CONCURRENT MIGRATION AND VELOCITY ESTIMATION IN A RANDOM, NON-GAUSSIAN EARTH

Sunil Mehta

The problem we have set out to solve is one of detecting variations in reflectivity in the earth from the waves that we record at the earth's surface during exploration seismology experiments. Since numerical techniques of projecting waves through the earth are now well developed, we want to project the observed waves back into the earth with a constant velocity and detect their point of origin. The number of z-steps necessary to move the waves back to their point of origin and the travel times of the waves to the surface can be used to estimate an rms velocity for the strata that the waves propagated through.

To begin, we will describe the amplitude distribution in a random wavefield by a probability density function (pdf) $p(x)$. Waves undergo space averaging (diffraction) and time averaging (filtering) during propagation. Therefore, we can expect, because of the Central Limit Theorem, that the amplitude pdf of a wavefield will resemble a Gaussian pdf more closely as the wave propagates. Consequently, the criterion that we will use to detect when the waves have been migrated to their point of origin is the following: *The amplitude pdf for the wavefield,* $p(x)$, *will deviate more from a Gaussian pdf at the source point than at any other point in the migration process.*

Such a velocity-independent criterion could be particularly useful in migrating reflection data generated by deep structures in the earth. A typical example of such data is the deep crustal reflection data recorded in Hardeman County, Texas, by the Consortium for Crustal Reflection Profiling Project (COCORP). In this case, velocity estimation from CDP gathers is difficult because the recording spread is very short compared to the long travel times. However, if diffracted events exist on the stacked section, it may be downward continued using the migration criterion. This should yield a better stacked section than would be obtained by using the conventional, and possibly inaccurate, velocity estimate. We would also obtain a refined velocity estimate in the process.

The focus of this paper will be the implementation of the criterion above. We will first show how to estimate an amplitude pdf for seismic data. Then we describe a normalization technique for pdfs so that they can be compared with a standard Gaussian pdf. Finally, we will examine a method of measuring the distance of an estimated pdf from the Gaussian pdf. We will present two synthetic models that we studied.

Our first problem is to estimate the pdf for a sample of seismic data. Claerbout (SEP-10, p. 101) has shown how to estimate the pdf for a sample of n independent, random data points. We will quickly summarize the procedure. To begin with, we order the data points in our sample and estimate a quantile value for each point. The quantile value corresponding to the i^{th} data point out of n data points will be represented as $\alpha(i,n)$. The estimated quantile, $\hat{\alpha}(i,n)$ is constructed by requiring that the probability that it overestimates the actual quantile $\alpha(i,n)$ be equal to the probability that it underestimates $\alpha(i,n)$. The probability that $\hat{\alpha}(i,n)$ underestimates $\alpha(i,n)$ is equal to the probability that there are $n[1 - \hat{\alpha}(i,n)]$ or less points in the upper tail area of the pdf, or that

$$(n-1)/n \leq 1 - \hat{\alpha}(i,n) \ . \tag{1}$$

Then, the quantile value corresponding to $(n-1)$ or less points of the sample being drawn from the upper tail area of the pdf is given by the solution to the equation

$$\sum_{k=i}^{n} \binom{n}{k} [\hat{\alpha}(i,n)]^k [1 - \hat{\alpha}(i,n)]^{n-k} = 0.5 \ . \tag{2}$$

Claerbout (SEP-10, p. 109) has shown how to get an approximate solution to the equation. By estimating the quantile value for each data point, we specify the probability mass between any two data points. To construct an estimate we assume a constant probability density between each pair of data points and write the estimated pdf as

$$\hat{p}(x) = \frac{\hat{\alpha}(i+1,n) - \hat{\alpha}(i,n)}{x_{i+1} - x_i} \ , \qquad x_i < x < x_{i+1} \ , \tag{3}$$

where x_i is the i^{th} ordered data point. To use every data point in a sample of seismic data is not practical, since a reasonable size for a sample is on the order of a thousand data points. Secondly, we are not able to solve Eq. (2) for

values of n greater than 100. Therefore, we selected a hundred evenly-spaced data points from the ordered sample, and estimated the pdf for this reduced set. This data-reduction procedure does not yield the results that we would obtain if we had used the complete sample. We did not determine the error associated with this approximation. We feel, however, that the final result is not very sensitive to the procedures used at this stage.

The next step is to estimate the mean and standard deviation of a Gaussian pdf that we will compare our estimated pdf with. Our tests on various data indicate that whereas the estimated pdfs have widely varying probability masses on their tails, they resemble a Gaussian pdf near the median. Therefore we will estimate the parameters of a Gaussian pdf that fits our estimated pdf near the median. An estimator of the standard deviation σ is suggested by the fact that for a Gaussian pdf, 40% of the probability mass lies within $\pm \sigma/2$ from the mean. We estimate σ as the distance between the 30^{th} and 70^{th} sample percentiles. The mean is estimated by the sample median. To normalize all estimated pdfs, we will scale the sample data points so that the median is zero and the distance between the 30th and 70th quantiles is unity. In every case, now, the corresponding Gaussian pdf is the standard (zero mean, unit σ) Gaussian pdf. Figure 1 shows a typical normalized pdf and the corresponding Gaussian pdf.

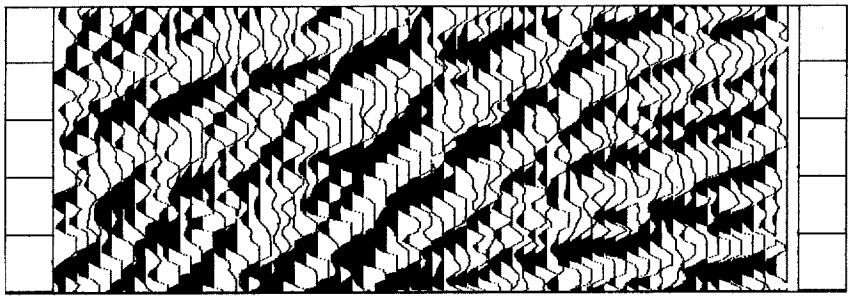
To determine the distance between the estimated pdf $\hat{p}(x)$ and $g(x)$, we use two measures of distance between probability density functions:
(1) The Directed Divergence [Mathai and Rathie, 1975]:

$$D_d\Big]_{x_1}^{x_{100}} = \int_{x_1}^{x_{100}} \hat{p}(x) \, \ell n \frac{\hat{p}(x)}{g(x)} \, dx \, . \tag{4}$$

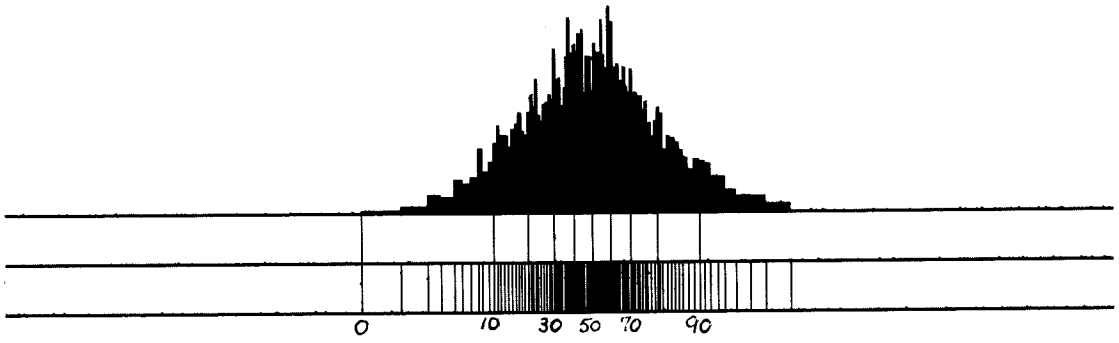
(2) The Bhattacharya Distance, which is a modified form of the Bhattacharya coefficient of affinity [Mathai and Rathie, 1975]:

$$D_b\Big]_{x_1}^{x_{100}} = \int_{x_1}^{x_{100}} \ell n[\hat{p}(x) \, g(x)] \, dx \, , \tag{5}$$

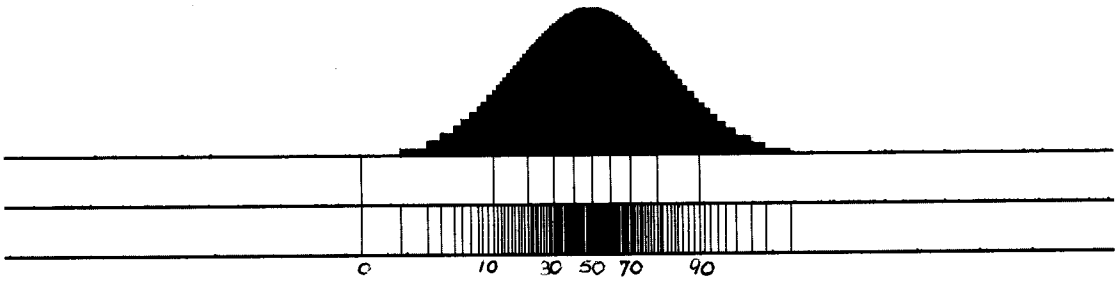
where x_α is the data value corresponding to the α^{th} percentile. The Bhattacharya distance is very sensitive to deviations from the Gaussian pdf on the tails of $p(x)$. This is because the geometric mean of two values is biased



(a)



(b)



(c)

FIGURE 1.—An example of probability density function estimation for seismic data. (a) A sample of seismic data from a near trace section, 64 traces by 128 msec (4096 data points). (b) The estimated pdf $\hat{p}(x)$ for the data sample in (a). The horizontal scale is marked in percentiles. (c) The Gaussian pdf $g(x)$ that fits $\hat{p}(x)$ near the median.

towards the lower value and because the logarithm of a small number is large and negative. There is a simple relation between the Bhattacharya distance and the Directed Divergence. Define

$$\varepsilon(x) = \hat{p}(x) - g(x). \tag{6}$$

Substituting (6) into (5) and simplifying,

$$D_b \Big]_{x_1}^{x_{100}} = \int_{x_1}^{x_{100}} \ell n[g(x)]\, dx + \frac{1}{2} \int_{x_1}^{x_{100}} \ell n[1 + \frac{\varepsilon(x)}{g(x)}]\, dx. \tag{7}$$

Similarly, substituting (6) into (4), we have

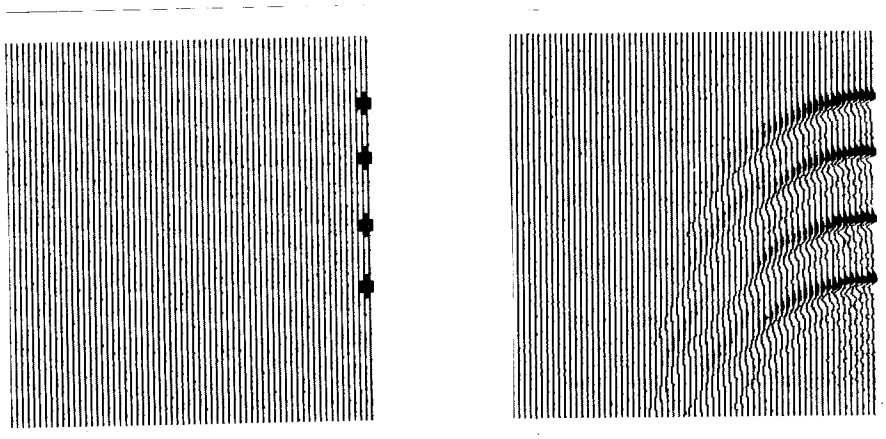
$$D_d \Big]_{x_1}^{x_{100}} = \int_{x_1}^{x_{100}} \hat{p}(x)\ \ell n[1 + \frac{\varepsilon(x)}{g(x)}]\, dx\ . \tag{8}$$

The first term on the right-hand side of (7) is independent of $\varepsilon(x)$. The right-hand side of (8) shows that the Directed Divergence is a weighted version of the second term on the right-hand side of (7). Comparing these two terms, we can see that the Directed Divergence will be less sensitive to deviations from $g(x)$ on the tails of $p(x)$, as it is heavily downweighted there. However, it will always be more stable than the Bhattacharya Distance.

At this point, we have all the tools to test the migration criterion proposed earlier. To restate the hypothesis, as we downward continue the waves observed at the surface, the estimated pdf for their amplitudes $\hat{p}(x)$ will deviate more from a Gaussian pdf at the point of origin of the waves than at any other point, whether the waves are over- or undermigrated. We did two tests on synthetic models. Model I was a flat layered earth with constant reflection coefficients. A data frame, Fig. 2(b), akin to a common shot profile, was generated by diffracting the frame in Fig. 2(a) for 60 z-steps. The dimensionless parameter

$$a = \frac{v\ \Delta t\ \Delta z}{8\ \Delta x^2}$$

in the 15° finite difference approximation to the wave equation that we used was 1/16. These data were then migrated back and overmigrated 20 steps. Amplitude probability density functions were calculated for the last 40



(a) (b)

FIGURE 2.—Model I. The initial data frame (b) was generated by diffracting frame (a) for 60 steps with a = 0.0625. Subroutine FAST-15 (SEP-1, p. 60) was used for diffraction.

migration steps. Nine steps of migration around the point of origin of the diffracted wave are displayed in Fig. 3.

Since the estimated pdf is constant between any two data points, the distance integrals (4) and (5) are evaluated very easily as summations. We evaluated the integrals in two intervals:

(1) On the tails of $\hat{p}(x)$, from the 1st to the 30th and 70th to the 10th percentiles;

(2) Near the median, from the 30th to the 70th percentiles.

The results are tabulated in Table 1 and graphed in Figs. 4 and 5. The Bhattacharya Distance showed a monotonic increase till the point of best migration and a monotonic decrease for overmigration. The Directed Divergence behaved erratically before the point of best migration. It did reach a global-maximum and decrease monotonically thereafter.

Model II represented four point scatterers in a homogeneous earth, two with a positive reflection coefficient and two with a negative reflection coefficient. The sequence of Figs. 6-9 and Table 2 is analogous to Figs. 2-5 and Table 1. The calculations for the two models are identical. In this case, also, the Bhattacharya Distance had a single well-defined maximum at the point of best migration. As in Model I, the Directed Divergence had an erratic behavior leading up to the point of best migration and decreased monotonically thereafter.

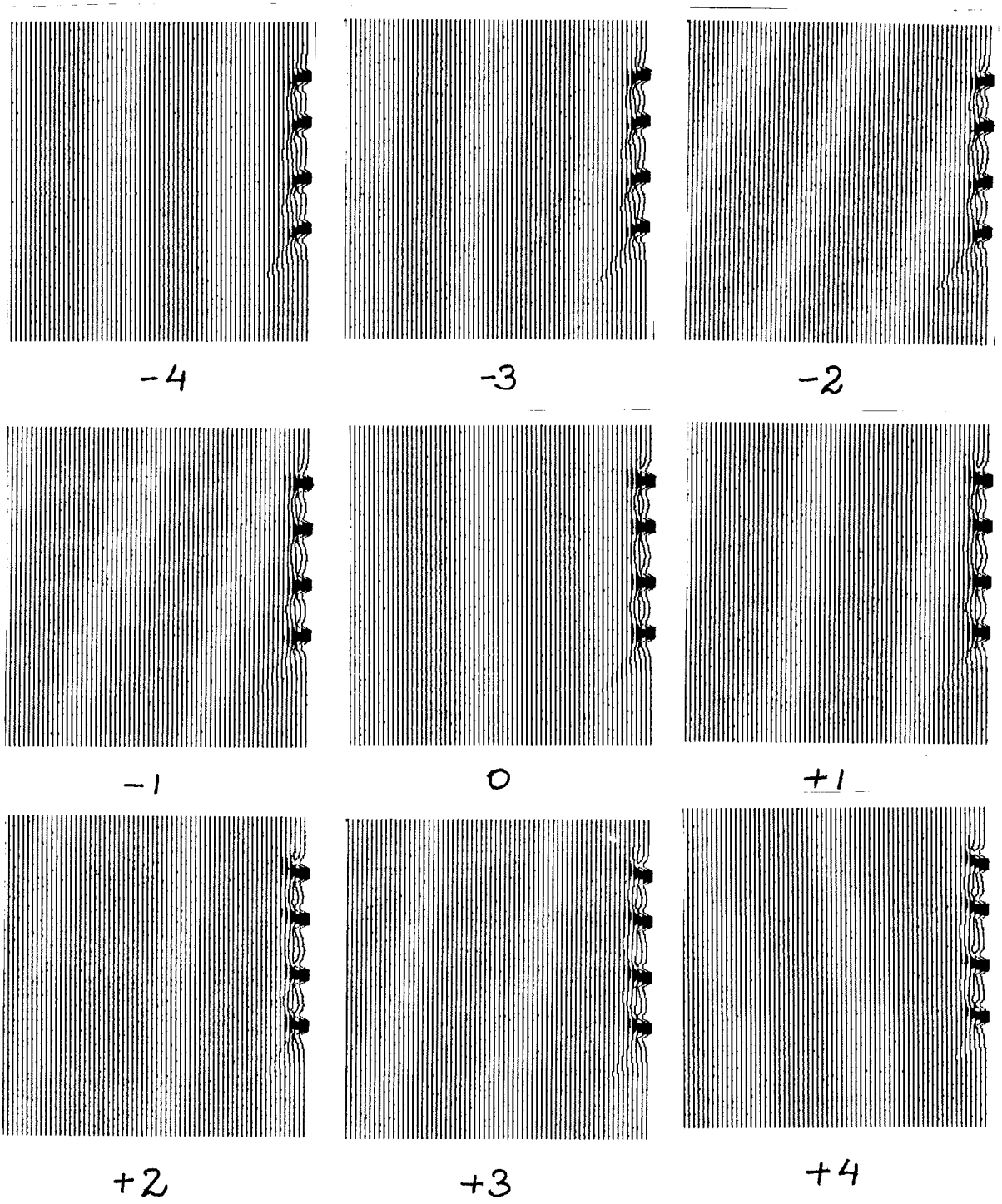


FIGURE 3.—Nine frames about the point of best migration labelled "0". Use of the migration criterion correctly indicated that this frame was best migrated. Note that the frames labelled "-1", "0" and "+1" are almost indistinguishable.

TABLE Ia

DIRECTED DIVERGENCE CALCULATIONS
FOR MODEL I

# of z-steps	Directed Divergence	Deviations on tails	Deviations near median
-19	9.403e+01	9.396e+01	6.875e-02
-15	8.144e+01	8.137e+01	7.344e-02
-11	8.317e+01	8.310e+01	6.598e-02
- 7	8.165e+01	8.160e+01	5.792e-02
- 3	9.687e+01	9.682e+01	4.501e-02
0	1.097e+02	1.097e+02	3.743e-02
5	9.005e+01	9.001e+01	3.349e-02
9	7.553e+01	7.550e+01	3.329e-02
13	6.691e+01	6.688e+01	3.430e-02
17	5.634e+01	5.631e+01	3.087e-02

TABLE Ib

BHATTACHARYA DISTANCE CALCULATIONS
FOR MODEL I

# of z-steps	Bhattacharya Distance	Deviations on tails	Deviations near median
-19	-3.387e+05	-3.387e+05	-10.29e-01
-15	-3.610e+05	-3.610e+05	-10.29e-01
-11	-4.561e+05	-4.561e+05	-10.25e-01
- 7	-6.166e+05	-6.166e+05	-10.11e-01
- 3	-9.704e+05	-9.704e+05	- 9.88e-01
0	-1.153e+06	-1.153e+06	- 9.76e-01
5	-9.362e+05	-9.362e+05	-9.727e-01
9	-4.940e+05	-4.940e+05	-9.630e-01
13	-2.903e+05	-2.903e+05	-9.674e-01
17	-1.892e+05	-1.892e+05	-9.652e-01

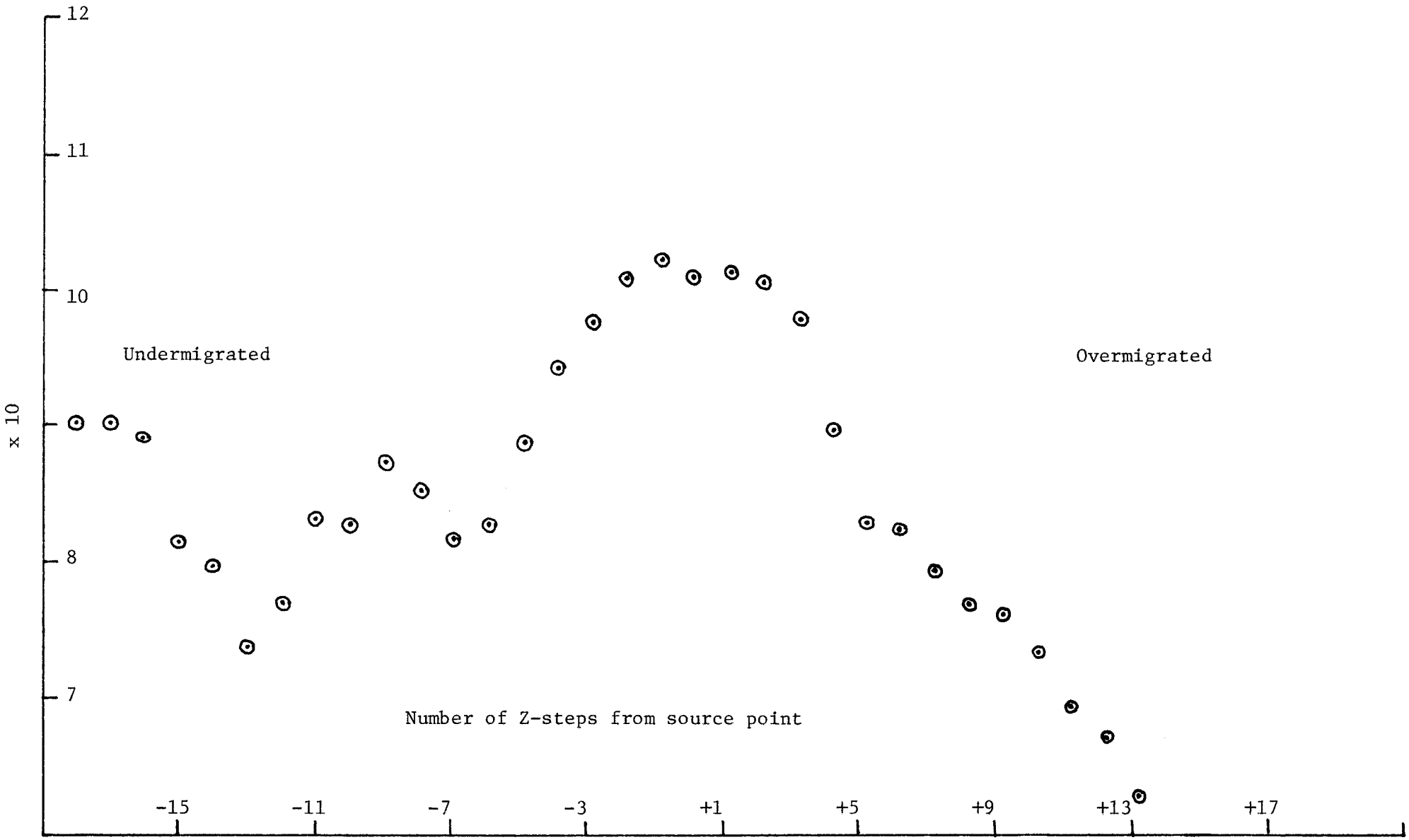


FIGURE 4.—Directed Divergence as a function of distance from the source point for Model I.

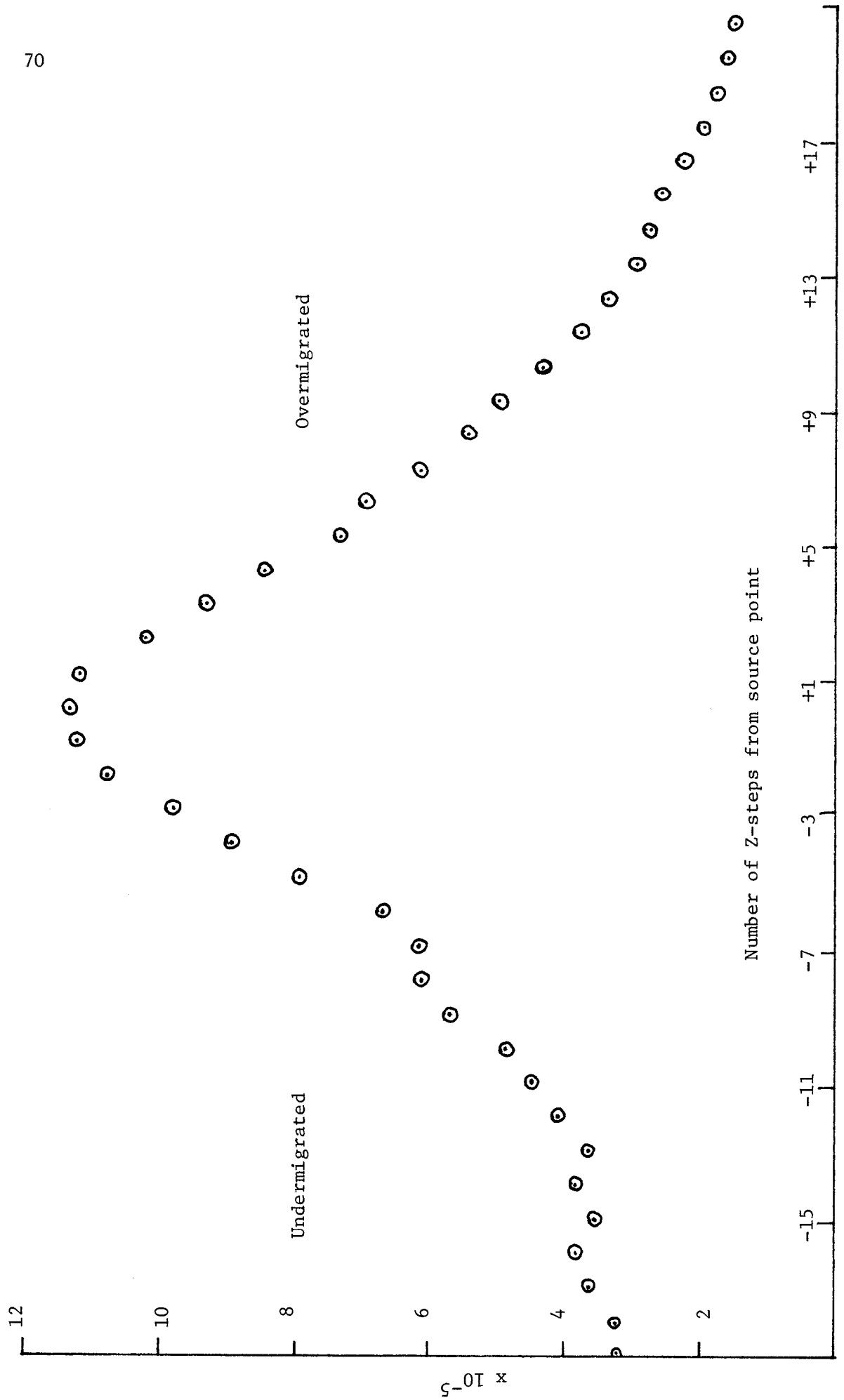
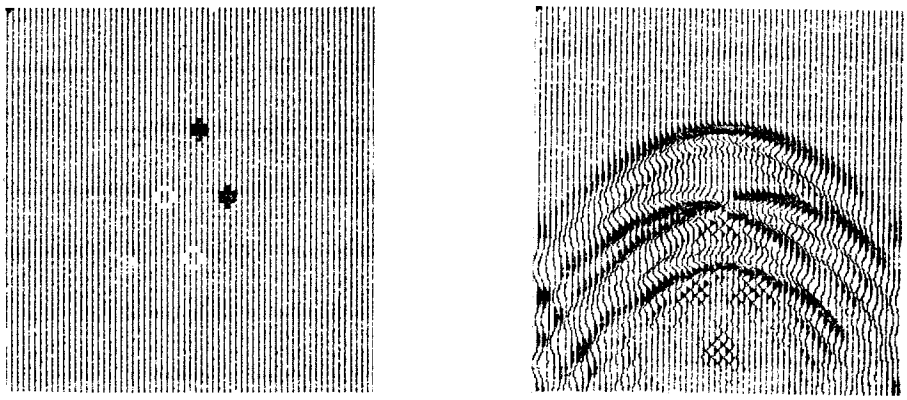


FIGURE 5. Bhattacharya Distance as a function of distance from the source point for Model I.



(a) (b)

FIGURE 6.—Model II. The initial data frame (b) was generated by diffracting frame (a) for 60 steps with a = 0.0625. Subroutine FAST-15 (SEP-1, p. 60) was used for diffraction.

We have noted that in both the models that we studied, there were migration artifacts present. However, we have ensured that as long as the artifacts were low-amplitude events, they would not affect the final result significantly. This is because the contributions to the distance integrals from the low-amplitude, or central area of the pdf are very small, as a look at Tables 1 and 2 will verify.

Let us review what we have achieved. We posed and implemented a criterion for determining the source point of surface-observed waves during migration. In the two preliminary tests, we successfully used the Bhattacharya distance to detect the point of origin of diffracted waves while downward continuing them into the earth. The Directed Divergence, being less sensitive, did not yield as good results.

On the basis of these two model studies, we believe that the migration criterion we have formulated can be used to migrate either stacks or profiles, without requiring a prior velocity estimate. One of the outputs of the migration process can be an rms velocity estimate. The only requirement for the criterion to be effective is that the data be such that downward-continuing it collapses energy along hyperbolic paths into foci.

Reference

Mathai, A. M. and P. N. Rathie, *Basic Concepts in Information Theory and Statistics* (New York: Wiley), 1975.

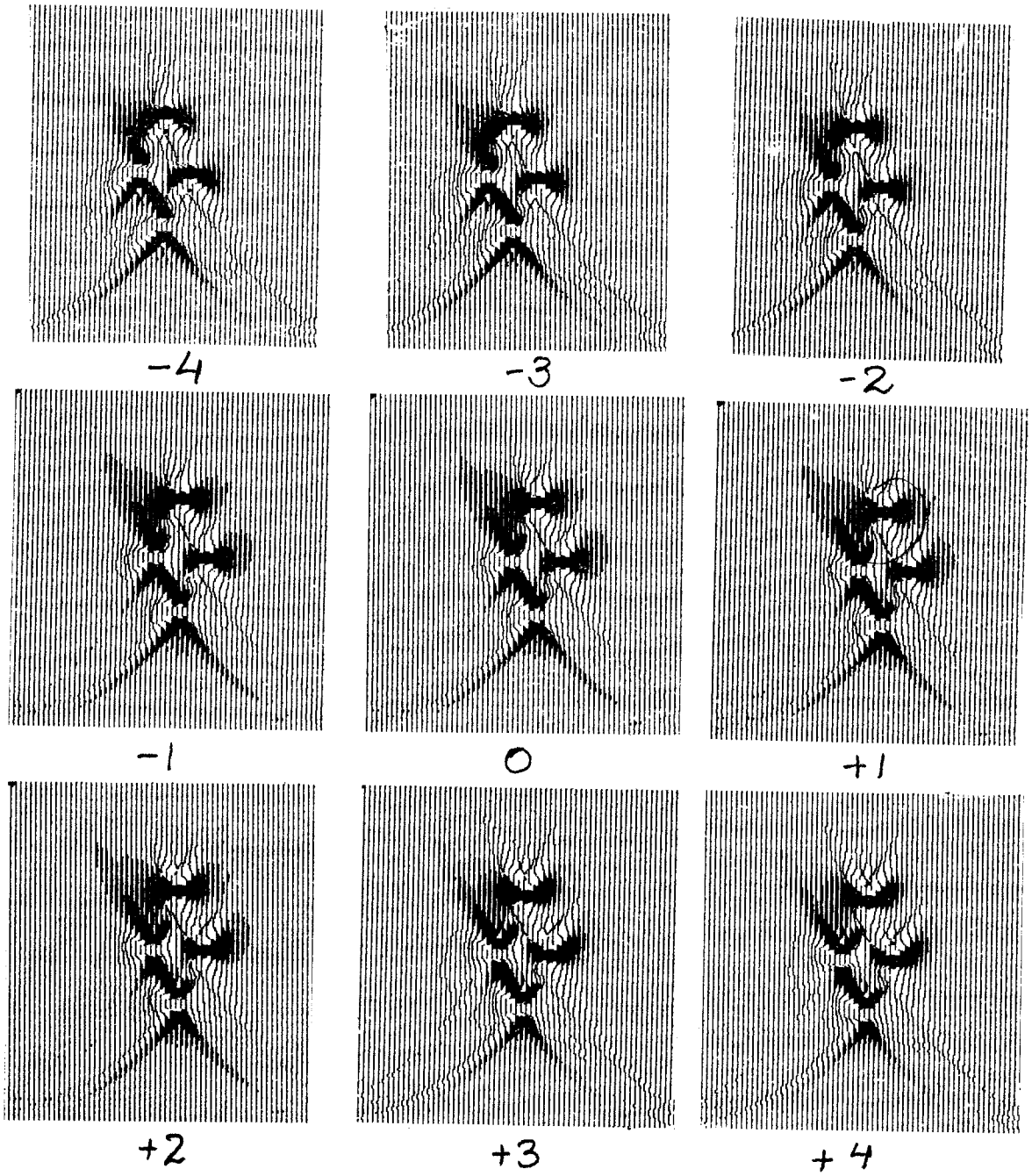


FIGURE 7.—Nine frames about the point of best migration labelled "0". The migration criterion indicated that frame "-1" should be the source point for the waves. Note that there are many migration artifacts in the frames. We do not feel that these have affected our calculations, since they are of low amplitude. The major contribution to the distance integrals comes from the high-amplitude events on the tails of $\hat{p}(x)$ as Table 2 shows. As noted for Model I, the difference in quality between frames labelled "-1", "0", and "+1" is minimal.

TABLE IIa

DIRECTED DIVERGENCE CALCULATIONS
FOR MODEL II

# of z-steps	Directed Divergence	Deviations on tails	Deviations near median
-19	8.909e+01	8.902e+01	6.472e-02
-15	8.594e+01	8.589e+01	5.796e-02
-11	8.499e+01	8.494e+01	5.310e-02
- 7	8.317e+01	8.312e+01	4.358e-02
- 3	8.887e+01	8.884e+01	3.311e-02
0	8.836e+01	8.833e+01	2.872e-02
5	6.513e+01	6.510e+01	2.957e-02
9	5.224e+0	5.221e+01	2.221e-02
13	4.175e+01	4.173e+01	2.209e-02
17	3.423e+01	3.421e+01	2.339e-02

TABLE IIb

BHATTACHARYA DISTANCE CALCULATIONS
FOR MODEL II

# of z-steps	Bhattacharya Distance	Deviations on tails	Deviations near median
-19	-2.347e+05	-2.347e+05	-1.009e+01
-15	-3.050e+05	-3.050e+05	-1.007e+01
-11	-3.994e+05	-3.994e+05	-0.962e+01
- 7	-5.145e+05	-5.145e+05	-0.979e+01
- 3	-7.741e+05	-7.741e+05	-0.964e+01
0	-7.817e+05	-7.817e+05	-0.957e+01
5	-4.304e+05	-4.304e+05	-0.956e+01
9	-2.224e+05	-2.224e+05	-0.948e+01
13	-1.324e+05	-1.324e+05	-0.948e+01
17	-0.741e+05	-0.741e+05	-0.951e+01

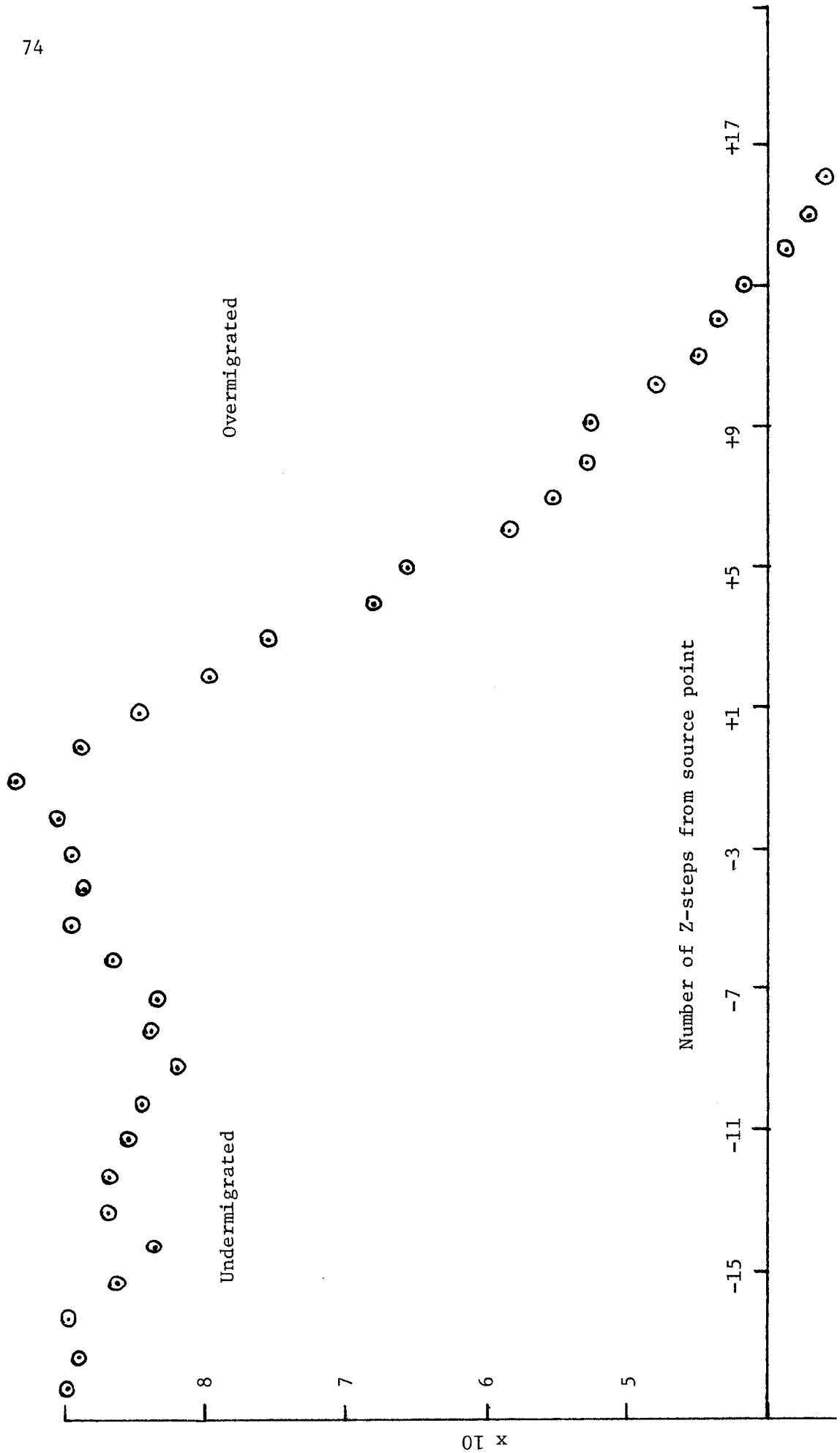


FIGURE 8.—Directed Divergence as a function of distance from the source point for Model II.

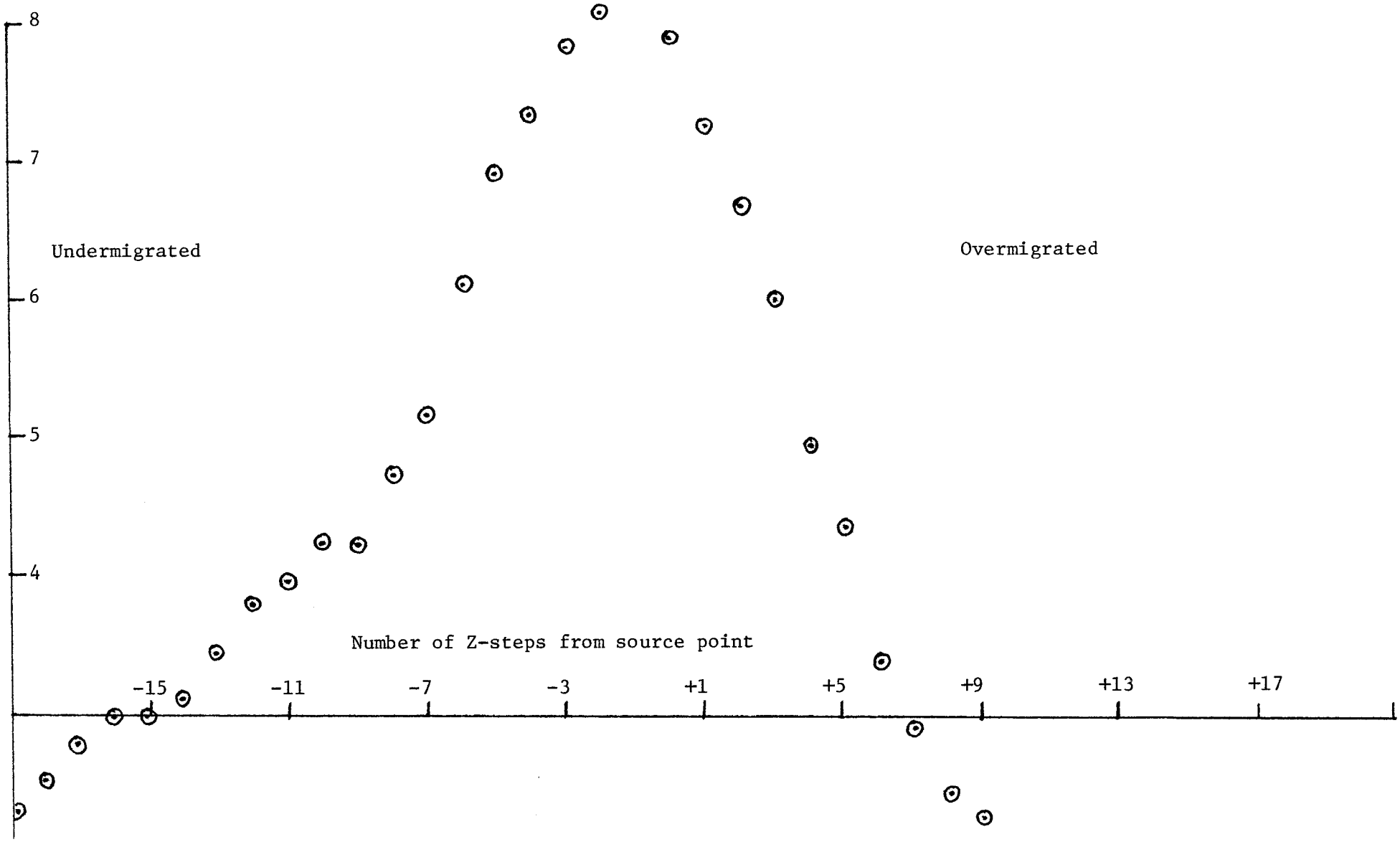


FIGURE 9.—Bhattacharya Distance as a function of distance from the source point for Model II.