OPTIMIZATION OF THE SEMBLANCE IN THE 2D CRS STACK USING A COARSE LINE SEARCH

P. Witte and D. Gajewski

email: *pwitte@eos.ubc.ca*
keywords: *CRS, semblance, optimization*

ABSTRACT

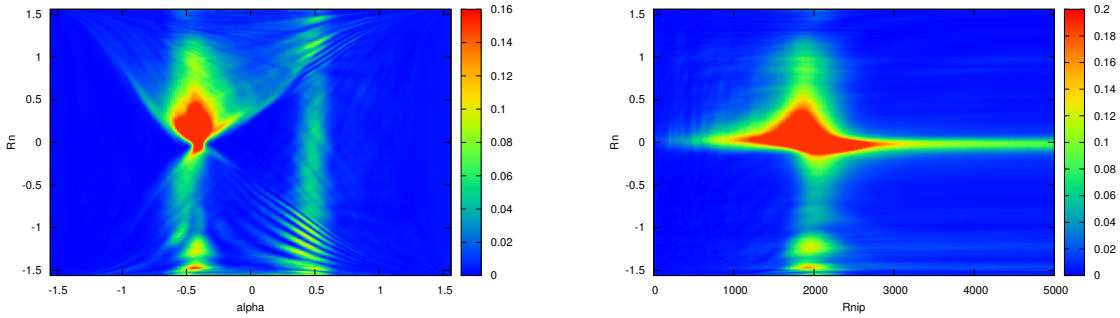
In this paper, we propose an alternative method for optimizing the semblance in the two-dimensional common-reflection-surface stack. An analysis of the objection function shows that the semblance generally provides a fairly broad and clearly distinguishable global maximum, which is of quadratic shape. We make use of this observation by coarsely scanning the parameter search space along a single direction to find the approximate location of the maximum. Then, the objective function is optimized with a local method, using the initial guess from the coarse line search. An application of this method to synthetic seismic data shows that this approach improves the semblance and the parameter sections in comparison to the results obtained with the so-called pragmatic approach. In particular, the strength of our new approach is that we are able to perform a fairly good parameter estimation even for extremely weak events that appear, for example, in sub-salt areas.

INTRODUCTION

For the simulation of a zero-offset stacked section using the two-dimensional common-reflection-surface stacking operator, the optimal stacking parameters α, R_N and R_{NIP} have to be obtained by means of coherency analysis using a multi-parameter optimization. Since global optimization methods are computationally expensive and do not guarantee convergence towards the global maximum of the objective function, a pragmatic approach for estimating initial values for the CRS parameters has been developed by Mann et al. (1999) and Jäger et al. (2001). The basic idea of the pragmatic approach is to split the simultaneous three parameter search into subsequent one-dimensional searches for one parameter at a time, and then to use these initial guesses for a local multi-parameter optimization. As an alternative and despite the relatively large computational cost, several authors (Santos et al., 2005; Minato et al., 2012; Garabito et al., 2012; Barros et al., 2014) have tested global optimization methods like simulated annealing or differential evolution for the CRS parameter estimation and compared them to the standard approach. From their results, it can be generally observed that global methods perform superior compared to local optimization methods, which strongly rely on the initial parameter values from the pragmatic approach.

However, as pointed out by Jäger et al. (2001), the semblance function, which is the objective function in the CRS workflow, has some distinct properties that suggest to utilize specialized optimization methods for this problem. In particular, it was observed that the semblance features a large number of local small scale extrema, but that it generally has a very broad and distinct global maximum. Furthermore, the optimization is actually strongly constrained by the physical meaning of the CRS parameters, i.e. the parameters have to lie in a range that make sense physically. Therefore, the motivation for the following study was to further analyze the properties of the semblance function and to introduce an optimization strategy that makes use of its special properties and physical constraints.

Following these two steps, we have applied the new optimization strategy to a complex synthetic data set. The result and comparison to the classical approach demonstrate the high potential of the new method.



(a) Clipped semblance in the α–R_N plane.

(b) Clipped semblance in the R_N–R_{NIP} plane.

Figure 1: Slices of the semblance function extracted at the position of the global maximum. Note that the N-wave radius depicted here is not R_N itself but its projection onto the Riemann sphere (see text).

VISUALIZATION OF THE SEMBLANCE FUNCTION

In the first section of this paper, we further investigate the observations by Jäger et al. (2001) concerning the overall shape and properties of the objective function associated with the CRS stack. Most commonly, the objective function chosen for estimating the optimal wave field attributes is the semblance (Taner and Koehler, 1969),

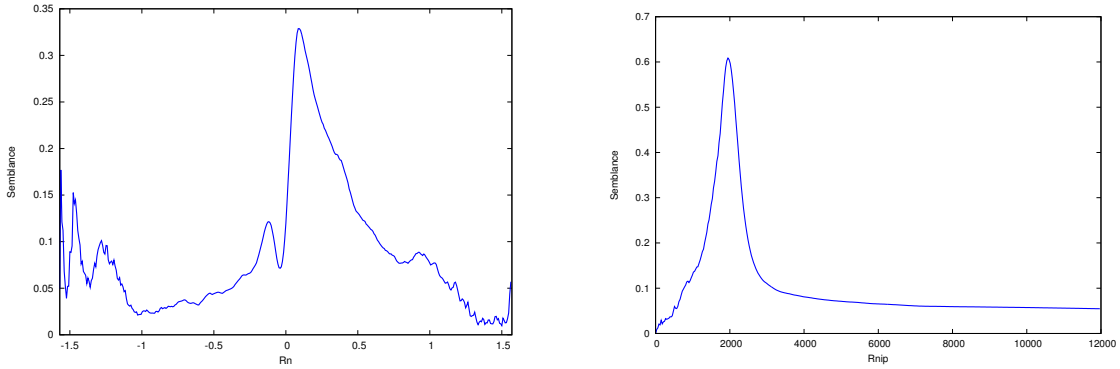
$$S = \frac{\sum_{t}^{t+\Delta m} \left(\sum_{i=1}^{N} g_{ti} \right)^2}{\sum_{t}^{t+\Delta m} \sum_{i=1}^{N} \left(g_{ti} \right)^2}, \tag{1}$$

where g_{ti} is the amplitude of the ith trace at time sample t, but other coherence measurements might be considered as well. The function depends on the CRS parameters implicitly through the data that is summed up over the common reflection surface for every time sample. Every evaluation of the semblance function within the optimization requires the calculation of the travel time in dependence of the (half) offset and midpoint coordinate using the CRS travel time formula,

$$t^2(x_m, h) = \left(t_0 + \frac{2 \sin \alpha}{v_0} (x_m - x_0) \right)^2 + \frac{2 t_0 \cos^2 \alpha}{v_0} \left(\frac{(x_m - x_0)^2}{R_N} + \frac{h^2}{R_{NIP}} \right), \tag{2}$$

and the data has to be summed up over the travel time surface. Therefore, every evaluation of the objective function is expensive and the number of function evaluations should be kept as low as possible.

Using a simple synthetic seismic data example of a circular reflector with a radius of 1 km, several plots of the semblance function were extracted, which correspond to different locations in the section. The selected figures for this paper were chosen in a way to be as representative for as many semblance plots as possible. The shown plots were generated on top of the reflector, where the reflector has an angle of approximately 30°. To emphasize the local extrema of the function, the amplitudes in the figures were clipped to 25 % of the maximum value. In addition, two one-dimensional plots (without clipping of the maximum amplitudes) were generated from the two-dimensional figures as well. Once more, both plots were extracted from the position of the true maximum. From Figures 1(a) and 1(b), the observations from Jäger et al. (2001) can be generally confirmed: even though the objective function features several local extrema, the function shows a clearly distinguishable global maximum in all of the planes and, even more importantly, the extension and magnitude of the global maximum is considerably larger than any of the local extrema. Furthermore, the one-dimensional plots (Figures 2(a) and 2(b)) suggest that the objective function is likely to be non-smooth in large parts of the search space, but the global maximum itself is of quadratic shape and appears to be smooth. The large number of local extrema that is evident in Figure 2(a), as well as large areas of the search space where the function values are close to zero (Figure 2(b)), emphasize why local optimization optimization methods without starting values in the vicinity of the maximum do not converge towards the true maximum.



(a) Semblance function in dependence of R_N.

(b) Semblance in dependence of R_{NIP}.

Figure 2: One-dimensional semblance plots, extracted from the three-dimensional semblance function at the position of the global maximum.

These observations suggest that an efficient strategy for optimizing the semblance function is to evaluate the objective function on a coarse grid and therefore approximately locate the position of the absolute maximum. Then, a local optimization technique, which assumes a quadratic shape of the objective function in the vicinity of the global extremum, can be applied for precisely isolating the maximum. In the following section, we describe one possible approach for this based on the conjugate direction method.

CONJUGATE DIRECTION METHOD WITH COARSE GRID LINE SEARCH

For obtaining the optimal triplet of stacking parameters that provide the highest semblance, the following unconstrained optimization problem needs to be solved for every time sample:

$$\max_{\alpha,R_N,R_{NIP}} f(\alpha, R_N, R_{NIP}) = \frac{\sum_t \left(\sum_i g_{ti}(\alpha, R_N, R_{NIP}) \right)^2}{\sum_t \sum_i \left(g_{ti}(\alpha, R_N, R_{NIP}) \right)^2}. \tag{3}$$

Here we chose a slightly different notation including the arguments α, R_N, R_{MIP} in order to emphasize the dependence of the objective function on the CRS parameters. To shorten the following expressions, we define the variable $\mathbf{x}$ of the objective function as $\mathbf{x} = (\alpha, R_N, R_{NIP})$. The function is optimized by performing subsequent one-dimensional line searches along a series of search directions $\mathbf{d}_n$. Instead of using the negative gradient as the search direction in the current iteration (i.e. the method of steepest descend), we use Powell's method of conjugate directions, which guarantees that a quadratic n-dimensional function is optimized after n iterations along conjugate directions (Powell, 1965).

We start by optimizing f using the unit vectors $\mathbf{e}_n$ as initial directions. By performing a line search in the current direction $\mathbf{d}_n$, the maximum of f in this direction is located at $\mathbf{x}_n$, so the gradient in the direction $\mathbf{d}_n$ vanishes at this position (Press et al., 2007):

$$\mathbf{d}_n \cdot \nabla f(\mathbf{x}_n) = 0, \tag{4}$$

which is equivalent to saying that the gradient of f at $\mathbf{x}_n$ is perpendicular to $\mathbf{d}_n$. In order to preserve the optimization along the first direction, we require that the gradient after the next line search in a new direction $\mathbf{d}_{n+1}$ is still perpendicular to the first direction:

$$\mathbf{d}_n \cdot \nabla f(\mathbf{x}_{n+1}) = 0. \tag{5}$$

Expanding f into a Taylor series around $\mathbf{x}_n$ and taking the derivative yields

$$\nabla f(\mathbf{x}_n + \delta\mathbf{x}) = \nabla f(\mathbf{x}_n) + \mathbf{A} \cdot \delta\mathbf{x}, \tag{6}$$

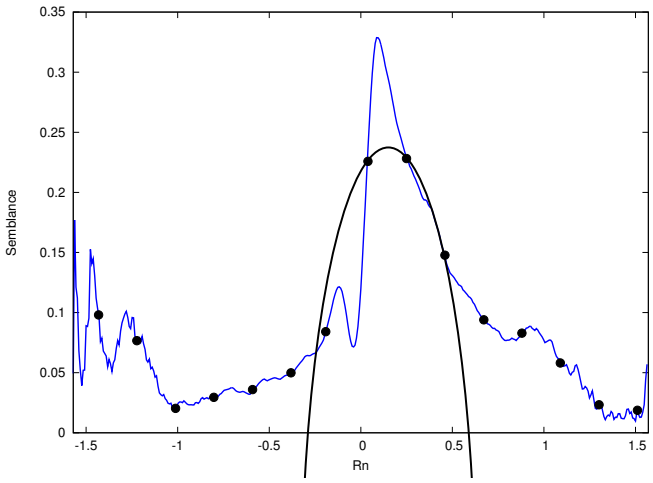


Figure 3: The semblance function evaluated along a single search direction $\mathbf{d}_n$ at evenly spaced grid points within the limited search interval. The grid point with the highest function value and its neighbouring points are used to estimate the position of the function's maximum by inverse parabolic interpolation. The maximum of the fitted parabola is used as the starting point for an exact line search.

where $\mathbf{A}$ is the Hessian matrix. Replacing $\mathbf{x}_{n+1}$ by this expression, we obtain

$$\mathbf{d}_n \cdot \mathbf{A} \cdot \mathbf{d}_{n+1} = 0. \tag{7}$$

If Equation 7 holds true, the directions $\mathbf{d}_n$ and $\mathbf{d}_{n+1}$ are conjugate. Since we do not want to compute $\mathbf{A}$ explicitly, we use Powell's method in which a set of subsequent conjugate directions is obtained by discarding the largest direction of the last k iterations and replacing it by the average direction travelled in these k iterations, where k is the number of dimensions of f. By doing this, the quadratic convergence of the method is lost, but neither the gradient nor the Hessian needs to be calculated and the algorithm converges superlinearly.

Using this strategy to calculate new sets of directions, f is optimized iteratively by taking steps along the current direction:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha \cdot \mathbf{d}_n \tag{8}$$

As mentioned, the step length α is generally calculated using a line search algorithm, such as Brent's method or an approximate line search where α has to satisfy some conditions, e.g. the Wolfe conditions (Nocedal and Wright, 1999). Considering our observations from the last section concerning the semblance function, however, we propose to use a different strategy for the line search.

First, we limit our parameter search space by introducing upper and lower bounds for the CRS parameters α, R_N and R_{NIP}. These bounds are chosen in correspondence to their physical properties, i.e. the angle of emergence of the normal ray has to lie in the range of $\pm\pi$, and the radius R_{NIP} of the eigenwave from the hypothetical point source experiment cannot be negative. However, the exact choice of the constraints may depend on the actual problem and choosing the bounds effectively is still subject to current research. By introducing these bounds, the search space in which the semblance maximum lies is reduced to a cube, or more generally to a hexahedron. The next step is to take the current search direction of the optimization algorithm $\mathbf{d}_k$ and to calculate the intersections of the direction vector with the boundaries of the search space. Then the objective function is evaluated along the current direction for a finite number of equally spaced nodes within the search space. This approach is shown in Figure 3, where the semblance function from the previous examples is coarsely evaluated along the R_N direction.

Since the global maximum in extent and magnitude is larger than any of the local extrema, 15 function evaluations in this case are enough to approximately locate the maximum. Because of the maximum's quadratic shape, we can use the node with the highest function value and its neighbouring points to perform an inverse parabolic interpolation and use the maximum of the fitted parabola as the initial guess for an accurate line search.

Summing up the optimization strategy, we use the coarse line search in every iteration to approximately locate the maximum and then perform an exact line search starting from the maximum of a parabola fitted through the nodes with the highest function values. For optimal convergence, the line searches are performed along subsequent conjugate directions, which are calculated using Powell's method. With this approach, no initial values for the CRS parameters are required, which means that no pragmatic approach needs to performed. In the following section, our method is tested on the synthetic Sigsbee 2A data set.

DATA EXAMPLE

To evaluate the performance of the proposed optimization method, we compare the results of processing the Sigsbee 2A data set, using the pragmatic approach together with Powell's method and a standard line search algorithm (Brent's method) against Powell's method with the coarse line search. Since for the latter, no starting values from the pragmatic approach are used, we need user supplied starting values and a set of initial search directions for the algorithm.

For the choice of the initial set of directions $\mathbf{d}_k$ (k=1,2,3), the strategy from Powell's method will be kept, in which the unit vectors are used as initial directions. The more crucial aspect is the choice of the initial guess for **x**, which is the initial set of CRS parameters. From the conclusions regarding the shape of the semblance function, it is clear that the direction $\mathbf{d}_1$ for the initial line search needs to lie in the vicinity of the global maximum. Even though it is not necessary that **d** cuts exactly through the maximum, it must not completely miss the elevation that is associated with the maximum either. However, in the following example we will show that the coarse line search optimization is not very sensitive to the initial parameters and works as long as the starting values are not totally unrealistic in a physical sense. For the angle of emergence α, a small value close to zero degree will be used as the initial parameter in the following tests. For a horizontal reflector, this guess will always be very close and for a dipping reflector, a small angle close to zero works just as well for negative as for positive dips. For the parameters R_N and R_{NIP}, the choice of the initial values is even less decisive. To constrain them for the optimization, both parameters are transformed using a tangent function (Mann, 2002)

$$R_{transform} = \arctan\left(\frac{r}{R}\right) \tag{9}$$

after which the semblance maximum is usually close to zero for these parameters (as is evident in Figure 1(a)). R is the input parameter R_N or R_{NIP} and r is the transformation radius. All processing parameters for the data example are shown in Table 1. In the data example, we use the same initial values for the CRS parameters at every sample.

A comparison of the semblance plots from the conventional CRS workflow (4) and the coarse line search optimization (5) reveals the superior performance of the new method. In the parts of the section with horizontal reflections, both methods perform similarly well, but in the areas of weak signal, the coarse line search optimization is able to obtain higher semblance values. One example for this observation is the part around CMP number 500 and four seconds TWT. In case of the semblance from the conventional CRS workflow, several samples that correspond to reflections and especially diffractions show semblance values of zero, where the optimization failed. For this reason, the section looks discontinuous and noisy. On the other hand, the semblance from the coarse line search optimization looks a lot smoother and the semblance along events is mostly continuous.

Another area where the CGS optimization was able to find more coherent events is the sub-salt part. Apparently both methods did not recover any reflections in this part (CMP number 500 – 2000 and travel times larger than 4–5 seconds), but our approach could recover more diffractions than the conventional workflow. Still, it has to be noted, that the semblance in this area is mostly zero.

The R_N parameter section from the conventional workflow corresponds mostly to the semblance section. This means that the parameter estimation worked well in areas with high semblance, but in the sub-salt area, hardly any reliable parameter values exist. In contrast to this, the estimation of R_N using the coarse line search optimization worked even in the sub-salt area, even though the obtained parameter values correspond to diffractions only and not to any reflections. Figure 7 shows that the coarse line search optimization was able to estimate parameters for several events that are entirely absent in the parameter section of the pragmatic approach. Apparently the differences between the two optimization methods is

Surface velocity	1500 m/s
Average velocity	2000 m/s
Coherence band	7 samples (28 ms)
Mean frequency	20 Hz
Minimum CMP aperture	2000 m at 0.0 s
Maximum CMP aperture	4000 m at 12.0 s
Minimum ZO aperture	200 m
Maximum ZO aperture	300 m
Optimization tolerance ϵ	10^{-2}
Max number of iterations	100
Number of nodes for CGS	100
R_N transformation radius	100 m
R_{NIP} transformation radius	10 km
initial α value	$-\pi/8$
initial R_N value	15 km
initial R_{NIP} value	5 km

Table 1: CRS processing parameters for the test runs in which the coarse line search optimization was compared to the pragmatic approach with Powell's conjugate direction method using a standard line search. Note that the R_{NIP} parameter transformation was only done for the coarse line search optimization.

more influential to the parameter section, than to the semblance section and our new optimization is able to perform a good parameter estimation even in areas with a very low semblance. Furthermore we observe that even though we use the same intial values for the CRS parameters throughout the entire section, the optimization provides better results than the pragmatic approach with Powell's method using a standard line search.

In terms of computational time, the coarse line search optimization took approximately ten times longer than the regular Powell's method together with the pragmatic approach. In contrast to the example figure in the previous section (Figure 3), the number of function evaluations along the search direction had to be larger and satisfying results could be obtained with function evaluations at approximately 100 grid points.

Method	Calculation time [hours]
Pragmatic approach + Powell's method	0.32 + 3.79 (total 4.11)
Coarse line search optimization	40.32

Table 2: Calculation times for simulating a ZO CRS stack with the conventional method and the coarse line search optimization for the Sigsbee data set. For the conventional CRS run, the calculation time for the initial parameter sections and the final optimization run are given.

CONCLUSIONS

In this work, we introduced an alternative strategy for optimizing the semblance function in the common-reflection-surface stack workflow. Instead of relying on initial parameters using the pragmatic approach by Mann et al. (1999) and Jäger et al. (2001), we use a direction set method that consists of a series of one-dimensional line searches. For every line search, the objective function is coarsely evaluated along the current direction within the bounded search space. As was shown with example plots of the semblance, the objective function generally features a broad global maximum, so that it can be approximately located using the coarse line search. Once we are in vicinity of the maximum where the objective function is smooth, an exact line search is used to precisely isolate the extremum.

One major drawback of the pragmatic approach is, that the initial parameter values are obtained by three one-parameter searches in a subspace of the common reflection surface operator only. In particular, this means that, e.g., α is estimated by performing a 1D search for a fixed half offset ($h = 0$ m), which means the semblance is not evaluated by stacking data along the entire CRS but only along a one-dimensional

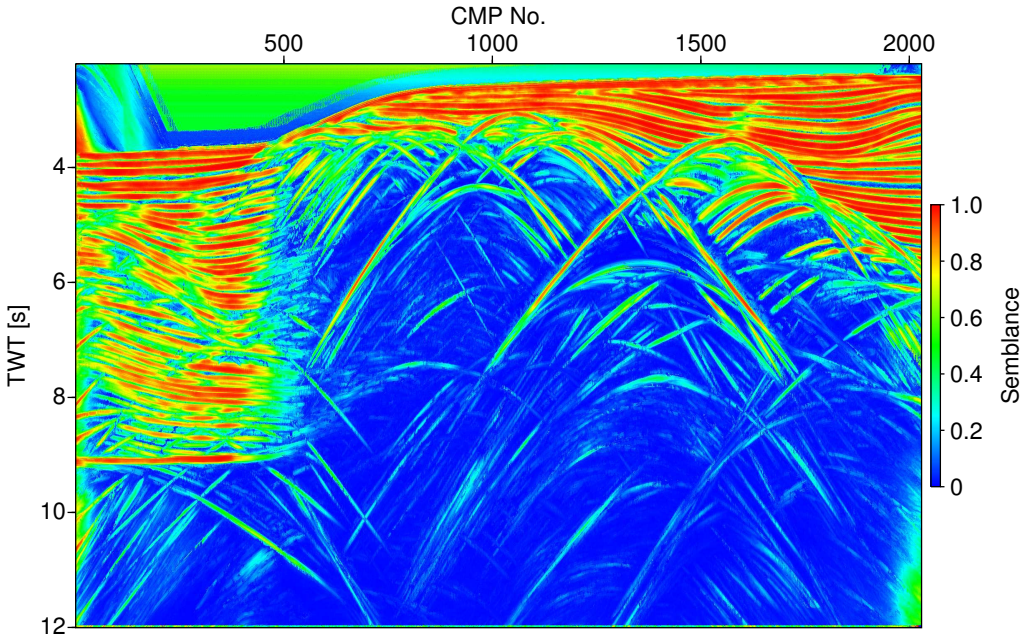


Figure 4: Semblance section that contains the maximum semblance value for every sample as obtained by the optimization. Here, the pragmatic approach was used to calculate initial values for the CRS parameters, which were used in a final multi-parameter optimization using Powell's method with a standard line search.

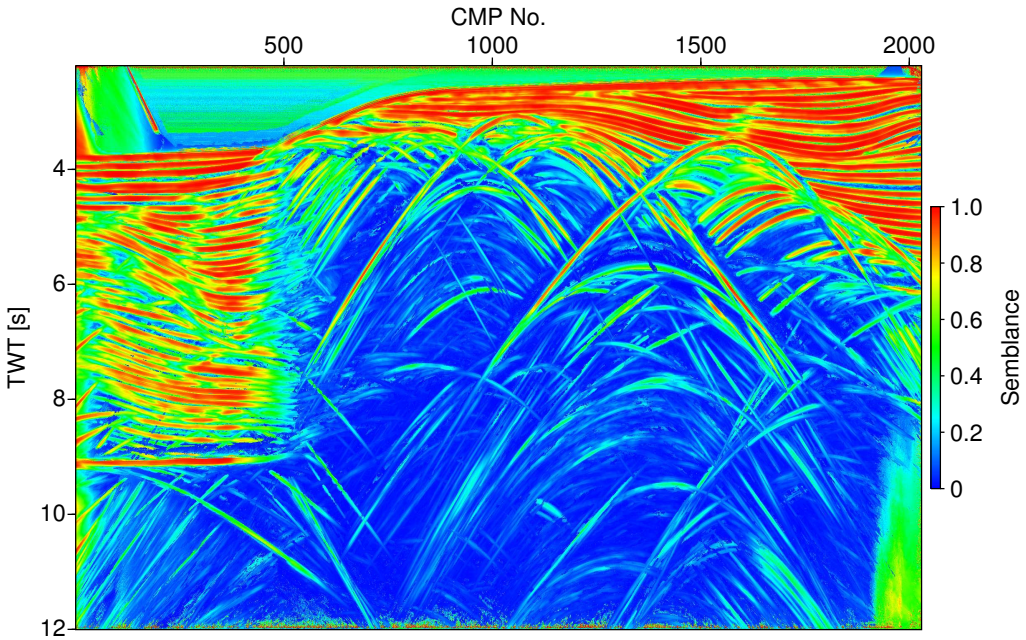


Figure 5: Optimization result of our proposed coarse line search method, in which the same constant starting values are used for every sample. We then perform coarse line searches to approximately localize the semblance maxima and use an exact line search for precisely isolating it. Once again, Powell's method is used to calculate the search directions.

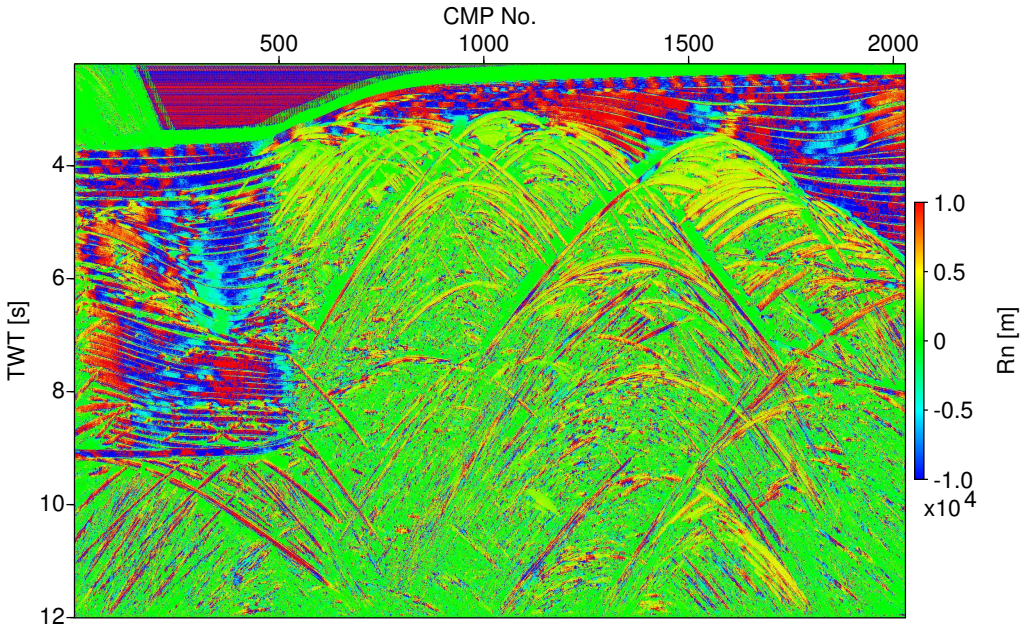


Figure 6: R_N section using the pragmatic approach and Powell's method with exact line search. Especially in the subsalt region, very little signal seems to be recovered and reliable values for R_N are obtained at very few samples corresponding to some of the stronger diffractions only.

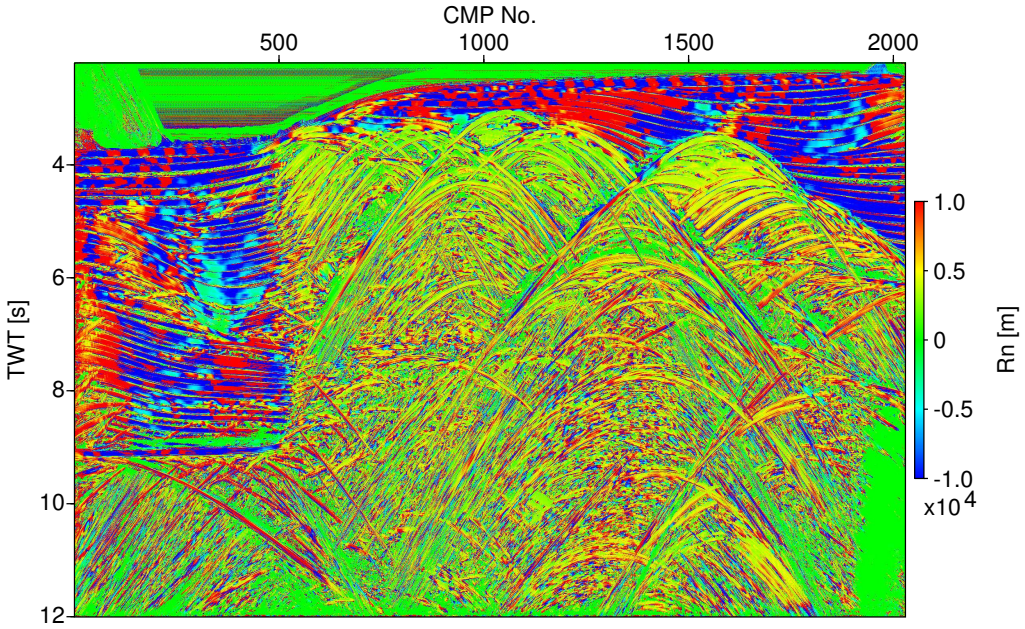


Figure 7: R_N section using the new approach with the coarse grid line search. Even though the semblance (see Figure 5) is very low in the sub-salt region, our method is able to estimate R_N fairly well and a lot more information is recovered in comparison to the R_N section from the pragmatic approach, Figure 6.

travel time curve. This way, a lot less traces contribute to the semblance value and very weak events are likely to be missed. If no good initial parameters are found for a certain sample, then the subsequent local optimization will in most cases converge towards a local maximum.

To fully utilize the strength of the CRS method, that is, to obtain a high data fold by stacking data along the two-dimensional travel time surface, we therefore propose to use an optimization method that not solely depends on initial values from the pragmatic approach, for which we introduced one possible strategy. Applying this method to the Sigsbee data set shows that this new strategy in fact provides better results than a local optimization with standard line searches using initial values from the pragmatic approach.

However, up to this point, our new method needs a fairly large number of function evaluations for the coarse line search to succeed. In the future, we would like to improve our method by coming up with techniques to further bound our parameter search space and a more efficient way of approximately locating the semblance maximum. Overall, we see a high potential in the special properties of the semblance function that should be taken advantage of in the optimization.

ACKNOWLEDGMENTS

We thank the Applied Seismics group in Hamburg for continuous discussion and Claudia Vanelle for proofreading. This work was kindly supported by the sponsors of the Wave Inversion Technology (WIT) Consortium.

REFERENCES

Barros, T., Krummernauer, R., Ferrari, R., and Lopes, R. (2014). Global optimization of the parameters of the common reflection surface traveltime using differential evolution. *76th EAGE Conference*, Exhibition 2014.

Garabito, G., Stoffa, P., Lucena, L., and Cruz, J. (2012). Part I – CRS stack: Global optimization of the 2d CRS-attributes. *Journal of Applied Geophysics*, 85:92–101.

Jäger, R., Mann, J., Höcht, G., and Hubral, P. (2001). Common-reflection-surface stack: Image and attributes. *Geophysics*, 66.

Mann, J. (2002). *Extensions and Applications of the Common-Reflection-Surface Stack Method*. PhD thesis, University of Karlsruhe.

Mann, J., Jäger, R., Müller, T., Höcht, G., and Hubral, P. (1999). Common-reflection-surface stack – a real data example. *Journal of Applied Geophysics*, 42:301–318.

Minato, S., Tsuji, T., Matsuoka, T., Nishizaka, N., and Ikeda, M. (2012). Global optimisation by simulated annealing for common reflection surface stacking and its application to low-fold marine data in southwest Japan. *Exploration Geophysics*, 42:59–69.

Nocedal, J. and Wright (1999). *Numerical Optimization.* Springer Verlag.

Powell, M. J. D. (1965). An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Computer Journal*, 7:155–162.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes, Third Edition.* Cambridge University Press.

Santos, L., Yano, F., Salvatierra, M., Martinez, J., Andreani, R., and Tygel, M. (2005). A global optimization algorithm applied to the Common Reflection Surface (CRS) problem. *Journal of Seismic Exploration*, 14:217–233.

Taner, M. T. and Koehler, F. (1969). Velocity spectra – digital computer derivation and applications of velocity functions. *Geophysics*, 56:654–663.