Restricted optimization: a clue to a fast and accurate implementation of the Common Reflection Surface Stack method

Ernesto G. Birgin, Ricardo Biloti, Martin Tygel, and Lúcio T. Santos

Department of Applied Mathematics,
IMECC-UNICAMP,
CP 6065, 13081-970 Campinas, SP, Brazil.
e-mail: {ernesto,biloti,tygel,lucio}@ime.unicamp.br.

November 10, 1999

Abstract

For a fixed, central ray in an isotropic elastic or acoustic media, traveltime moveouts of rays in its vicinity can be described in terms of a certain number of parameters that refer to the central ray only. The determination of these parameters out of multi-coverage data leads to very powerful algorithms that can be used for several imaging and inversion processes. Assuming two-dimensional propagation, the traveltime expressions depend on three parameters directly related to the geometry of the unknown model in the vicinity of the central ray. We present a new method to extract these parameters out of coherency analysis applied directly to the data. It uses (a) fast one-parameter searches on different sections extracted from the multi-coverage data to derive initial values of the sections parameters, and (b) the application of a recently introduced Spectral Projected Gradient optimization algorithm for the final parameter estimation. Application of the method on a synthetic example shows an excellent performance of the algorithm both in accuracy and efficiency. The results obtained so far indicate that the algorithm may be a feasible option to solve the corresponding, harder, full three-dimensional problem, in which eight parameters, instead of three, are required.

Keywords: Multi-coverage data, Common Reflection Surface method, hyperbolic traveltime, coherency function, optimization, Spectral Projected Gradient method.

1

1 Introduction

In the framework of zero-order ray theory, traveltimes of rays in the paraxial vicinity of a fixed central ray can be described by a certain number of parameters that refer to the central ray only. The approximations are correct up to the second order of the distances between the paraxial and central ray at the corresponding initial and end points. They are, thus, valid independently of any seismic configuration.

Assuming the central ray to be the primary zero-offset ray, the number of parameters are three and eight, for two- and three-dimensional propagation, respectively. For two-dimensional propagation, the parameters are the emergence angle of the normal ray and the wavefront curvatures of the normal and normal-incident-point eigenwaves, as introduced in Hubral (1983). All parameters are defined at the point of emergence of the central ray, called the central point. This point coincides with a common midpoint (CMP), where the simulated zero-offset trace is to be constructed.

The use of multi-parametric traveltime approximations for imaging purposes is a well-investigated subject. Main contributions are *Multifocusing* (see, e.g., Gelchinsky et al., 1997, for a recent description), *Poly Stack^{TM}* (see, e.g., de Bazelaire et al., 1994) and the very recent *Common Reflection Surface* (CRS) method (see, e.g., Hubral et al., 1998, and Perroud et al., 1999). These methods vary in general on two aspects, namely, the multi-parametric traveltime moveout formula that is used, as well as in the strategy employed to extract the traveltime parameters from coherency analysis applied on the multi-coverage data.

The basic lines of the CRS approach are the choice of the hyperbolic traveltime function (see Tygel et al., 1997), and the strategy of breaking the original three-parameter estimation problem into simpler ones involving one or two unknowns. As shown in Müller (1999), quick estimations for the three parameters can be obtained by simple one-parameter searches performed on CMP and CMP-stacked sections out of the multi-coverage data. Direct use of these parameters in the CRS stacking algorithms leads to very acceptable imaging results. As reported in Müller (1998), application of the algorithm to real-data examples have produced imaging results comparable and, in many cases, even superior to those obtained by conventional NMO/DMO.

To improve the accuracy of the parameter estimations, as needed, e.g., for the construction of velocity models, a natural idea is to use the previously obtained parameter estimations as initial values for an optimization scheme directly applied to the multi-coverage data problem. Following this philosophy, Müller (1999) and Jäger (1999) were able to obtain significantly better results on synthetic data examples, however, at a high computational cost.

In this work, we present a new optimization strategy so as to achieve more accurate results than the ones derived by purely one-parameter searches, while maintaining the computational effort at a reasonable level. This becomes a crucial matter when real-data applications are envisaged. The method is illustrated by its application on a synthetic example, where the various aspects of the algorithm can be better understood.

2 Hyperbolic traveltime expansion

As shown in Figure 1, let us assume a fixed target reflector Σ in depth, as well as a fixed *central point* X_0 on the seismic line, considered to be the location of a coincident source- and -receiver

pair $S_0 = G_0 = X_0$. The corresponding zero-offset reflection ray, X_0 NIP X_0, will be called from now on the *central ray*. It hits the reflector at the *normal-incident-point* (NIP). For a source-receiver pair (S, G) in the vicinity of the central point, we consider the primary reflected ray SRG relative to the same reflector Σ. We use the horizontal coordinates x_0, x_S and x_G to specify the location of the central point X_0, the source S and the receiver G, respectively. We find it convenient to introduce the *midpoint* and *half-offset* coordinates

$$x_m = (x_G + x_S)/2 - x_0 \quad \text{and} \quad h = (x_G - x_S)/2. \tag{1}$$

We consider the hyperbolic traveltime expression as in Tygel et al. (1997)

$$T^2(x_m, h; \beta_0, K_N, K_{NIP}) = \left(t_0 + \frac{2x_m \sin \beta_0}{v_0}\right)^2 + \frac{2t_0 \cos^2 \beta_0}{v_0}\left(K_N \, x_m^2 + K_{NIP} \, h^2\right), \tag{2}$$

where t_0 is the zero-offset traveltime and β_0 is the angle of emergence at the zero-offset ray with respect to the surface normal at the central point. The quantities K_N and K_{NIP} are the wavefront curvatures of the *normal* N-wave and the NIP-wave, respectively, measured at the central point.

The N- and NIP-waves are fictitious eigenwaves introduced by Hubral (1983) for the analysis of the actual propagation of the zero-offset ray, as well as for its corresponding paraxial rays. Their wavefront curvatures at the central point carry important information about the velocity model in which the wave propagation takes place. The N-wave can be conceptually visualized as the one that starts as a wavefront that coincides with the reflector and travels to the surface with half of the medium velocity. It arrives at the central point at the same time as the zero-offset ray. The NIP-wave can be visualized as starting as a point source at the reflection point (NIP) of the zero-offset reflection ray and propagates upwards with half of the medium velocity. It arrives at the central point at the same time as the zero-offset ray, too.

For particular source-receiver gathers, the hyperbolic traveltime formula (2) can be simplified. The most used configurations are:

The common-midpoint configuration: Setting the fixed midpoint to coincide with the central point, the CMP-traveltime expression can be readily obtained from the hyperbolic traveltime (2) by simply placing $x_m = 0$ in that formula. We find the one-parameter expression

$$T_{CMP}^2(h; q) = t_0^2 + \frac{2t_0 \, h^2 q}{v_0}, \tag{3}$$

on the combined parameter $q = \cos^2 \beta_0 \, K_{NIP}$.

The zero-offset configuration: The zero-offset traveltime expression is readily obtained setting $h = 0$ in the hyperbolic traveltime (2). We find the two-parameter expression

$$T_{ZO}^2(x_m; \beta_0, K_N) = \left(t_0 + \frac{2x_m \sin \beta_0}{v_0}\right)^2 + \frac{2t_0 \cos^2 \beta_0}{v_0} K_N \, x_m^2, \tag{4}$$

on the original parameters β_0 and K_N.

3

The common-shot configuration: Placing the common source to coincide with the central point, the common-shot traveltime expression is derived by setting $x_m = h$ in the hyperbolic traveltime (2). As a result, the common-shot traveltime becomes the two-parameter expression

$$T_{CS}^2(h; \beta_0, \mu) = \left(t_0 + \frac{2h \sin \beta_0}{v_0}\right)^2 + \frac{2t_0 \cos^2 \beta_0}{v_0} \, \mu \, h^2 \, , \tag{5}$$

depending on the original and combined parameters β_0 and $\mu = K_N + K_{NIP}$, respectively.

The common-offset configuration: The expression of the common-offset traveltime coincides with the general hyperbolic traveltime (2) upon the consideration of $h = constant$.

3   Formulation of the problem and its solution

The data obtained by a multi-coverage seismic experiment, performed on a given seismic line, consists of a multitude of seismic traces $U(x_m, h, t)$ corresponding to source-receiver pairs located by varying coordinate pairs (x_m, h) and recording time $0 < t < T$. The basic problem we have to solve is the following:

> Consider a dense grid of points (x_0, t_0), where x_0 locates a central point X_0 on the seismic line and t_0 is the zero-offset traveltime. For each central point X_0, let the medium velocity $v_0 = v(x_0)$ be known. From the given multi-coverage data, determine the corresponding parameters β_0, K_N and K_{NIP}, for any given point (x_0, t_0) and velocity v_0.

One approach to solve this problem could be the application of a multi-parameter coherency analysis to the data, using the traveltime formula (2) to a number of selected traces $U(x_m, h, t)$ in the vicinity of the central ray X_0 and for a suitable time window around the time t_0. Depending on the choice of seismic traces and gathers involved, the desired values of the sought-for parameters are expected to be very close to the ones for which the maximum coherence is achieved when applying the traveltime (2) to the data.

Given the seismic traces $U(x_m, h, t)$, and the vector of parameters $P = (\beta_0, K_N, K_{NIP})$, the coherency measure called *semblance*, as introduced by Taner and Koehler (1969), is given by

$$S = \frac{\sum[\sum U(x_m, h, T(x_m, h; P))]^2}{M \sum \sum[U(x_m, h, T(x_m, h; P))]^2} \, , \tag{6}$$

where $T(x_m, h; P) = T(x_m, h; \beta_0, K_N, K_{NIP})$ is the traveltime expression (2) and M is the total number of selected traces. The inner summation is performed over all selected traces, and the outer one is performed over a given time window around t_0. For each given pair (x_0, t_0), the objective is to find the global maximum of the semblance function (6) with respect to the parameters β_0, K_N and K_{NIP}. These parameters are restricted to the ranges $-\pi/2 < \beta_0 < \pi/2$ and $-\infty < K_N, K_{NIP} < \infty$.

To compute the global maximum of the semblance function, we propose the strategy described by the flow chart in Figure 2. In the first part we obtain initial values of the parameters.

In the second part, an optimization process employs these parameters as initial values to produce the final estimations. Following the same lines as Müller (1999), the first part consists of two steps, namely, (a) a one-parameter search of the combined parameter q, performed on the CMP sections with the help of the traveltime expression (3), and (b) two one-parameter searches for β_0 and K_N, performed on the CMP-stacked section realized using the previous q-parameter. The CMP-stacked section is considered as an approximate zero-offset section, so the traveltime expression (4) is used in these computations.

The optimization process of the second part determines the two parameters β_0 and $\mu = K_N + K_{NIP}$. For this purpose, we use the recently introduced Spectral Projected Gradient (SPG) method applied to common-source sections extracted from the multi-coverage data. The explanation of the SPG method exceeds the scope of this work. We refer the interested reader to Birgin et al. (1999) for a comprehensive description of the method, including convergence results and numerical experiments that prove the efficiency of the method. We use the traveltime expression (5) to obtain the original parameter β_0 and the combined parameter μ. Finally, using the relationships $K_{NIP} = q/\cos^2\beta_0$ and $K_N = \mu - K_{NIP}$ all the desired parameters can be determined.

The SPG method maintains the basic features of gradient-type methods (easy-to-implement, low-memory requirements) and is also naturally suited to bound constrained problems. These good features lead us to use the SPG method in the present work.

4   A synthetic example

Referring to Figure 3, we consider the synthetic two-dimensional model of three smoothly curved reflectors separating different homogeneous acoustic media. Assuming unit density, the constant velocities are: $c_1 = 1400$m/s above the first reflector, $c_2 = 2000$m/s between the first and the second reflector, $c_3 = 3400$m/s between the second and the third reflector, and, finally, $c_4 = 5500$m/s below the deepest reflector.

The input data for our experiment are a collection of 334 CMP seismic sections, centered at coordinates x_0 varying from 3010m to 13000m. Each CMP gather has 84 traces with half-offsets varying from 0m to 2490m. All traces are sampled within the range of $0s \leq t \leq 6s$, at a sample rate of 4ms. Noise was added to the data with a ratio signal:noise of 7:1. Figure 4 shows the CMP gather centered at coordinate $x_0 = 9910m$.

Initial estimation – the combined parameter q: Our process starts with the estimation of the combined parameter q. This is performed as a one-parameter search on the CMP gathers extracted from the multi-coverage data. The situation is similar to a conventional NMO-velocity analysis. For each midpoint x_0, taken as a central point, we consider its corresponding CMP gather. Setting the central point x_0 fixed, we determine, for each time sample t_0, the value of q that yields the best semblance in the CMP gather. For this computation, we use the CMP-traveltime formula (3) that depends on the q-parameter only. The above procedure leads to the construction of two auxiliary CMP-related sections, namely, the *q-section*, which consists of assigning to each (x_0, t_0) its corresponding q-parameter, and the *semblance section* in which the semblance values, instead of the q-values, are assigned. An extensive use of semblance sections,

as well as other auxiliary sections, is described in Gelchinsky et al. (1997).

The q-search may be refined for greater accuracy. We consider the estimated q-parameters for which the current semblance values exceed a threshold that is interactively selected by the user. This provides an ensemble of q-values concentrated on a smaller range (in our case three orders of magnitude less than the original range search). It allows us to perform a new search, restricted to this smaller range divided into a much finer grid. As a consequence, for comparable computational costs, we gain orders of magnitude in the accuracy of the estimated parameter. Figure 5 shows the semblance section obtained after the use of the above-described refinement strategy. The employed threshold semblance values were 0.13 and 0.15 for the time intervals $0s < t_0 < 2.5s$ and $2.5s < t_0 < 6s$, respectively. The very clear semblance section of Figure 5 can be looked upon as a simulated zero-offset section. The theoretical and estimated values of the combined parameter q along the reflectors are shown in Figure 6. The highly accurate results confirm the expectations of employing an exhaustive search to solve a one-parameter problem. As a consequence, the obtained values of the q-parameter will be retained during the whole process.

Initial estimation – the parameters β_0, K_N and K_{NIP}: Using the just estimated q-values in the CMP-traveltime formula (3), we construct (like in conventional NMO-stacking) the corresponding CMP-stacked section (see Figure 7). This will now be used as an approximation of a zero-offset section.

To extract the emergence angles β_0 and the N-wave curvatures K_N, we proceed as follows: (a) Using the zero-offset traveltime expression (4), we first set $K_N = 0$ and perform, for each pair (x_0, t_0), a one-parameter search for β_0 between $-\pi/2$ and $\pi/2$; (b) Setting the obtained value of the β_0 parameter in the same zero-offset traveltime expression (4), we perform a further one-parameter search, this time for the parameter K_N. Use of the above results, together with the relationship $K_{NIP} = q/\cos^2 \beta_0$, completes the initial estimations of the three parameters β_0, K_N and K_{NIP}.

Optimization procedure – Final estimations: The second part of our method consists of the application of an optimization algorithm to common-shot sections extracted from the multi-coverage data. For that matter, we use the common-shot traveltime formula (5), depending on the two parameters β_0 and $\mu = K_N + K_{NIP}$. Using the initial estimation of the parameters obtained in the first step, we apply the Spectral Projected Gradient method (see Birgin et al., 1999) to achieve the final estimations.

Figures 8, 9 and 10 show the comparison between the theoretical and optimized parameters. We can recognize that the method provides generally accurate parameter estimations in most of the section. We note, however, that the method also yields inaccurate results at various points within the range [6000m,8000m]. These points are characterized by small coherence measures and, for that matter, have not been displayed in Figures 8, 9 and 10. The reasons for those small coherence values may be (a) Lack of illumination: The use of end-on, common-shot gathers may not be the most adequate choice of illumination for the whole section. This holds, in particular, for the first, shallowest reflector, where accentuated dips are present in the "problematic region". (b) Caustics: The same region contains a caustic due to the second reflector. This is expected to introduce problems, not only to this reflector, but also to the third one below it.

A possible improvement of the results could be obtained upon the combined use of traces that belong to different gathers (e.g, split-spread common-shot and common-offset gathers) that are available in the multi-coverage data. The use of additional gathers may be recommended to overcome these difficulties. These important aspects are under current investigation.

5   Some considerations and remarks

The application of the proposed method relies on some choices to be made by the user on a more or less a priori manner. We mention these choices and also pose some related questions.

1. The optimization procedure of our implementation could be alternatively carried out on common-offset or split-spread common-shot sections extracted from the multi-coverage data, instead of the common-shot sections as actually performed. A natural question to be posed is what are the differences in the obtained results.

2. Concerning the optimization solver, another alternative could be to combine SPG with a Newton-type method to speed up the convergence. The idea is that SPG (as well as any gradient-projection type algorithm) can be used to get a closer approximation to the solution, so that a Newton-type method can be applied ignoring the constraints. This implicitly assumes that the stationary point is an interior point of the feasible region.

3. The strategy of using particular configurations to reduce the number of parameters to be estimated has definite advantages, but may have also hide some disadvantages. The main advantage is the significant reduction of computational effort and simplification of the optimization procedures. A possible disadvantage is the use of less redundancy, because many available traces that do not conform to the selected configuration have to be left out. In some situations, this could lead to a decrease in the accuracy of the final estimation of the parameters. An alternative to overcome this possible shortcomings would be to consider more configurations, e.g., common-shot and common-offset sections, and average the corresponding parameter estimation results.

4. Given a pair (x_0, t_0), the selection of the traces and the time window required for the evaluation of the coherency function is an important a priori decision to be made by the user. The situation is similar to the one encountered in the application of velocity analysis in conventional NMO/DMO stacking. Investigations on the accuracy of the hyperbolic and other traveltime moveout approximation expressions would certainly help to get the most of these formulas.

5. The semblance section obtained by the one-parameter search for the combined parameter q (see Figure 5) provides a valuable initial indication of the regions where primary zero-offset reflections are to be located. As a consequence, for each central-point location x_0, the number of zero-offset traveltimes t_0 connected to actual reflections can be dramatically reduced. This means that, in principle, we could restrict the subsequent searches and optimization procedures to those "promising" points only, leading to a significant reduction of computational costs. If the number of points is sufficiently reduced, maybe an

optimization procedure can be applied to an ensemble of sections of various configurations or even to the whole multi-coverage data.

6 Conclusions

We have proposed a new algorithm to determine the traveltime parameters out of coherency analysis applied to 2-D multi-coverage seismic data. Following the general philosophy of the CRS approach, we used the hyperbolic traveltime moveout together with a sequential application of one-parameter searches, followed by a two-parameter optimization scheme. The restriction of the two-parameter optimization to common-shot sections, together with the use of the recently introduced Spectral Projected Gradient method, lead us to a fast and generally accurate estimation of all three parameters. We also observed that some inaccuracies on the estimations may be due to non-sufficient illumination from the selected gathers, as well as from usual complications associated with caustical regions. We have addressed a number of questions concerning alternatives to improve the accuracy, as well as to reduce computational costs.

The proposed approach to perform the three-parameter estimation is the main contribution of the present work. We applied the algorithm to a three-reflector synthetic example. Although this is a simple model, it presents already some of the basic complications of more realistic situations. The obtained results were very encouraging, confirming our expectations concerning accuracy improvements at reasonable computational costs, as compared with currently available search methods. Next steps will be to test the new algorithm on more complex models and to real data sets.

7 Acknowledgments

References

[1] Birgin, E. G., Martínez, J. M. and Raydan, M., 1999, Nonmonotone spectral projected gradient methods on convex sets. To appear in *SIAM Journal on Optimization*.

[2] de Bazelaire, E. and Viallix, J. R., 1994, Normal moveout in focus: *Geophys. Prosp.*, **42**, 477–499.

[3] Gelchinsky, B., Berkovitch, A. and Keydar, S., 1997, Multifocusing homeomorphic imaging: Parts I and II. Presented at the special course on Homeomorphic Imaging by B. Gelchinsky, Seeheim, Germany.

[4] Hubral, P., 1983, Computing true amplitude reflections in a laterally inhomogeneous earth. *Geophysics*, **48**, 1051–1062.

[5] Hubral, P., Höcht, G., and Jäger, R., 1998, An introduction to the common reflection surface stack, EAGE 60th meeting and technical exhibition, Leipzig, *Extended Abstracts* 1–19.

[6] Jäger, R., 1999, The Common reflection surface method. M. Sc. Thesis submitted to the Geophysical Institute, Karlsruhe University, Germany.

[7] Müller, T., 1998, Common reflection surface stack versus NMO/stack and NMO/DMO/stack. 60th Mtg. Eur. Assoc. Expl Geophys., *Extended Abstracts*, 1–20.

[8] Müller, T., 1999, The Common Reflection Surface Stack method – Seismic imaging without explicit knowledge of the velocity model. Ph. D Thesis submitted to the Geophysical Institute, Karlsruhe University, Germany.

[9] Perroud, H., Hubral, P., and Höcht, G., 1999, Common-reflection-point stacking in laterally inhomogeneous media. *Geophysical Prospecting*, **47**, 1–24.

[10] Taner, M. T. and Koehler, F., 1969, Velocity spectra – Digital computer derivation and applications of velocity functions. *Geophysics*, **34**, 859–881.

[11] Tygel, M., Müller, T., Hubral, P. and Schleicher, J., 1997, Eigenwave based multiparameter traveltime expansions. Ann. Internat. Mtg., Soc. Expl. Geoph., *Expanded Abstracts*, 1770–1773.
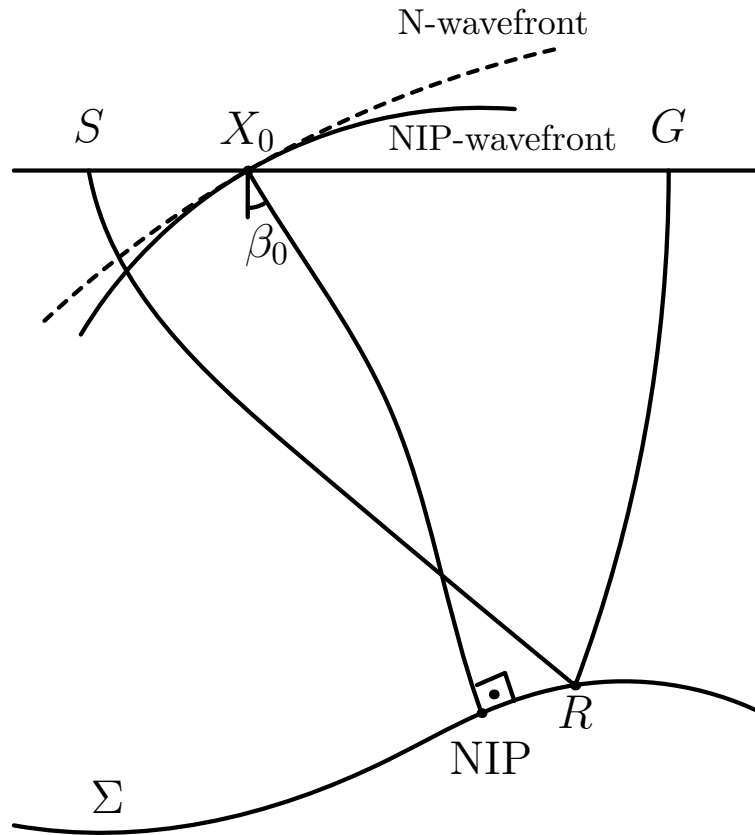
Figure 1: Physical interpretation of the hyperbolic traveltime formula parameters: Emergence angle, β_0, normal-wave curvature, K_N, and normal-incident-point-wave curvature, K_{NIP}.
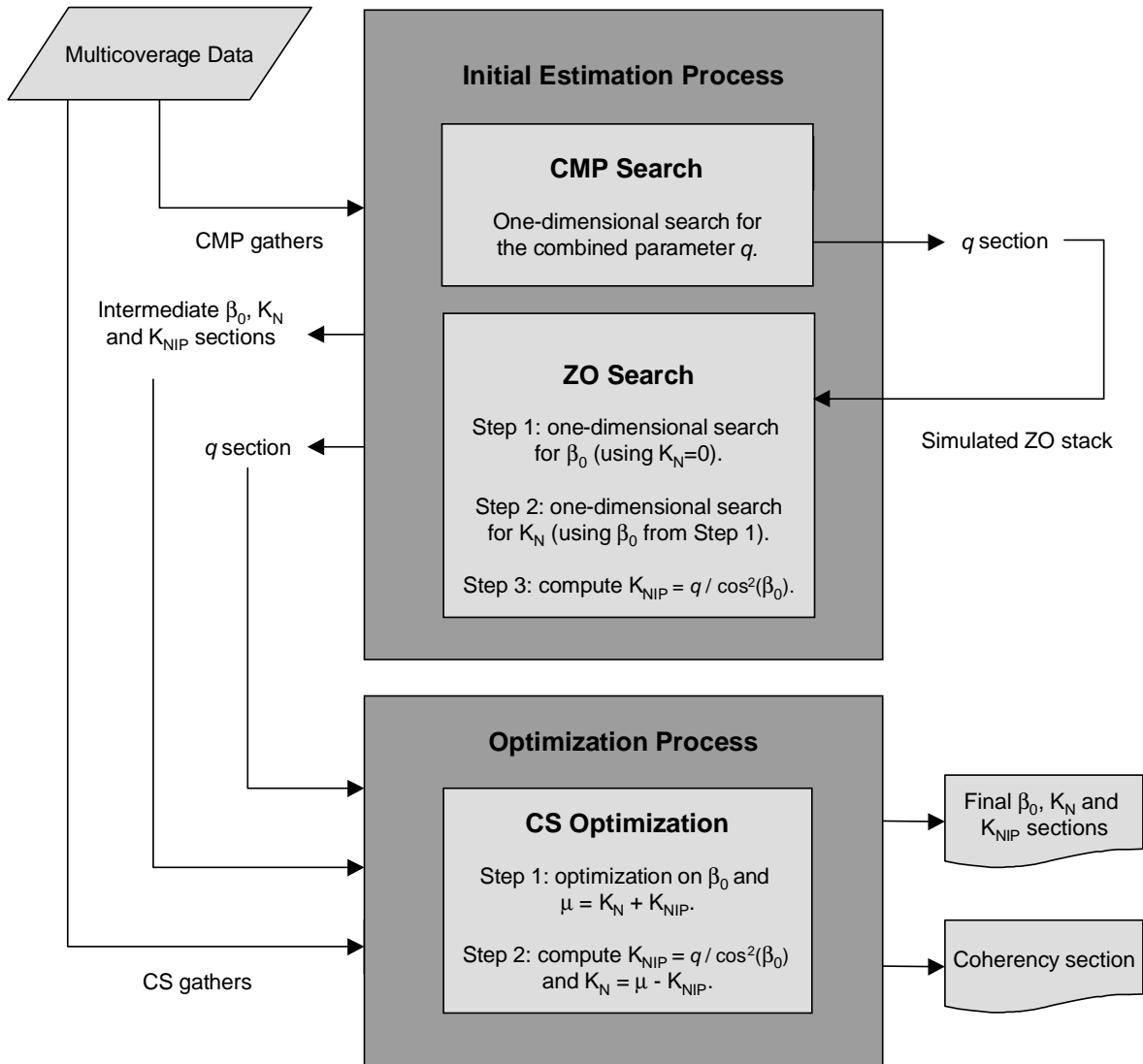
Figure 2: Flow-chart description of the parameter-estimation strategy. First part: Computation of initial estimations by one-parameter searches. Second part: Optimization method applied to common-shot sections for final parameter estimation.
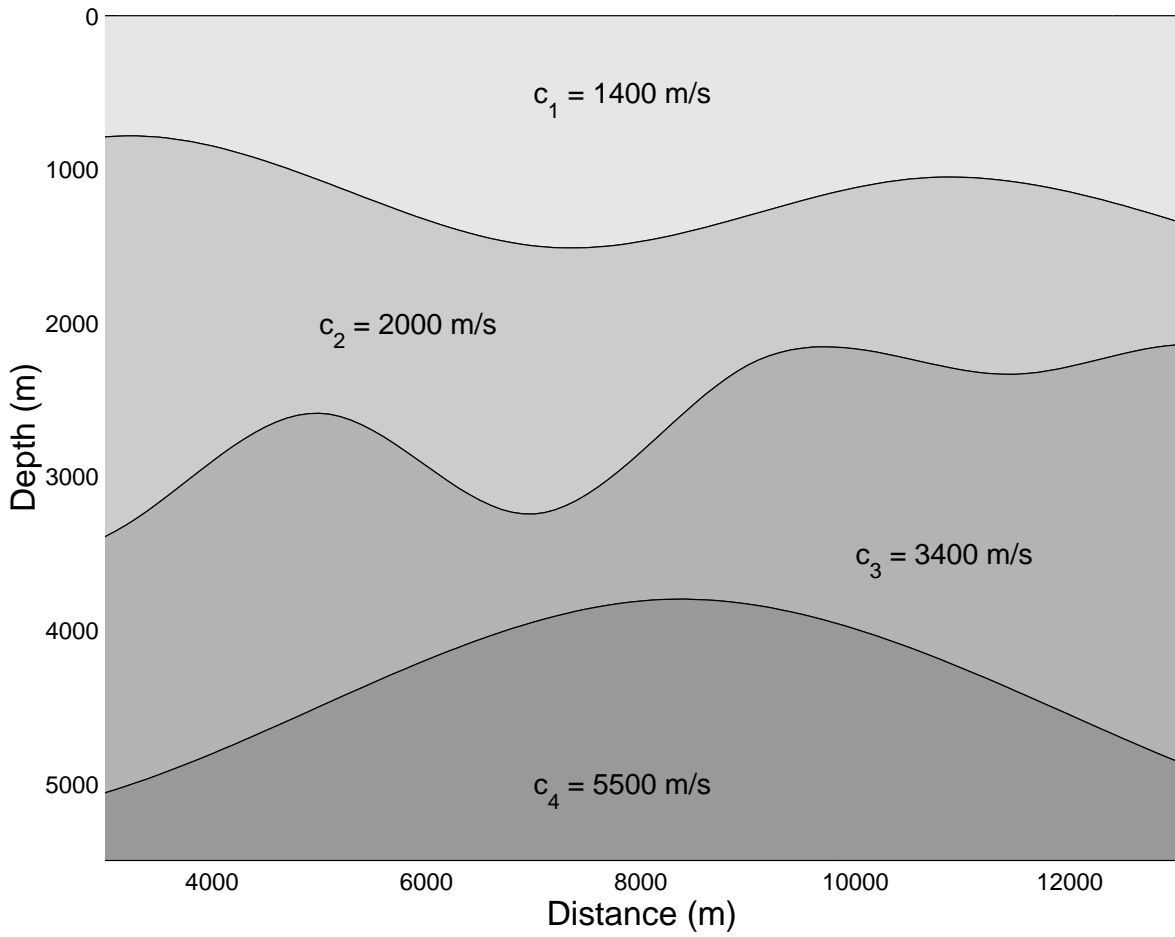
Figure 3: Synthetic two-dimensional model of four homogeneous acoustic layers separated by smooth curved interfaces. Unit density is assumed in all media.
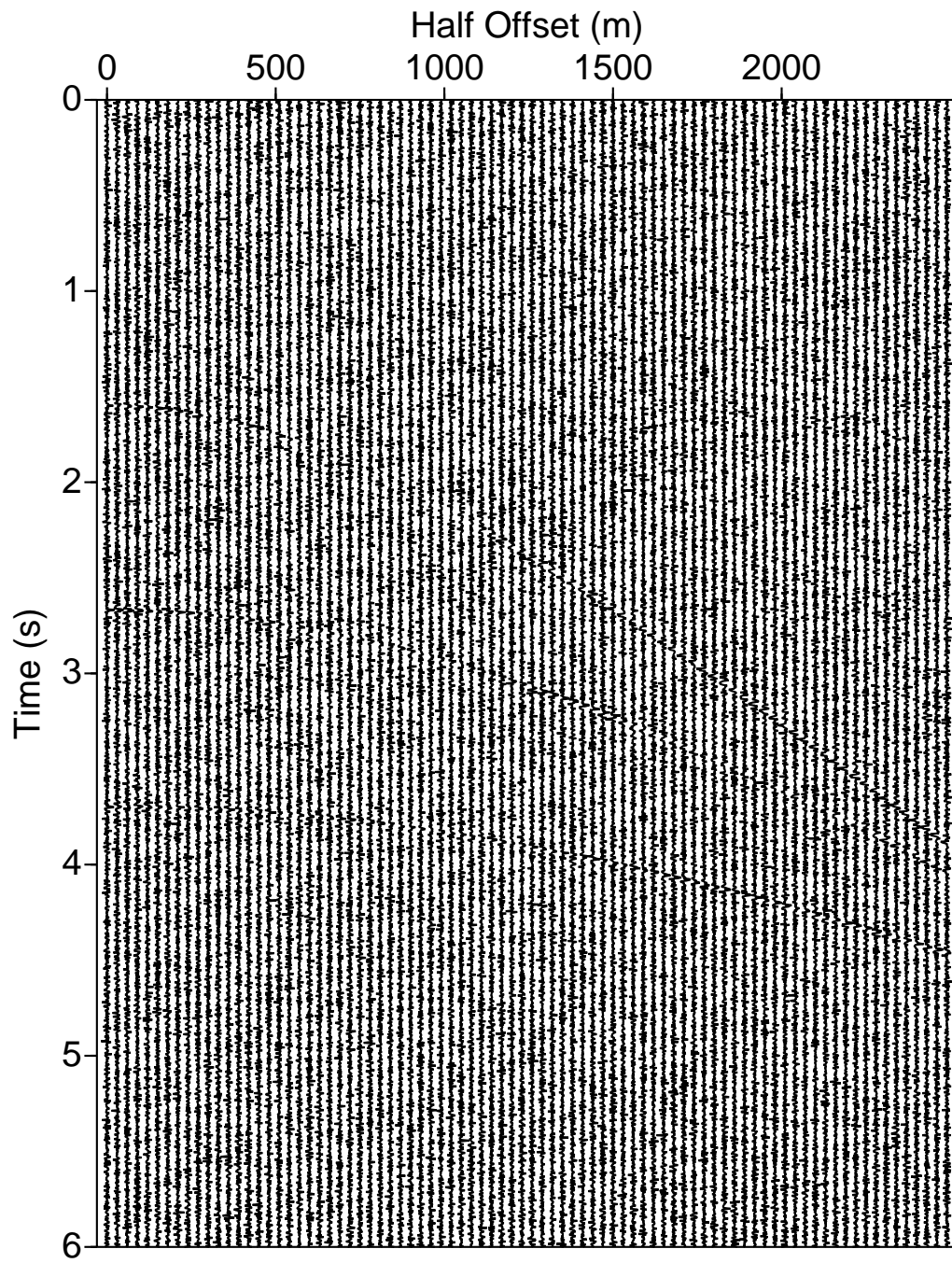
Figure 4: A typical CMP gather extracted from the multi-coverage data.
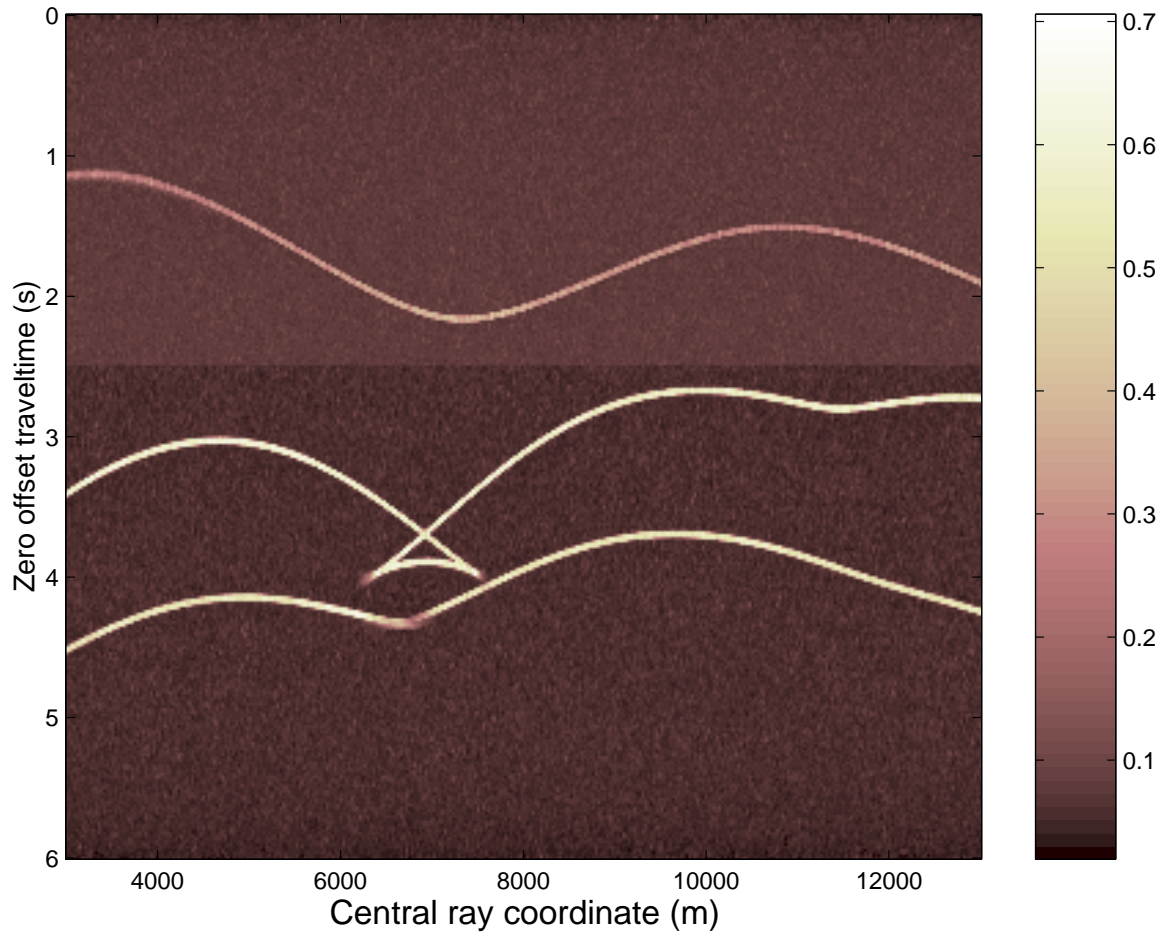
Figure 5: Semblance section obtained as a result of the one-parameter search of the combined parameter q. Note the excellent resolution of the section.
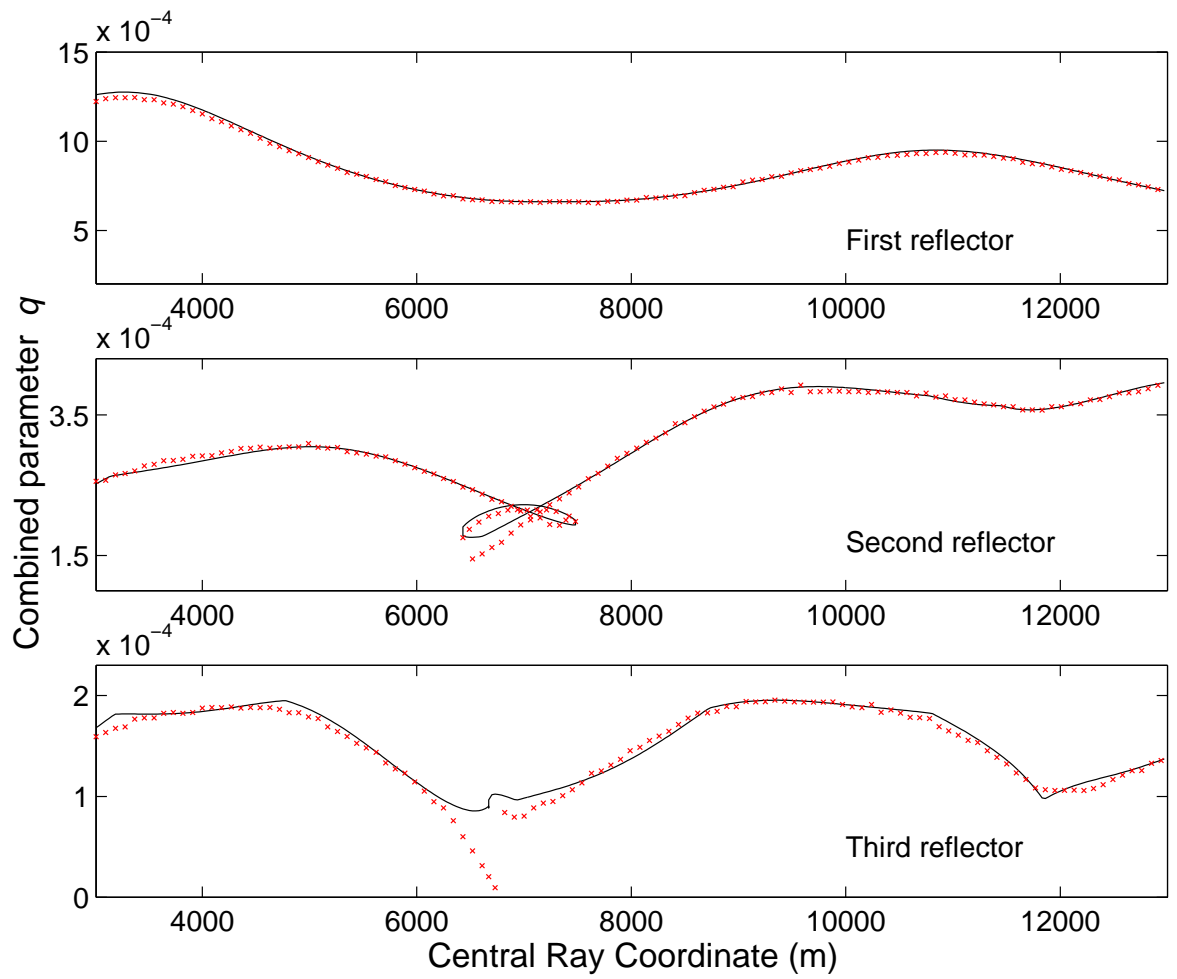
Figure 6: Combined parameter q: Theoretical curve (solid line) and estimated values (small x) obtained after the one-parameter search on the CMP sections.
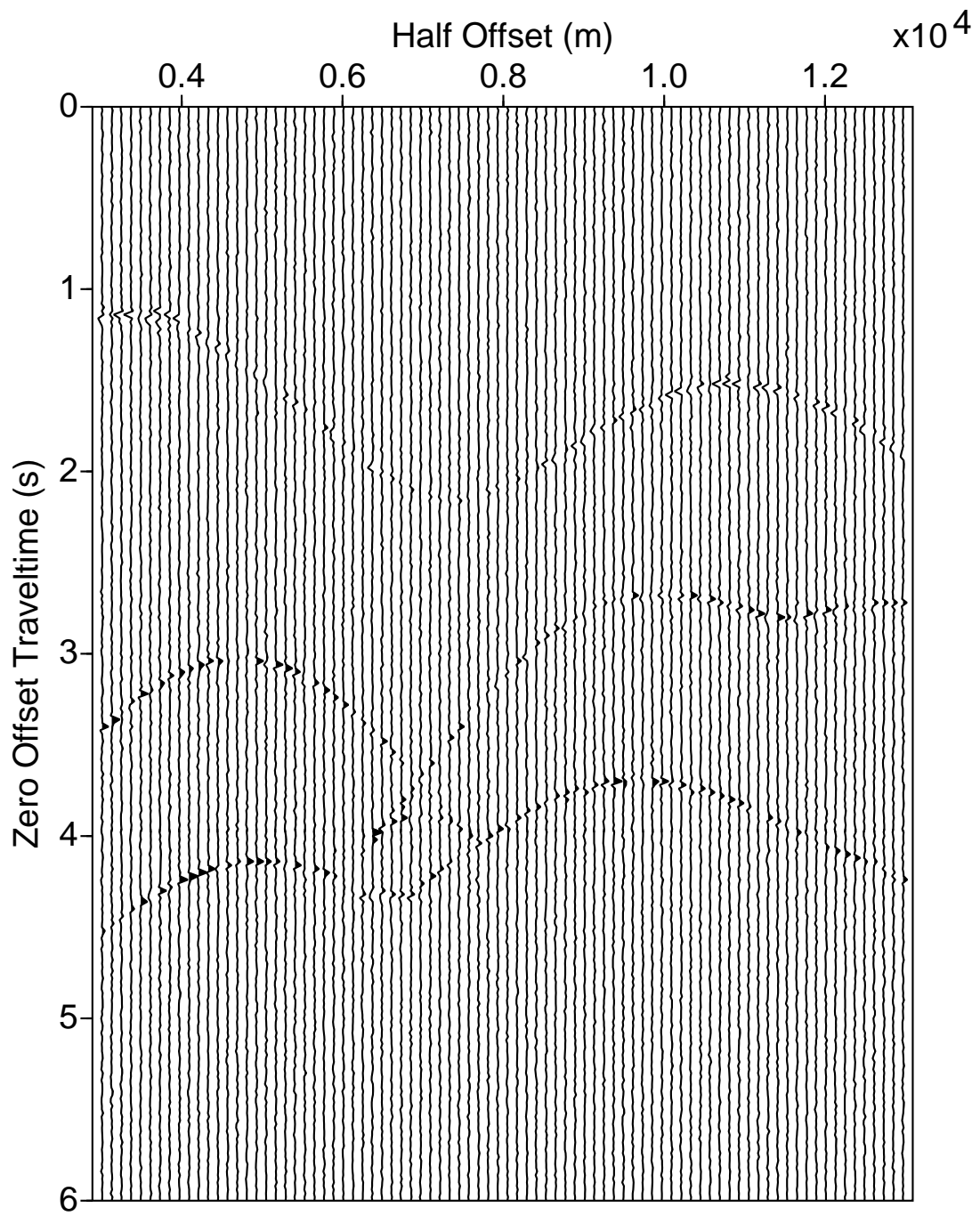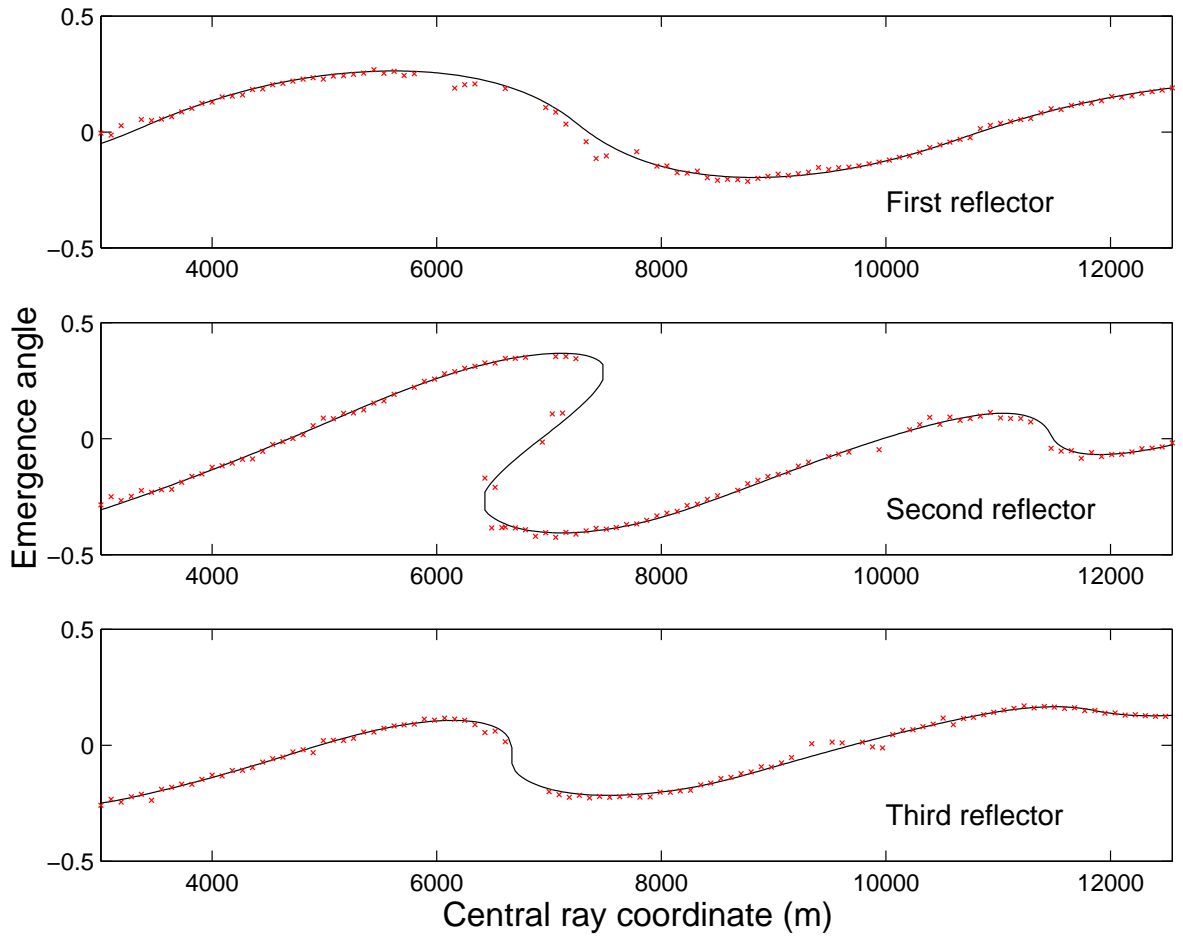
Figure 7: CMP Stacked section using combined parameter q.

Figure 8: Emergence angle β_0: Theoretical curve (solid line) and estimated values (small x) obtained after the two-parameter optimization on the common-shot sections.
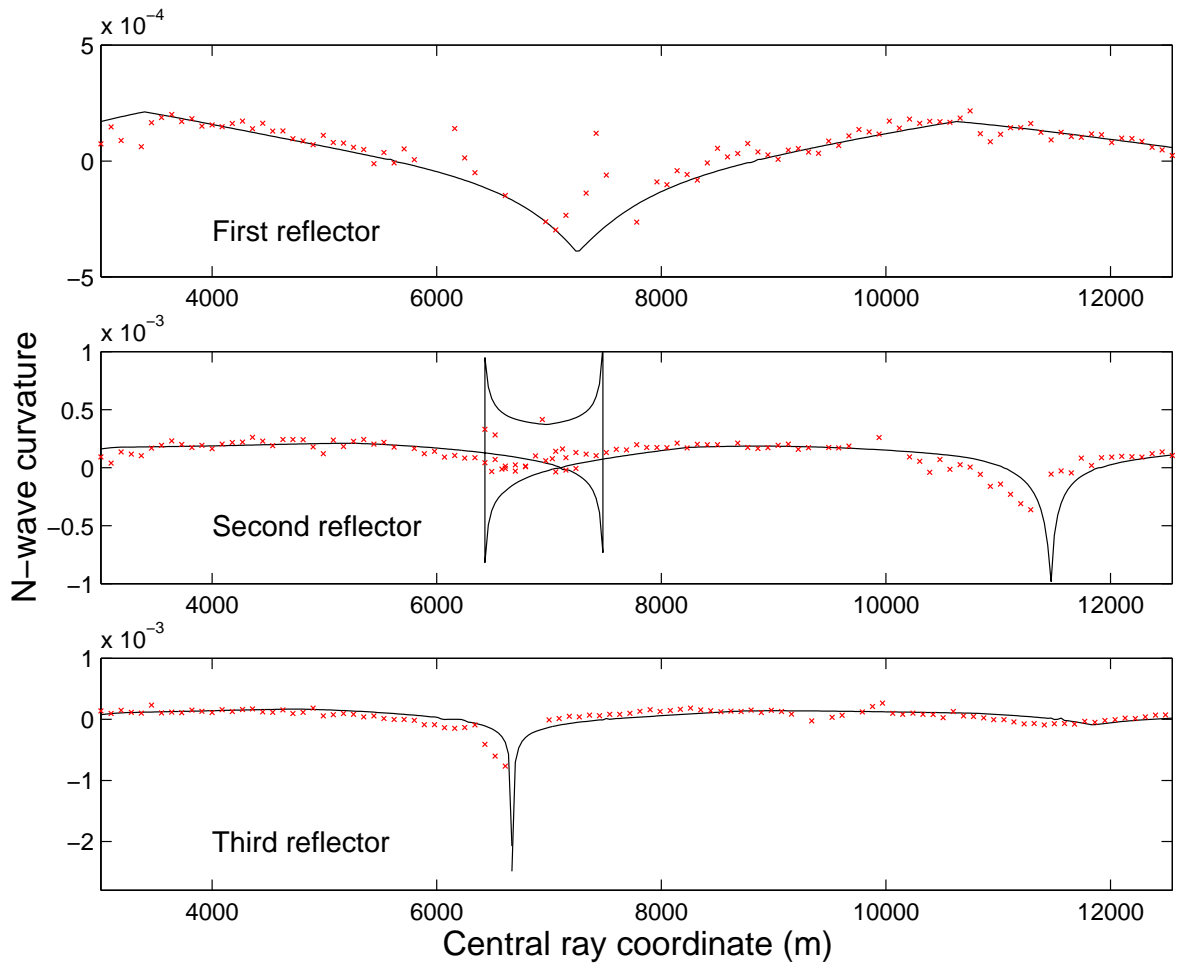
Figure 9: N-wave curvature K_N: Theoretical curve (solid line) and estimated values (small x) obtained after the two-parameter optimization on the common-shot sections.
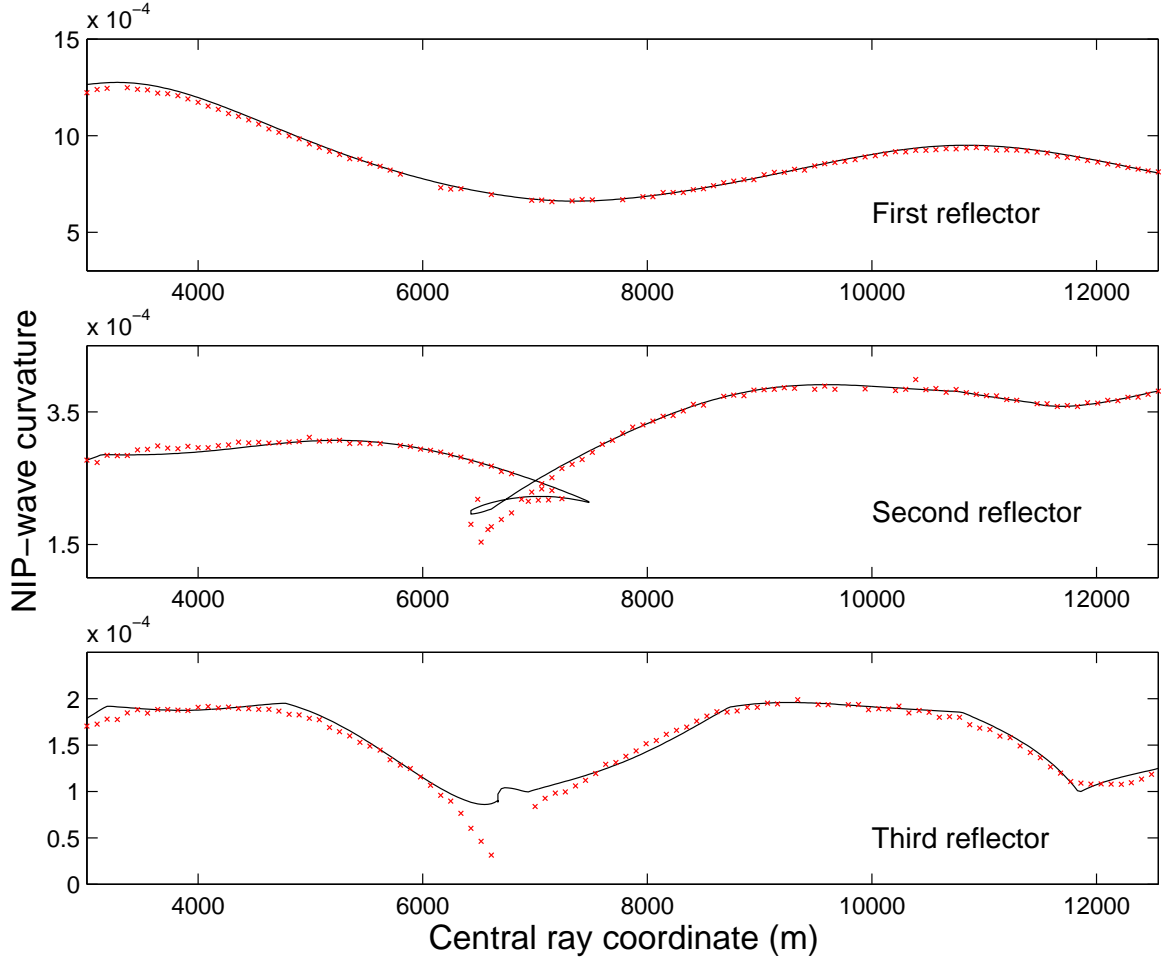
Figure 10: NIP-wave curvature K_{NIP}: Theoretical curve (solid line) and estimated values (small x) obtained after the two-parameter optimization on the common-shot sections.