Estimating residual statics by optimizing a complexity-reduced stacking-power function

H. Lydia Deng  
*Center for Wave Phenomena and Department of Mathematical & Computer Sciences*  
*Colorado School of Mines*  
*Golden, Colorado 80401, USA*

Bin Wang & Keh Pann  
*Mobil Exploration & Producing Technical Center,*  
*Dallas, Texas 75247, USA*

**ABSTRACT**

Near-surface weathering layers introduce time anomalies in seismic data, which cause poor quality in stacked seismic sections. The ultimate goal of statics estimation in seismic data processing is to obtain high-quality stacked seismic sections. One measure of the stacking quality is stacking-power function. Therefore, statics estimation is a mathematical optimization problem where we look for the time-shifts among traces such that the stacking power is maximized.

We begin by reviewing both linear and non-linear approaches of conventional residual statics estimation under the “umbrella” of this optimization. The time-picking approach is an oversimplified linearization of the stacking-power function where only one travel time in a correlation is used, while the stacking-power approach uses all trace information in correlations for complicated non-linear optimizations.

We then propose using partial information from the correlations, envelope functions, to formulate a new objective function. We conjecture that this new objective function approximates the global structure of its corresponding stacking-power function, and that the number of local maxima is significantly reduced. Statics are estimated by optimizing this simplified objective function using local-search algorithms.

This new algorithm is tested on two field data sets where routine processing failed to give good answers. It shows that this approach is particularly appropriate for data where large statics are involved.

**Key words:** Residual Statics, Objective Function, Stacking Power, Optimization, Envelope

**Introduction**

Land seismic data are generally contaminated by lateral variation of near-surface irregularities, which cause misalignment of reflection events prior stacking. This misalignment of events can usually be corrected by estimating surface-consistent static-shifts for each source and receiver location. The ultimate goal of statics correction is to obtain high-quality stacked seismic sections. One measure of the stacking quality is the sum of squares (power) of the stacked sections. Therefore, residual statics estimation can be formulated as a mathematical optimization problem where we look for the time-shifts among traces such that the stacking-power function is maximized.

Conventionally, residual statics estimation by solving a linear system (Tanner et al., 1974; Wiggins et al., 1976;
Lamer et al., 1979) has been widely used by industry due to its computational simplicity. However, the only information used from data in this approach is one time-shift in each correlation. When data have poor signal-to-noise ratio, this time-picking may be difficult due to the presence of ambiguities. Choosing wrong peaks of correlations causes the alignment of wrong events in the seismic section, which is known as “cycle skips”. This problem is especially serious where large statics (say, more than half a dominate period) exist in data. We recognize that the conventional linear-system can be considered as a result of linearization from the general algorithm of maximizing the stacking power. The condition of this approximation is that all seismic traces are identical except for small statics shifts.

Ronen and Claerbout (1985) proposed using stacking-power function directly as a measure for estimations. Using the stacking power as an objective function, seismic trace information is used rather than only one picked time from each correlations. A commonly acknowledged problem of the stacking-power criterion is that the objective function to be optimized is a high-dimensional, complex surface with many local maxima. Global search algorithms, such as simulated annealing (SA) (Rothman, 1985; Rothman, 1986) and genetic algorithms (GA) (Stork & Kusuma, 1992; Smith et al., 1992; Whitley et al., 1995b), are needed in the optimization of stacking-power functions. However, these global search algorithms tend to be computationally expensive; there is no guarantee that these global searches would converge to global maxima within a finite computation time. Unsuccessful global search may find local maxima, which causes “cycle skips” on the static corrected section.

Efforts have been made to simplify objective functions in order to avoid local optima in global searches for such difficult optimization problems. A multi-grid scheme was suggested to sub-sampled, low-frequency versions of data in seismic waveform inversion (Saleck et al., 1993; Chen, 1994; Bunks et al., 1995). Multi-resolution analysis (MRA) (Deng, 1995) using a shift-invariant wavelet base (Saito & Beylkin, 1993) has been also applied to seismic data for the purpose of simplifying objective functions. Although some successes have been reported, it has not been carefully studied that whether the objective functions were so severely distorted that they lose their qualification as a measure of stacking quality. Furthermore, Elston (1992) pointed out that band limiting data cause values of local optima in stacking-power function approaching that of global optimum, therefore, ambiguity in optimization could be increased when searching on such a high-dimensional surface. Refer to Deng & Scales (1996) for a study of such issues.

In this paper, we propose estimating residual statics by optimizing a simplified objective function, the reduced stacking-power. This new objective function approximates global feature of its corresponding stacking-power function by using a smooth approximation for each correlation functions.

This residual statics estimation technique is applied on two field data sets, which have serious statics problems. After the estimated statics are corrected, results of stacking show better lateral coherence than the output from ProMax.

### Complexity of Residual Statics Estimation

#### Stacking power as an objective function

Generally speaking, the ultimate goal for statics corrections is to improve stacked images; it is reasonable to use a measure of stack quality as an objective function in residual statics estimations. One measure of stack quality is the sum of squares of the stacked traces – stacking power (Neidell & Taner, 1971; Ronen & Claerbout, 1985; Rothman, 1985). A stacking power of N seismic traces may be written as

\[
\sum_{t} \left( \sum_{i=0}^{N-1} d_i(t + \tau_i) \right)^2,
\]

where \(d_i(t)\) is the \(i\)th trace and each component of vector \(\vec{d}\) is the time-shift for a trace. It can be easily shown that maximizing the function in equation (1) is equivalent to maximizing a summation of cross-correlation functions,

\[
F(\vec{\tau}) = \sum_{i=0}^{N-1} \sum_{j=i+1}^{N} \Phi_{ij}(\tau_j - \tau_i),
\]

where \(\Phi_{ij}(\tau)\) is the cross-correlation of the \(i\)th and \(j\)th traces evaluated at \(\tau = \tau_j - \tau_i\). Equation (2) is mostly used in implementations because cross-correlations \(\Phi_{ij}(\tau)\) may be pre-computed and stored.

In most practical residual statics problems, each component of \(\vec{d}\) is not completely independent to each other. For example, under the surface-consistent assumption, each trace shift \(\tau_i\) is a linear combination of unknown statics-shifts of the corresponding source and receiver. Furthermore, when the residual normal moveout and subsurface shifts are considered, \(\tau_i\) should be the linear function of all these contributing factors (Taner et al., 1974; Wiggins et al., 1976). In essence, a general residual statics problem is a mathematical optimization problem in which we look for parameter vectors \(\vec{\tau}\) so that

\[
\max_{\vec{\tau}} F(\vec{\tau}) = \sum_{i} \sum_{j \neq i} \Phi_{ij}(\tau_j(\vec{\tau})),
\]

where \(\Phi_{ij}(\tau)\) is the cross-correlation of the \(i\)th and \(j\)th traces evaluated at \(\tau = \tau_j(\vec{\tau})\).
where \( \tau_j(\vec{x}) \) is a linear function of the unknown parameters. The function \( F(\vec{x}) \) in equation (3) is referred to as the stacking-power function.

**Behavior of stacking-power functions**

In practice, stacking-power functions are generally highly multi-modal, high-dimensional surfaces with many local maxima. Smith et al. (1992) shows a two-dimensional projection of one such objective function, which presents an irregular shape with many local maxima.

By observing equation (3), however, we see that stacking-power functions have some special properties:

(i) Each term \( \Phi_{ij} \) in the summation of \( F \) is a one-dimensional function, whose values are pre-computed. The function \( F \) has maxima at points where all correlation functions exhibit minima (see Appendix A).

(ii) \( F \) is separable, i.e. there are no non-linear interactions between variables. Therefore, a change of variable can be found so that the optimal value for each parameter can be determined independently of all other parameters (Whitley et al., 1995a).

(iii) Not all maxima of \( F \) are produced by maxima of one-dimensional function \( \Phi_{ij} \) because the arguments of \( \Phi_{ij}(\tau) \) are not completely independent (see Appendix A). However, for the purpose of statics estimation, we are not interested in those local-maxima that not caused by the alignment of traces.

To demonstrate these observations, we start with a simple, schematic statics problem: consider three identical seismic traces which can be shifted independently of each other, and look for the independent time-shifts \( \{\tau_i, \ i = 0, 1, 2\} \) that can best align the traces. Without loss of generality, we use the first trace as a reference \( \{\tau_0 \equiv 0\} \) and look for time-shifts of other two traces \( \{\tau_1 \text{ and } \tau_2\} \). The stacking power is a two-dimensional function,

\[
F(\tau_1, \tau_2) = \Phi_{01}(\tau_1) + \Phi_{02}(\tau_2) + \Phi_{12}(\tau_2 - \tau_1).
\]

If the three traces are identical, all correlation functions are auto-correlations. The left of Figure 1 shows one example of such correlations. The two-dimensional stacking-power function in this case is shown on the right of Figure 1, which demonstrates that some local maxima exist even in such a simple case. The converged models of local-search algorithms will strongly depend on the initial models, and the estimated statics may easily be at a local maximum if a local search is used. On the other hand, Figure 1 shows that the observable maxima are scattered regularly on the function surface. The global maximum is located at the point where maximum positive peaks of all three correlation functions coincide (the correct time shift). And the observable local maxima are regularly distributed at locations where positive peaks of all three correlation functions coincide. Among these local maxima, those that coincide with the maximum positive peak of one correlation function have significant values.

In Figure 1, the dominate stripes (vertical, horizontal, and diagonal) are located at areas where either of the three correlation functions is at the highest maxima.

Observations from Figure 1 suggest that contributions to significant maxima of stacking-power function are mostly due to positive peaks of the correlations from which the stacking-power function is made of. Small and negative correlation values are of little interest to us.

**Linearization of stacking-power functions**

Equation (3) generally a high-dimensional function in residual statics estimation of seismic data, it can be simplified by an approximation. Let \( T_{ij} \) be the maximum-correlation time between the \( i \)th and \( j \)th traces. When \( \tau_{ij}(\vec{x}) - T_{ij} \) is small, equation (3) can be approximated with the Taylor expansion,

\[
F(\vec{x}) \approx \sum_i \sum_{j \neq i} \left[ \Phi_{ij}(T_{ij}) + \frac{\Phi_{ij}'(T_{ij})(\tau_{ij}(\vec{x}) - T_{ij})}{2} \right],
\]

where \( \tau = T_{ij} \) corresponds the maximum of correlation function \( \Phi_{ij}(\vec{x}) \) with \( \Phi_{ij}'(T_{ij}) = 0 \) and \( \Phi_{ij}''(T_{ij}) < 0 \). With the above linearization, maximization of function \( F(\vec{x}) \) is approximated by a least-squared problem,

\[
\min_{\vec{x}} \sum_{i=0}^{N-1} \sum_{j=i+1}^{N-1} [ \Phi_{ij}'(T_{ij}) | (\tau_{ij}(\vec{x}) - T_{ij}) |^2.]
\]

If all traces are identical except for static-shifts, correlation function \( \Phi_{ij}(\tau) \) is a shifted version of an auto-correlation \( \Phi(\tau) \), with the highest maximum at \( T_{ij} \), i.e. \( \Phi_{ij}(\tau) = \Phi(\tau - T_{ij}) \). Therefore, the above least-squared problem becomes,
\[
\min_{\vec{x}} \sum_i \sum_{j \neq i} (\tau_{ij}(\vec{x}) - T_{ij})^2. \tag{5}
\]

Let \( A \) be the linear operator representing the linear combination \( \tau_{ij}(\vec{x}) \), equation (6) is equivalent to solving the following linear system,

\[
A \vec{x} = \vec{T}, \tag{6}
\]

where \( \vec{T} \) is a vector whose components are the maximum correlation time \( T_{ij} \). Equation (6) is a general formulation of the traditional linear residual statics estimation. When the surface consistent are assumed as well as the residual normal moveout and the subsurface structural variations, equation (6) is the tradition linear system (Tanner et al., 1974; Wiggins et al., 1976).

According to the observations from the previous section, the stacking-power function is separable. Optimization of this high-dimensional function can be reduced to a series of one-dimensional optimization. The travelt ime picking in the conventional statics estimation is, in effect, the procedure of these one-dimensional optimization. For the three-trace example, this approach is equivalent to isolating the three dominant stripes (horizontal, vertical and diagonal) in Figure 1, and looking for the intersection of them, which is at the origin for this example.

This approach is the most commonly used in industry for residual statics estimation due to its computational efficiency. However, this is an over-simplified approach. The only information used from the seismic data is the time corresponding to the maximum peak of correlations. Results of statics estimation rely entirely on the time-pickings \( T_{ij} \). When seismic data are contaminated by severe noise and large statics, time-picking becomes difficult and ambiguous. If wrong peaks are picked from the correlation function, cycle skips would occur and seismic events would be misaligned.

For problems with severe noise and large statics, Ronen and Claerbout (1985) proposed using equation (3) as the objective function for the optimization. However, global searches, such as simulated annealing (SA) and genetic algorithms (GAs), are generally needed for searching the global maximum of stacking-power functions (Rothman, 1985; Rothman, 1986; Smith et al., 1992). These global searches require intensive computation, and the behavior is unpredictable.

**Envelope approach of the stack-power function**

Looking for a compromise between using “too little” (linearized approach) and “too much” (global optimization) information from correlations, we seek ways of using “partial” information given by correlation function. Since the complexity of stacking powers is mostly due to the complexity of correlation functions, suppressing high-frequency components of the oscillatory correlations will result in a complexity-reduced stacking-power function. For obtaining smooth approximation of a signal, we can have the following three options:

- applying low-pass filtering to the signal, including using multi-resolution analysis (Deng, 1995). This approximation is to cut off high-frequency components, and preserving low-frequency components;
- replacing the signal with a slowly varying one by using Hilbert transform; this approximation is, in effect, to shift the spectrum to be centered at 0 frequency by removing the carrier frequency of the signal;
- interpolating positive peaks of the signal, formulating a smooth signal by preserving large positive information.

Using high-cut filter could be dangerous when the correlation is narrow-band or frequency band of noise is lower than that of the signal. On the other hand, computing the envelope by Hilbert transform is heavily influenced by large negative values in correlations. As we see from the previous section, these values are of little interest to us in residual statics problems. We choose the last option for achieving a simplified objective function.

Finding all positive peaks of a one-dimensional correlation function, we can obtain a slowly varying function by interpolating these peak points. This function represents global feature of large correlation values; we refer this smooth approximation of a one-dimensional signal as the *envelope*. Using the envelope functions, we define a similar function as the stacking power,

\[
\hat{F}(\vec{T}) = \sum_{i=0}^{N-1} \sum_{j=i+1}^{N-1} \left[ \Phi_{ij}(\tau_j - \tau_i) \right], \tag{7}
\]

where function \( \Phi_{ij}(\tau) \) is the envelope of correlation \( \Phi_{ij}(\tau) \). The function in equation (7) uses “partial” information from each correlation functions; that is this function ignores all information given by correlations except for all positive peaks. Figure 2 shows a correlation function of two identical traces and its envelope, which is obtained by using the cubic-convolution interpolation scheme (Keys, 1981; Keys & Pann, 1993). Envelope functions contain fewer peaks than does correlation function themselves. Therefore, one expect that the summation of these envelope functions, equation (7), contains fewer local maxima than does the stacking-power function, i.e. complexity of this function is reduced from that of the original. We call this function *reduced stacking-power function*.

In order for equation (7) to be a useful objective function, it needs to be a qualified measure for the stack-
Figure 2. A correlation function between two identical seismic traces and its envelope function.

Figure 3. The correlation function, \( \exp[-(x/10)^2] \cos(x) \) and its envelope are shown on the left. The two-dimensional stacking-power function \( F(r_1, r_2) \) formed by this correlation is shown on the right.

Figure 4. The contour plot of the interpolated function from positive maxima of the original stacking-power function \( F(r_1, r_2) \) is shown on the left. And the contour plot of the reduced stacking-power function is shown on the right.

Conjecture 1. Reduced stacking-power functions have their complexity reduced from the stacking-power functions; they contain fewer local maxima than do stacking-power functions.

Conjecture 2. Reduced stacking-power functions maintains the major global/local maxima from those of the corresponding original stacking-power functions.

Conjecture 3. Reduced stacking-power maintains the global feature of its corresponding stacking-power function.

Figure 3 shows a narrow-band correlation function,

\[
\Phi(r) = \exp[-\frac{x^2}{100}] \cos(x),
\]

and its envelope. For simplicity in this test, the envelope function is obtained by linearly interpolating local peaks of the correlation function, both of them are shown on the left of Figure 3. Using equation (8) as the correlation functions for the three-trace alignment problem, the stacking power function \( F(r_1, r_2) \) is shown on the right of Figure 3. Figure 4 shows the corresponding reduced stacking-power function the right. The left of Figure 4 shows the interpolated contours from all positive maxima of the stacking-power function in Figure 3. It can be seen that these two plots have the same global feature.

Surface consistent reduced stacking-power function

For most residual statics problems, it is often assumed that for waves traveling through the near-surface, only those with approximately vertical raypaths with respect the interface can be transmitted to deep subsurface structure, therefore can be recorded in our reflection seismic data. Although not being always true, it is a good enough assumption for many practical problems. Therefore, residual statics are surface consistent, that is each source or receiver contributes the same statics shift to all traces it involves, regardless of the ray paths connected with it. Suppose the ith shot is fired at time \( t = 0 \) and let the data recorded at the jth receiver, after offset-dependent moveout is removed, be denoted by \( d_{ij} \). Let \( f_{ij}(t) \) be the same data but without noise \( n_{ij} \) and without static shifts. Ignoring the lateral variation of the sub-surface structure and residual moveout, the recorded trace can be written as

\[
d_{ij}(t) = f_{ij}(t - s_i - r_j) + n_{ij}(t),
\]
where \( s_i, r_j \) are unknown static-shifts need to be estimated from seismic traces. In equation (9), source and receiver statics generally cannot be recovered by using the conventional linearized approximation unless both statics and noise contamination to the data are not severe.

In solving residual statics problems, seismic data are often sorted into a common-midpoint (CMP) domain; equation (3) is in a form of

\[
F(\mathbf{s}, \mathbf{r}) = \sum_{y} \sum_{h_1 \neq h_2} \Phi_{h_1, h_2}^{y}(\tau(\mathbf{s}, \mathbf{r})),
\]

where \( \mathbf{s}, \mathbf{r} \) are shot and receiver estimated statics vectors, \( y \) denotes the midpoint index, \( t \) is the time index over a specified window, and \( h \) is the offset index, and \( \Phi_{h_1, h_2}^{y}(\tau) \) is the cross-correlation between traces \( d_{y, h_1}(t) \) and \( d_{h_2}(t) \) and \( \tau(\mathbf{s}, \mathbf{r}) \) is a linear function of source- and receiver-statics. When ignoring the residual normal-moveout and time-shift caused by subsurface structure, function \( \tau(\mathbf{s}, \mathbf{r}) \) is determined by the recording geometry,

\[
\tau = s_i(y, h_1) + r_j(y, h_1) - s_i(y, h_2) - r_j(y, h_2),
\]

and \( (y, h) \) and \( (y, h) \) are the source and receiver indices for midpoint \( y \) and offset \( h \), respectively. Therefore, solving a residual statics problem is a nonlinear optimization where we look for \( \mathbf{s}, \mathbf{r} \) that

\[
\max_{\mathbf{s}, \mathbf{r}} F(\mathbf{s}, \mathbf{r}).
\]

Following the same analysis as the previous section, the reduced stacking-power function with the surface-consistent constraints can be defined as,

\[
\tilde{F}(\mathbf{s}, \mathbf{r}) = \sum_{y} \sum_{h_1 \neq h_2} \tilde{\Phi}_{h_1, h_2}^{y}(\tau(\mathbf{s}, \mathbf{r})),
\]

where \( \tilde{\Phi}_{h_1, h_2}^{y}(\tau) \) is the envelope of correlation functions of \( \Phi_{h_1, h_2}^{y}(\tau) \). To keep the smoothness and prevent the overshoot of the interpolated envelope functions, the cubic convolution interpolation scheme originally developed by Keys (1981; 1993) are used throughout the project.

**A new residual-statics estimation algorithm**

Following the analysis and conjectures of the previous section, we design an algorithm for estimating residual statics for NMO corrected, CMP data contaminated by surface-consistent statics as follows,

**Algorithm 1. (Residual Statics using Reduced Stacking Power)**

(i) (Pre-compute and store:) For all cross-correlation \( \Phi(\tau) \), find positive-peaks on all cross-correlation between traces for all CMP. Use cubic convolution (Keys & Pann, 1993) to form envelope of the sequences \( \Phi(\tau) \).

(ii) (Gross estimation of large magnitude statics:) Local optimization on the reduced stacking-power function \( \tilde{F}(\mathbf{s}, \mathbf{r}) \) for the initial model \( \tilde{s}_0, \tilde{r}_0 \), correct data for the estimated statics.

(iii) (Refinement:) After large magnitude statics are corrected in step (ii), results can be refined by optimizing the stacking-power function using local optimization with 0 initial guess, or the conventional linearized statics estimation. Depending upon results from step (ii), this step may not be necessary.

In case of loss of resolution of the reduced stacking-power function, the last step of refinement is important for further improvement of the stacking quality.


**Figure 7.** Contour plots of stack powers as a function of static time shift at the 10th source and the 20th receiver points, with all other statics held to be correct. The left figure shows that of the original stacking-power function, while the right figure shows that of the reduced stacking-power function (using envelopes).

**Figure 8.** The left shows the stacked section with existence of surface-consistent and noise. The right figure shows that stacked section after the source and receiver statics are estimated and corrected. Two figures are plotted with the same scale.

**Figure 9.** Comparison of the added statics and estimated statics. Horizontal axis is the index of unknowns: 0–19 are sources and 20–54 are receivers. The solid-thin line shows the added statics, while the thick-dashed line shows the estimated statics.

Figure 6 and 7 show behavior of the original stacking power function and its reduced form from slices. Similar to Figure 4, the original stacking power appears to have many local maxima, which are scattered regularly on the 2-D slices. On the other hand, complexity of the reduced stacking power function is largely reduced observing from these 2-D slices.

When the above dataset are contaminated by severe noise and large source and receiver statics, the traces are no longer identical. The left of Figure 8 shows the stacked section of this data set contaminated by surface-consistent statics and 50% additive random noise (signal-to-noise ratio is 2). The noise is band-limited Gaussian noise which has the same bandwidth as the original signal. The added statics are generated by a uniformly distributed random number generator between [−100, 100] ms. The added and estimated statics are shown in Figure 9. The estimated statistics in general agree well with the added statics; the observed large error at beginning and end of receiver statics are caused by the edge effect — low folds. The stacked section after estimated residual statics are corrected is shown on the right of Figure 8, where horizontal events show up except those at the beginning and end of CMPs.

**Synthetic Examples**

Figure 5 shows the stacking chart of the synthetic data set used for testing the residual statics estimation algorithm. This data set has 20 shots, 35 distinctive receivers and 320 traces; all traces are generated by copying and shifting of a field trace. Correlations of these traces are, therefore, shifted auto-correlation of the field trace. This auto-correlation function and its envelope were shown in Figure 2. For this problem, the stacking-power function and its reduced form both have dimension of 55. When no statics are added to the data, the global maximum of the objective function is at the origin of the model space (all statics are zero).

Not being able to visualize the 55-dimensional objective function in this problem, we can observe slices where we hold all unknowns being correct except two.

**Alberta foothill data set**

We now test Algorithm 1 on a set of field data from Alberta foothill. This data set is contaminated by large statics which cause the cycle skips of strong events in stacked sections. Figure 10 shows stacked sections, where residual statics are corrected using ProMax. A large cycle skip can be observed at about CMP 420. This data set has an average of 60 folds. We choose 50-CMP gathers from this dataset, which was obtained by 88 sources and 88 receivers. Therefore, this statics
Figure 10. Alberta foothill data: output of statics-corrected stacked sections from ProMax. An obvious cycle skip shows at the circled area.

Figure 11. Left figure shows 50 stacked CMPs from the Alberta foothill data set before residual statics are corrected. The right figure shows the same stacked section after residual statics are corrected using the proposed algorithm.
Figure 12. Paradox basin data: output of statics-corrected stacked sections from ProMax. An obvious cycle skip shows at the circled area.

Figure 13. A stacked section from Paradox data before residual statics are corrected is shown on the left. And the stacked Paradox data after statics are estimated and corrected using the proposed algorithm is shown on the right.
estimation problem is an optimization of 176 dimensions. The stacked section for CMP 401-450 are shown on the left of Figure 11. Discontinuity of events are severe in this section. Especially, events between time 1.5-2.0 s show an obvious cycle skip. The statics-corrected stacked section using the proposed residual statics correction algorithm is shown on the right of Figure 11.

The difference of event-shapes between Figures 10 and 11 is caused by an elevation correction while doing stacking in ProMax. For solving such a 176-dimensional global optimization problem, running time is around 48 minutes on an SGI Indy.

Paradox basin data set

Another data set was tested on the Algorithm 1. Figure 12 shows a ProMax output of statics-corrected stacked section. This output shows a strong cycle skip at around CMP 2110, especially for the event of 0.8 s.

This data set has an average fold of 15, we processed 100 CMPs with 35 sources and 146 receivers. A stacked section containing 100 CMPs before the residual statics are corrected is shown on the left of Figure 13. The residual statics estimation is an optimization problem of 181 dimensions, for which we used a non-linear Conjugate Gradient algorithm. The right figure in 13 shows the stacked section after residual statics are estimated and corrected using the proposed algorithm. The strong event at around 0.8 s are almost continuous across the section and some stratigraphic details can be seen.

However, continuity of the 0.8 s is not satisfactory at around CMP 2110. This indicates the loss of resolution in reduced stacking-power function. The second iteration of optimizing the stacking-power function is needed for refinement. Figure 14 shows the result of this second iteration using non-linear conjugate-gradient search with 0 initial guess. The lateral correlation of the strong event at around 0.8 s has notably improved.

Each iteration of the statics estimation for this data set is about 25 minutes on an SGI Indy.

![Second iteration](image)

**Figure 14.** A stacked section from Paradox data after the second iteration of residual statics correction. The statics are estimated by optimizing the stacking-power function formed from the right of Figure 13 by a non-linear conjugate-gradient search with 0 initial guess. The continuity of events at around 0.8 s has notably improved.

Conclusion & Discussion

Residual statics estimation is an important stage in seismic data processing. As this procedure is designed for improving high quality stacks, the problem of residual statics estimation can be formulated as an optimization problem which looks for time-shifts among traces (or sources and receivers) that gives the maximum value of the sum of squares (power) of the stacked section. We have studied both linear and non-linear conventional approaches for residual statics estimation. We notice that the conventional linear approaches have over-simplified the problem by using one traveltime datum for each correlation of two traces, and conventional non-linear approach had complicated the problem by using all information given by such correlations.

By studying the nature of stacking-power function, we recognize that this function can be simplified by using partial information given by correlations. Because of the special objective in this non-linear optimization problem, only positive peaks are interesting to us. In this paper, we propose using the envelope of correlations to form a reduced stacking-power function as the objective function for optimization. We also observe that the proposed objective function has fewer local maxima than its corresponding original stacking power, and this complexity-reduced function approximates the global structure of the original.

Using the surface-consistent constraint, we have applied the algorithm to a set of synthetic data which is generated by one field recorded trace, but contamina-
ated by large statics (up to $\pm 100\text{ms}$) and 50\% band-
limited Gaussian noise. The frequency band of the noise is
the same as that of the signal. We demonstrate that
Algorithm 1 deals well with data contaminated by such
noise and large statics. We have also applied the new al-
gorithm to two sets of field data, where ProMax residual-
correction output show cycle-skips on stacked sections.
The new algorithm was able to resolve the residual stat-
icss correctly.

The effectiveness of residual statics algorithms usu-
ally depend on the redundancy of data; CMPs with low-
fold result in poor estimated statics. Long-wavelength
components of statics are generally either in the null
space or are poorly constrained by residual-statics al-
gorithms. We do not address above issues.

The basic assumption of the stacking-power criterion
is that offset-dependent moveout has been correctly ad-
justed before the optimization. However, this assumption
may not be valid if the sub-surface structure is complex
or velocity models are poor. In addition, the assumption
of surface consistency may not be appropriate in many
cases. The robustness of the proposed algorithm needs
to be studied when these conditions are not ideal.

There are several alternatives in the implementation
of the Algorithm 1. For saving memory used to hold all
correlation values, we could only store the picked maxima
for all correlations, and they can be interpolated on the fly
while evaluating the reduced stacking-power function or
its derivative. Another approach is to use coarser grids of
the optimization due to smoothness of envelope functions.

Acknowledgment

This project was initially supported by the summer pro-
gram for university students at Mobil Exploration & Pro-
ducing Technical Center during the summer of 1995. We
appreciate Dave Hinkley and Rachel Ho for their gener-
ous help throughout the project. Our thanks also go to
Bob Evans and Bob Peterson for their support and per-
mission for publishing this work. We are also grateful to
owners of the two dataset, Thistle Inc. and Mobil Canada,
for their generous offer of the data and their permission for
release.

The first author also appreciates her adviser, Dr.
John Scales, who has pushed the direction of studying
and simplifying objective functions in optimization prob-
lems. She was partially supported by the sponsors of the
Consortium Project on Seismic Inverse Methods for
Complex Structures at the Center for Wave Phenomena,
Colorado School of Mines and Oak Ridge National Lab-
oratory.

References

Bunks, C., Saleck, F. M., Zaleski, S., & Chavent, G.
1995. Multiscale seismic waveform inversion. Geo-
physics, 60(5), 1437–1473.

Chen, T. 1994. Multilevel Differential Semblance Op-
timization for Waveform Inversion. Tech. rept.
CWP-153. Center for Wave Phenomena, Colorado
School of Mines.

Deng, H. L. 1995. Using multi-resolution analysis to
study the complexity of inverse calculations. Tech.
rept. CWP-183. Project Review, Center for Wave
Phenomena, Colorado School of Mines.

Deng, H. L., & Scales, J. A. 1996. Entropy-Based Com-
plexity of Optimization Problems. Tech. rept.
CWP-203. Center for Wave Phenomena, Colorado
School of Mines.

in geophysical inverse theory. Ph.D. thesis, Prin-
ceton University.

digital image processing. IEEE Trans. on Acoustics,
Speech and Signal Processing.

Keys, R. G., & Pann, K. 1993. Geophysical Applica-
tions of Cubic Convolution Interpolation. Pages

Simultaneous estimation of residual static and cross-

Nedell, N. S., & Taper, M. T. 1971. Semblance and
other coherency measures for multichannel data.
Geophysics, 36, 482–497.

Ronen, J., & Claerbout, J. 1985. Surface-consistent res-
dual statics estimation by stack-power maximiza-

Rothman, D. H. 1985. Nonlinear inversion, statistica-
cl mechanics, and residual statics estimation. Geo-
physics, 50, 2797–2807.

residual statics corrections. Geophysics, 51, 332–
346.

Saito, N., & Beylkin, G. 1993. Multiresolution rep-
resentations using the auto-correlation functions of
compactly supported wavelets. IEEE Transactions
on Signal Processing, 41, 3585–3590.

Saleck, F. M., Bunks, C., Zaleski, S., & Chavent, G.
1993. Combining the multigrid and gradient meth-
ods to solve the seismic inversion problem. Pages

search and genetic algorithms. The Leading Edge of
Stork, C., & Kusuma, T. 1992. Hybrid Genetic Auto-
statics: New Approach for Large-Amplitude Statics
with Noisy Data. Pages 1127–1131 of: 62nd An-
Taner, M. T., Koehler, F., & Alhilali, K. A. 1974. Esti-
mation and correction of near-surface time anomalies.
Geophysics, 39, 441–463.
Whitley, D., Mathias, K., Rana, S., & Dzubera, J.
1995a. Evaluating Evolutionary Algorithms: The
Perils of Poor Empiricism. Department of Com-
puter Science, Colorado State University.
Whitley, D., Mathias, K., & Stork, C. 1995b. Global Op-
imization for Geophysical Applications using Ge-
etic Algorithms. Tech. rept. CASI-TR-95-09. De-
partment of Computer Science, Colorado State Uni-
versity.
Residual statics analysis as a general linear inverse

APPENDIX A: Maxima of stacking-power functions

For a generic stacking-power function

\[ F(\vec{x}) = \sum_i \sum_{j \neq i} \Phi_{ij}(\tau_{ij}(\vec{x})) \]

(A1)

each component of its gradient vector can be written as

\[ \frac{\partial F(\vec{x})}{\partial x_k} = \sum_i \sum_{j \neq i} \Phi'_{ij}(\tau_{ij}(\vec{x})) \frac{\partial \tau_{ij}}{\partial x_k} \quad \text{for } \forall k. \]

(A2)

Let us study the relationship between the maxima of the
stacking-power function and those of the correlation func-
tions. We can prove the following theorem.

Theorem 1. If there exists \( \vec{x}_0 \) so that

\[ \nabla \Phi_{ij}(\vec{x}) \bigg|_{\vec{x}=\vec{x}_0} = 0, \quad \forall i, j, \]

(A3)

then,

\[ \nabla F(\vec{x}) \bigg|_{\vec{x}=\vec{x}_0} = 0. \]

(A4)

Theorem 1 states that if a point in the domain is
the maximum of all correlation functions simultane-
ously, then it must be a maximum of the stacking-power
function. This theorem can be proved as follows.

Since \( \nabla \Phi_{ij}(\vec{x}) \bigg|_{\vec{x}=\vec{x}_0} = 0 \), for \( \forall i, j \), we can also
write that for each \( \Phi_{ij} \),

\[ \frac{\partial \Phi_{ij}(\vec{x})}{\partial x_k} \bigg|_{\vec{x}=\vec{x}_0} = \Phi'_{ij}(\vec{x}) \frac{\partial \tau_{ij}}{\partial x_k} = 0, \quad \text{for } \forall k. \]

(A5)

From equation (A2), we have that \( \nabla F(\vec{x}) \bigg|_{\vec{x}=\vec{x}_0} = 0 \).
Therefore, \( \vec{x}_0 \) is a maximum point of the stacking-power
function equation (A1).

However, the converse of Theorem 1 is not true. That
is, not all maxima of the stacking-power function are
the maxima of correlation functions. It is easy to see that if
equation (A4) is true, equation (A3) is not necessarily
true.

Therefore, while oscillations in seismic data cause
global/local extrema of stacking-power functions, they
are not the only contributing factor. There exist local
maxima on stacking-power functions that are not caused
by stationary points of each correlations. These spurious
maxima are those we are not interested in residual statics
estimation.

To demonstrate this point, Figure A1 shows a 2-D
projection of a 55-dimensional stacking-power function in
which all parameters are fixed except two components.
The left figure shows the slice when the components are
fixed at correct values. In this case, the apparent maxima
are regularly distributed on the function surface. On the
other hand, the right figure shows the projection when
the fixed components are chosen at random. In this case,
the structure is evidently more irregular.

Figure A1. Contour plots of 2-D projections of a 55-dimensional
stacking-power function with all other statics fixed but two
components. All other components are fixed to their correct
values for the left figure, and those components are chosen at
random for the right one.