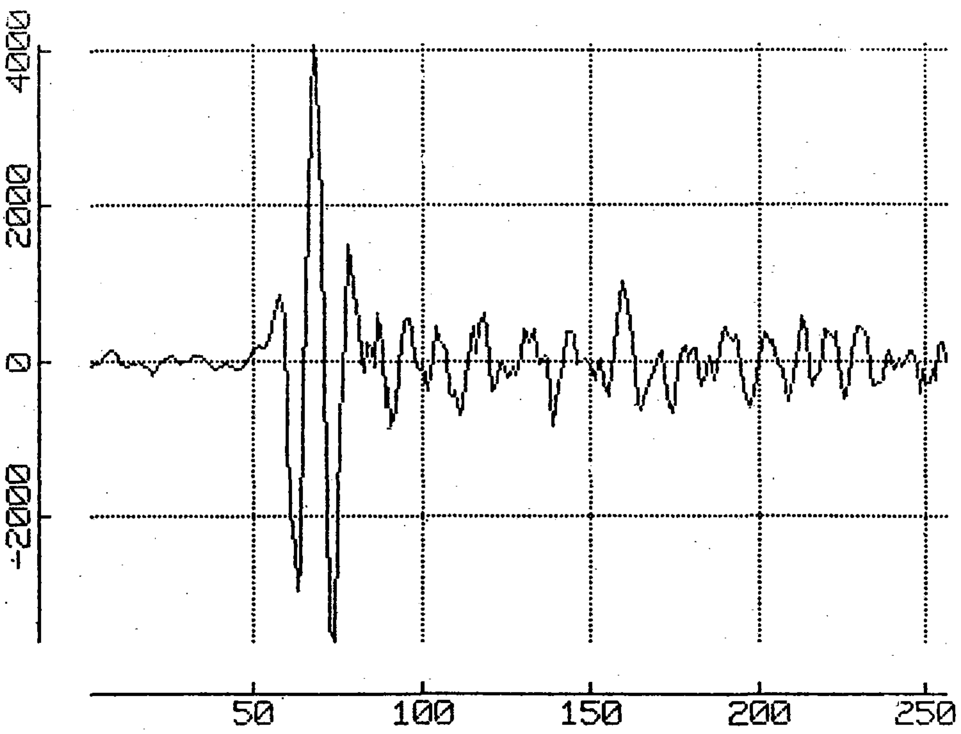

Fig. 7.1.2. Hindu Kush earthquake recordings at subarray 01A.

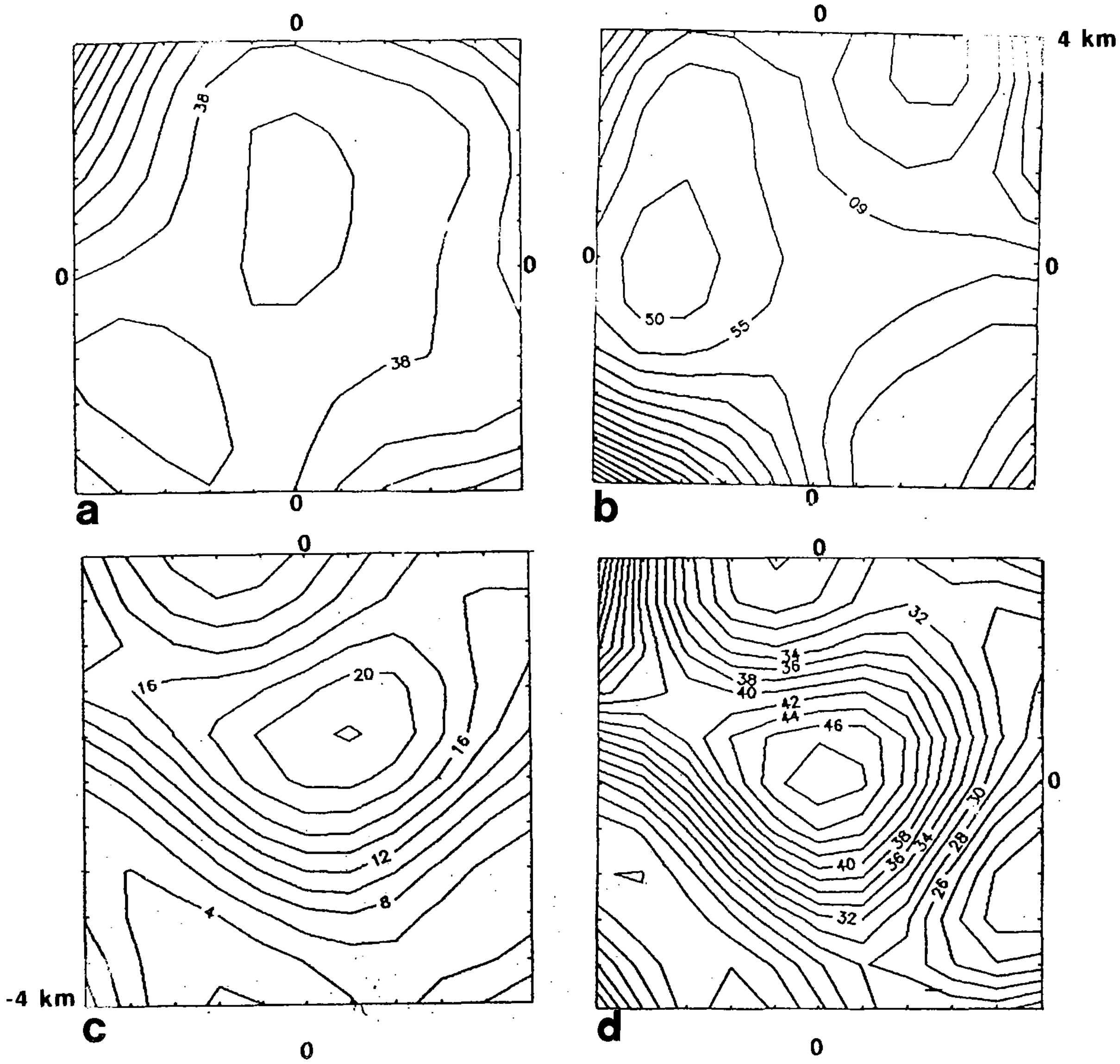

Fig. 7.1.3. Array data processing using conventional holographic group filter. (a) Simulated plane wave from earthquake; (b) simulated earthquake plane wave + simulated scattered wave; (c) P-wave of the Hindu Kush earthquake; (d) Hindu Kush earthquake P-wave + simulated scattered wave. The contour levels reflect the estimated power of waves arriving from different points of the scanned plane. The power estimates are *not* normalized, hence only relative values within each map should be considered.

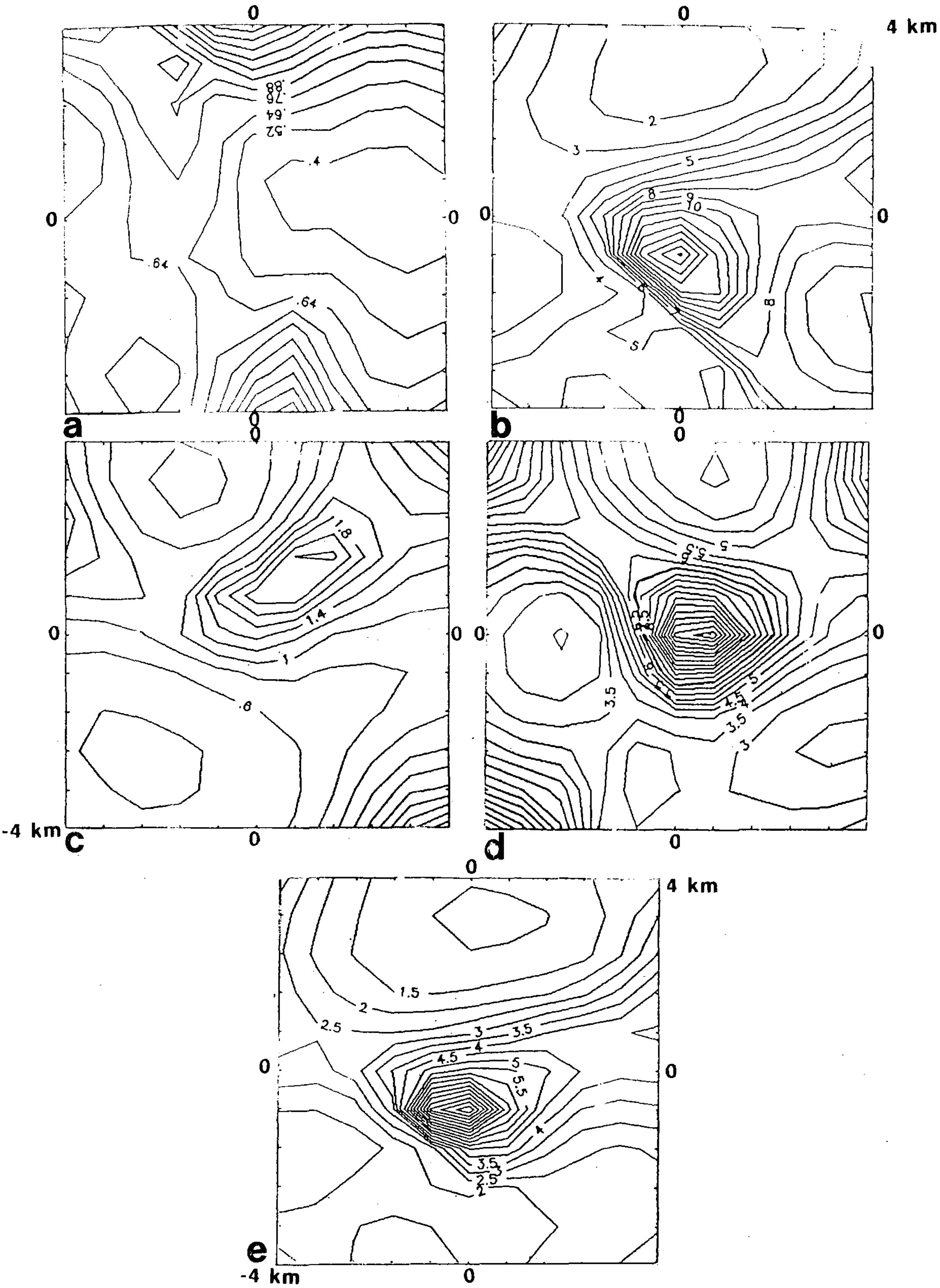

Fig. 7.1.4. Array data processing using adaptive optimal group filter: (a) Simulated plane wave from earthquake; (b) Simulated earthquake plane wave + simulated scattered wave; (c) P-wave of the Hindu Kush earthquake; (d) Hindu Kush earthquake P-wave + simulated scattered wave; (e) the same as (b) but for N=1024, p=10.

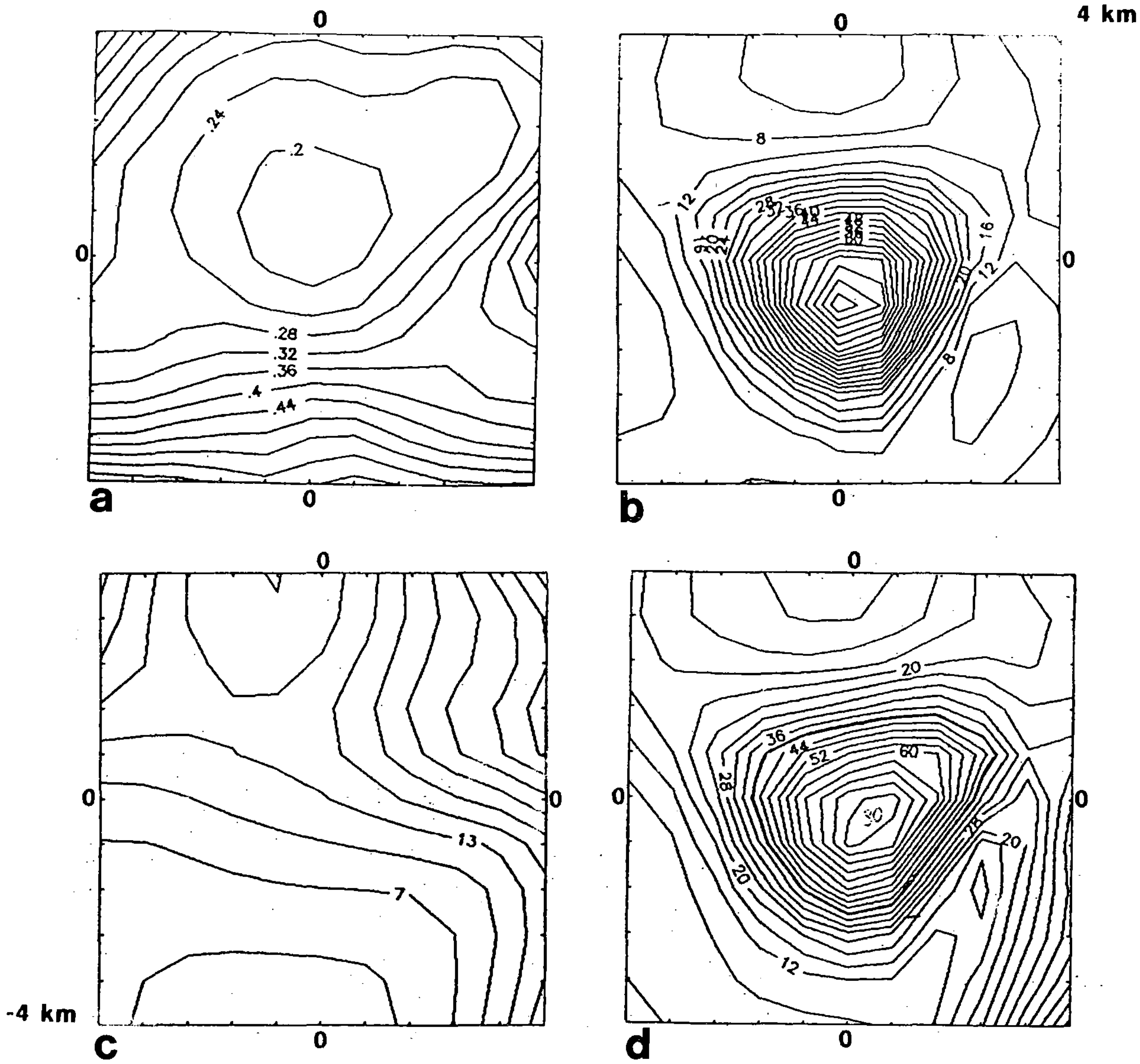

Fig. 7.1.5. Array data processing using spatial rejecting filter: (a) Simulated plane wave from earthquake; (b) Simulated earthquake plane wave + simulated scattered wave; (c) P-wave of the Hindu Kush earthquake; (d) Hindu Kush earthquake P-wave + simulated scattered wave.