Chapter 3

Interpolation of irregularly sampled data

Most modern seismic acquisition methods aim to sample data regularly along all axes. Deviations from this sampling happen for various reasons. On land and ocean-bottom cable data, obstacles and terrain cause both sources and receivers to be moved. An example of the distribution of sources in an ocean-bottom cable (OBC) survey is shown on Figure 3.1a. The source positions are not evenly distributed because of the source ships both firing irregularly and not sailing in straight lines. There is also a large gap in the center of the survey because of an ocean platform. The ocean-bottom receiver cable positions in Figure 3.1b show that the receiver cables are not straight and are unevenly spaced in the crossline. Some acquisition designs intentionally acquire data with irregular sampling. One example of this is an ocean-bottom node array, where the nodes may be randomly placed on the sea floor. Land data can be randomly sampled to reduce acquisition footprint (Zhou and Schuster, 1995).

While data are often irregularly sampled, many of the algorithms that we wish to apply to seismic data, such as a spatial Fourier transform or a finite-difference migration, are best-suited for data on a regular grid. Interpolation methods for irregular data typically are based upon either move-out or partial prestack migration

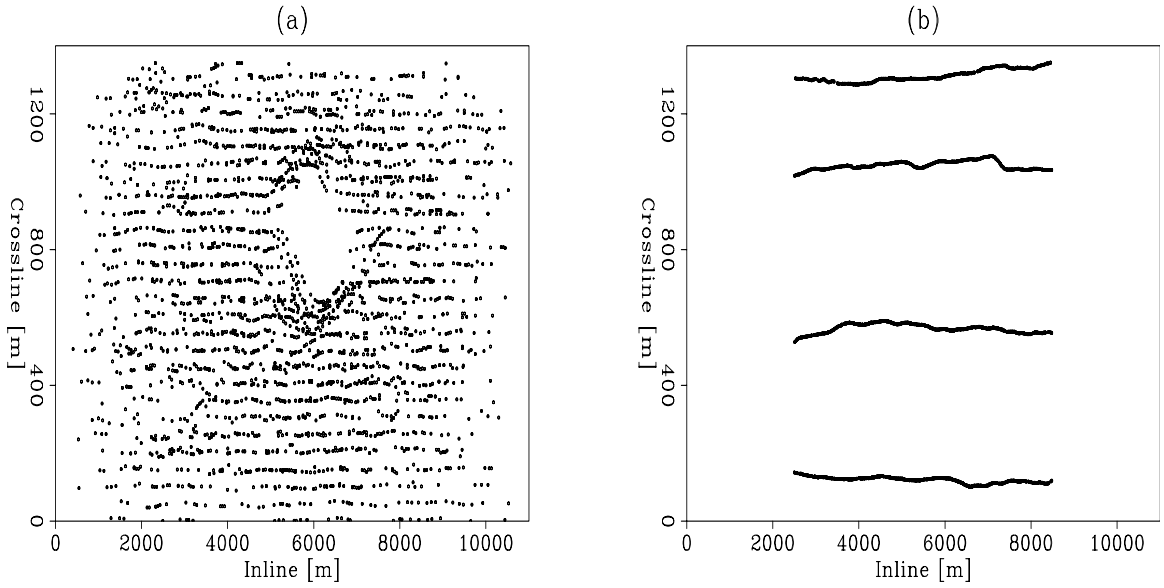

Figure 3.1: An example of source positions (left) and receiver positions (right) from an OBC survey. Figure courtesy of Daniel Rosales. **NR** MSPEF/. acquisition

based methods (Chemingui, 1999; Fomel, 2001; Clapp, 2003) that require a root-mean-square velocity model, or Fourier or Radon transform-based methods (Gulunay and Chambers, 1996; Schonewille and Duijndam, 1998; Liu and Sacchi, 2004; Xu et al., 2005), that typically require unaliased data. Aliased data may be interpolated in f-x (Spitz, 1991), but this method requires regularly-sampled data. As described in Chapter 2, a prediction-error filter can be used to interpolate data on a regular grid with missing samples. This PEF can interpolate many simultaneous slopes that can be aliased, which is often the case with irregularly sampled data. However, a prediction-error filter is also dependent upon regularly-sampled training data. In practice, we use nearest-neighbor interpolation to place the irregular sampled data on to a regular grid with both known and unknown values. When estimating a PEF, the rows of the autocorrelation matrix that depend on missing samples in the training data are weighted to zero. With many samples missing from the data, we encounter a problem of too many rows being weighted to zero, leaving insufficient information to estimate a PEF.

Previously, in order to add more information to the PEF estimation, the prediction-error filter was spaced out over the data, so that the values of the filter sample every second point of the data along each axis (Crawley, 2000; Claerbout, 2004). This spacing of the filter can be changed to other integer values, and a single multi-scale PEF can be estimated from one copy of the input data. This method requires somewhat regular sampling so that the PEF falls upon entirely known data at some filter spacing. Also, care must be taken to ensure that the filter spacing in time is not too coarse to alias the data. Alternatively, instead of changing the spacing of the filter on the original data, I generate training data with fewer holes by regridding the sparse data onto a coarser grid so that more values are nonzero (known), so fewer rows of the autocorrelation matrix will be zeroed (Curry and Brown, 2001). Since the sampling varies with position, different areas are best captured by different grids, both in cell size and position. In order to use as much of the data as possible, I estimate a single PEF from multiple regridded copies of the data, and then use this PEF to interpolate the missing data. I test this on a random sampling of traces from the two-slope example introduced in Chapter 2.

This multi-grid approach to training data can also be used for nonstationary PEF estimation (Curry, 2002). Multiple grids are even more useful here as many more unknown coefficients must be estimated than in the stationary case. Also, because different areas have different sampling and contain unique information because of this nonstationarity, multiple grids are needed to extract this information. To test this approach, I randomly sample the quarter-dome synthetic data from the previous chapter, and estimate a nonstationary PEF from multi-gridded data, with more regridded copies of data producing better results. I then perform this multi-grid estimation on 2D land data in the source-offset domain to produce a PEF that fills most gaps and improves velocity supergathers.

PREDICTION-ERROR FILTER ESTIMATION WITH MISSING DATA

In Chapter 2, I estimate a multi-dimensional prediction-error filter, $\mathbf{f}$, of length n_{f} (including the leading 1) using regularly-sampled training data, $\mathbf{d}$, of length n_{d} by solving a least-squares problem,

$$\mathbf{Kf} = -(\mathbf{D}^{\dagger}\mathbf{D})^{-\mathbf{1}}\mathbf{D}^{\dagger}\mathbf{d}, \tag{3.1}$$

where the unknown filter coefficients, $\mathbf{Kf}$, are estimated from the training data, a convolutional matrix, $\mathbf{D}$, and its adjoint , $\mathbf{D}^{\dagger}$, constructed from elements of $\mathbf{d}$. When data values are unknown, we premultiply $\mathbf{D}$ with a $n_{\mathrm{d}} \times n_{\mathrm{d}}$ diagonal matrix, $\mathbf{W}$, that multiplies rows with unknown data by zero and all others by one. Replacing $\mathbf{D}$ in equation 3.1 with $\mathbf{WD}$ and recognizing that $\mathbf{W}$ is a symmetric, idempotent matrix ($\mathbf{W}^2 = \mathbf{W}^{\dagger} = \mathbf{W}$), equation 3.1 becomes

$$\mathbf{Kf} = -(\mathbf{D}^{\dagger}\mathbf{WD})^{-\mathbf{1}}\mathbf{D}^{\dagger}\mathbf{Wd}. \tag{3.2}$$

For example, if a nine-point data vector ($n_{\mathrm{d}} = 9$) with a missing third data point, ($d_3 =?$) were used to estimate a four-component PEF ($n_{\mathrm{f}} = 4$), matrices $\mathbf{WD}$ would be written as

$$\mathbf{WD} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_1 & 0 & 0 & 0 \\ d_2 & d_1 & 0 & 0 \\ ? & d_2 & d_1 & 0 \\ d_4 & ? & d_2 & d_1 \\ d_5 & d_4 & ? & d_2 \\ d_6 & d_5 & d_4 & ? \\ d_7 & d_6 & d_5 & d_4 \\ d_8 & d_7 & d_6 & d_5 \\ d_9 & d_8 & d_7 & d_6 \end{bmatrix}. \tag{3.3}$$

In addition to weighting equations with missing data to zero, we can also weight equations where the filter rolls off the known data to zero, so that for the same example

$$\mathbf{WD} = \left[\begin{array}{ccccccccc} 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ \hline 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ \hline 0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&1 \end{array}\right] \left[\begin{array}{cccc} d_1 & ? & ? & ? \\ d_2 & d_1 & ? & ? \\ ? & d_2 & d_1 & ? \\ \hline d_4 & ? & d_2 & d_1 \\ d_5 & d_4 & ? & d_2 \\ d_6 & d_5 & d_4 & ? \\ \hline d_7 & d_6 & d_5 & d_4 \\ d_8 & d_7 & d_6 & d_5 \\ d_9 & d_8 & d_7 & d_6 \end{array}\right]. \tag{3.4}$$

Introducing a single unknown point in the training data in equation 3.3 causes n_{p} rows of the **D** matrix to be ignored, as the unknown point is multiplied by each filter coefficient once. The **W** matrix in equation 3.4 also includes a mute for areas where the filter rolls off of the data boundary. An example of this in two dimensions with one unknown data value and a 7×3 2D PEF is shown in Figure 3.2.

Because each unknown training data value greatly decreases the number of usable rows of the **D** matrix, sparsely sampled training data, where the gridded data have many unknown cells, can have an all-zero **W** matrix, making PEF estimation impossible. For such sparsely sampled data, we can regrid the data with coarser bins so that the number of missing data points (and the number of zero-values in **W**) is reduced, as in Figures 3.3a and 3.3c. We can also shift the origin point of this grid, as in Figures 3.3d-f. Depending on the origin point and the cell size of the grid, the gaps present in the data will differ. If many different grids are used, more of these gaps will be filled so that more rows of the **W** matrix are non-zero. The regridding procedure used to generate these multiple grids is described next.

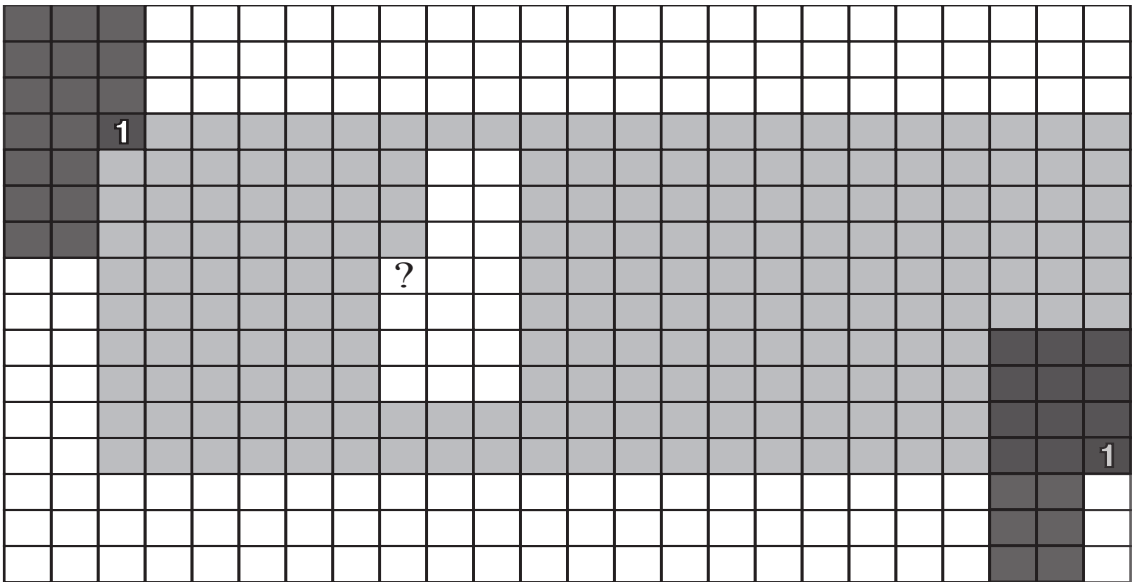

Figure 3.2: An example of missing data and boundary roll-off when estimating a 7×3 (18 coefficient) PEF, shown in black. The 24×16 grid has three rows at both the top and bottom as well as two columns on the left where the PEF falls of the boundary of known data, so those output locations are ignored, denoted with white cells. The one unknown data value, denoted with a ?, causes $n_{\mathrm{p}} = 18$ output points, also shown in white, to be ignored. The remaining usable output locations, where the PEF falls entirely upon known data, are shaded gray. **NR** MSPEF/. PEF-boundaries

ESTIMATION OF A PREDICTION-ERROR FILTER WITH MULTI-GRID DATA

Sparsely sampled data can be represented as a combination of known and unknown values on a desired regularly sampled grid as shown in Figure 3.3a. Mapping from one grid to another is a two-stage process. The first stage is mapping the known values of the original grid to a vector of known points, as in Figure 3.3b. We do this with a sampling matrix $\mathbf{B}$ that is n_{d} elements horizontally and n_{k} elements vertically, so it maps from the n_{d}-length vector containing both known and unknown points to a n_{k}-length vector. These data points can be placed on a different grid with a normalized linear interpolation matrix that maps the known data points onto the $n_{\mathrm{d}_{\mathrm{r}_i}}$-length output grid vector with a matrix, $\mathbf{L}_i$. This new grid is generally coarser and also can have different origin points, as the grids in Figures 3.3c-f do. Cascading these two operators maps data from a fine grid to a coarser grid, so that

$$\mathbf{d}_{\mathrm{r}_i} = \mathbf{L}_i \mathbf{B} \mathbf{d}. \tag{3.5}$$

Applying equation 3.5 to the illustrated example in Figure 3.3a with varying grid origins and sizes produces Figures 3.3c-i, with Figure 3.3b as the intermediate $\mathbf{Bd}$ result. These different grids each provide additional information to the PEF estimation. This regridding is more useful in higher dimensions, where the $\mathbf{L}$ matrix is a bilinear (2D) or trilinear (3D) interpolation matrix. A coarsened version of the data is useful as training data for a PEF because plane waves are somewhat scale-invariant (Claerbout, 2004), so that a PEF estimated on coarser scale data typically contains the same slopes as do the original scale of data. This assumption is not completely satisfied for the frequency content of the data of curved events, where slope is scale dependent.

Regularly-sampled data have a single grid that optimally captures the data: one datum within each bin. Because this chapter addresses irregularly sampled data, no single grid is ideal because different regions of the training data are captured best with different grids. To use irregularly-sampled data effectively, we can estimate a

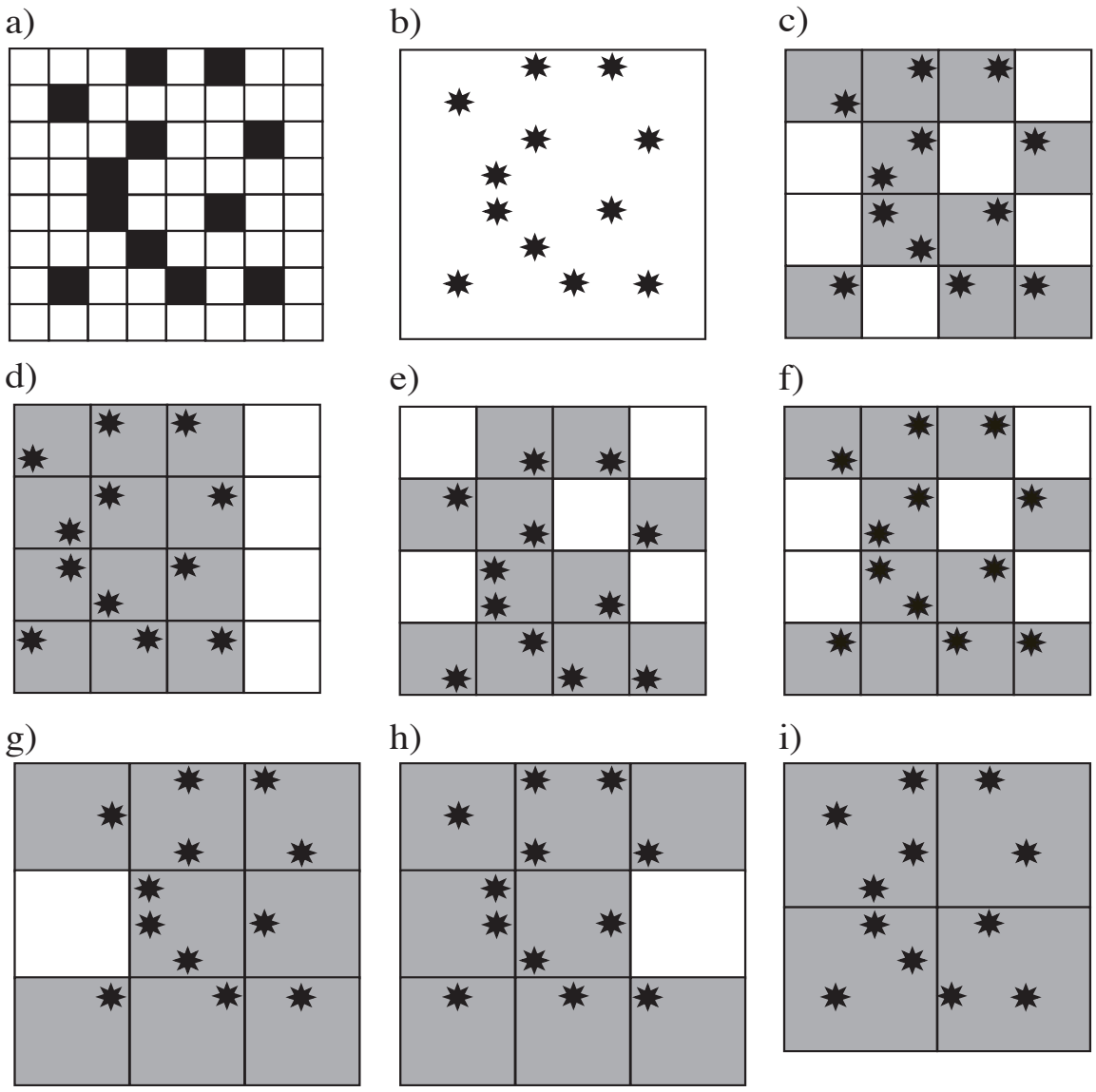

Figure 3.3: Regridding of data with $n_{\mathrm{k}} = 12$ known points on a $n_{\mathrm{d}} = 64$-cell grid. (a): original data on a finely-sampled grid; known data are shaded. (b): data from known cells are sampled with matrix $\mathbf{B}$ and their locations stored. (c)-(i): multiple different grids with varying origin points for multiple cell sizes generated by applying different $\mathbf{L}_i$ matrices. The different grids produce different distributions of known grid values. **NR** MSPEF/. Regridding

single set of PEF coefficients on multiple regridded copies of the data, each grid with a different cell size or origin. Taking equation 3.2 and replacing the $\mathbf{W}$, $\mathbf{D}$, and $\mathbf{d}$ matrices and vectors with their multi-grid equivalents (denoted by a subscripted m), we have

$$\mathbf{Kf} = -(\mathbf{D}_{\mathrm{m}}^{\dagger}\mathbf{W}_{\mathrm{m}}\mathbf{D}_{\mathrm{m}})^{-1}\mathbf{D}_{\mathrm{m}}^{\dagger}\mathbf{W}_{\mathrm{m}}\mathbf{d}_{\mathrm{m}}. \tag{3.6}$$

The data vector $\mathbf{d}_{\mathrm{m}}$ is a concatenation of n regridded versions of the data, $\mathbf{d}_{\mathrm{m}} = [\mathbf{d}_{\mathrm{r}_1}|\mathbf{d}_{\mathrm{r}_2}|\cdots|\mathbf{d}_{\mathrm{r}_n}]^T$, for a total length of $n_{\mathrm{d}_m} = \sum_i n_{\mathrm{d}_{\mathrm{r}_i}}$. Similarly, $\mathbf{D}_{\mathrm{m}}$ is a concatenation of n convolutional matrices, totaling $n_{\mathrm{d_m}}$ rows and n_{p} columns, so that

$$\mathbf{D}_{\mathrm{m}} = \begin{bmatrix} \mathbf{D}_{\mathrm{r}_1} \\ \mathbf{D}_{\mathrm{r}_2} \\ \vdots \\ \mathbf{D}_{\mathrm{r}_n} \end{bmatrix}. \tag{3.7}$$

$\mathbf{W}_{\mathrm{m}}$ is a $n_{\mathrm{d_m}} \times n_{\mathrm{d_m}}$ matrix constructed from multiple versions of the $\mathbf{W}$ matrix in equation 3.2 for each of the rescaled copies of the data, so that

$$\mathbf{W}_{\mathrm{m}} = \begin{bmatrix} \mathbf{W}_{\mathrm{r}_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{\mathrm{r}_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{W}_{\mathrm{r}_n} \end{bmatrix}. \tag{3.8}$$

This approach uses multiple regridded copies of the sampled data as training data, as data on a coarser grid should have fewer holes and better constrain a PEF.

STATIONARY PEF ESTIMATION ON SYNTHETIC DATA WITH MULTIPLE GRIDS

Figure 3.4a is the same two-plane data used as the stationary test case in Chapter 2, and Figure 3.4b is a random sampling of 30 percent of the data traces. Interpolating these data using the same (10×3) 2D PEF estimated in Chapter 2 from fully-sampled

data in Figure 3.4a results in the successful interpolation in Figure 3.4c, where the interpolated data contain the same spectrum as the recorded data, with slightly-lower amplitude where the large gaps were. We now attempt to estimate a PEF using only the sampled data by using the multiple-grid approach to generating training data.

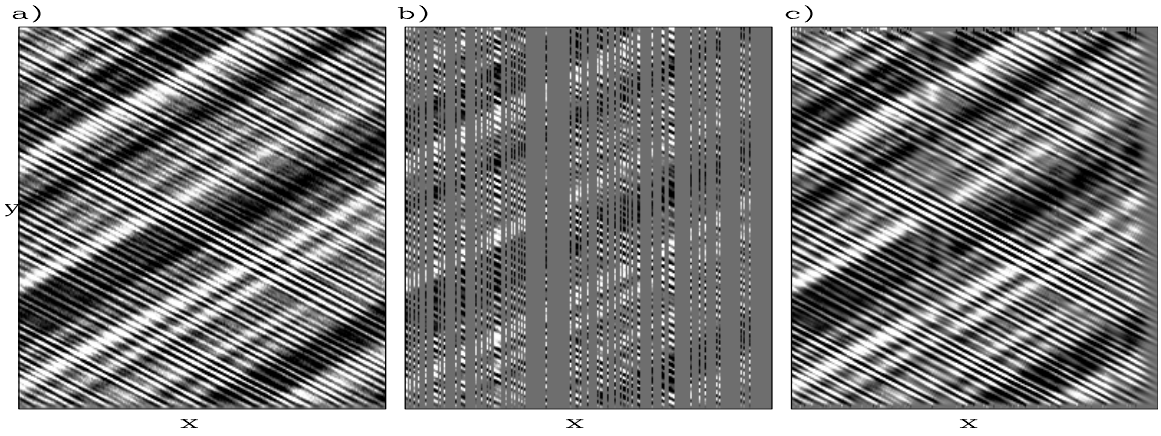

Figure 3.4: The two-plane data from Figure 2.2: (a) fully sampled; (b) with 30% of traces randomly sampled; (c) interpolation of (b) using a PEF estimated on (a). **ER** MSPEF/. idealtraceann

Multiple regridded versions of the sampled data are shown in Figure 3.5, where I vary the coarseness of the grid from the original 256×256 down to 64×64. The higher-frequency event is not present in the coarser grids, as the regridding has effectively acted as a high-cut filter in both dimensions.

I generate seven regridded copies (one 192×192, two 128×128 and four 64×64) of the sampled data using equation 3.5, and place those regridded data along with the original 256×256 data into a 150 thousand-element vector $\mathbf{d}_{\text{m}}$ and the matrix $\mathbf{D}_{\text{m}}$ in equation 3.6. I then use all of these scales of training data to generate a single PEF. I use 100 iterations of a conjugate-direction solver on equation 3.6, and with the PEF thus obtained, I use it to interpolate the missing data using equation 2.18. Comparing this result in Figure 3.6c with both an interpolation using an isotropic Laplacian filter in Figure 3.6a and a PEF estimated on a single 64×64 regridded data in

Figure 3.6b, the multiple-grid result interpolates both the low-frequency and the high-frequency events, whereas the Laplacian filter (minimum curvature interpolation) only interpolated the low-frequency event and the single-grid PEF incorrectly interpolated the high-frequency event with the slope of the low-frequency event.

The multiple-grid PEF, however, produced a poorer result than the interpolation with the ideal PEF (Figure 3.4c), which more completely interpolated the high frequency event. This is not surprising: Figure 3.4c was generated using a PEF estimated on the fully-sampled data, in effect providing the answer to the problem. The differences between Figures 3.4c and 3.6c lie in the approximations used to obtain the training data; that a regridded version of the data could act as a surrogate for the fully-sampled data. The finer of the grids were sufficient for the PEF to capture the high-frequency event, but not with the fidelity of the fully-sampled data, meanwhile the low-frequency event was sufficiently captured in both the multi-grid and the single 64×64 grid, as it was present in all of the grids in Figure 3.5. This example was stationary, with two constant dips. I next adapt nonstationary PEF estimation to use multiple grids of data.

ESTIMATION OF A NONSTATIONARY PREDICTION-ERROR FILTER WITH MULTI-GRID DATA

Multi-grid training data can also be used to estimate nonstationary prediction-error filters, in order to extract as much local information as possible, that is crucial for generating a nonstationary PEF. The multiple grid estimation equations are similar to those for nonstationary PEF estimation in equation 2.21, so

$$\min_{\mathbf{f}_{\rm ns}} \|\mathbf{r}_{\rm d}\|^2 + \epsilon^2 \|\mathbf{r}_{\rm f}\|^2 \tag{3.9}$$

$$\mathbf{r}_{\rm d} = \mathbf{W}_{\rm m}\mathbf{D}_{{\rm ns}_{\rm m}}\mathbf{K}_{\rm ns}\mathbf{f}_{\rm ns} + \mathbf{d}_{\rm m} \tag{3.10}$$

$$\mathbf{r}_{\rm f} = \mathbf{R}\mathbf{f}_{\rm ns}. \tag{3.11}$$

Here, I do not change the nonstationary filter, rather the data being used in the estimation, so the estimation is still the minimization of a nonstationary convolution. In equation 3.11, the data vector, $\mathbf{d}_{\mathrm{m}}$, is now the same $n_{\mathrm{d_m}}$-element vector of concatenated regridded versions of the data from equation 3.6, and the nonstationary data convolutional matrix $\mathbf{D}_{\mathrm{ns_m}}$ is again a concatenation of multiple data-convolutional matrices (this time the nonstationary convolutional matrices from equation 2.21).

One complication to this approach is that each of the component matrices in the $\mathbf{D}_{\mathrm{ns_m}}$ matrix are mapping between the non-stationary filter of length $n_{\mathrm{d}} \times n_{\mathrm{f}}$ to a component of the rescaled data vector $\mathbf{d}_{\mathrm{m}}$, which is of a differing length, $n_{\mathrm{d_{r_i}}}$. I use a subsampling matrix, $\mathbf{S}$, that I apply to the nonstationary filter, $\mathbf{f}_{\mathrm{ns}}$, and subsample it from a $n_{\mathrm{d}} \times n_{\mathrm{f}}$-length vector to a $n_{\mathrm{d_{r_i}}} \times n_{\mathrm{f}}$ length vector, so that $\mathbf{D}_{\mathrm{ns_m}}$ is

$$\mathbf{D}_{\mathrm{ns_m}} = \begin{bmatrix} \mathbf{D}_{\mathrm{ns_{r_1}}} \mathbf{S}_{\mathrm{r_1}} \\ \mathbf{D}_{\mathrm{ns_{r_2}}} \mathbf{S}_{\mathrm{r_2}} \\ \cdots \\ \mathbf{D}_{\mathrm{ns_{r_n}}} \mathbf{S}_{\mathrm{r}_n} \end{bmatrix}. \tag{3.12}$$

Now we have replaced all of the data from equation 3.11 with multiple regridded versions of that data. We also have to include the matrix $\mathbf{W}_{\mathrm{m}}$ from equation 3.6 that zero-weights rows of $\mathbf{D}_{\mathrm{ns_m}}$ containing unknown data.

MULTIGRID ESTIMATION OF A NONSTATIONARY PEF ON SYNTHETIC DATA

Now that nonstationary PEF estimation from multiple grids of data has been covered, I test this approach on the 3D synthetic quarter-dome data from Chapter 2. Recall that I estimated a nonstationary 3D PEF on a fully-sampled version of the quarter-dome synthetic dataset and used it to interpolate a near-perfect result from a severe random sampling of 30 percent of the traces from the data; that was the result, however, of using ideal training data. Now, instead of using ideal fully-sampled training data we start with only the randomly-sampled traces from the data, and we

generate multiple regridded training data from these sparse data.

The quarter-dome data are again shown fully-sampled in Figure 3.7a and with only 30 percent of traces randomly selected in Figure 3.7b. A PEF cannot be estimated on the data in Figure 3.7b because of the lack of contiguous data, making $\mathbf{W}$ zero *everywhere* in equation 3.11. To overcome this problem, the randomly-sampled data are regridded to four different bin sizes in Figures 3.8a-d, where the bins are anywhere from 25 to 400 percent larger along each axis. As the size of the bins increases, the holes in the data become smaller, and more contiguous regions appear from which we can estimate prediction-error filter coefficients. Eventually, the quality of the regridded data degrades, such as the $40 \times 20 \times 10$ data in Figure 3.8d where the sampled data have destructively interfered within the large grid cells.

Figure 3.9a presents a baseline comparison generated by interpolating the randomly sampled data in Figure 3.7b by solving equation 2.18, but using a three-dimensional Laplacian filter rather than a prediction-error filter. The stationary regions at shallow and great depths are properly interpolated as the slopes in both of these regions are small. The areas with large slopes are poorly reconstructed. This is because the Laplacian filter is isotropic and does not interpolate well along the steep events. I then estimated a $10 \times 3 \times 3$ nonstationary PEF that varied every second point on each axis, for a total of over 9 million coefficients. I estimate this PEF on the four regridded copies of the data shown in Figure 3.8, and then use it to interpolate the data to produce the result in Figure 3.9b. This result is much better than the Laplacian result, but is still unacceptable over the portions of the data with high curvature. Given the sparsity of the data and the rapidly-changing slope of the underlying model, there is no hope of reconstructing this region without external information.

I generate more regridded copies of the data by not only varying the size of the cells in the grid, but also their origin point. Figure 3.9c was generated by estimating the same nonstationary PEF, but now on an additional 28 different grids where the origin of the grid has been variously shifted. As grid bin size increases, the number of possible shifts does likewise, so there is only one $160 \times 80 \times 40$ or $120 \times 60 \times 30$

grid, but there are four possible $80 \times 40 \times 20$ grids and 27 possible $40 \times 20 \times 10$ grids.

This result is noticeably improved over that with four scales of data in Figure 3.9b. The slowly-varying areas at the deep right of the depth slice are not noticeably improved, as they were reasonably interpolated with four scales. The center-left of the depth slice is noticeably improved, where the coherent energy introduced by the additional shifted grids of data appears in the difference panel between the four-grid and 34-grid PEF in Figure 3.9d. The multiple-grid approach still cannot reconstruct the rapidly-changing area of the model with the quality of the result of the ideal nonstationary PEF in Figure 2.17b, but this goal is unrealistic given the small amount of input data. The improvement of the result when more gridded datasets are added to the estimation shows how additional regridded data contributes to the result.

Both the stationary result and the nonstationary quarter-dome result are based upon random sampling of traces. Most data are not acquired in such a manner, such as the 2D land field data shown in the next section.

MULTIGRID ESTIMATION OF A NONSTATIONARY PEF ON 2D LAND FIELD DATA

While the previous examples are synthetic data with random sampling, field data are typically not randomly sampled, nor so sparsely sampled. The Hulia dataset is a 2D split-spread land survey acquired in Colombia. The sampling of the data, shown in source-offset coordinates in Figure 3.10 has occasional gaps on the 790-element source axis as well as both a near-offset gap and periodic gaps in the 420-element offset axis, both sampled in 20-m intervals. With the reciprocal traces included, the missing sources now contain some offsets from these reciprocal traces, with roughly half of the grid unsampled.

The recorded data in Figure 3.11 have a low signal-to-noise ratio, with few discernible reflectors in the constant-offset section (520-m) on the left panel. Both the shot gather at 6220 m in the right panel and the time slice at 1.68s in the top panel

show little usable signal. The gaps in the recorded data can be seen in all three of the panels, the missing sources in the constant-offset panel, the missing offsets from the reciprocal traces of these shots in the shot gather, and the near-offset gap in the time slice.

I break the data into 10 overlapping windows along the source axis that I treat as independent problems because of memory restrictions. I generate four regridded versions of the data, where the cells are 25, 66, and 150 percent larger along each axis than in the original 20-m $\times$20-m, 4-ms cell. Once these regridded copies of the data are generated, I estimate a $10 \times 3 \times 3$ PEF that varies every 40 time samples, every 10 source locations, and every 15 offsets, for a total of 375000 filter coefficients. I applied 900 iterations of a conjugate-direction solver on equation 3.6 to estimate these PEF coefficients. This lower number of coefficients, compared to the quarter-dome example is possible because of the more gradual changes in dip in the field data.After estimating the PEF from the multiple grids of data, I apply it in 400 iterations of a conjugate-direction solver on equation 2.29 to reconstruct the missing data (Figure 3.12). The total computational cost of this interpolation is on the order of one day for each window on a single core of a machine from 2006.

The interpolation in Figure 3.12 is a mixed result. The shot gather on the right panel of Figure 3.12 contains multiple dips, many aliased, which were correctly interpolated. The near-offset gap was only interpolated successfully at later times, which I attribute to three factors. First, there is a lack of out-of-plane information; the near-offset gap is present for all sources. Second, the recorded data at later times extends to further offsets, which implies that more data contribute to estimate a PEF at near offsets and later times. Finally, the frequency content of the data is much lower at later times, and would be less attenuated in the coarser grids of data.

The constant-offset section on the left panel of Figure 3.12 contains relatively few gaps compared with many other offsets, but the gaps in the region with the most coherent reflections, at roughly 13000 m, are correctly interpolated. Conversely, the interpolated gap at roughly 4000 m is weak below the first arrival. This is because the recorded data surrounding the gap is almost incoherent, as seen in Figure 3.11. The

time slice in the top panel of Figure 3.12 is more coherent, even at the near offsets, as this slice is at 1.68 s.

This interpolation can also be judged based on the results of further processing of the data. One such example is migration, which requires a velocity model that is not available for this dataset. One way to generate a root-mean-square velocity model is through semblance scans. Figure 3.13a is a semblance scan for an area of the survey with higher signal-to-noise, where I calculated semblances from ten adjoining common midpoint gathers and summed the results. Figure 3.13b shows the same semblance scan using the interpolated data, which, containing less noise, is easier to interpret, particularly at early times. The differences are slight but still noticeable, and given that this image is the result of combining ten CMP gathers the benefits of the interpolation are not totally negligible.

CONCLUSIONS AND FURTHER WORK

Multi-grid training data can be used to estimate a prediction-error filter on data that are insufficiently sampled. It was useful for both stationary and nonstationary cases, in that a PEF could be estimated from data that were insufficiently sampled to estimate a PEF in the conventional manner. While the PEF produced is superior to a PEF produced from a single grid of data, this PEF is inferior to one estimated from ideal training data. The basic assumption used is that a regridded version of the data contains the same dip spectra as the data on the original grid, essentially the familiar assumption that the slopes at lower frequencies are the same as those at higher frequencies. This assumption was reasonably valid for the stationary and quarter-dome cases, and somewhat less accurate for the field data.

Results were improved by increasing the number of regridded data sets used. The size and location of the grids were manually chosen and the cell sizes were evenly distributed, which is reasonable for randomly-sampled data. An improvement to this method is to choose scales more intelligently so that they best fit the data; the portions of the data that are better-captured by one grid could also be weighted higher

relative to the other scales by using a more sophisticated $\mathbf{W}_{\mathrm{m}}$ matrix. However, field data are rarely randomly-sampled, as ships travel in straight lines and receivers are typically attached to cables. Increasing the size of grid cells in this scenario is a poor choice, since the cell size from one grid to another requires the same aspect ratio for scale-invariance so the recorded data will be placed on too coarse of a grid. For example, receivers in a 3D land survey are well-sampled along the cable, but are poorly-sampled between cables. Since the desired output grid cell is (roughly) square, the only grid that would provide contiguous data from cable to cable would drastically undersample the axis along the cable, so that higher frequency and higher-wavenumber events would be penalized.

This method is not ideal for other systematic gaps in data acquisition, such as the near-offset gap present in the field data. These gaps require external information instead of only information from the nearest traces. One example of such information is pseudo-primary data, discussed in the next chapter.

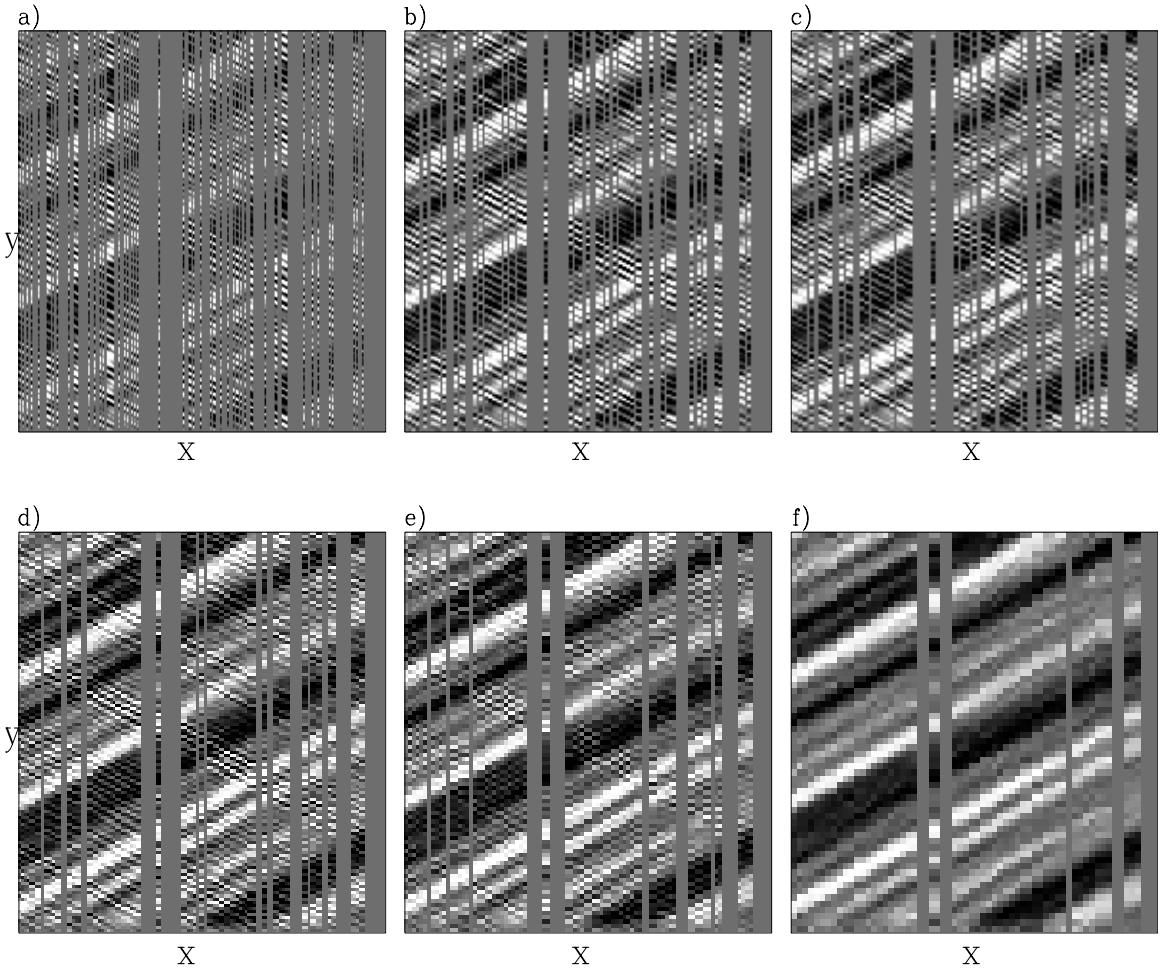

Figure 3.5: Six coarsening regridded versions of the (256×256) sampled data: a) 256×256; b) 192×192; c) 160×160; d) 128×128; e) 96×96; f) 64×64. At the coarser scales the high-frequency event disappears. **ER** MSPEF/. tracescalesann

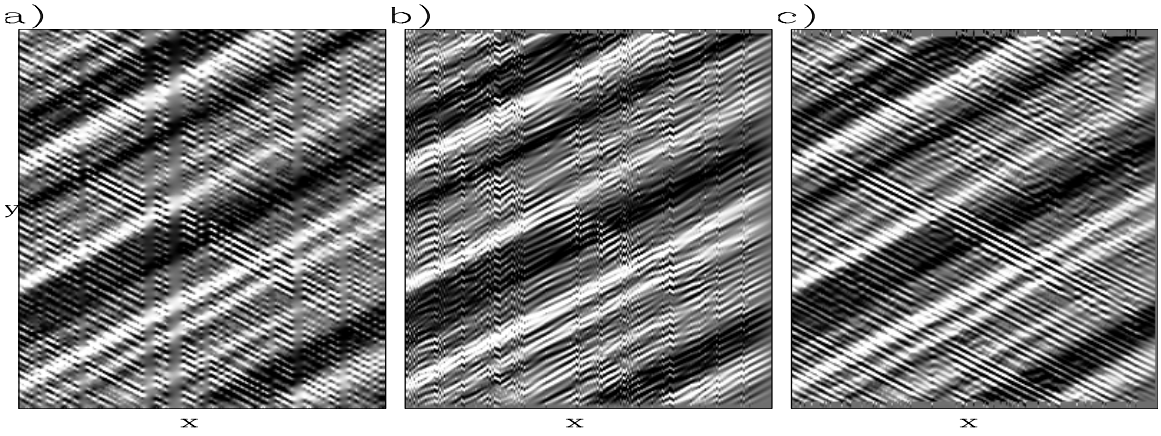

Figure 3.6: Sampled data from Figure 3.4b interpolated three ways using equation 2.18 with the filter $\mathbf{f}$ as a: (a) Laplacian filter; (b) PEF estimated from a single 64×64 regridded version of 3.4b; (c) PEF estimated from nine different regridded copies of the data. The high-frequency dip is correctly interpolated only with the PEF estimated on multiple-grid data. **ER** MSPEF/. multitracefillann

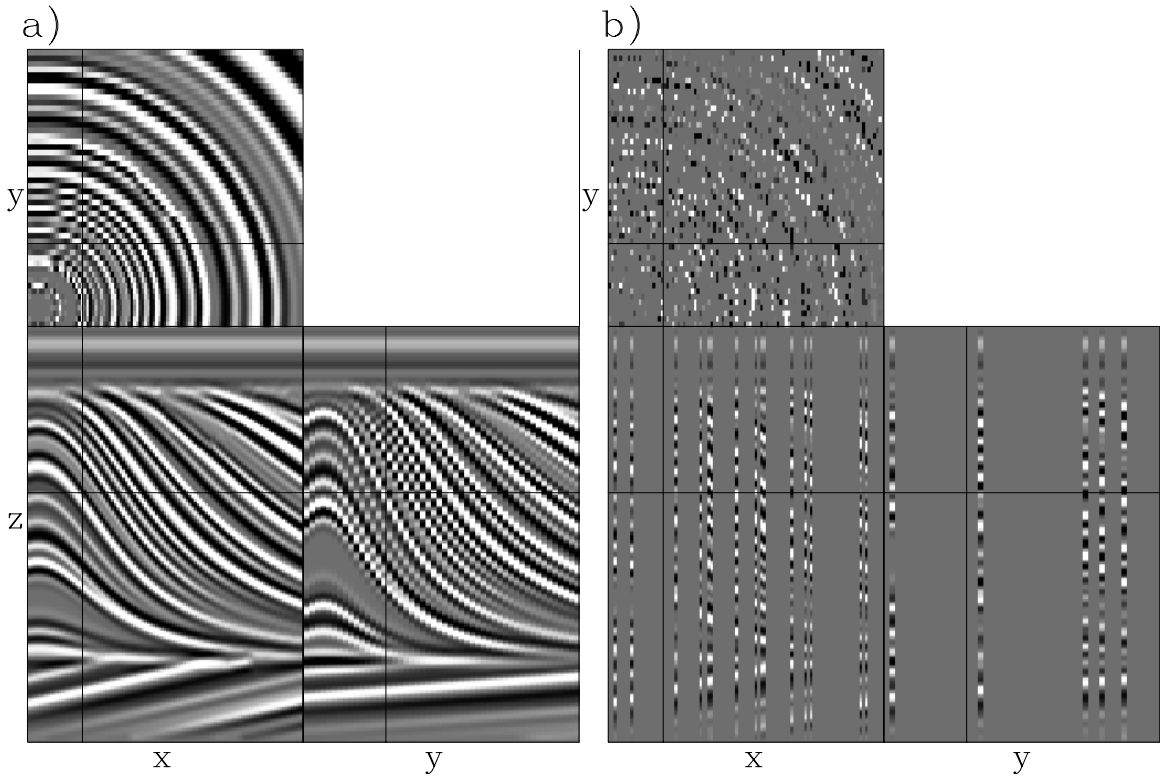

Figure 3.7: The quarter-dome synthetic: fully-sampled (a) and with 30 percent of traces randomly-sampled (b). **ER** MSPEF/. qdomeann

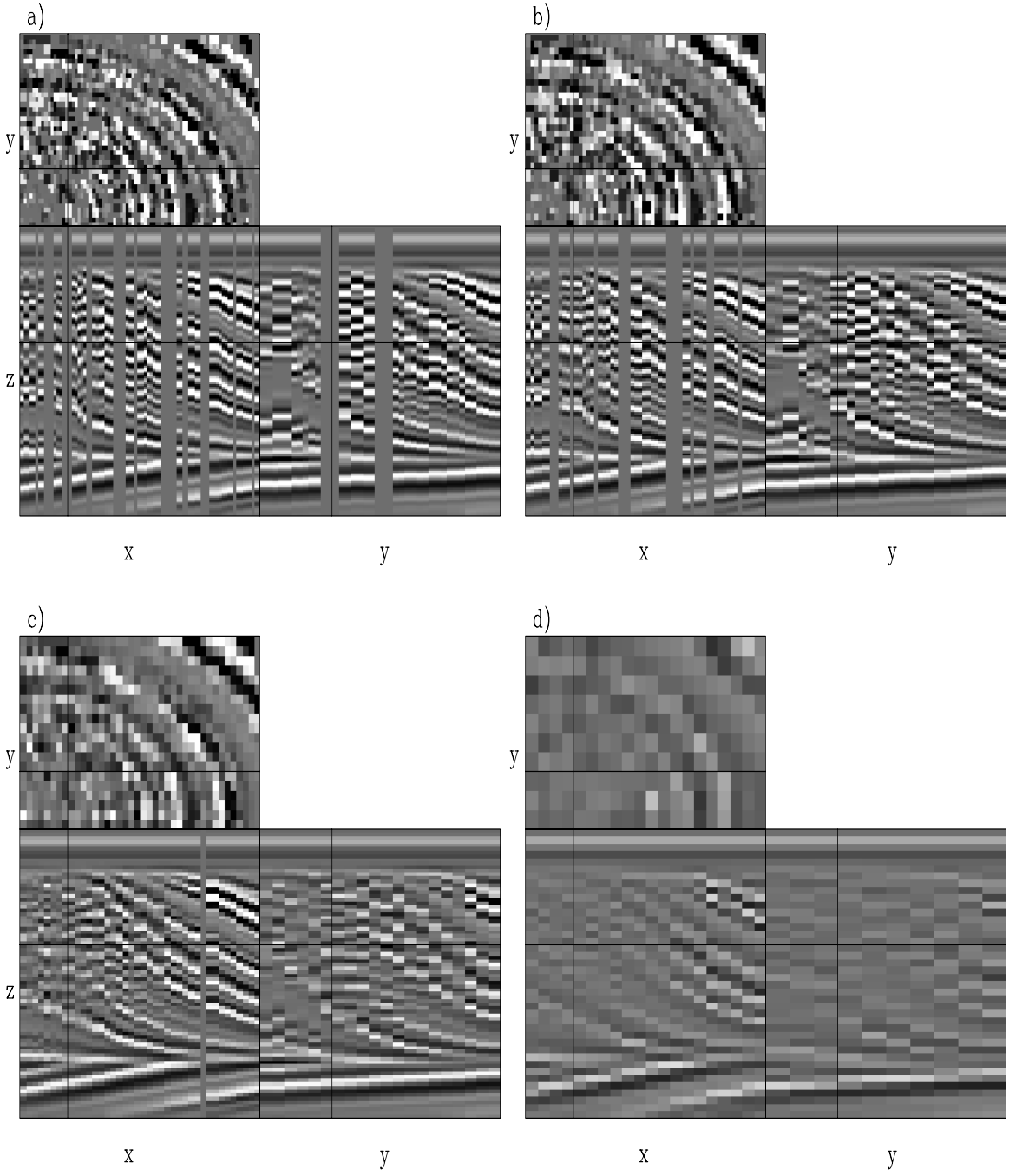

Figure 3.8: The randomly-sampled quarter-dome data (200×100×50) on four different grids: (a) $160 \times 80 \times 40$, (b) $120 \times 60 \times 30$, (c) $80 \times 40 \times 20$, (d): $40 \times 20 \times 10$. As the number of cells in the grid decreases, the number of unknown cells also decreases. **ER** MSPEF/. qdomescaledann

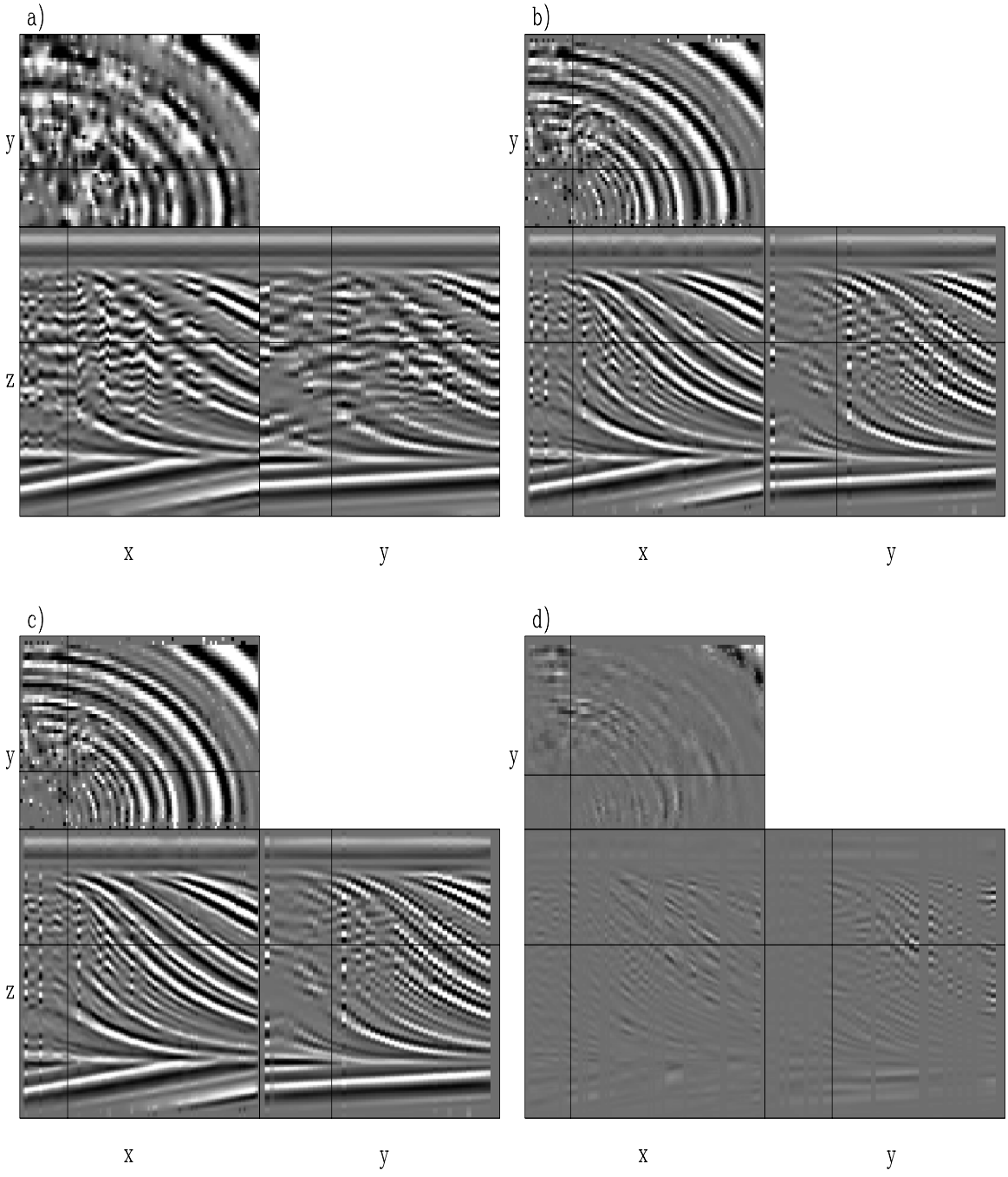

Figure 3.9: Interpolations of the randomly sampled quarter-dome data from Figure 3.7b: a): Laplacian interpolation; b): Interpolation with a PEF generated from four grids of data; c): Interpolation generated from 32 grids of data; d) Difference between b and c. The added grids of data improve the range over where the interpolation succeeds, but all methods fail at the highly nonstationary region. **CR** MSPEF/. qdomeinterpann

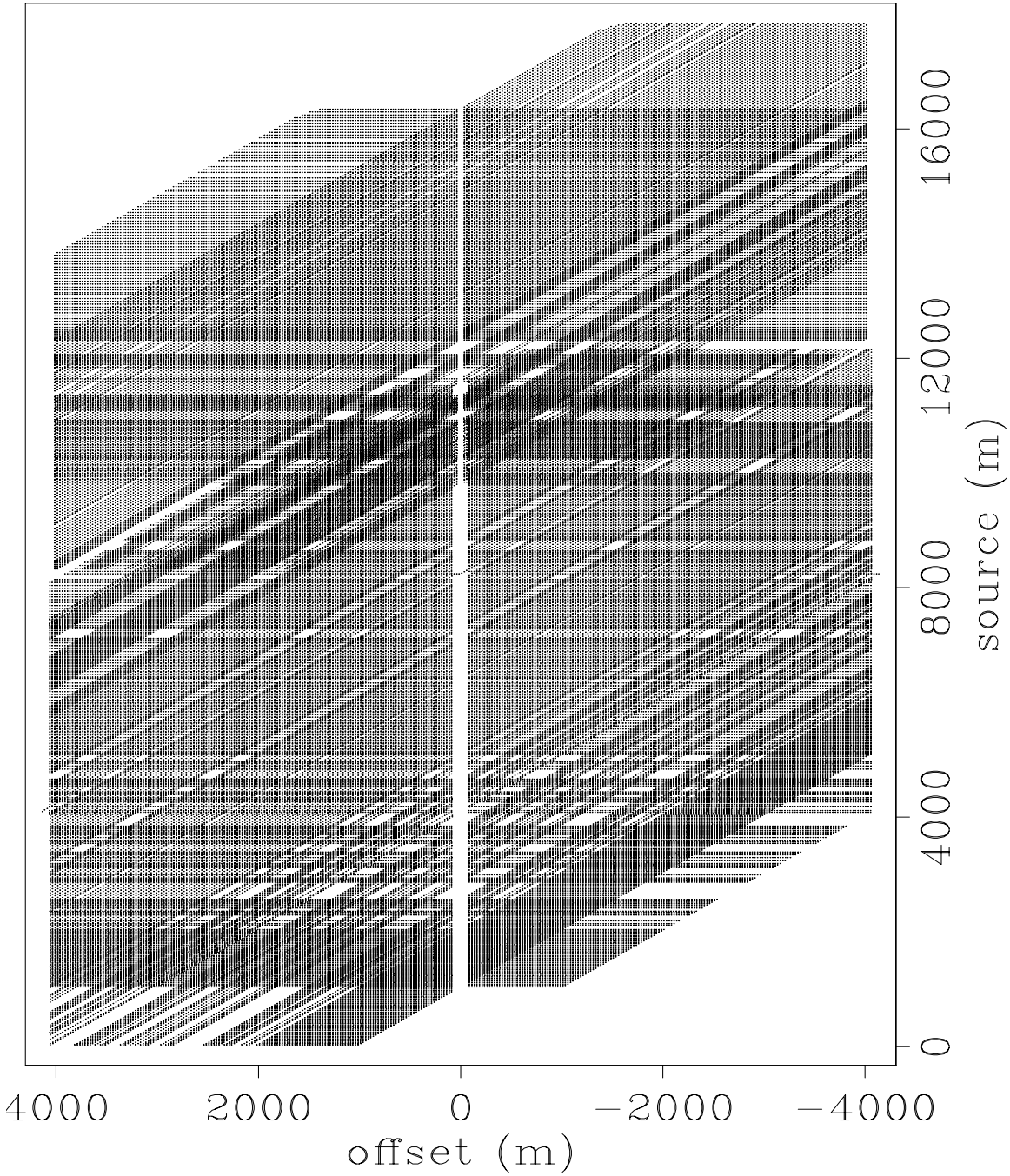

Figure 3.10: The source-offset locations of the Hulia dataset, with points representing the positions of traces in source-offset space. Both original recorded data coordinates as well as traces predicted by source-receiver reciprocity, which appear as diagonal lines, are included. **ER** MSPEF/. huliamap

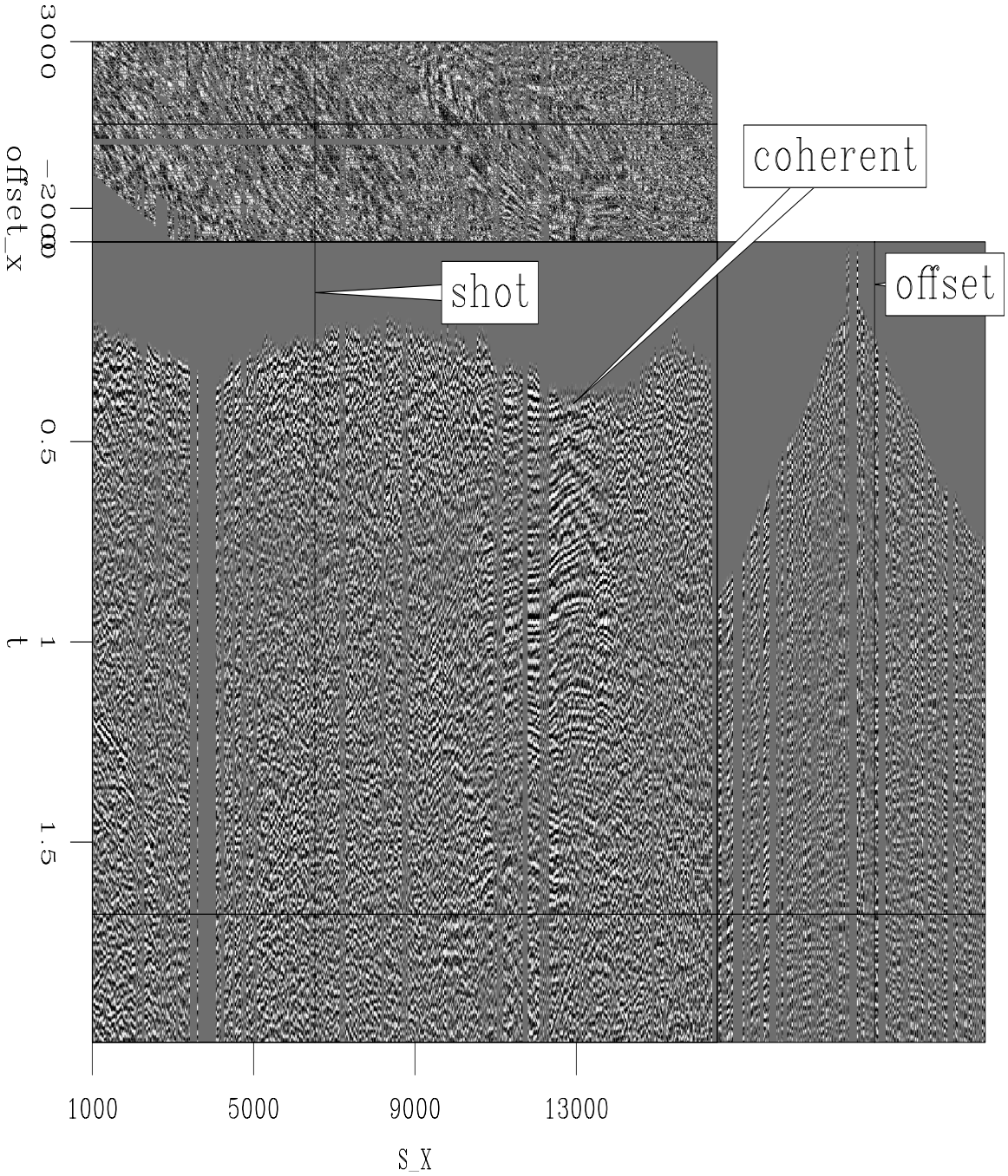

Figure 3.11: The Hulia dataset in source and offset. As seen, the data contains gaps visible in the common-offset gather at 520 m on the left, the shot gather at 6220 m on the right, and the time slice at 1.68 s on the top. **ER** MSPEF/. huliaann

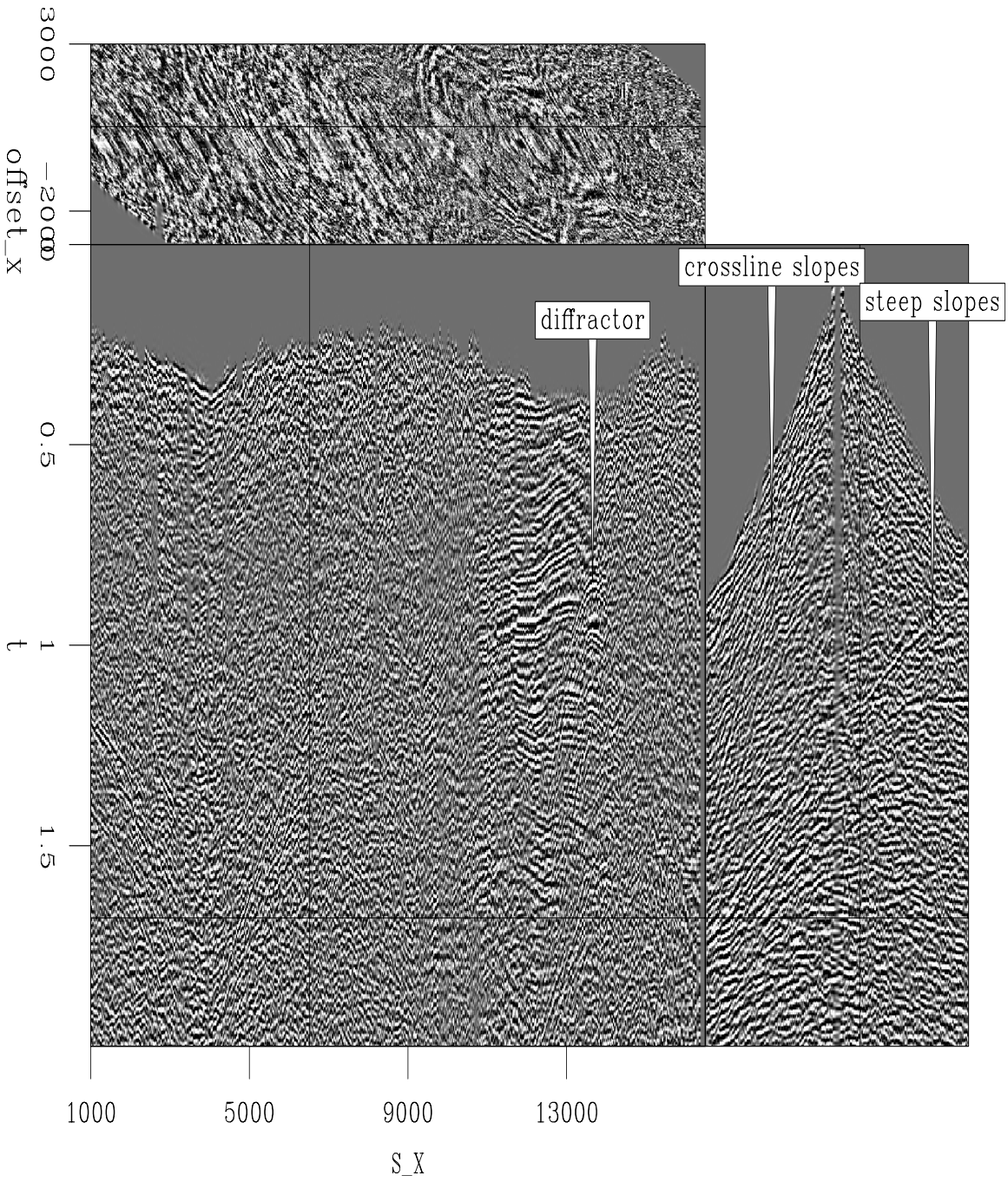

Figure 3.12: The Hulia dataset interpolated with nonstationary PEFs estimated from 4 regridded versions of the data. The shot gather contains multiple conflicting dips, as well as spatially-variable slopes that are successfully interpolated. **CR** MSPEF/. huliainterpann

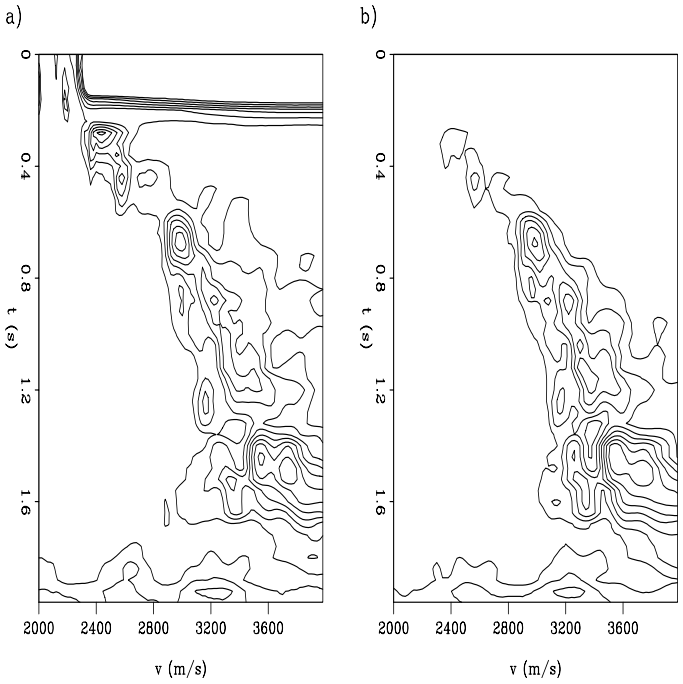

Figure 3.13: Two semblance scans generated from super-gathers. (a) semblance generated from 10 adjacent CMP gathers from original data. (b) semblance generated from same data after interpolation. The supergather in (b) is slightly more focused and less noisy than the original super-gather in (a). **CR** MSPEF/. huliascans