Surface-consistent residual statics estimation by stack-power maximization*

Joshua Ronen and Jon Claerbout

ABSTRACT

Application of incorrect static shifts will decrease the power of the CDP-stack. Conversely, static shifts can be estimated by maximizing the stack power. We tried it and it worked well enough that we recommend it in routine practice for data with low signal-to-noise ratio.

Likewise, stack power was maximized by adjustment of surface-consistent phase correction.

INTRODUCTION

Review of travel-time picking methods

Near-surface lateral velocity variations and topographical changes cause time anomalies that can be approximated as surface-consistent static time shifts. This approximation has obvious limitations; nevertheless, it is adequate often enough that time anomalies are routinely corrected by static time shifts. Methods of automatic estimation of the near-surface anomalies, based on static time shifts, have been developed (Hileman et al., 1968; Disher and Naquin, 1969; Taner et al., 1974; Wiggins et al. 1976; Kirchheimer, 1983). These methods fit a surface-consistent model to time anomalies of a particular event on various traces:

$$\Delta t_i = S(s_i) + G(g_i) + Y(y_i) + R(y_i) \times h_i^2 \ . \quad (1)$$

* Presented at the 54th SEG Meeting in Atlanta, Georgia, December 1984. Accepted for publication; will appear in the golden anniversary issue of Geophysics, December 85.

Δt_i is the time anomaly for trace i, having the shot station s_i, geophone station g_i, midpoint station $y_i = (s_i + g_i)/2$, and shot-receiver offset $h_i = (s_i - g_i)/2$. $S(x)$ and $G(x)$ are shot and geophone static corrections as a function of station x, $Y(x)$ is a structure term, and $R(x)$ is a residual NMO correction.

Travel-time-picking methods obtain the time anomalies by picking crosscorrelations maxima: the maximum of a crosscorrelation is the most likely relative time shift between two similar traces (Foster and Guinzy, 1967). Once the time anomalies, Δt_i, are picked, the functions $S(x)$, $G(x)$, $Y(x)$ and $R(x)$ are found by solving an over-determined, under-constrained system of equations.

Picking peaks of crosscorrelations is a non-linear operation. Non-linear operations are susceptible to failure in the presence of ambiguities or noise. This led us to hypothesize that static shifts should be determined so that the power in the final stack is maximized. Our implementation of this hypothesis does not avoid picking crosscorrelations but it incorporates the picking at a later stage as part of the model estimation.

Model estimation by optimization

Consider the estimation of surface-consistent static corrections as an optimization problem: the stacked section is a function of the statics model; if we change the statics, the CDP stack will be different. To find the best static model we can, in principle, try all possible models and choose the one that yields the best stack.

Model estimation by direct optimization has been used in various inversion problems: maximizing semblance in velocity analysis (Koehler and Taner, 1967; Neidell and Taner, 1971; Toldi, 1985), maximizing parsimony in missing data (Thorson, 1984) and extremizing entropy or maximizing likelihood in deconvolution (Burg, 1972; Wiggins, 1978; Chi et al., 1983). We suggest estimation of surface-consistent statics by maximization the sum of squares (power) of the stacked section.

The power of the stacked section is a good measure of quality because if all the traces are the same except for static time shifts, then the stack is most powerful when all the traces are aligned with no relative shift. (The proof is by using Cauchy-Schwartz inequality.)

Travel-time-picking methods for statics use indirect optimization when fitting the travel times to a surface-consistent model (usually by least squares). The optimization we suggest is different: it is based directly on the seismic data and not on picked travel times.

Super-trace crosscorrelation

Unstacked data is a function of midpoint y, offset h, and time t. In the volume of this unstacked data, a static time shift of a particular shot is a uniform shift of the plane containing that shot profile (Figure 1). Common midpoint stacking is summing along the offset direction. The power in the resulting stack is a function of the static shifts of every shot and every geophone. For a single shot station, one could try every shift, stack and sum the squares of the stack; then one could choose the shift that gave the highest power. What we actually do is equivalent but more efficient: a super trace built from all the traces of the shot profile in sequence (trace F in Figure 1) is crosscorrelated with another super trace of all the traces in the relevant part of the stack in sequence without the contribution of that shot (trace G in Figure 1). We then pick the maximum of that crosscorrelation. Zero segments are included between traces so that end effects are avoided. The equivalence of maximizing the power of the sum and maximizing a crosscorrelation is shown by:

$$\begin{aligned} \mathbf{Power}(\Delta t) &= \sum_t \left[F(t-\Delta t) + G(t) \right]^2 \qquad (2) \\ &= \sum_t \left[F^2(t-\Delta t) + G^2(t) \right] + 2\sum F(t-\Delta t) G(t) \\ &= \mathbf{Constant} + 2 \times \mathbf{Crosscorrelation} \, . \end{aligned}$$

To estimate a shot-consistent static shift, the shot profile is crosscorrelated with the relevant part of the stack and the maximum is picked. The stack is corrected accordingly. This procedure is repeated in a cyclical coordinate ascent:

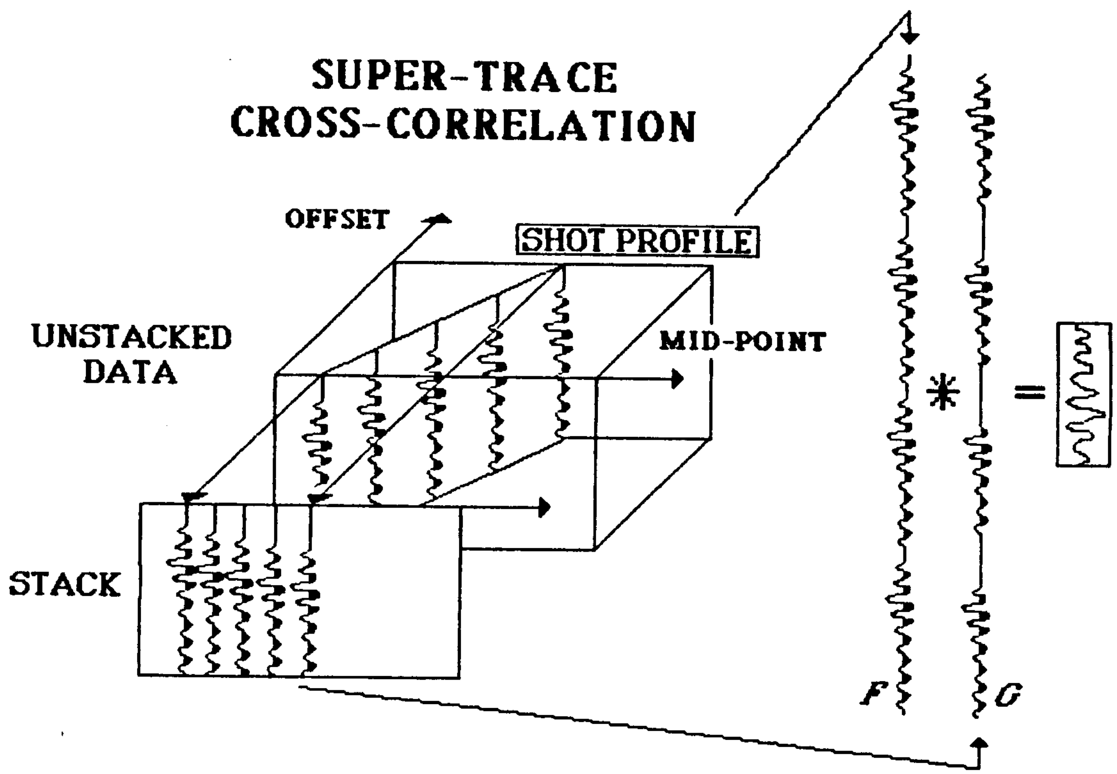

FIG. 1. Super trace crosscorrelation: The plane containing the shot profile in the unstacked data volume is moving up or down according to the static shift of the shot. The CDP stack is changing as a function of that static shift. Maximizing the power of the CDP stack as a function of that particular shot static is equivalent to maximizing the crosscorrelation between two super traces built from the shot profile (trace F) and the relevant part of the CDP stack (trace G). The procedure is repeated for every shot and geophone. Convergence is usually achieved within 5 to 20 iterations.

Repeat {
    For every shot and every geophone {
        (1) Form the super traces
        (2) Cross-correlate them
        (3) Pick the maximum
        (4) Correct the stack
    }
    (Optional) Constrain the statics
} until convergence

The optional constraining routine removes undesired features from the model: subtracting a regression line to remove a linear trend, and occasionally a running median to extract glitches. (A running median is a non-linear filter in which the output at each point, $out(x)$, is a median of the input, $in(x')$, over some range $\left| x' - x \right| < \Delta x$).

Modified objective function

Our objective function has been so far the power in the CDP stack. A good way to incorporate constraints is to modify the objective function:

$$\underset{\mathbf{m}}{\mathrm{MAX}} \left\{ \text{Power}\,(\mathbf{m},\mathbf{d}) - F(\mathbf{m}) \right\}$$

where $\mathbf{d}$ is the data, $\mathbf{m}$ is the model. $F(\mathbf{m})$ is a penalty function, which is easily incorporated.

The simplest F is parabolic,

$$F(\mathbf{m}) = \mathbf{m}^T \, C \, \mathbf{m},$$

where C is a scalar. This corresponds to damping in least-squares. F can be designed to softly constrain the shot statics so that they resemble the receiver statics and to reject undesired features of the model.

DATA EXAMPLES

The method was tested on several data sets; two of them, are shown here. The first example includes a slowly varying, near-surface velocity anomaly (Figures 2-5). The reflectors of the stacked section without static correction have gaps at 3.7 seconds, and sags at 0.5, 1.4, 2.0 and 2.6 seconds under midpoint 100, (Figure 2). The gaps and sags

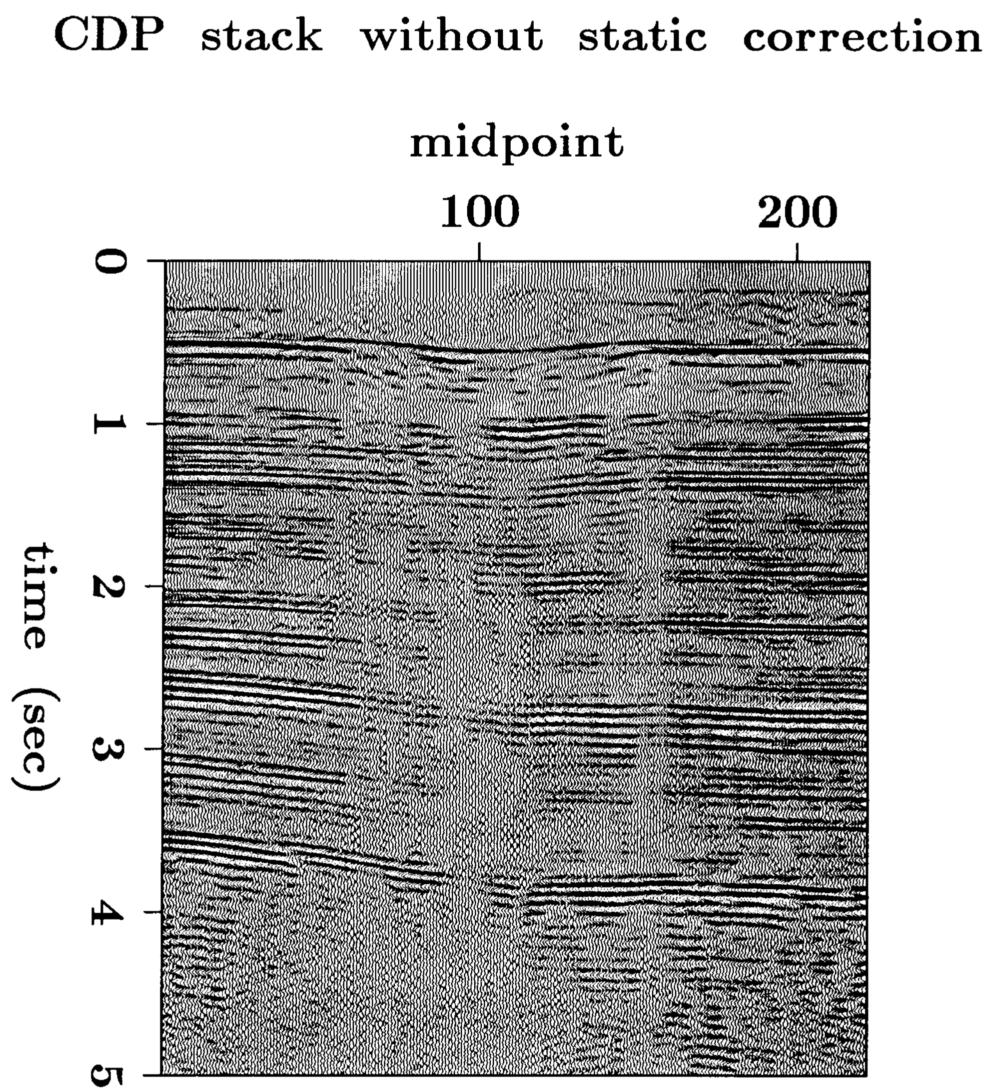

FIG. 2. Example 1: CDP stack without static correction. The sags and gaps in the reflectors are due to a large near-surface, low-velocity anomaly.

are repaired when the static correction are made (Figure 3). The statics program converged to a model that shows a low velocity anomaly in the center (Figure 4). The high-spatial-frequency statics converged within a few iterations, the low-frequency statics took about 20 iterations to converge (Figure 5). The faster convergence of the high frequencies is similar to what Wiggins et al. (1976) found using a travel-time-picking method. In this example, the long-wavelength statics could be independently estimated from velocity analysis (Toldi, 1984) with excellent agreement.

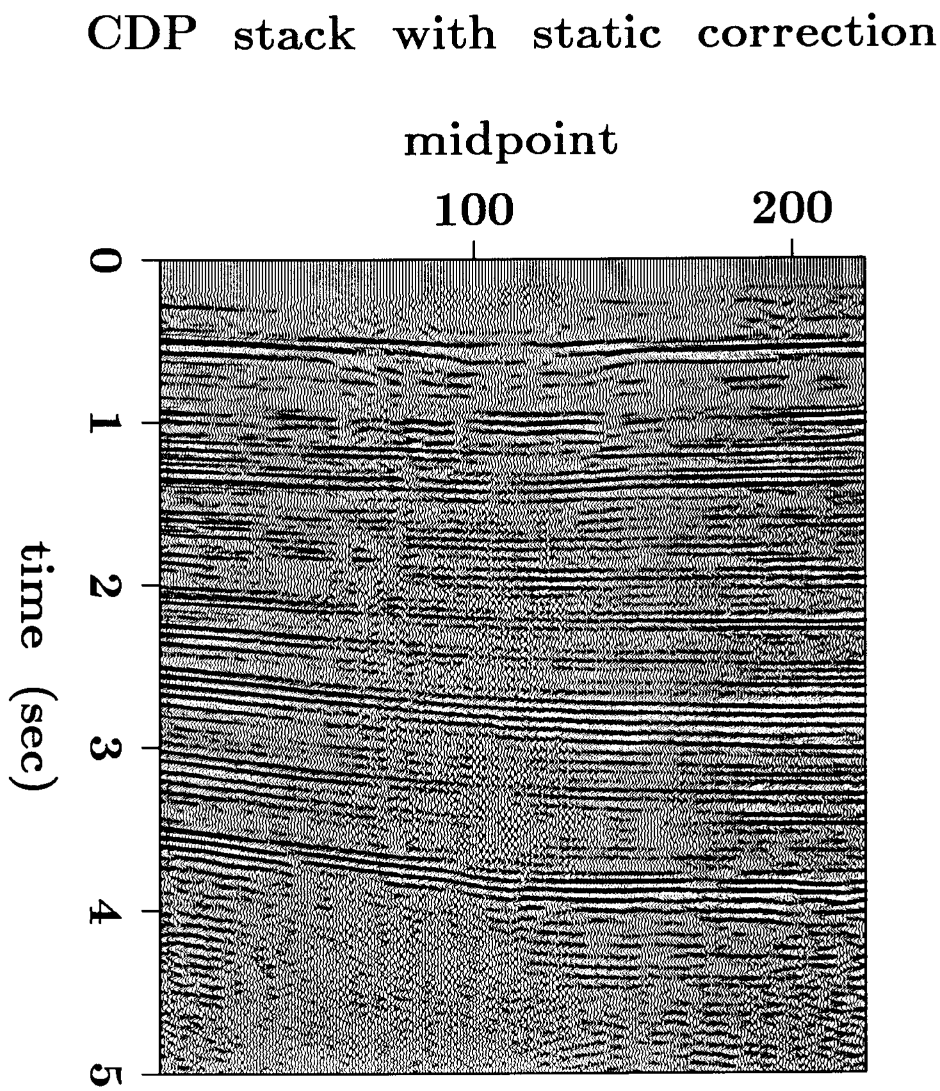

FIG. 3. Example 1: CDP stack with static correction. The statics were estimated on the reflector at 3.5 to 4 seconds.

The data of the second example have a poor signal-to-noise ratio: lightning bolts and air waves are much stronger than the reflections. The noise and near surface velocity anomalies are the reasons why the uncorrected stack is uninterpretable (Figure 6). The static correction improved the upper 2 seconds of the left part of the CDP stack (Figures 7-9). The time window for the statics estimation was 0.8 to 1.4 seconds. Runs with deeper time windows were unsuccessful, possibly because of low penetration of the reflector that produced the multiples on the left side of Figure 7.

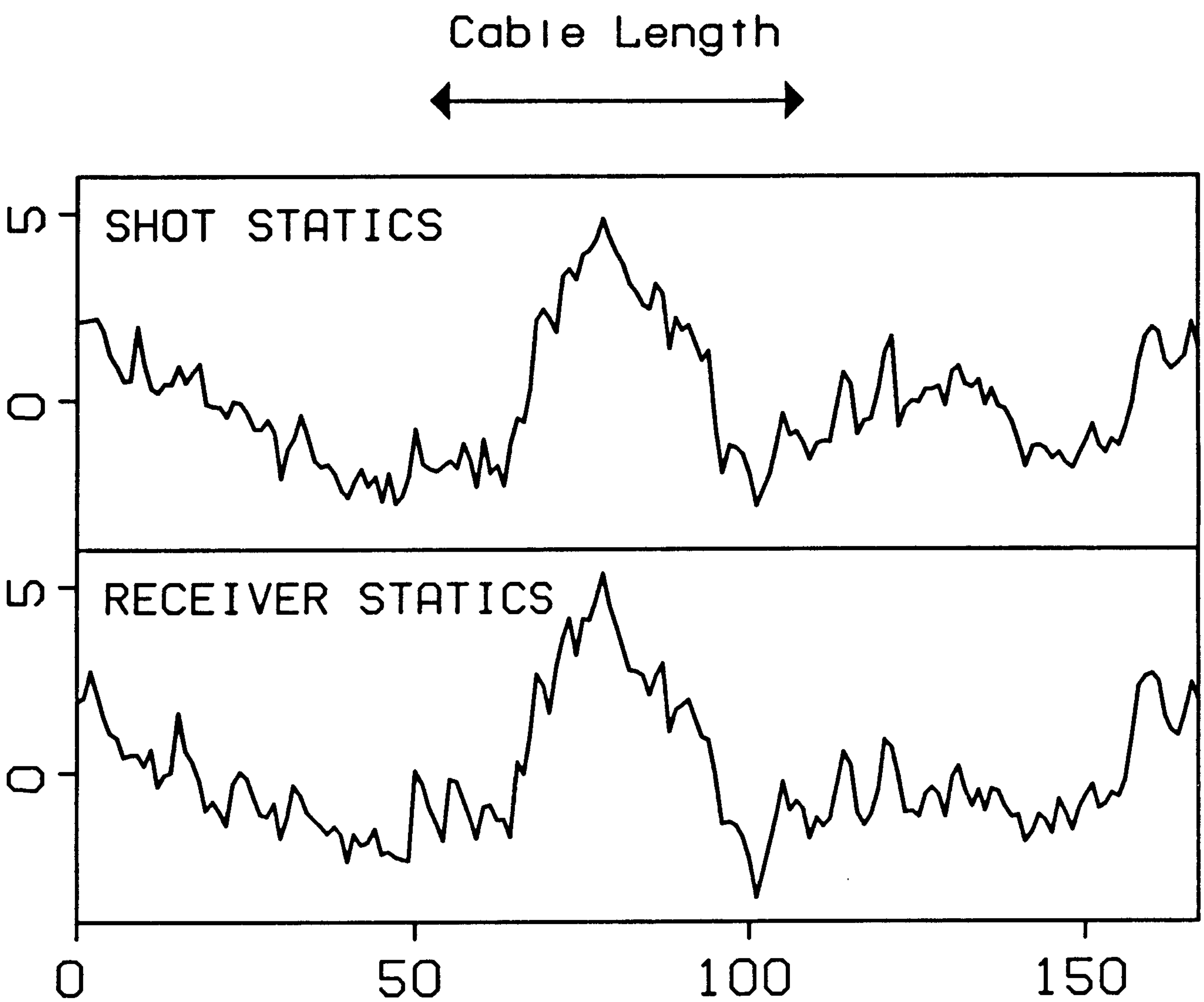

FIG. 4. Example 1: Shot and receiver statics. The similarity is expected in Vibroseis * data.

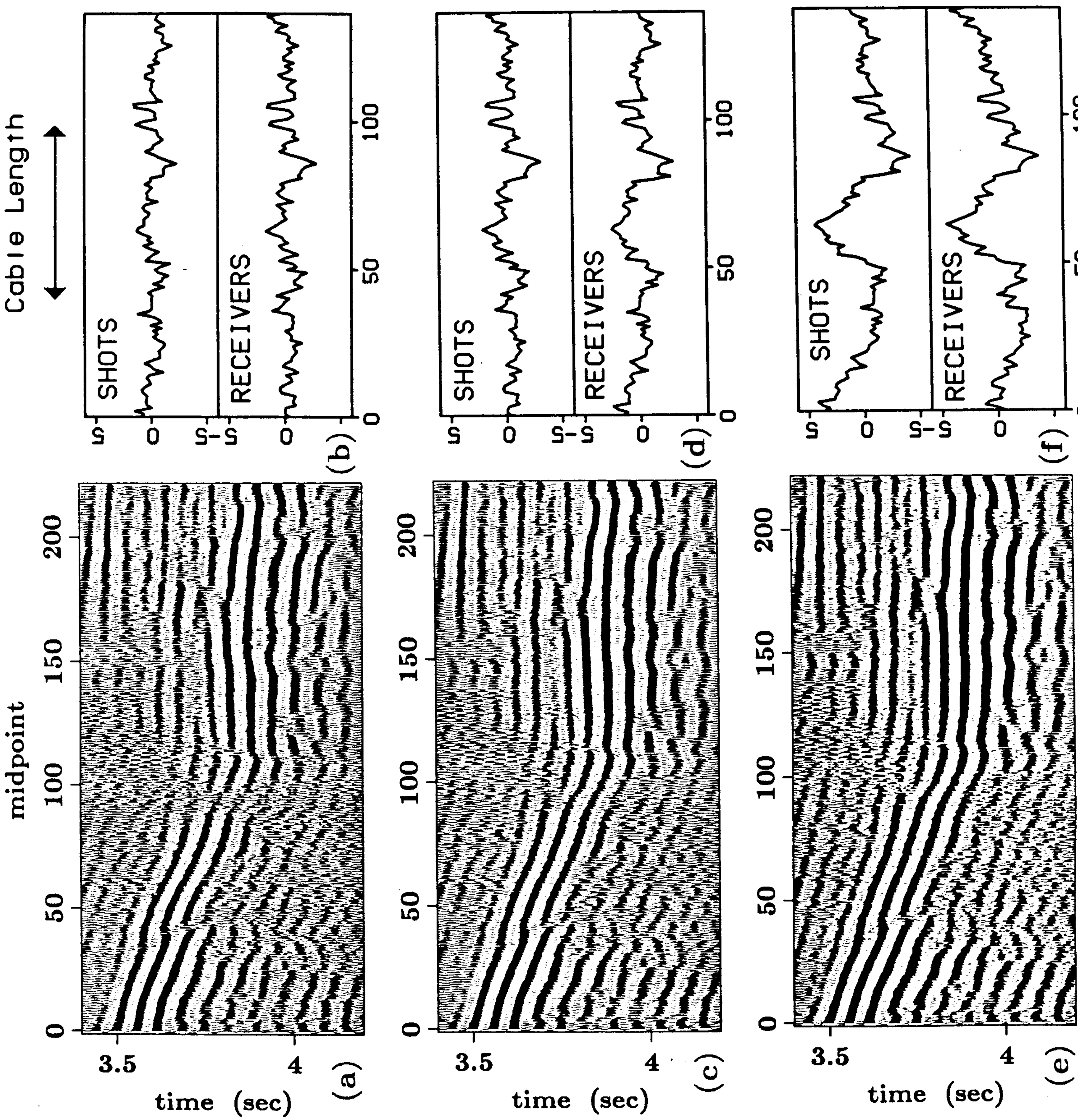

FIG. 5. Example 1: The fast convergence of short-wavelength statics. (a) Stack after one iteration. (b) Statics after one iteration: the short-wavelength statics already converged. (c) and (d) After three iterations. (e) and (f) After twenty-two iterations.

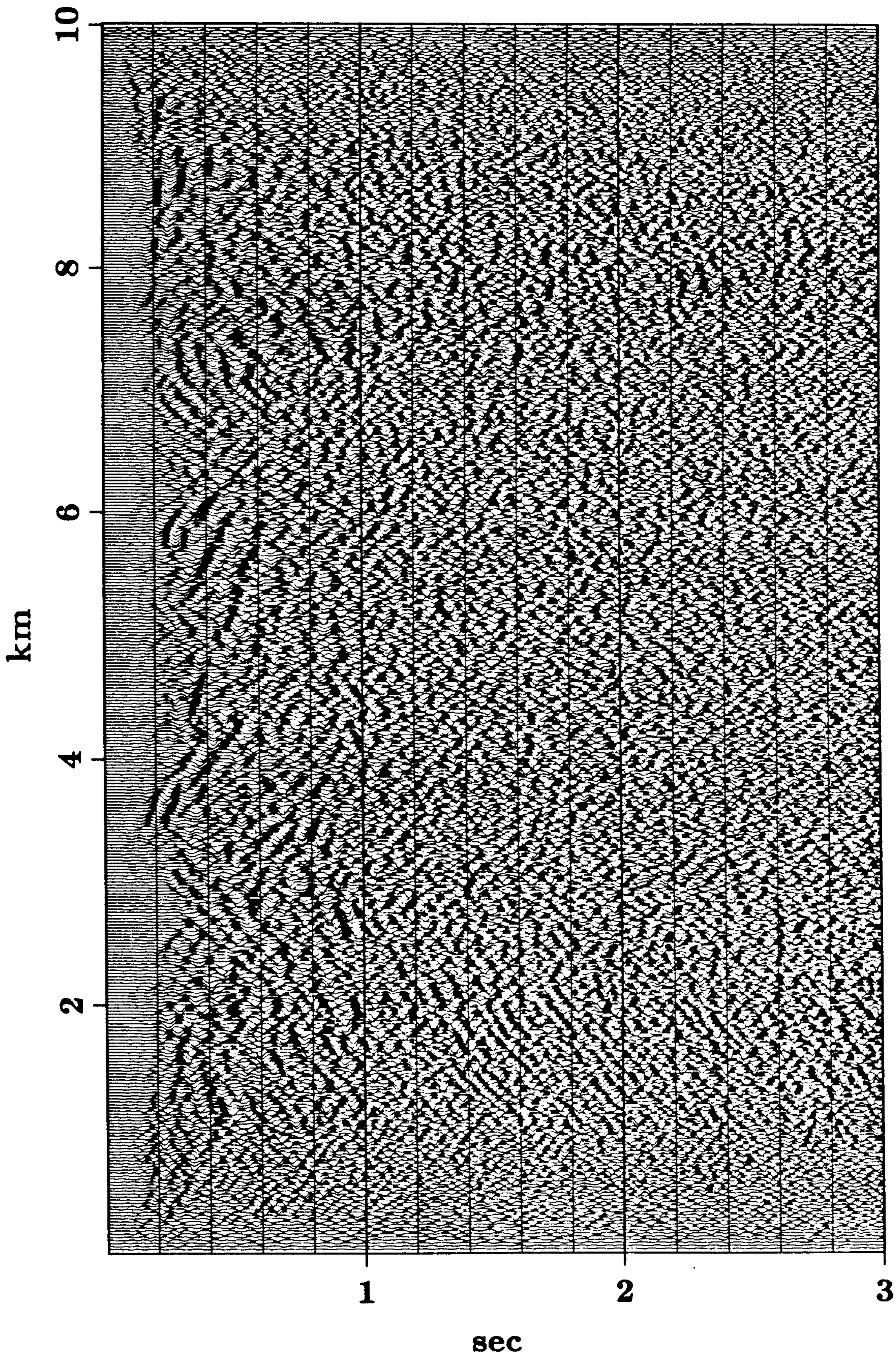

FIG. 6. Example 2: CDP stack without static correction.

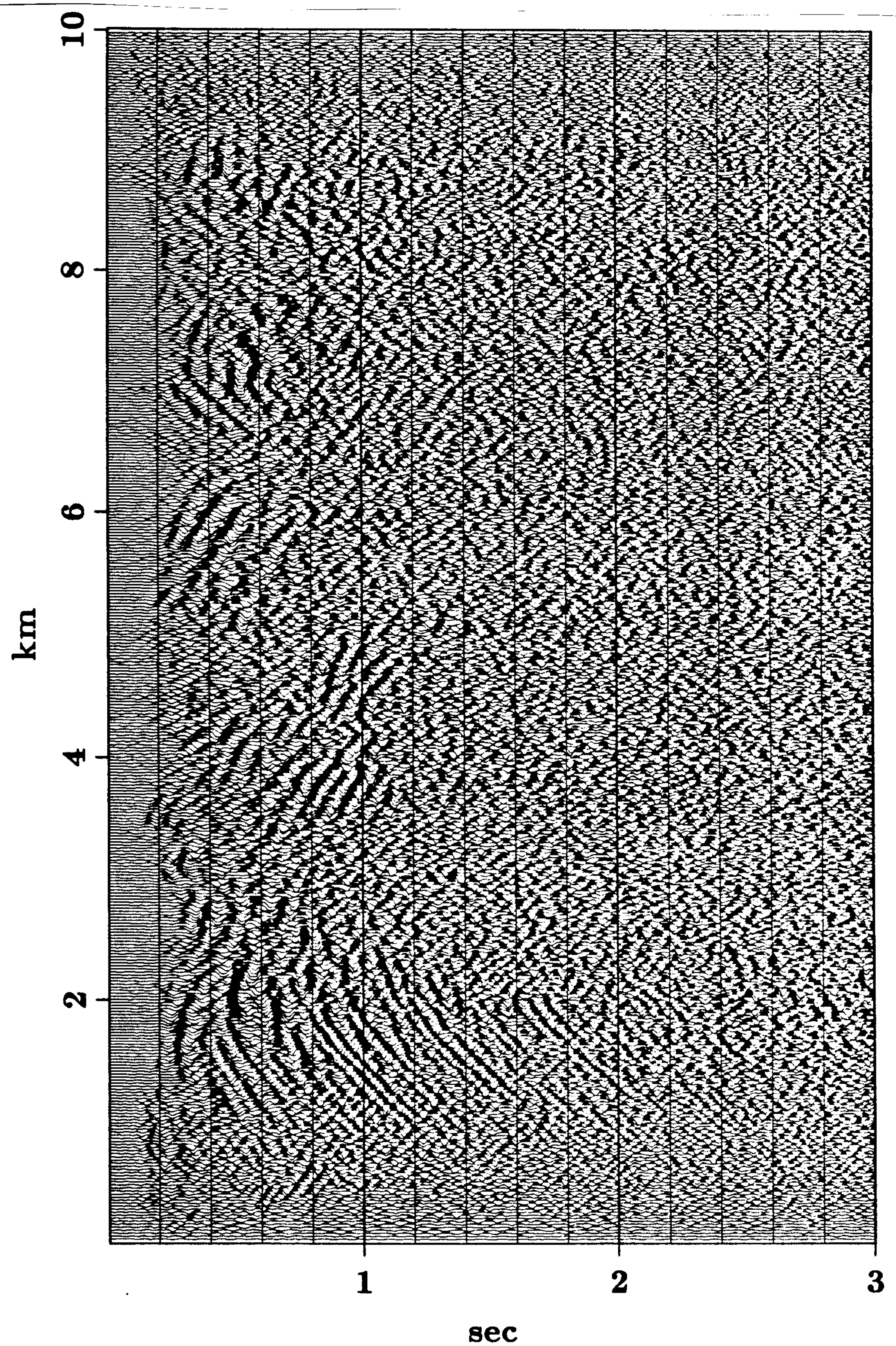

FIG. 7. Example 2: CDP stack with static correction.

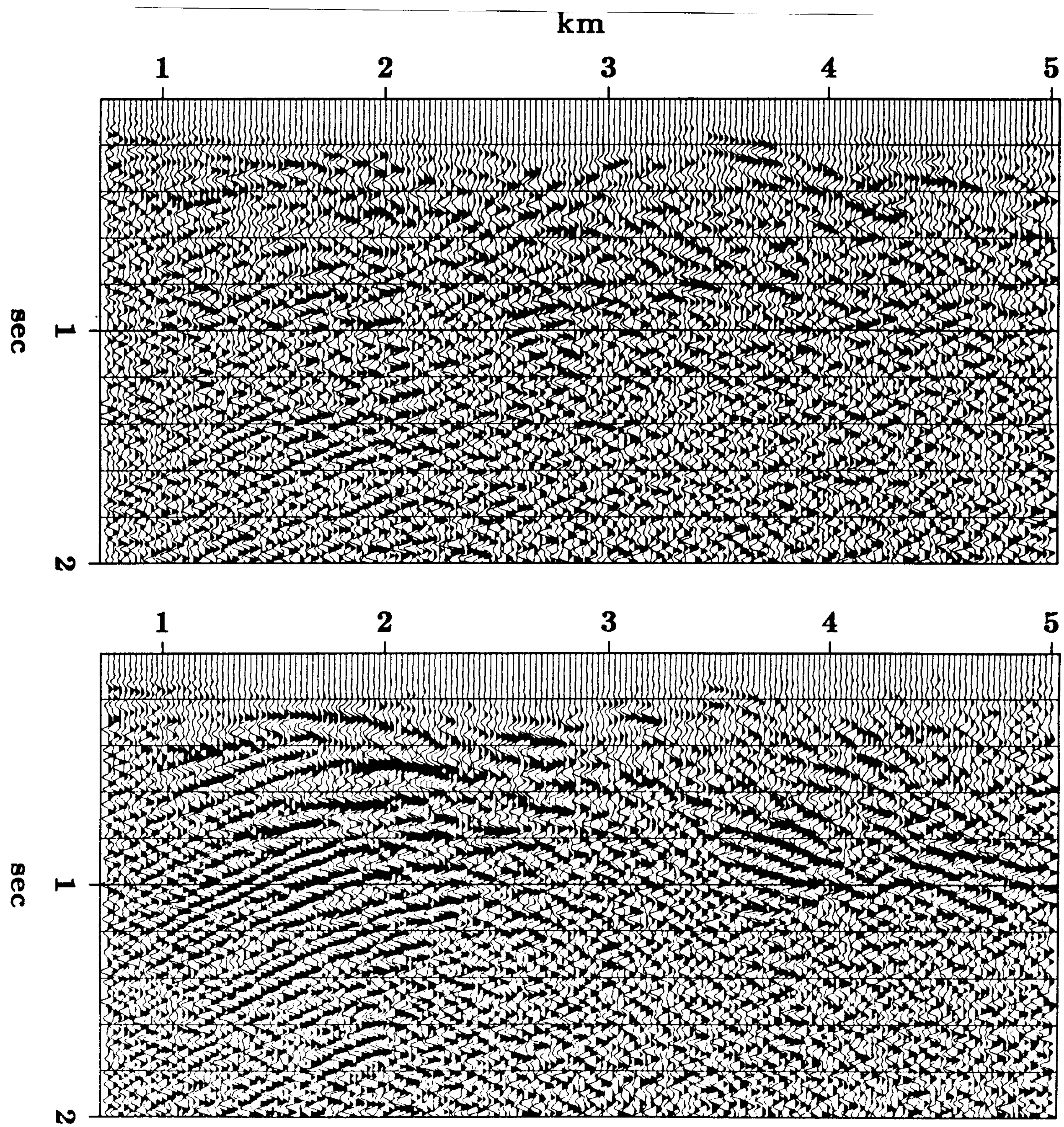

FIG. 8. Example 2: Stack without (above) and with (below) statics correction. Upper 2 seconds of left part of Figures 6 and 7.

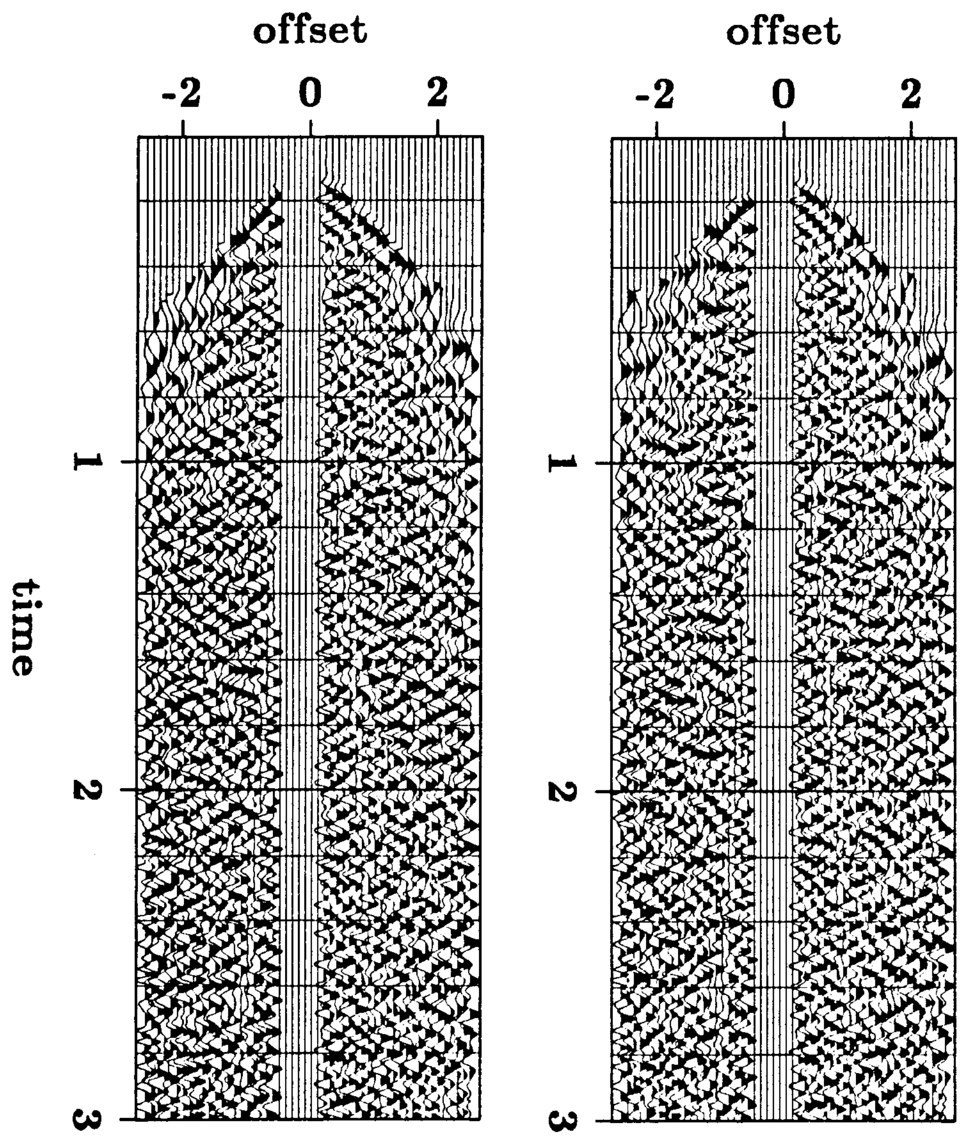

FIG. 9. Example 2: CDP gather before (left) and after (right) static correction.

DISCUSSION

Comparison to travel-time-picking methods

Travel-time-picking methods usually generate pilot traces by shifting and stacking neighboring traces. Single traces are then crosscorrelated to the pilot traces. Crosscorrelating single traces with pilot traces is more reliable than crosscorrelating single traces with single traces because of the improvement in signal-to-noise ratio in the pilot traces. Our method carefully generates the CDP stack which is the final product, but is also used as a pool of pilot traces. The shifts that we use to generate the pilot traces become the static model; we base the statics estimation directly on the unstacked data and not on previously picked travel times. The unreliable picking is eliminated as a separate step.

The cost of the stack-power maximization method may be higher than a travel-time-picking method if many iterations are needed (if the model contains slowly varying anomalies). The cost may be lower if convergence is achieved within few iterations. In any case it is more reliable in the presence of noise. When the signal-to-noise ratio is high, travel-time-picking should be used if it is more efficient.

The limitations of static time shifts

In making the static time shift approximation we assume vertical rays in the near surface, and that the near surface affects only the travel time. In general, the anomalies have some finite depth, and they change wavelets as well as travel times. Generalization of the static model can be done in the directions of velocity analysis and of deconvolution.

Statics and velocity analysis

Statics estimation is effectively a velocity analysis of the near surface. Ideally, statics and velocity analysis would be done together. In practice, however, the near surface anomalies are analyzed by static shifts, and velocities are analyzed by NMO. Statics cannot be estimated before NMO is removed, but sometimes velocity cannot be accurately estimated before statics are corrected. Some work has to be done, to either combine the two processes or at least make the static estimation less sensitive to inaccuracies in NMO velocities. Our method is sensitive to errors in NMO velocities, but it is fast enough to be used iteratively with velocity analysis: run the statics for some possible NMO velocities and choose the velocity that gives the best result, analyze the velocity again and if necessary, estimate the statics again, with the improved velocity model.

Velocity analysis by NMO cannot resolve features much smaller than the cable length. Statics estimation is unreliable for anomalies much larger than the cable length. There is an area of overlap: anomalies that can be detected by either statics or velocity analysis. The data set of Figures 2-5 has one of these overlap cases.

The static time-shifts approximation is more adequate for deep reflectors than for the shallow ones. Dynamic correction (time- and offset-variable, or downward continuation) is necessary for the shallow reflectors. Also, the sensitivity to pre-statics velocity analysis is smaller for deep reflectors. The correction can be *estimated* statically on a deep reflector and *applied* dynamically to the whole section.

Statics and deconvolution

Static shifts are phase corrections that should, in principle, be part of deconvolution. Our method does not combine static shifts with deconvolution, but it is possible to combine the two in an iterative processing sequence, similar to the sequence of statics-velocity analysis. With this iterative sequence, some cases will require the statics program be run many times. It substantiates the importance of the method being fast. Restartability is also important because results of previous runs can then be used as starting points.

Deconvolution before statics might be helpful in reducing ringing and easing the cycle skips problem but it should be done with a surface-consistent scheme for the following reason. Most deconvolution techniques assume a minimum-phase wavelet, but they work anyway even for Vibroseis* which does not have a minimum-phase wavelet (Gibson and Larner, 1984). The minimum phase of the deconvolution operator is subtracted from the non-minimum phase of the wavelet; the residual phase should be corrected separately. If the deconvolution is not surface consistent, we cannot correct that residual phase by surface-consistent statics correction (or by any other surface-consistent phase correction).

Phase deconvolution without phase unwrapping

The estimation and correction of static shifts by stack optimization can be generalized to phase deconvolution (Muir, pers. comm.): the goal is to find the surface-consistent phase that will maximize the stack power. Solving for the phase function $\Phi(\omega)$ in all frequencies is frustrated by the overwhelming number of parameters: the

* Conoco Inc.

number of frequencies times the number of shots and receivers. Instead, the phase can be decomposed to components by the series:

$$\Phi(\omega,p) = \sum_n \Phi_n(\omega,p) = \sum_n \phi_n(p)\ sgn\ \omega \left|\omega\right|^n . \tag{3}$$

ω is the frequency and p is a surface consistent variable, i.e. shot station. $sgn\ \omega$ is the sign function:

$$sgn\ \omega \equiv \begin{cases} 1 & \omega > 0 \\ 0 & \omega = 0 \\ -1 & \omega < 0 \end{cases}$$

$\Phi(\omega)$ is an odd function because the operator $e^{i\Phi(\omega)}$ is Hermitian.

We now want to find the phase coefficients $\phi_0(p)$, $\phi_1(p)$ and maybe $\phi_2(p)$ that will maximize the stack power. It can be done similarly to the estimation of the statics by stack power maximization: Phase unwrapping is not required.

The most important component of the phase correction in equation (3) is the static shift, $\phi_1(p)\omega = \Delta t\, \omega$. The $n = 0$ term is the bulk-phase correction:

$$e^{i\ \phi_0\ sgn\ \omega} = \cos\phi_0 + i\ \sin\phi_0\ sgn\ \omega\ ,$$

which inverse-Fourier transforms to

$$\cos\phi_0 + i\ \sin\phi_0\ sgn\ \omega\ \subset\ \cos\phi_0\delta(t) + \sin\phi_0\ H\ .$$

H is the Hilbert transform operator. The other terms of the series of equation (2) are dispersive and more complicated than the first two. The $n = 2$ term,

$$F_2(\omega) = e^{i\ \omega^2\phi_2\ sgn\ \omega}\ ,$$

Fourier-transforms to

$$f_2(t) = \frac{\cos\alpha^2}{\sqrt{\phi_2}}\left[\sqrt{\frac{\pi}{2}} - 2\int_0^\alpha \cos x^2 dx\right] + \frac{\sin\alpha^2}{\sqrt{\phi_2}}\left[\sqrt{\frac{\pi}{2}} - 2\int_0^\alpha \sin x^2 dx\right]\ ,$$

where $\alpha = t/(2\sqrt{\phi_2})$. Impulse responses of the first three terms are shown in Figure (10).

Applying phase deconvolution by stack optimization, in the one case we tried, gave disappointing results: the power of the stack was improved compared to a stack produced with only static correction, but the lateral continuity of the stack degraded and the near surface model did not make any sense (Figures 11-12).

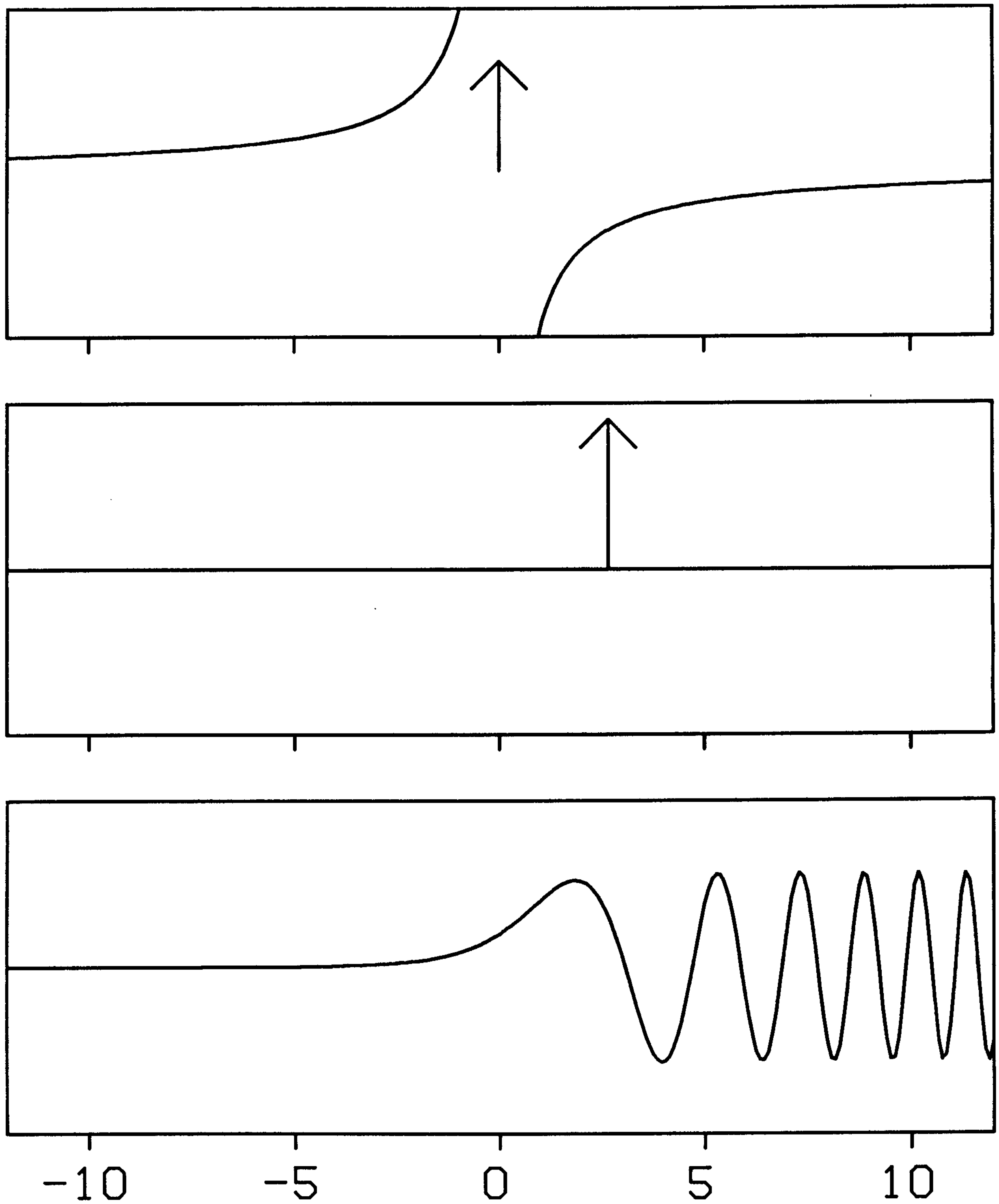

FIG. 10. Impulse responses of the first three phase terms: (a) The bulk phase $\Phi(\omega)=sgn\ \omega\ \phi_0$ with $\phi_0=\pi/4$ (b) The static shift $\Phi(\omega)=\omega\Delta t$ with $\Delta t=2.5$ seconds. (c) The square phase $\Phi(\omega)=sgn\ \omega\ \phi_2\omega^2$ with $\phi_2=1$ seconds^2.

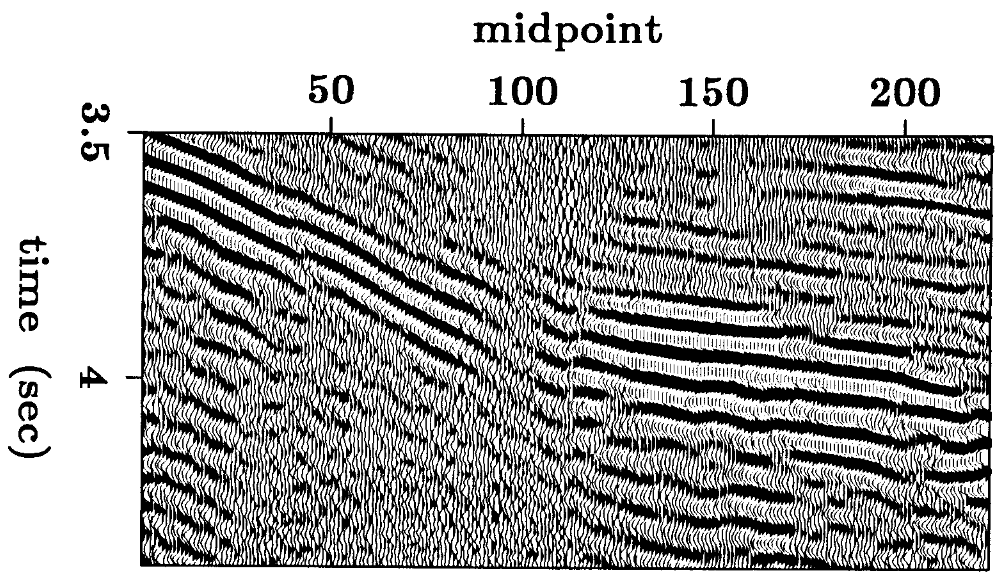

FIG. 11. Stack after unconstrained phase deconvolution. Comparing to static correction (Figure 5), the power is higher but the lateral continuity is lower.

Power maximization and optimization

Our failure to make the phase correction by stack-power maximization shows that stack power does not equal stack quality. We chose to maximize stack power in this initial study because a simple sum of squares is readily understood and economically computed. Other objective functions might be more sensitive to lateral continuity and less sensitive to NMO velocity than stack power. Energy can be maximized in a filtered stack: the stack can be filtered over midpoint (to increase the sensitivity to lateral continuity) or the data can be filtered or dip-filtered over offset (to decrease the sensitivity to NMO).

Null space

The static model that maximizes the stack power cannot be unique: we can always add a constant to all the shot statics and subtract that constant from the geophone statics and no time shifts will result. This is an example of a solution that is in the null space. Analysis of the null space for travel-time picking methods was given by Taner et al. (1974) and by Wiggins et al. (1976).

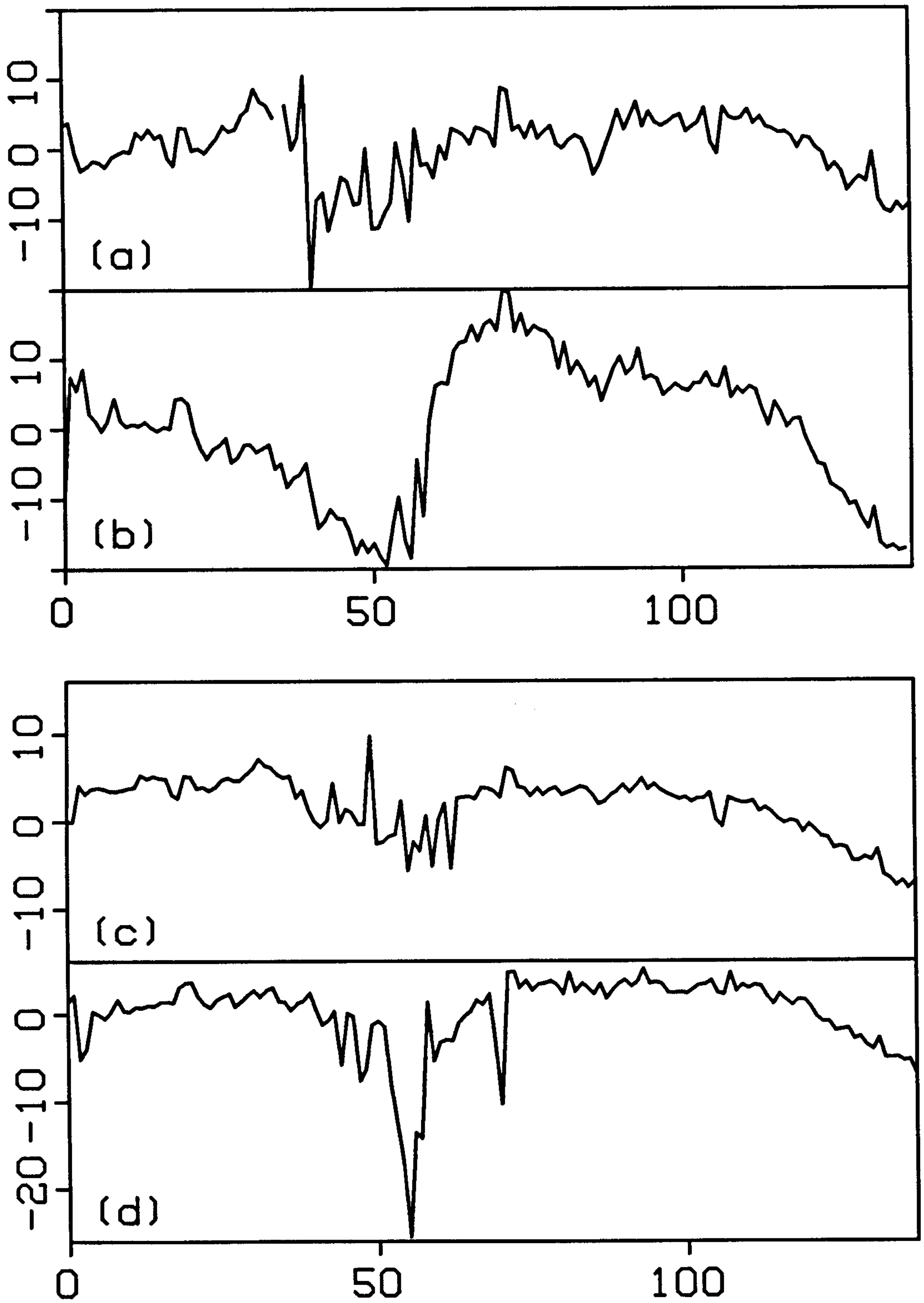

FIG. 12. The phase corrections: (a) Shot static shifts. (b) Receiver static shifts. (c) Shot bulk phase. (d) Receiver bulk phase.

In the context of stack-power optimization, the null space is defined as any component of the model that does not change the power of the stack. The null space contains, in addition to the models that do not cause any time shifts, all models that cause shifts of CDP gathers but do not cause relative shifts within CDP gathers. For example the model

$$S(x) = G(x) = a \times x + b ,$$

will cause the shifts

$$\begin{aligned} \Delta t_i &= S(s_i) + G(g_i) \\ &= a \times s_i + b + a \times g_i + b \\ &= a \times (s_i + g_i) + 2b \\ &= a \times 2y_i + 2b \ . \end{aligned}$$

The time shifts are uniform within every CDP gather (because $y_i = (s_i + g_i)/2$); therefore they cannot change the power of the stack, although they may be important in interpretation.

Slowly varying corrections, which change little within a cable's length, hardly affect the stack power. The power-maximization method, therefore, has the usual problem with the long-wavelength corrections, they remain to be determined by velocity analysis.

Local extrema

The direct ascent maximization converges to a local maximum which depends on the starting point. If the starting point is far from the global maximum, the process may end in a local maximum. Dan Rothman (1984) developed a method of stack-power maximization by stochastic relaxation that takes more computation time but has higher probability of finding the global maximum. An alternative approach is to compose a model from results of previous iterations that produced local maxima. It is also possible to run the program from randomly chosen starting points.

CONCLUSIONS

We suggested a new approach to the statics problem: estimating the statics by maximizing the power in the CDP stack. The method we presented is capable of estimating statics in the presence of a low signal-to-noise ratio by simultaneous model fitting and travel-time picking.

ACKNOWLEDGMENTS

The authors would like to thank Bill Ostrander from Chevron U.S.A., and Ron Chambers from Western Geophysical for releasing data that were used. Dan Rothman who worked with us in the early stages of this project, John Toldi, Rob Schreiber, and Francis Muir made helpful suggestions. This work was supported by the sponsors of the Stanford Exploration Project.

REFERENCES

Burg, J.P., 1972, The relationship between maximum entropy and maximum likelihood spectra: Geophysics, v. 37, p. 375-376.

Chi C., Mendel, J.M., and Hampson, D., 1983, A computationally fast approach to maximum-likelihood deconvolution: Geophysics, v. 49, p. 550-565.

Disher, A.D., and Naquin, P.J., 1969, Statistical automatic statics analysis: Geophysics, v. 35, p. 574-585.

Foster, M.R., and Guinzy, N.J., 1967, The coefficient of coherence: its estimation and use in geophysical data processing: Geophysics, v. 32, p. 602-616.

Gibson, B. and Larner, K., 1982, Predictive deconvolution and the zero phase source: Geophysics, v. 49, p. 379-397.

Hileman, J.A., Embree, P., and Pflueger, J.C, 1968, Automated static corrections: Geophysical Prospecting v. 16 p. 326-358.

Kirchheimer, F., 1983, Long period static analysis by trigonometric approximation: Presented at the 53rd Annual International SEG Meeting, Las Vegas, September 1983.

Koehler, F., and Taner, M.T., 1967, Velocity spectra - Digital computer derivation of velocity function: 37th SEG Meeting, Oklahoma City, Oklahoma.

Neidell N.S., and Taner, T.M., 1971, Semblance and other coherency measures for multichannel data: Geophysics, v. 36, p. 482-497.

Rothman, D., 1984, Non linear inversion by stochastic relaxation with application to residual statics: Presented at the 54th Annual International SEG Meeting, Atlanta, Georgia, 5 December 1984.

Taner, M.T., Koehler, F., and Alhilali K.A., 1974, Estimation and correction of near surface time anomalies: Geophysics, v. 39, p. 441-463.

Toldi, J., 1985, Velocity analysis without picking: Ph.D. thesis in progress, Stanford University.

Toldi, J., 1984, Estimation of a near surface velocity anomaly from stacking velocities: Stanford Exploration Project report, v. 41 p. 99-120. Presented at the 46th EAEG Meeting, London, England, June 1984.

Thorson, J.R., 1984, Velocity stack and slant stack inversion methods: Ph.D. thesis, Stanford University.

Wiggins, R.A., Larner, K.L. and Wisecup, R.D., 1976, Residual statics as a general linear inverse problem: Geophysics, v. 41, p. 922-938.

Wiggins, R.A., 1978, Minimum entropy deconvolution: Geoexploration, v. 16, p. 21-35.