Stanford Exploration Project, Report 110, September 18, 2001, pages 1–140

Whitening track residuals with PEFs in IRLS approach to the Sea of Galilee

Andrey Karpushin and Morgan Brown[1]

ABSTRACT

We applied an Iteratively Reweighted Least Squares (IRLS) approach to create a map of the Sea of Galilee. We use a bank of Prediction Error Filters (PEFs) as a residual whitener to reduce the acquisition footprint and map artifacts caused by non-Gaussian noise in the data.

INTRODUCTION

The bottom sounding survey data set collected by Zvi Ben-Avraham in 1986-87 on the Sea of Galilee in Israel (Ben-Avraham et al., 1990) is often used at SEP to test new algorithms for data interpolation. The main difficulties associated with interpolating the Galilee data set onto a regular grid are inconsistencies in the data due to random spikes, and an acquisition footprint (ship tracks) left in the final image caused by systematic error in the data.

There are different approaches developed at SEP to confront these difficulties (Claerbout, 1999; Fomel and Claerbout, 1995; Fomel, 2001; Brown, 2001). Fomel and Claerbout (1995) suggested to use a derivative operator to filter out low-frequency components of the residual, and the Iteratively Reweighted Least Squares (IRLS) technique to suppress non-Gaussian spikes in the data. However, they showed that while the acquisition footprint and the noisy portion of the model disappear, the price of the improvement is a loss of image resolution.

In this paper we implement an idea from Claerbout (1999), and use a bank of a Prediction Error Filters instead of the derivative operator to whiten the residual on individual data tracks before implementing IRLS. We show that our algorithm leads to reduced artifacts in the final map without an apparent decrease in map resolution.

DATA DESCRIPTION

The Galilee data set includes more than 131,500 (x_i, y_i, z_i) triples where x_i varies over about 12 km and y_i varies over about 20 km. The water depth was measured using an Odom Echtrack DF3200 Echosounder, and the position of the boat at the time of the measurement was determined using a Motorola Miniranger system, which is composed of radio stations on shore

[1]**email:** andrey@sep.stanford.edu, morgan@sep.stanford.edu

and a recorder on the boat. Navigation was recorded using "event marks" rather than time, and no information on time is available. Although the precision of the acquisition system is high, there are some defects in the data caused by an occasional malfunction of the equipment. This creates artifacts in the final map, and makes processing of this dataset a challenging problem. Another source of artifacts in the model is the changing conditions of the acquisition, which create a strong footprint in the final map.

METHOD

The problem of interpolating irregularly sampled data, like the Galilee data set, onto a regular grid to produce a map can be written in terms of the fitting goals of an inverse linear interpolation problem (Claerbout, 1999):

$$\begin{aligned} \mathbf{0} &\approx \mathbf{Lm} - \mathbf{d} \\ \mathbf{0} &\approx \epsilon \mathbf{Am} \end{aligned} \tag{1}$$

L is the linear operator which maps data onto the map. Usually **L** is either a binning or a bilinear interpolation operator. The second equation in system (1) is a regularization term where **A** is a roughening operator that imposes smoothness of the model in this underdetermined problem, at the price of fitting data exactly. Everywhere below, **L** is a binning operator (**B**) and **A** denotes a gradient filter consisting of two first order derivatives.

To suppress the artifacts caused by non-Gaussian noise in the data, Fomel and Claerbout (1995) introduced a weighting operator **W**:

$$\begin{aligned} \mathbf{0} &\approx \mathbf{W}(\mathbf{Bm} - \mathbf{d}) \\ \mathbf{0} &\approx \epsilon \mathbf{Am} \end{aligned} \tag{2}$$

The choice of the weighting operator **W** follows two formal principles:

1. Statistically bad data points (spikes) are indicated by large values of the residual $\mathbf{r} = \mathbf{Bm} - \mathbf{d}$
2. Abnormally large residuals attract most of the conjugate gradient solvers effort, directing it the wrong way. The residual should be whitened to distribute the solvers attention equally among all the data points to emphasize the role of the "consistent majority"

Based on these principles operator **W** in equation (2) was chosen by Fomel and Claerbout (1995) to include two components: the first derivative filter **D** taken in the space along the record tracks and the diagonal weighting operator $\tilde{\mathbf{W}}$.

$$\begin{aligned} \mathbf{0} &\approx \tilde{\mathbf{W}}\mathbf{D}(\mathbf{Bm} - \mathbf{d}) \\ \mathbf{0} &\approx \epsilon \mathbf{Am} \end{aligned} \tag{3}$$

Fomel and Claerbout (1995) chose weighting function to be $\tilde{\mathbf{W}}_\mathbf{i} = \tilde{\mathbf{W}}(\mathbf{r_i}) = \frac{2\bar{\mathbf{r}}}{|\mathbf{r_i}|+\overline{\mathbf{r_i}}}$, where $\bar{r}$ stands for the median of the absolute values from the whole dataset, and $\overline{r_i}$ is the median in a small window around a current point r_i. Since $\tilde{\mathbf{W}}$ depends upon the residual, the inversion problem becomes non-linear and system (3) can be solved using a piece-wise linear approach (Fomel and Claerbout, 1995). However, Fomel and Claerbout (1995) showed that while the noisy portion of the model disappeared, the price of the improvement is a loss of the image resolution.

In our approach we use a bank of PEFs to decorrelate the residual. Using Prediction Error Filters as a residual whitener better satisfy the second of the formal principles used by Fomel and Claerbout (1995). Since the character of the systematic errors in the data may vary in time and upon the location of the ship, an individual PEF $\mathbf{P_i}$ is estimated for each data track from the residual obtained after solving system (1). In this case we defined a data track as a series of measurements recorded with a distance less than 100 meters between consecutive data points. System (3) then becomes:

$$\begin{aligned} \mathbf{0} &\approx \tilde{\mathbf{W}}\mathbf{P}(\mathbf{Bm}-\mathbf{d}) \\ \mathbf{0} &\approx \epsilon\mathbf{Am} \end{aligned} \tag{4}$$

$\mathbf{P}$ is an operator composed from PEFs $\mathbf{P_i}$. After a bank of PEFs is estimated we solve this non-linear problem [equation (4)] in the manner of piece-wise linearization similar to Fomel and Claerbout (1995). The first step of the piece-wise linearization is the conventional least squares linearization. The next step consists of reweighted least squares iterations made in several cycles with reweighting applied only at the beginning of each cycle. We chose the weighting function $\tilde{\mathbf{W}}$ to be $\tilde{\mathbf{W}}_\mathbf{i} = \tilde{\mathbf{W}}(\mathbf{r_i}) = \frac{\mathbf{1}}{(\mathbf{1}+\mathbf{r_i}^2/\overline{\mathbf{r_i}}^2)^{1/4}}$ (Claerbout, 1999).

In the next section we compare the maps obtained by the three different methods.

RESULTS

Figure 1 shows a roughened image of the lake bottom constructed by using the three methods discussed above. All of the three models were computed with the same value of damping parameter (ϵ) in a linear step of the IRLS iterations to make a fair comparison.

The top panel of Figure 1 shows the result corresponding to system (2), implementing IRLS without whitening the residual. This panel has a strong acquisition footprint; some "spikes" are still present in the model, even though IRLS managed to attenuate most of them. At the same time this panel has a good resolution and we can see some fine features of the bottom, especially in the southern part of the lake where the acquisition footprint is weaker. The central panel of Figure 1 shows the result of solving system (3) with a derivative operator along the tracks as a residual whitener. This panel has no acquisition footprint but has a lower resolution as was observed by Fomel and Claerbout (1995). Although this method preserved main features of the bottom, some features such as a ledge marked in the image are not present in the model. The third panel presents the result of our algorithm. As we can see almost all of the ship's tracks and spikes are suppressed without a visible decrease in resolution. For example, the ledge we observed before is still present in the model. It is interesting to notice that the scarp is even easier to see here than in the first image, where it is obscured by the acquisition footprint.

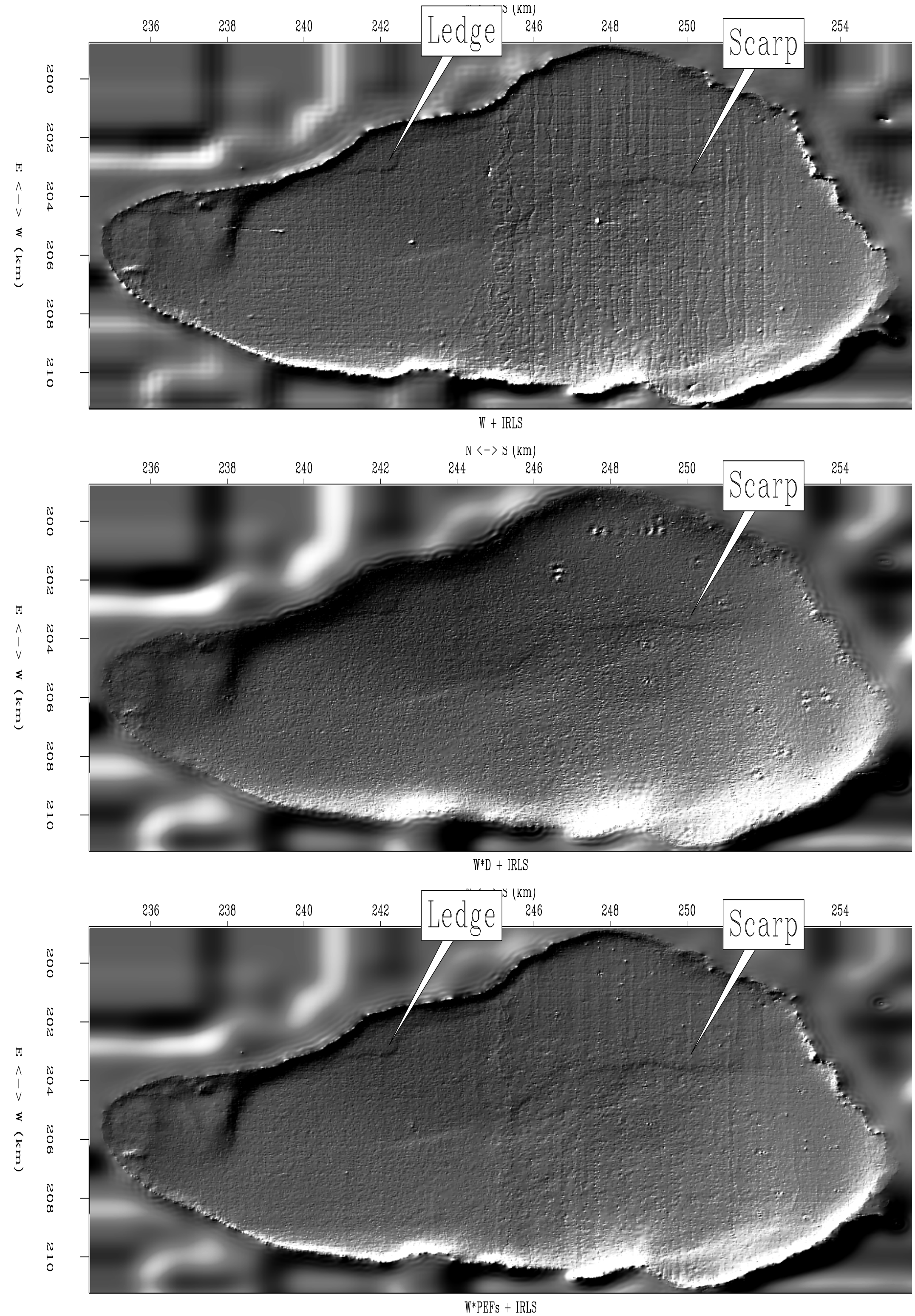

Figure 1: Comparison of the roughened results obtained by the three methods. Top: Result of IRLS with no residual whitener (system 2). Central: IRLS with a derivative operator as a residual whitener (system 3). Bottom: Our algorithm with a bank of PEFs as a residual whitener (system 4) andrey-comp_1_ann [ER]

Figure 2 shows the difference among the results computed with the three algorithms discussed above.

The top panel shows the difference between the maps computed with IRLS without whitening the residual and IRLS with a derivative operator as a residual whitener. On this panel we can see ship tracks and features of the bottom which were attenuated from the final map. The central panel shows the difference between the maps computed with IRLS without whitening the residual and IRLS with a bank of PEFs as a residual whitener. Geological features are much less visible in this panel, but it is easy to see the acquisition footprint that our method removed. The third panel shows the difference between the final maps generated with the use of a derivative and a bank of PEFs to whiten the residual. On this panel we can still see some traces of the acquisition footprint which were left in the final map by our method. It is also easy to see the geological features of the bottom of the lake which were preserved by our method.

CONCLUSIONS

We presented and tested a new method of interpolation of tracked datasets with non-Gaussian errors onto a regular grid. We showed that with the Galilee dataset, our method was able to remove artifacts caused by irregular spikes in the data as well as the acquisition footprint caused by systematic errors in the data. We think that this method is directly applicable to other problems where measurements are taken along the tracks and data have errors due to equipment malfunctioning and systematic errors due to slowly changing acquisition conditions.

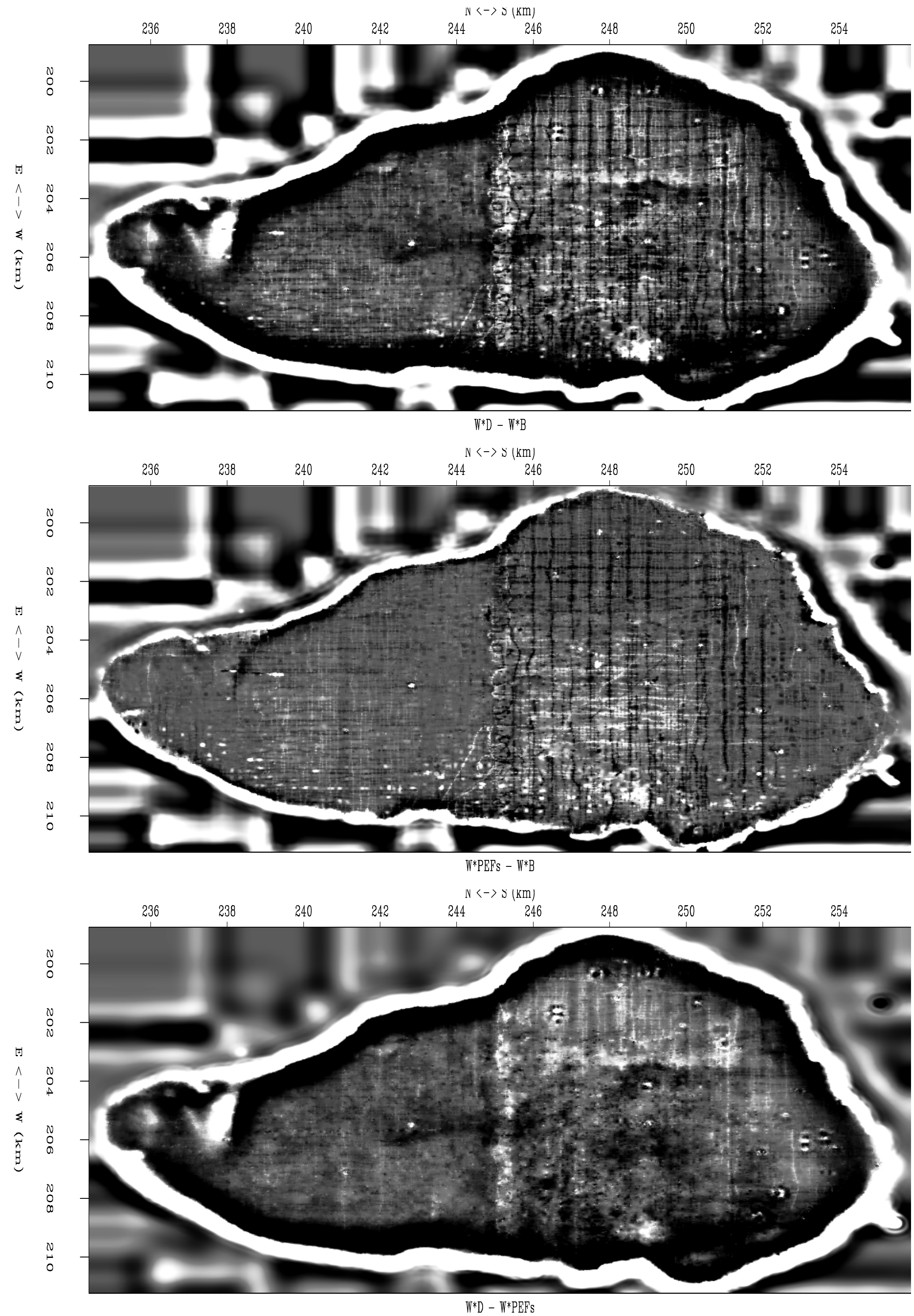

Figure 2: Difference in the results obtained by the three methods. Top: Difference between IRLS without whitening the residual and with a derivative operator as a whitener. Central: Difference between IRLS without whitening the residual and IRLS with a bank of PEFs as a residual whitener. Bottom: Difference between IRLS with a derivative and a bank of PEFs as a whitener. andrey-comp_2 [ER]

REFERENCES

Ben-Avraham, Z., Amit, G., Golan, A., and Begin, Z., 1990, The bathymetry of Lake Kinneret: Israel J. Earth Sci, **39**, 77–84.

Brown, M., 2001, Estimation of systematic errors in tracked datasets using least squares crossing point analysis: SEP–**110**, 123–132.

Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, `http://sepwww.stanford.edu/sep/prof/`.

Fomel, S., and Claerbout, J., 1995, Searching the Sea of Galilee: The splendors and miseries of iteratively reweighted least squares: SEP–**84**, 259–270.

Fomel, S., 2001, Three-dimensional seismic data regularization: Ph.D. thesis, Stanford University.