Chapter 1

Introduction

INTRODUCTION

Passive seismic imaging is a processing methodology which attempts to extract subsurface information due to unconventional sources. In this thesis, the goal is the production of subsurface structural images by recording the ambient noisefield of the Earth with surface arrays of seismometers or geophones. The images produced with this technique are directly analogous to those produced with conventional reflection seismic data. Manipulating wavefields with such unconventional sources has been given the general descriptor in the literature "interferometry". Passive imaging is not new science. This statement is true for two reasons.

First, the idea of imaging the subsurface by correlating the recordings from subsurface sources was first introduced by Claerbout (1968). That work provided a one-dimensional proof that the auto-correlation of time series collected by a receiver array can produce the equivalent to a zero-offset time section. Subsequently, Zhang (1989), used plane wave decomposition to prove the result in 3D over a homogeneous medium. Derode et al. (2003) developed the Green's function of a heterogeneous medium with acoustic waves via correlation and validated the observational theory with an ultrasonic experiment. Wapenaar et al. (2004), through one-way reciprocity, proved that cross-correlating traces of the transmission response of an arbitrary medium synthesizes the complete reflection response, in the form of

shot-gathers, collected in a conventional active source experiment. In fact, in recent years, there has been a sudden increase in interest around interferometry that has generated a wealth of literature (Wapenaar et al., 2006).

Second, interferometry is simply a generalization of the seismic problem formulated by earthquake and active-source exploration seismology. On the of the continuum of purposefulness, active experiments occupy an endpoint where humans initiate an experiment over a specific patch of Earth that is of interest. In this case the source function is (hopefully) zero phase, impulsive, and time zero is precisely known. Toward the middle of the continuum, humans install receivers over an area of interest and capitalize on naturally occurring far-field (teleseismic) earthquake energy to probe the subsurface (Rondenay et al., 2001; Shragge et al., 2006). Here, the source energy is areal, and time zero is defined by windowing data beginning at the onset of the earthquake arrival. At the opposite end of this spectrum is this work, which asks if the complicated, low amplitude ambient noisefield in the Earth, about which one has no knowledge, can provide sufficient energy to image the subsurface (Artman, 2006). However, the physics, and mathematics used to manipulate these various data sets are the same. The more purposeful the experiment, many of the details of seismic interferometry can be simplified or are implicitly included in the form of the acquisition. The less one knows about the source energy probing the subsurface, the more precise the processing needs to be in terms of the rigorous formulation of the more general theory.

Passive seismology is a very broad field that basically recapitulates the entire field of seismology with unconventional sources. To date, the literature concerning subsurface structural imaging at the reservoir scale has been largely relegated to theoretic and modeling pursuits due largely to a lack of suitable data acquisition. This thesis adds to the body of work in this field by presenting results from the processing of a large volume of passive records with both the standard processing sequence and a new methodology developed herein: direct migration. This new algorithm produces a subsurface image directly from the raw passive data without first synthesizing shot gathers through correlation.

DEVELOPMENT OF THE CORRELATION THEORY

Figure 1.1, simplified for clarity, shows the basic kinematics exploited in processing trans-
mission data. The figure includes two recording stations capturing an approximately planar
wavefront emerging from a two-layer subsurface. Panel (a) shows the raypaths associated
with the direct arrival and one reflected both at the free surface and a subsurface interface. The
second travel path (labeled reflection ray) has the familiar kinematics of the reflection seismic
experiment if a source were excited at receiver one. The transmission wavefield is depicted in
panel (b). Wavelet polarity is appropriate for direct arrivals and reflection. The three main fea-
tures of transmission wavefields can be appreciated here. First, the exact timing of the arrival
is unknown. Second, the phase and duration of transmitted energy are unknown and likely
complicated. Third, if the incident wavetrain coda is long, arrivals in the transmission record
can interfere.

Figure 1.1: (a) Approximately pla-
nar arrival with rays showing im-
portant propagation paths for passive
imaging. (b) Idealized traces from
a transmission wavefield. (c) Shot-
gather (reflection wavefield) synthe-
sized using trace $r1$ as the source.
Intro-sketchP [NR]
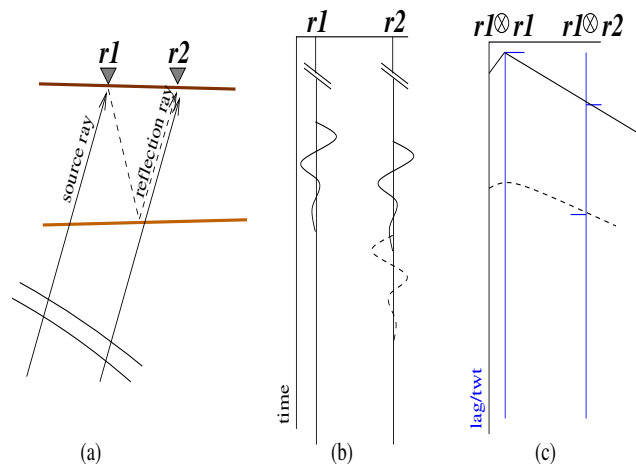
Choosing trace $r1$ as the comparison (source) trace, panel (c) depicts the correlation spikes
associated with the arrivals in the data panel (b), where $\otimes$ is correlation. A solid line with lin-
ear move-out is super-imposed across the correlated traces corresponding to the direct arrival
recorded at each receiver location. The dashed line on panel (c) has hyperbolic moveout.
However, no correlation peak exists on the $r1 \otimes r1$ trace under the hyperbola. Not drawn, the
second arrival on $r2$ will have a counterpart on $r1$ from a ray reaching the free surface farther
to the left of the model.

In fact, the correlations produced from a single plane wave will produce another plane wave. However, each planar reflection is moved to the lag-time associated with a two-way trip from the surface to the reflector. Correlation removes the wait time for the initial arrival and maintains the time differences between the direct arrival and reflections. Summing the correlations from a full suite of plane waves builds hyperbolic events through constructive and destructive interference. Analyzing seismic data in terms of plane wave constituents is a commonly invoked tool in seismic processing. Summing the correlations from incident plane waves is a plane wave superposition process that builds hyperbolas.

Cross-correlation of each trace with every other trace handles the three main difficulties of transmission data: timing, waveform, and interference. First, the output of the correlation is in lag units, that when multiplied by the time sampling interval, provide the time delays between like events on different traces. The zero lag of the correlation is the zero time for the synthesized shot-gathers. Second, each trace records the character and duration of the incident energy as it is reflected at the surface. This becomes the source wavelet analogous to a recorded vibrator sweep. Third, overlapping wavelets are separated by correlation.

Natural phase encoding

A simple framework within which to view correlation processing of transmission wavefields is phase-encoded migration (Romero et al., 2000). The motivation for phase encoding is to manipulate the independent shot gathers in a predictable manner such that the entire data volume can be processed (in this case migrated) simultaneously after a summation to reduce dimensionality. The reduction in dimensionality similarly reduces the cost of the processing. To that end, each shot gather is convolved with a unique function that will minimize cross-talk interference between the shots during a summation over the shot axes in a 3D seismic survey. Random time series are often used as effective encoding functions. Plane wave migration strategies can also be viewed as a phase encoding migration operation (**?**). Consider a set of arbitrary encoding functions $a(\mathbf{x}; \mathbf{x}_s, \omega)$ for each shot $\mathbf{x}_s$ over the map coordinates $\mathbf{x}$ with which shot records are convolved. Normally, source functions for active seismic data are bandlimited (in time) impulses, $\delta(\mathbf{x} - \mathbf{x}_s, \omega)$, at the source location, and zero elsewhere.

After convolving the impulses with encoding functions the source wavefield, $\hat{S}$, required for simultaneous migration is

$$\hat{S}_{z=0}(\mathbf{x},\omega) = \sum_{\mathbf{x}_s} a(\mathbf{x};\mathbf{x}_s,\omega)\delta(\mathbf{x}-\mathbf{x}_s,\omega) \, . \tag{1.1}$$

The five dimensional active seismic data, $R(\mathbf{x},\mathbf{x}_s,\omega)$, are similarly convolved with encoding functions and summed to reduce dimensionality

$$\hat{R}_{z=0}(\mathbf{x},\omega) = \sum_{\mathbf{x}_s} a(\mathbf{x};\mathbf{x}_s,\omega)R_{z=0}(\mathbf{x};\mathbf{x}_s,\omega) \, . \tag{1.2}$$

If the encoding functions do not correlate at all, migrating the sum of the shot gathers, $\hat{R}$, would reproduce the subsurface image desired at a substantial cost savings. Unfortunately, perfect encoding functions do not exist, and steps must be taken to minimize the cross-talk introduced during summation. This is normally done by summing the results of i encoded migrations with different function sets, a_i. In this manner, the cross-talk incurred by each migration will be out of phase and stacking the images will reduce their effects.

For passive seismic records, we can assume unknown and arbitrarily complicated encoding functions arrive from the subsurface. Therefore a transmission record $T(\mathbf{x},\omega)$ recorded by an array at the surface is $\hat{R}_{z=0}$. To recover $R_{z=0}(\mathbf{x};\mathbf{x}_s,\omega)$, we need to invert equation 1.2. For a single source, $\mathbf{x}_s$, the auto-correlation of the LHS would turn the encoding function a to a scalar amplitude term. The final result is the square of the reflection data, $\tilde{R}$. Unfortunately, after the summation in the equation, this is not possible. The rigorous development of the correlation theory developed in the literature also relies on single source processing, but now in terms of subsurface source index, ξ, whose locations need not be known. If it is possible to process a suite of transmission wavefields parameterized as a function of ξ, it is possible to synthesize R from T.

To calculate the Fourier transform of the reflection response, $R(\mathbf{x}_r,\mathbf{x}_s,\omega)$, of a loss-less (to neglect amplitude and attenuation issues) subsurface, Wapenaar et al. (2004), proved

$$2\Re[R(\mathbf{x}_r,\mathbf{x}_s,\omega)] = \delta(\mathbf{x}_s-\mathbf{x}_r) - \int_{\partial D_m} T(\mathbf{x}_r,\xi,\omega)T^*(\mathbf{x}_s,\xi,\omega)\,\mathrm{d}^2\boldsymbol{\xi} \, , \tag{1.3}$$

where the symbol * represents conjugation, and subscripts r and s indicate different station locations from a transmission wavefield. After correlation r and s acquire the meaning of receiver and source locations, respectively, associated with an active survey. The RHS represents summing correlations of windows of passive data around the occurrence of individual sources, ξ. The symbol ∂D_m represents the domain boundary that surrounds the subsurface region of interest on which the sources are located. Alternatively, many impulses can be substituted for a full suite of plane waves emerging from all angles and azimuths as in the kinematic explanation above.

The two main assumptions within the equation are: 1) processing a complete set of transmission wavefields (integral), and 2) processing wavefields due to individual sources, ξ. If these requirements are met, the sum of the many correlation volumes does undo the convolutions in equation 1.2 to provide a volume of data that synthesizes the kinematics of active seismic shot gathers. I will use the notation $\tilde{R}$ to distinguish the synthesized gathers from the active seismic experiment.

If too few sources are captured (and the integral above not completed), $\tilde{R}$ may not accurately reproduce the kinematics of $R(\mathbf{x}_s)$. For some applications, completing the integral of sources around the volume of interest may not be required. If one is specifically interested in body waves probing subsurface structure, only point sources below the domain (or relatively flat plane waves) are required to image the structure, but the synthesized gathers will contain some artifacts. Alternatively, Louie (2001), described using surface waves generated by cultural noise for near-surface refraction velocity analysis, and **?** extends a similar analysis to crustal scale observations.

The phase encoding experiment requires convolving both S and R with encoding functions. The source wavefield $\tilde{S}$ is the sum of all the encoding functions. To numerically calculate $\tilde{S}$, it would be necessary to separate the initial upcoming energy, $a(\mathbf{x}, \xi, \omega)$, from the recorded data, T, for correlation. However, the autocorrelations of the encoding functions will produce the bandlimited impulse functions, $\delta(\mathbf{x} - \mathbf{x}_s, \omega)$, in equation 1.1. Therefore, this procedure need not be performed, but instead a zero-phase wavelet can be modeled that shares the bandwidth of the recorded data. Therefore, if sufficient source energy is captured and the resulting transmission wavefields are correlated and stacked, the appropriate source functions

for migration are impulses at every recording station (Schuster et al., 2004). If the transmission data are incomplete or improperly processed however, it is impossible to know what source function would be required for migration.

Viewing passive seismic processing as an inverse phase-encoding problem reinforces the intuitive conception that a transmission wavefield is a convolution of the incoming energy from the subsurface and the reflection response of the Earth. Correlation finds similar energy packets at all times that correspond to the initial upcoming energy, and redefines time zero at each receiver location to correspond to the two-way traveltime to subsurface reflectors: Initial waiting time for the arrival of energy at any location is removed as the correlation calculates time differences between the initial arrival and subsequent reflections. Viewed in this light, passive imaging is an identical operation, with a possibly more complicated wavelet, to modeling primaries from multiples (Berkhout and Verschuur, 2006). The correlation removes the time in a record associated with waiting for the arrival of a multiple (or subsurface source), across the entire **x**-axis for every active shot gather (or subsurface source).

The new methodology for processing passive seismic data I present in this thesis falls immediately from the analogy to a phase encoding experiment presented above. Phase encoding surface seismic shot gathers was developed in the literature as a methodology to achieve the imaging accuracy of shot profile migration at the lower computational cost associated with survey-sinking algorithms. When we identify passive data as an example of a phase-encoded wavefield, shot profile migration of raw passive records is immediately suggested. Chapter 5 presents a full development of direct migration of passive data and compares images produced with the new strategy to images created by migrating synthesized gathers. Direct migration of passive data skips the generation of synthesized gathers via correlation in the time domain. Without producing shot gathers, a velocity model for migrating the data must be provided from another source. Removing the generation of the shot gathers from the processing flow can substantially reduce the computational cost of processing a large passive seismic recording campaign.

Chapter description

Chapter 2 will analyze the effectiveness of correlation processing using a two-layer, 2D synthetic model. Issues such as source character, distribution, and the inclusion of multiple source functions within a transmission wavefield will be illustrated with examples.

Chapter 3 introduces the Valhall field data and describes and interprets the shot gathers generated by correlation and spectral whitening. Chapter 4 extends the analysis of the data by further processing the gathers with NMO and stacking.

Chapter 5 Presents migrations of the correlated gathers shown in Chapter 2 to explore the effects of the artifacts within synthesized shot gathers in the image space. I then introduce the concept of passive seismic processing as an imaging problem within the course of shot profile migration: direct migration. The 2D synthetic data is used to illustrate the procedure with the suite of modeled data used previously.

Chapter 6 presents migration results from the Valhall passive data set produced by both correlation/imaging and direct migration strategies. An image provided by BP, produced with data from an active seismic survey, is used as ground truth to interpret the subsurface content of the final results produced with the passive data.

Chapter 7 extends the concept of data processing at the imaging condition of shot profile migration presented in Chapter 5 where the transmission wavefield was used as both the upcoming and downgoing wavefields. In this case, shot gathers are used for both wavefields, and the resulting image is a prediction of the surface-related multiples in the depth domain. The prediction volume is equivalent to migrating the multiple prediction produced by surface-related multiple prediction (SRMP) in the data domain. The prediction can be calculated during the course of traditional shot profile migration at minimal extra computational cost.

Chapter 2

Correlation processing experiments

SUMMARY

Experiments with 2D transmission wavefields modeled with sources across the bottom of a model domain show that correlation processing can synthesize the important kinematics in active seismic surveys. These transmission wavefields are limited to nearly vertical body waves, and result in some noise events in the synthesized gathers due to incomplete destructive interference required for a perfect result. Correlating transmission wavefields due to multiple subsurface sources leads to cross-talk that interferes with the desired result linearly with the number of sources contained within the recording.

INTRODUCTION

Cross-correlation of traces from transmission wavefields, followed by summing the individual results, can synthesize a data volume with the same kinematics as active seismic experiments. The two main assumptions invoked in developing the procedure are concerned with the temporal and areal distribution of the subsurface sources which enervate the subsurface domain beneath the array of recording stations. The success of the effort to image the subsurface with

truly passive data is predicated on the strength, number and distribution of the sources captured during the recording campaign. Therefore, the most important investigations into how well such an experiment can work in practice deal with the distribution and character of the subsurface sources within the data.

I use a suite of 2D synthetic transmission wavefields to analyze the efficacy of correlation processing under various source configurations. Imaging the subsurface structure with body waves is the goal of this exploration into the details of processing transmission wavefields. Therefore, only modeled transmission wavefields due to sources located directly beneath the model domain are used herein. All surface-related and internal multiples are included, and the transmission wavefields have been flux-normalized (Draganov et al., 2003) (to neglect attenuation). The model contains two synclines in an otherwise flat three layer velocity model. Impulsive subsurface sources were modeled every 25 m across the bottom of the model at the same depth level. The data contain 280 receivers at the surface, spaced 20 m apart. For comparison, Figure 2.1 shows two modeled surface seismic shot gathers at locations in the model from which the forthcoming synthesized gathers are extracted. For clarity, I will use the term "source" to describe the subsurface sources, and "virtual shot" in reference to the synthesized shot axis developed by the correlation processing of transmission wavefields.

Figure 2.2 shows two representative transmission wavefields and the velocity model with which they were generated. The left panel shows data from a source at the left edge of the model domain, while the center panel is a transmission wavefield due to a source directly beneath the center of the model domain. The first hyperbolas are direct arrivals and the rest are due to the structure within the velocity model.

SYNTHESIZED GATHERS

Summing the correlations of traces within a transmission wavefield synthesizes active surface data through destructive and constructive interference of the individual volumes. At the limit of processing only a single transmission wavefield due to one source, the product is incomplete and the gathers produced do not represent active surface data. As more transmission wavefields are processed, the synthesized shot gathers more accurately model active seismic data.
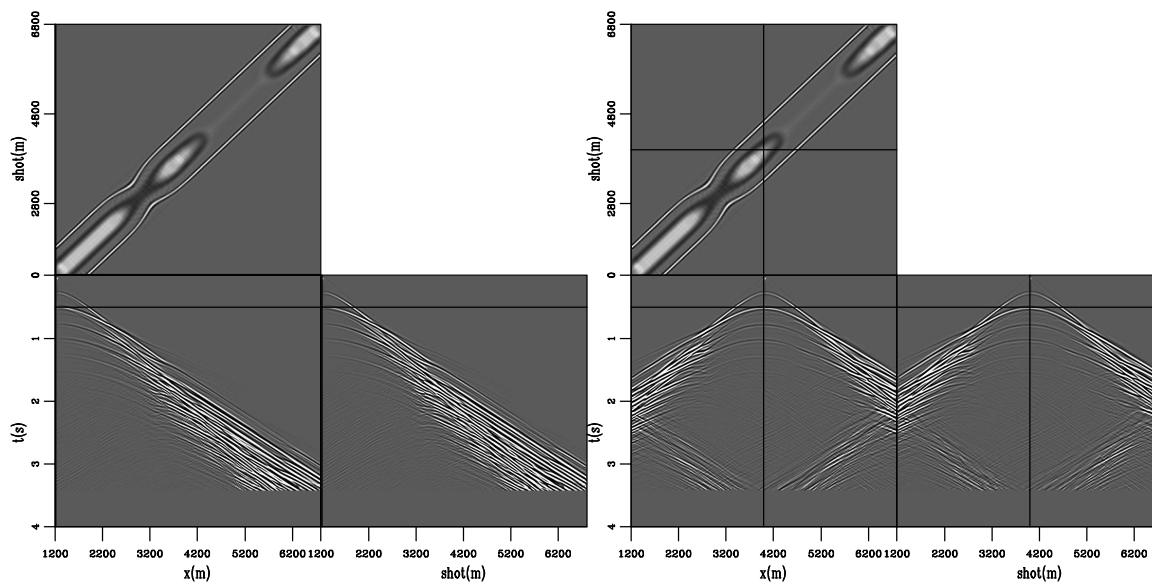
Figure 2.1: Two active reflection seismic shot gathers. These same shot locations, out of the total volume, are used in the forthcoming figures showing synthesized gathers. TimeSynth-actshots [CR]
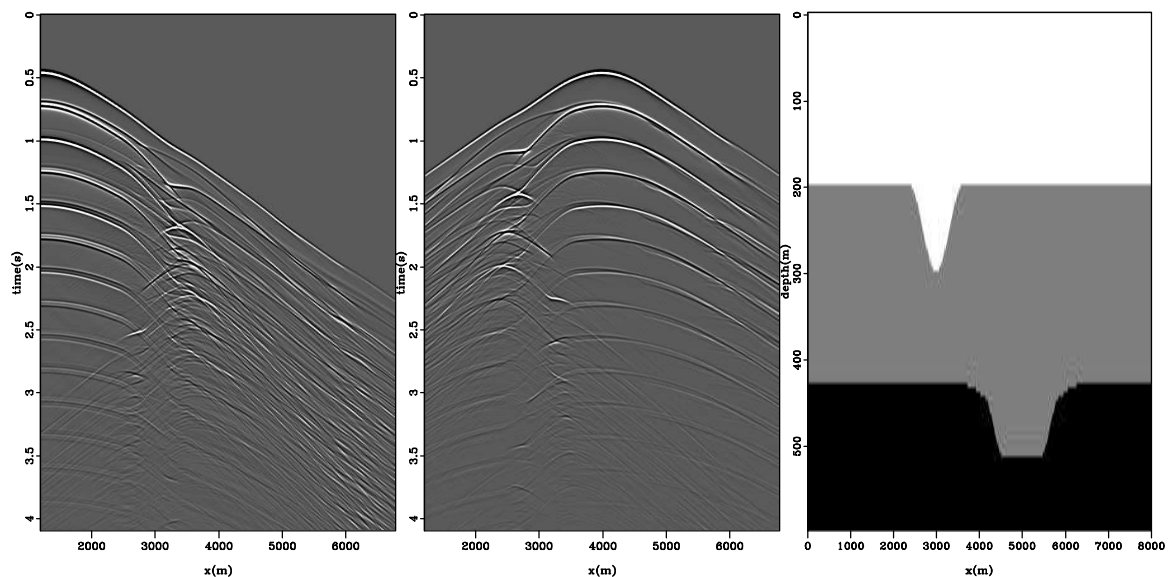

Figure 2.2: Representative transmission wavefields from impulsive subsurface sources through a three layer velocity model with two synclines (right). All surface-related and internal multiples are included. TimeSynth-data [ER]

Figure 2.3 shows two slices through a 3D volume of synthesized gathers produced by correlating all traces using a single transmission wavefield (left panel of Figure 2.2). Thus, this volume represents a partial processing of the source summation required to synthesize active data from transmission data. The time line at $t = 0.5$ s is at the zero-offset two-way traveltime to the second reflector. The front face of the left cube shows the first trace in the transmission data correlated with all other traces. The correlation does not dramatically change the relative kinematics of the events seen in Figure 2.2. The first-order effect is a simple subtraction of the time to the direct arrival in the trace used as a virtual source function from each subsequent trace. Notice that the second primary event (beginning at 0.5 s on the left edge) crosses the right edge of the gather at about $t = 2$ s. The same event in Figure 2.1 reaches the right edge of the gather at about $t = 3$ s.

The first event, that intersects the x, t axes at $(1200\text{m}, 0\text{s})$, is the manifestation of the delta function, $\delta(\mathbf{x}_r, \mathbf{x}_s)$, in equation 1.3. The next two events are the synthesized primaries. The right face of the panels have axes time and virtual shot position. The reflections show roughly hyperbolic acausal moveout on this section. As traces further to the right in the transmission wavefield are used as source functions for correlation, the hyperbolic increase in the delay time of the direct arrival moves the two primary events to negative time after virtual shot position $x = 3300$ m.

The trace under the vertical black line in the correlated shot gathers is the zero-offset trace in the synthesized gather shown and the location of the trace from the transmission wavefield used as a source function, or master trace, used to correlate all of the other traces in the wavefield. The front face of the right panel of Figure 2.3 contains only the late-time multiples from the two layer model. In real data, it is likely that these high-order multiples would be much weaker than in this model or completely absent.

Figure 2.4 shows synthesized gathers, at locations corresponding to Figure 2.4, produced by summing the correlations from subsurface sources across the entire bottom of the model. The left front face is not substantially different from the correlations of a single transmission wavefield in Figure 2.3. This gather suffers significantly from the limited distribution of subsurface sources used to synthesize the surface seismic volume of data. The quality of the result
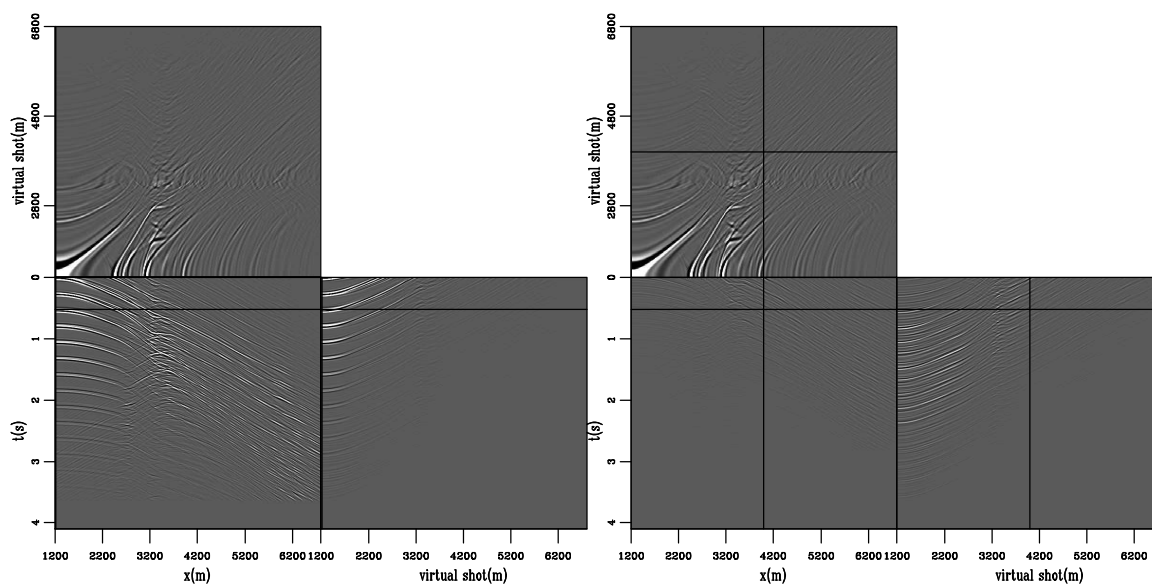
Figure 2.3: Two correlated shot gathers from the volume generated with a single source below the left edge of the model. TimeSynth-shot1200 [CR]

is predicated on upcoming energy reflecting down at a surface receiver location, and then returning back to the surface after investigating the subsurface. With most of the subsurface sources located to the right of the shot location, only the source directly beneath the shot location contributes to the synthesis of the gather. This is why the gather is of no better quality than that in the left panel of Figure 2.3. Notice that on the right face of the right panel (time-shot axes), the events develop a slight concave-up moveout toward the edge of the domain rather than an asymptotic approach to linearity.

The right panel of Figure 2.4 shows a gather from the center of the model domain, which should be the most complete gather in the synthesized volume. There is substantial noise at early time. Despite using transmission wavefields from 225 sources across the bottom of the model, the summation has not perfectly canceled this noise through destructive interference. Modeling transmission wavefields only from the bottom of the model ignores the energy contributions from locations at the sides of the model needed for a perfect result. However, the receiver and shot axes are much more symmetric than previously, which is appropriate for this model with synclines in otherwise flat layers.
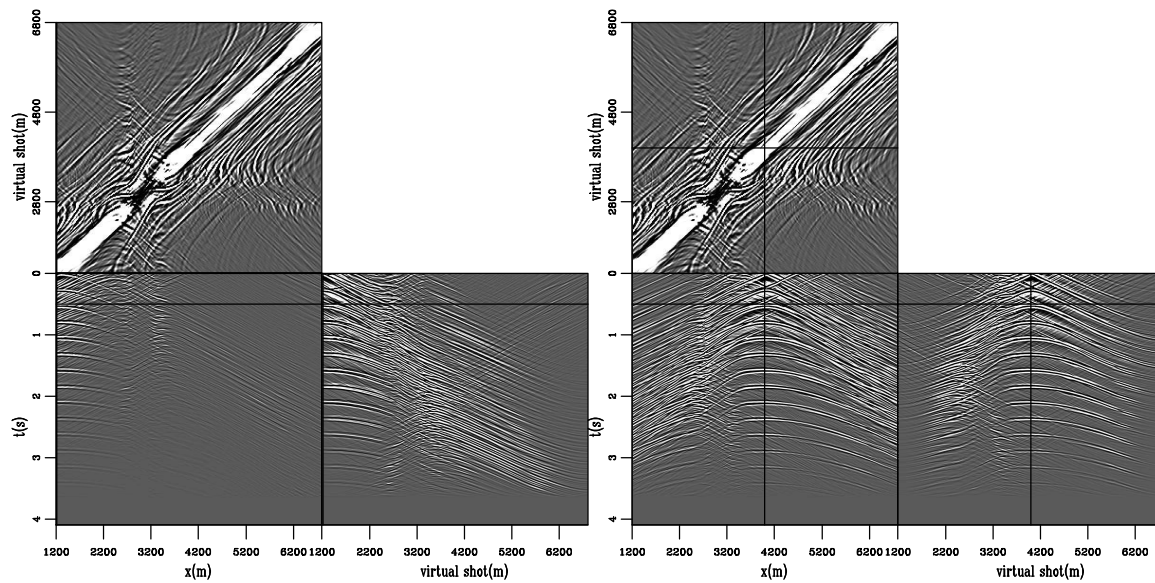
Figure 2.4: Two correlated shot gathers from the volume generated by summing correlations from 225 sources across the bottom of the model. TimeSynth-allshots [CR]

Comparing Figure 2.1 with Figure 2.4 highlights several key differences between surface seismic data and the synthesized gathers. Obvious amplitude differences should probably be ignored, as this work is focused on recovering the kinematics of the active experiment. The direct arrival in the surface seismic gathers interferes with events after about 2000 m offset, and is largely absent in the synthesized gathers. Because the transmission wavefields do not contain much energy with high ray parameters, the direct arrivals and refractions are not well represented in the synthesized gathers. The correlations above also have events that would be acausal relative to the direct arrival of the surface seismic. This energy can easily be removed by muting before further processing. The main events in Figures 2.4 and 2.1 look very similar. However, the synthesized gathers suffer from substantial coherent noise due to not using subsurface sources completely surrounding the domain of interest. In Chapter 5, migrated images produced with these data volumes are presented to explore these effects in the image space.

MULTIPLE SOURCES

If multiple sources are captured in a passive recording, processing them together violates the rigorous theory of interferometry, and the resulting volume of correlations can be more complicated than the desired shot gathers. It may be impossible to separate passive data into transmission wavefields due to only a single subsurface source if sources are weak and unidentifiable, fire simultaneously, or when attempting to process long passive records as a single wavefield.

The same wavelet was used to model the transmission wavefields in every case. Therefore, events in summed transmission wavefields are perfectly correlable. Because the sources were placed at the same depth in the model domain, there is perfect constructive and destructive interference when events overlap. Consider field data, T_f, composed of transmission wavefields $a(\mathbf{x}, t)$ and $b(\mathbf{x}, t)$ from individual sources. These individual transmission wavefields are placed on the field record at unknown times τ_a and τ_b such that

$$T_f(t) = a * \delta(t - \tau_a) + b * \delta(t - \tau_b). \tag{2.1}$$

In general, capital variables represent Fourier domain wavefields, and lowercase variables represent their time domain equivalents. The equation above violates this convention simply to avoid confusion between the transmission wavefield and the time variable. Time correlation of the transmission wavefield defined above, performed in the Fourier domain, is

$$T_f T_f^* = AA^* + BB^* + AB^* \mathrm{e}^{-i\omega(\tau_a - \tau_b)} + BA^* \mathrm{e}^{-i\omega(\tau_b - \tau_a)}. \tag{2.2}$$

The sum of the first two terms is a partial result of the summation over subsurface sources in equation 1.3. The last two terms are cross-talk. Though passive recordings are likely very long, the two-way traveltime to the deepest reflector of interest, $\max(t)$, is the latest lag that need be maintained after correlation. If $|\tau_b - \tau_a| > \max(t)$, one cross-talk term is acausal, and the other is at late lags that can be windowed away in the time domain. If $|\tau_b - \tau_a| < \max(t)$, the cross terms are included in the correlated gathers. For the modeled transmission wavefields due to impulsive sources at the same depth level, the phasors in the cross-talk terms simplify to 1 since $|\tau_b - \tau_a| = 0$.

Figure 2.5 represents the case where the energy from two subsurface sources sources ($x =$ 1200, 3500 m) is processed together. In contrast, Figure 2.6 shows the sum of correlations of the two transmission wavefields. Thus, Figure 2.6 is a partial result of the summation of correlations presented in Figure 2.4.
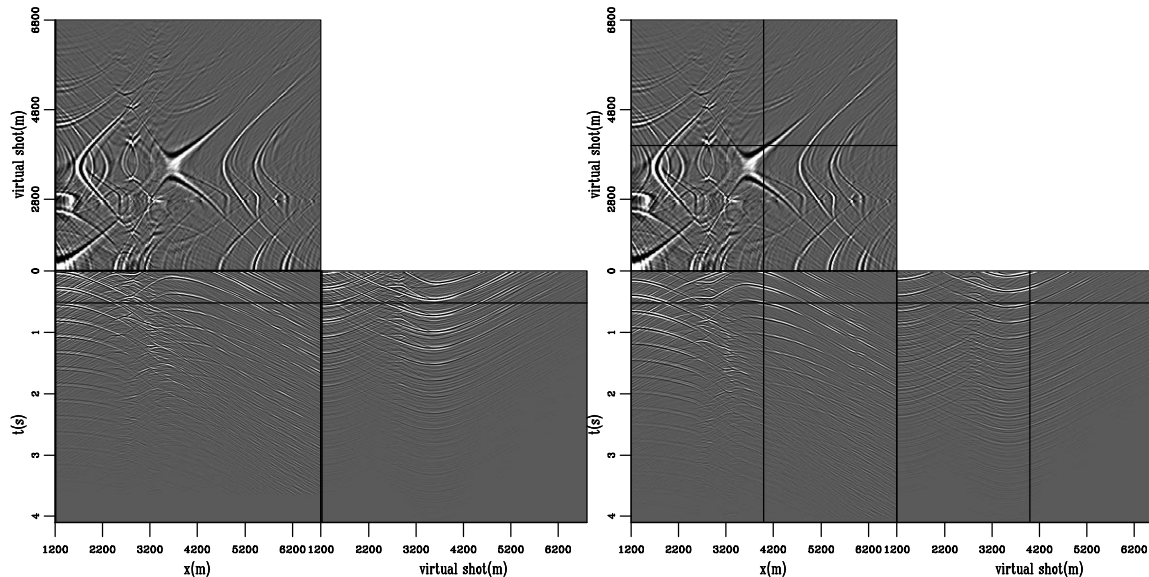

Figure 2.5: Synthesized shot gathers produced by correlating the sum of transmission wavefields due to sources beneath the left edge and middle of the model. TimeSynth-shot1200.3500.shift [CR]

Both the time sections and time slices in Figures 2.5 and 2.6 show the differences associated with the improper processing of correlating the sum of two transmission wavefields. There are ringing concave-up half hyperbolas in the top-left quadrant of the time-slices in Figure 2.5 that are not present in Figure 2.6, as well as a strong concave-right hyperbola at shot position $x = 3000$ m. There are also strong half hyperbolas on the left edge of the front face of the right panel in Figure 2.5 that are not present in Figure 2.6. Correlating the sum of the transmission wavefields produces a result that contains all the events of both transmission wavefields spread throughout the volume. Both shot positions in Figure 2.5 look very similar, and the noise is symmetric across the zero-offset diagonal in the time-slices. This observation is strengthened by noting the similarities in Figures 2.5 and the correlations of a single subsurface source, Figure 2.3.
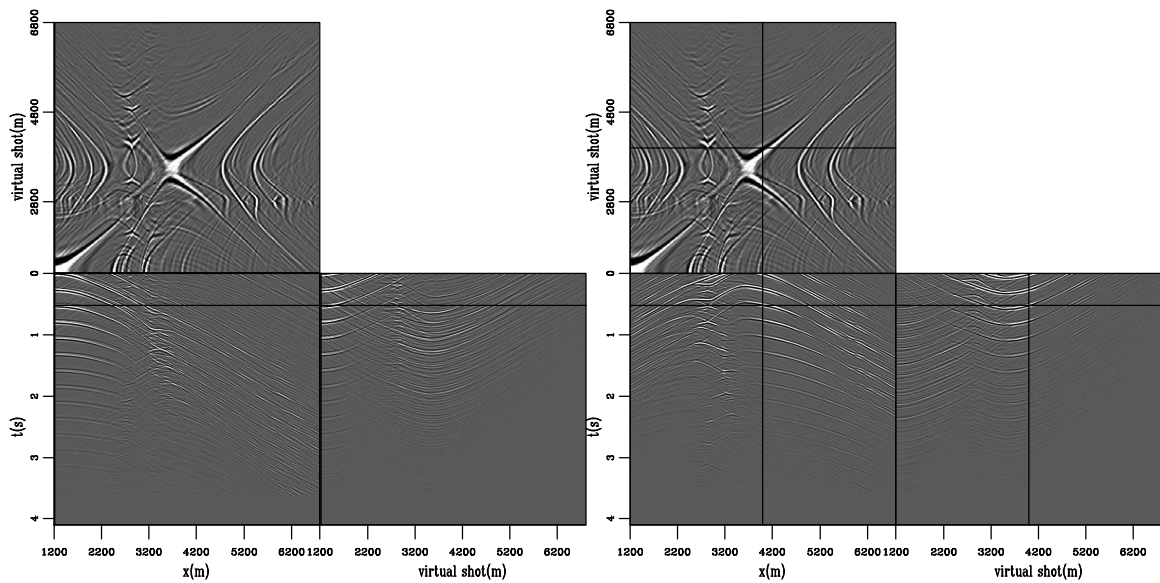
Figure 2.6: Synthesized shot gathers produced by summing the correlations from sources beneath the left edge and middle of the model. TimeSynth-shot1200csum3500 [CR]

As more sources are processed together, such as a storm of small subsurface sources during natural or induced fracturing, this problem could overwhelm the correlations. The left panel of Figure 2.7 shows the result of simultaneously processing 25 sources spaced every 125 m across the bottom of the model. Although the entire model is illuminated by this distribution, correlating the sum of transmission wavefields leads to so much cross-talk noise that the gathers are uninterpretable. The right panel of Figure 2.7 was produced with 30 shots every 25 m clustered around the center of the model. The gather in the midst of the localized swarm of sources has an undesirable ringing character. Concave-up moveout on the right face is indicative of the incomplete illumination of the domain by sources across only 600 m of the model.

INTRODUCING RANDOMNESS

In the previous examples concerning processing transmission wavefields containing multiple subsurface sources, the same wavelet was used in every case. Therefore, events in summed
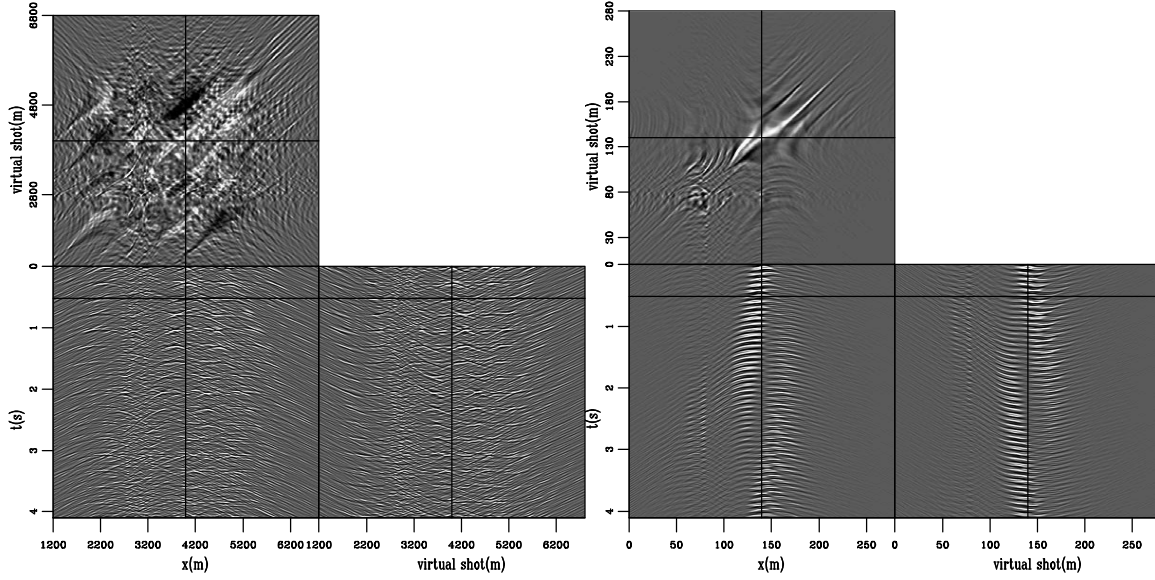
Figure 2.7: Left: Synthesized shot gathers with 25 subsurface sources spaced evenly across the bottom of the model. Right: Synthesized shot gathers with 30 subsurface sources clustered under the center of the model. TimeSynth-manysrc [CR]

transmission wavefields are perfectly correlable. While local sources probably share similar wavelet characteristics, it is probably more realistic to vary the source wavelets in the transmission wavefields. Therefore, before summing multiple wavefields before correlation processing, I convolve the individual transmission wavefields with a random number sequence before correlation. Random source wavelets will act to decrease the correlation cross-talk when it is not possible or practical to perfectly separate the passive data into individual transmission wavefields.

Define two transmission wavefields A and B as the convolution of the impulse response of the Earth, I_e, and source functions, $F_{a,b}$, which contain bulk phase delays and wavelet characteristics. As such, the cross terms of equation 2.2 are

$$AB^* = (F_a I_e)(F_b I_e)^* = F_a F_b^* \, I_e^2 = F_c \, I_e^2 \, . \tag{2.3}$$

We see that the cross-talk is the convolution of the Earth's impulse response (all associated reflections and multiples), with the correlation function, F_c, of the two source functions. The

correlation of the two source functions is not zero phase and could be a complicated function of space and time. The cross-talk events due to n_s sources will likely be more problematic than conventional multiples as every multiple and reflector can be repeated $n_s!/(n_s - 2)!$ times. The ratio of desirable zero phase terms, I_e, to cross terms, $F_j I_e^2$, decreases as $1/(n_s - 1)$ if the source terms are maximally correlable.

If the source functions are random series, the cross-talk terms $(F_c I_e^2)$ within the gathers will decorrelate and diminish in strength as the length of $f_i(t)$ and the number of cross terms increases. Designing uncorrelable functions F_i is the chief concern to effectively process sums of phase-encoded gathers. For passive recordings however, the summation of transmission wavefields can be a natural phenomenon. While we may hope to collect a large number of sources, it is probably unreasonable to expect many of them to be perfectly random series of great length. This is not probable if the noise sources have similar source mechanisms.

Previous authors have convolved impulsive transmission wavefields with long random time series to simulate passive data. Such modeling implies long, continuously ringing source functions, whose energy is distributed evenly over the time axis. I believe it is more appropriate to restrict the duration of the wavelet (no longer than 1 s), use random amplitude terms, and apply a bulk phase shift to the entire record. This modeling simulates finite duration, random phase source functions, that can happen at any moment within the entire recording interval.

Figure 2.8, shows the correlation of two summed transmission wavefields with randomized wavelets. The two sources were located below the left edge and the center of the model domain. The front face of the left panel looks like a weak version of the same gather in Figure 2.3. The right panel, however, now contains the information due to the source below the center of the domain. Note the half hyperbolas on the left edge of the right panel after 2 s. These are the time shifted events from the source at the left edge of the model domain seen in the left panel. Before 2 s, this result is nearly the same as that produced by processing the transmission wavefield due to the source located below the center of the domain. Therefore, the cross-talk introduced by summing the transmission wavefields is at comparatively late time. This could be windowed away to produce another incomplete result parallel to Figure 2.3 since the zero offset two-way traveltime to the deepest reflector in this model is only 0.5 s. Removing the

cross-talk by discarding the late correlation lags ignores the potential contribution to synthe-
sizing correct shot gathers by the source at the left edge of the domain. Interpreting this gather
without such apriori knowledge, one might assume that there is a diffractor (or a series of
diffractors if not aware of the many orders of multiples included in the data) on the left edge
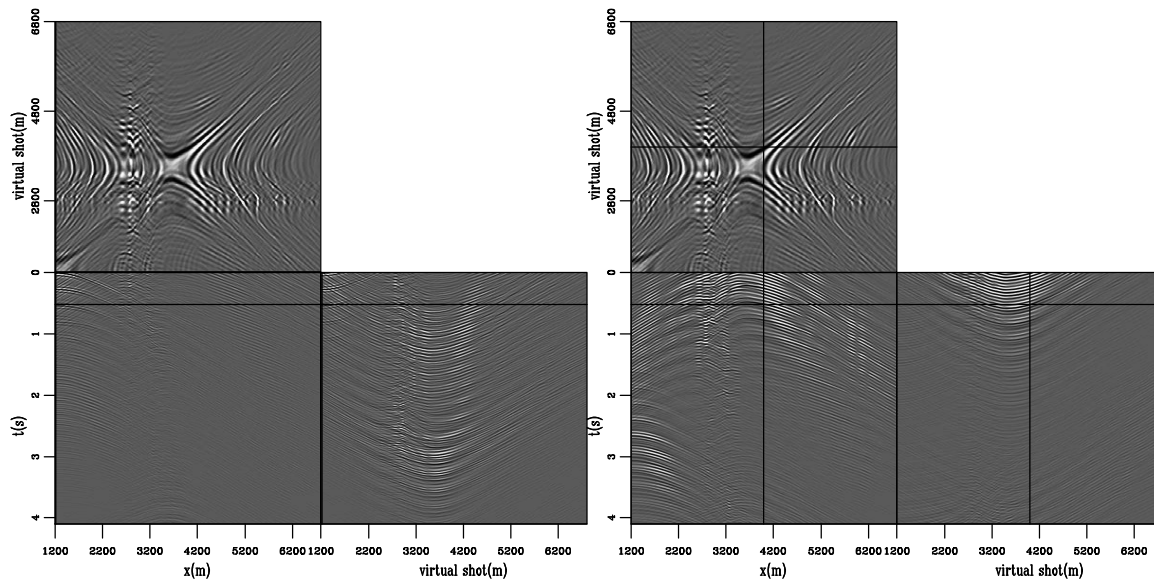of subsurface domain.

Figure 2.8: Synthesized shot gathers produced by correlating the sum of transmission
wavefields due to randomized sources beneath the left edge and middle of the model.
TimeSynth-shot1200.3500.rand [CR]

Figure 2.9 is the sum of correlations of two transmission wavefields to be compared with
Figure 2.8 as an example of the appropriate processing strategy of the two wavefields. The
figures are comparable since the two sources had a large time separation in the summed trans-
mission wavefield. The cross-talk below 2 s on the right panel is not present. Compared to the
gathers modeled as an active surface seismic experiment, each event has an extended wavelet
coda that leads to the events overlapping and an overall ringing character that diminishes the
quality of the gathers. Randomized source functions result in overall less energy within their
correlation functions, but this energy is distributed in time as diminishing correlation side-
lobes. Therefore, though the randomized wavelets decreased the total power of the cross-talk,
it has also created more events. If the wavelet codas are very long, the ringing in this result

would be attenuated by the square root of the length of the wavelet. Limiting the random wavelet to 1 s in duration does not allow for much cancellation of the correlation side-lobes.
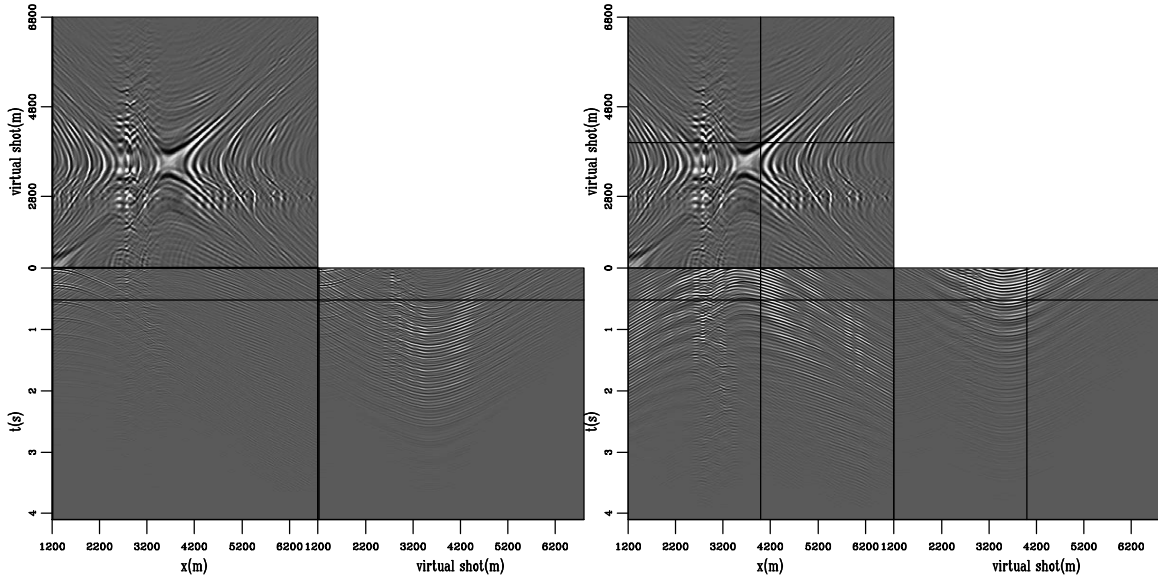

Figure 2.9: Synthesized shot gathers produced by summing the correlations from randomized sources beneath the left edge and middle of the model. TimeSynth-shot1200csum3500rand [CR]

The left panel of Figure 2.10 shows the result of simultaneously processing 25 sources spaced every 125 m across the bottom of the model. Although the entire model is illuminated by this distribution, processing a transmission wavefield due to so many sources leads to so much cross-talk noise that the gathers are completely uninterpretable. This result, is not any better than that produced without the randomized wavelet (left panel Figure 2.7). The synthesized gathers including randomized wavelets contain more events, and are not interpretable.

The right panel of Figure 2.10 was produced with 30 randomized subsurface sources clustered within 600 m beneath the center of the model. This result contains higher frequency, more distinguishable events than that produced without randomized wavelets (right panel Figure 2.7). Whether these more clear events are appropriately interpreted at subsurface information will be explored in the next chapter in which the correlation volumes above are migrated
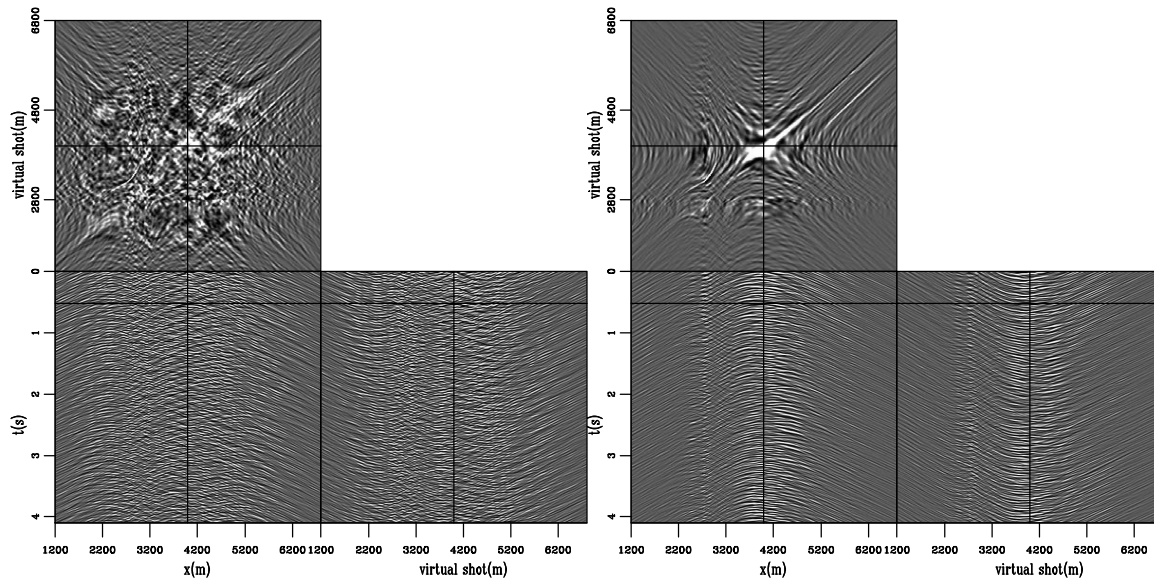
Figure 2.10: Left: Synthesized shot gathers with 25 randomized subsurface sources spaced evenly across the bottom of the model. Right: Synthesized shot gathers with 30 subsurface sources clustered under the center of the model. TimeSynth-manysrcrand [CR]

assuming that the synthesized shot gathers contain the kinematics of the active seismic reflection survey.

CONCLUSION

Synthetic transmission wavefields due to impulsive sources across the bottom of an arbitrary velocity model show that correlation of every trace with every other trace does synthesize a volume close to that produced by modeling the active surface seismic experiment. However, if only sources directly beneath the array are captured during the recording, some coherent noise events are included in the result due to incomplete destructive interference that would be enjoyed by a complete distribution of source locations surrounding the subsurface volume of interest. Positively however, direct arrivals and most of the refracted energy is not synthesized, which is often problematic in the conventional processing strategies used for active seismic data. Synthesized gathers with virtual shot locations a large distance from the areal coordinates

of the subsurface sources are very poorly reconstructed.

If it is not possible to correlate transmission wavefields from individual subsurface sources before summation, the resulting volume of synthesized shot gathers will contain coherent artifacts. These artifacts can not be distinguished as such and add coherent noise to the result that can quickly overwhelm the events from subsurface reflectors. For the scenarios of many clustered, or well distributed subsurface sources correlated simultaneously, the resulting volumes contain contain so much cross-talk that they can not be reliably interpreted.

Modeling the potential differences between local source functions as finite, random perturbations, is not a sure way to alleviate the cross-talk introduced by processing summed gathers. The model I have chosen to vary the source parameters does not produce good encoding functions. However, I do think that they represent a more reasonable likelihood in the real Earth than has been thus far reported in the literature.

Chapter 3

Synthesized shot gathers at Valhall

SUMMARY

Passive recordings from an array of 2500 four component receiver stations at the Valhall reservoir in the Norwegian North Sea were made available from February 2004 and January 2005 by BP. Analysis of some of the raw hydrophone records shows that the bulk of the records do not yield obvious, crisp events. Hyperbolic events with water velocity are sometimes seen which are centered at the location of the platforms within the array. Also, there are occasionally noise trains in the raw data with an appearance of ringing, noisy shot gathers approximately 7 s long. These are also sourced from the production facilities at Valhall.

Simple correlation does not yield interpretable synthesized shot gathers from the raw data. Spectral whitening of the correlations reveals several interesting features. There are a few complete hyperbolas in the data that may correspond to the desired events analogous to active seismic data. Dominating the gathers however, is a single event from a source probably 40 km to the S. West of the array traveling at roughly 1450 m/s. The same event is present in both the 2004 and 2005 data. The inverted source location for the event is exactly over the Ardmore field in British waters, operated by Tuscan Energy (Scotland) Limited.

INTRODUCTION TO THE VALHALL ARRAY

BP, with partners Hess, Shell, and Total, has donated approximately 43 hours of continuous passive seismic records from the permanent sensor installation above the Valhall reservoir in the Norwegian North Sea. Figure 3.1 is a map provided by the Norwegian Oil Directorate which shows the production infrastructure between the coasts of England and Norway showing the many developments in the region. It is 630 km from the Southern tip of the Norway shown and the city of Teeside in England. The Valhall development is roughly in the center of the line connecting those two geographic points.

Figure 3.1: English, Norwegian, and Danish coasts surrounding the North Sea oil development area. The Valhall reservoir is in the South-central area of the map and only 20 km from English-controlled waters to the West, and Danish waters to the South. TimeShot-areamap [NR]
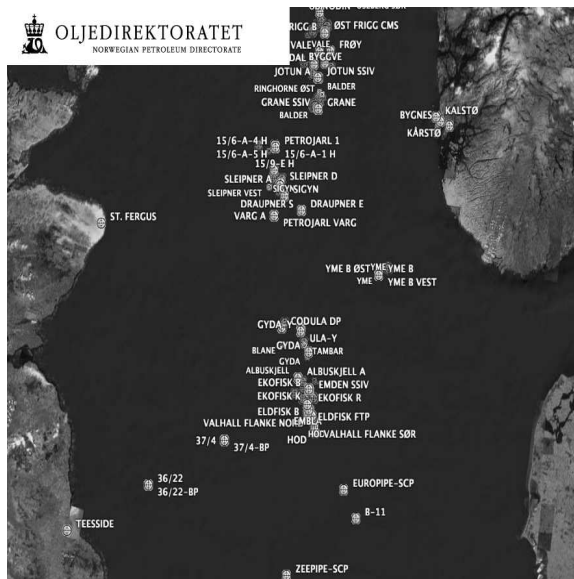
There are five surface structures (roughly over the center of the reservoir), and two well-head platforms at the North and South ends of the NW-SE elongate elliptical reservoir (both roughly 6 km from the central platforms). The main fluid export pipeline runs NNW to the Ekofisk reservoir, about 30 km distant. Figure 3.2 shows the location of the 13 lines of OBC 4C geophones installed in the summer of 2003. The +10,000 sensors are permanently installed on the sea floor, in roughly 70 m of water, forming a roughly 5 by 10 km rectangle. Nominal in-line spacing of receiver stations is 50 m, while the cross-line separation is 300 m.

Approximately half-way down the North axis of Figure 3.2, the cables on the Western side of the array are truncated. This is where the production facilities are located to which

Figure 3.2: Valhall permanent subsea geophone array. Each station measures 4 components (4C) of the seismic wavefield: Pressure with a piezo-electric sensor, and three orthogonal velocities of ground displacement with coil-spring geophones. TimeShot-actg [NR]
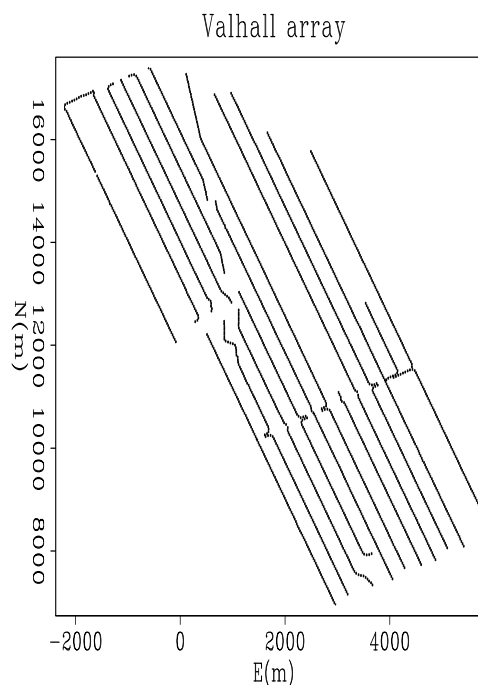
the instruments are telemetered. Also noticeable is the deviation of the ends of cables on the North side of lines 5-7 (E=1000m) to the NNW. This is the corridor for the export lines to Ekofisk. Production facilities are continuously manned and operated, to include: Over 150 total well-bores into the reservoir section, current production of approximately 80,000 b/d oil and gas from 43 wells, water injection through (at least) 3 wells, generation of almost 80 MW of electrical power with gas turbines, export lines and the continual presence of multiple modes of transportation[1]. Valhall is located in a heavily developed area, within kilometers of developments in the English and Norwegian controlled waters and is about half-way between the two countries. Also important to the chalk oil trend of this part of the North Sea is the rate of subsidence associated with pressure withdraw during production. In its first 15 years of production (to 1997), the sea-floor subsided 3.5 meters at the crest of the reservoir structure (Gebara et al., 2000).

---

[1]www.npd.no/engelsk/

PASSIVE SEISMIC RECORDING

Passive recordings were made available from 30 hours February 15, 2004, and 14 hours January 19, 2005. Of the former, some 7 hours was lost due to a failed disk-drive. Of the 2500 4C stations, this work uses the hydrophone measurement from each location. The data are sampled at 0.004 ms, and written as contiguous files between 12-20 s long. Figure 3.3 shows an example of the character of most of the data. The traces are balanced against each other, but are otherwise raw recordings. The most likely teleseismic arrival recordable at the array, as determined by querry of the National Earthquake Information Center event catalog[2], arrived approximately 5 hours after the last records donated by BP. Scanning through a subset of the raw data, I was not able to identify strong planar arrivals, but several interesting features were found that will be explored below.

Figure 3.4 shows a conspicuous noise-train in a section of raw data. The figure shows all of the traces within the array, progessing (roughly) North to South down each cable, and West to East across the cables. An inline section from a single cable contains roughly 220 traces, and a crossline section contains only 11 traces on average. While there are no coherent events in the figure, a noise-train about 7 s long within the background chatter is reminiscent of a noisy shot-gather. The minimum travel time of the noise-train is at trace 650. This trace is located at the South end of the top half of the fourth receiver line. The various jogs and cable terminations of this region are associated with the surface facilities on location. Plotted over the data are circles corresponding to a direct arrival from the location of the platforms in the plane of the receivers. These were shifted to align with the top and bottom of the noise train. The velocity used to model this event is 1450 m/s. The fact that there is no velocity increase for the bottom of the envelope suggests that this feature does not contain reflections from the subsurface. Similar examples of this type of noise-train are periodically recognizable throughout data recordings. The source of the energy has a complex coda and a finite duration.

Figure 3.5 shows a short time window when a crisp series of hyperbolic events is captured by the array. The top panel is the Eastern half of the array and the bottom is the Western

---

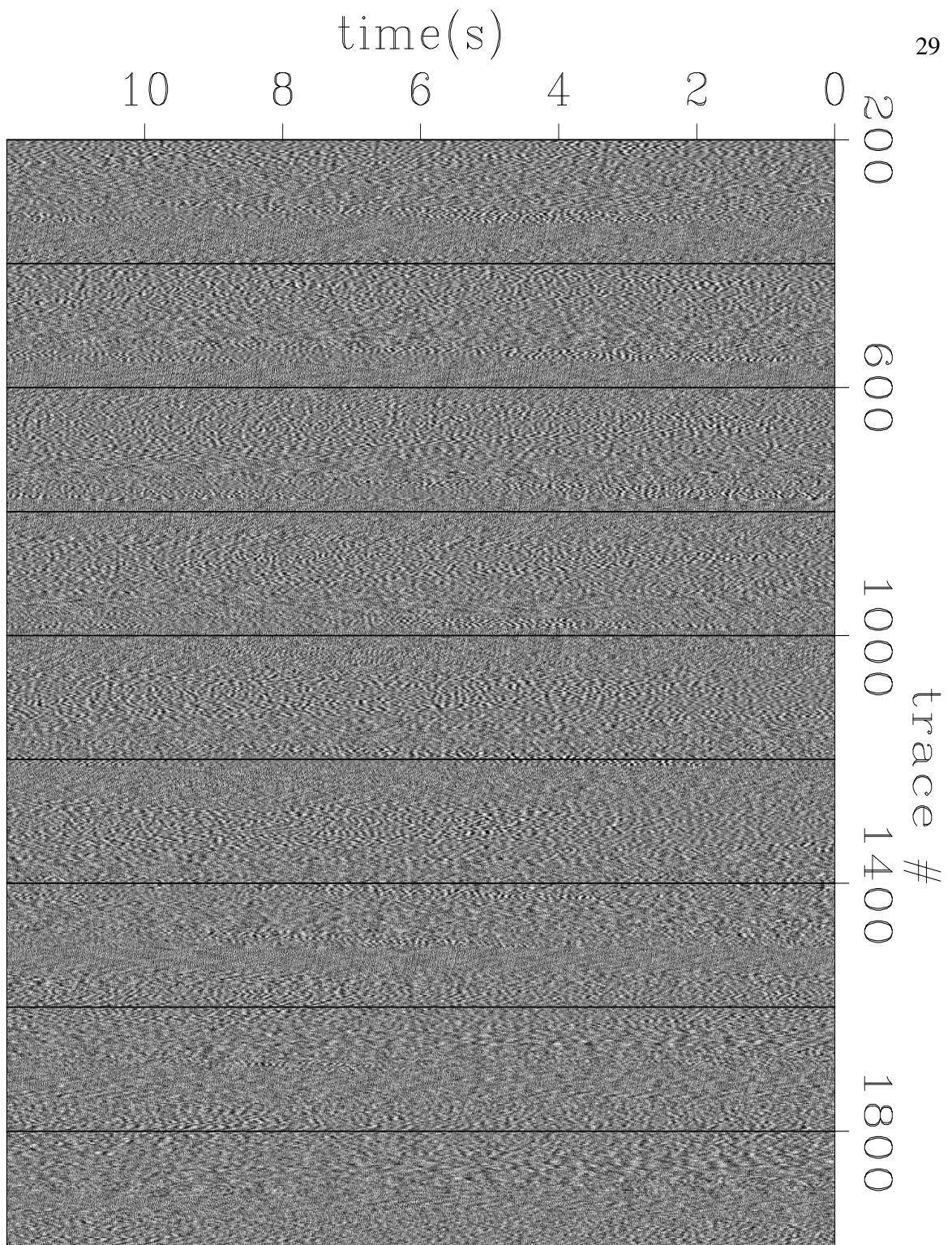[2]http://earthquake.usgs.gov/regional/neic/

Figure 3.3: Twelve seconds of representative passive hydrophone data from February 16, 2004. Trace number increases from the top left to bottom right corners of the array shown in Figure 3.2. TimeShot-raw1 [ER]
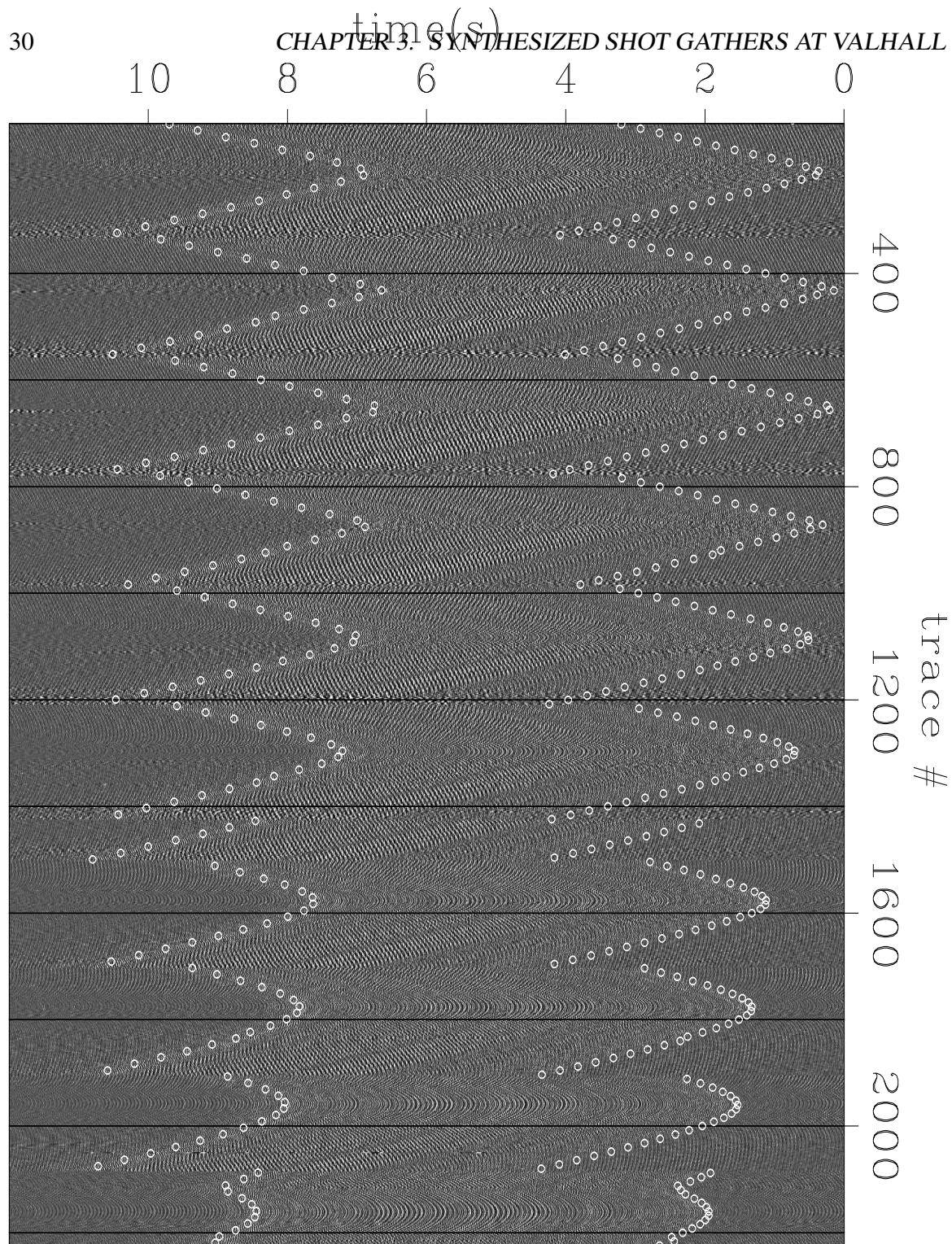
Figure 3.4: Raw data with a powerful noise train envelope. The 13 lines of the array show as the event is repeated and shifted over cross-line offset. Circles are co-planar direct arrivals from the center of the array traveling at 1450 m/s. TimeShot-raw2 [ER]

half. The minimum traveltime of the hyperbolas decreases to the West of the top panel in a similar fashion to the onset of the noise section in the previous figure. This suggests that the production facilities are again the source of the energy. The circles overlaying the data in Figure 3.5 are the kinematics of a a co-planar direct arrival from the 'O' in Figure 3.6 traveling at 1450 m/s. The modeled arrival times are plotted on every third trace to avoid clutter. However, even when plotting all picks in high resolution, the forward modeling clearly shows that the data are aliased at this slowness value over the center of the array where the event is not visible. Similar events are recognizable throughout the records, though this is a particularly clear example.
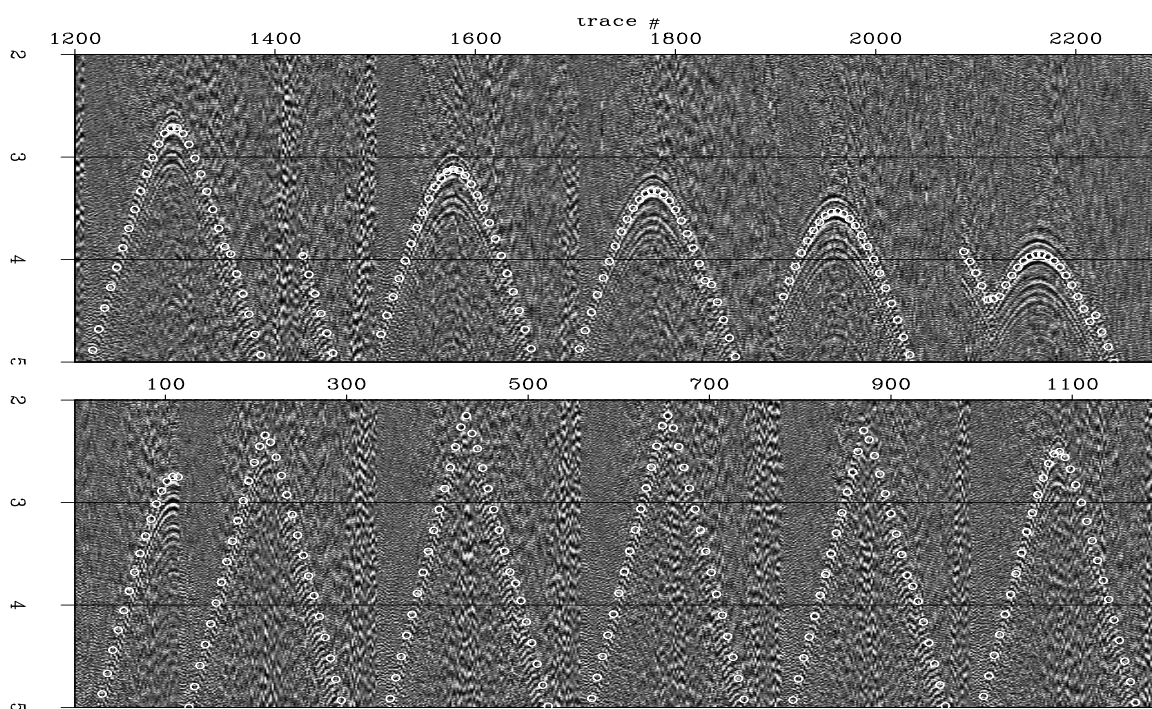
Figure 3.5: Raw data with a clear hyperbolic event annotated with the time calculated for a coplanar direct arrival from the array center traveling at 1450 m/s are. TimeShot-raw3ev [ER]

Figure 3.6 is a map of the array rotated to convenient field coordinates. The 'X' symbols mark the locations of the traces with minimum hyperbolic traveltimes that are obvious in the previous figures. The line containing the minimum traveltime picks shows no inline offset variation and includes the area around $y = 6250$ m, which is the location of the platforms

(presumably near the 'O' symbol). This figure provided the inline location for the picks plotted in Figures 3.4 & 3.5. The cross-line coordinate was deduced by matching the picks to the various observations.

Figure 3.6: Receiver array in field coordinates.   Marked with 'X' are the locations of traces with minimum traveltimes picked from the hyperbolas in Figure 3.5. Even though the event is not clear on the central lines of the array, the line connecting the traces where it is visible intersects the location of the surface facilities. TimeShot-minloc [ER]
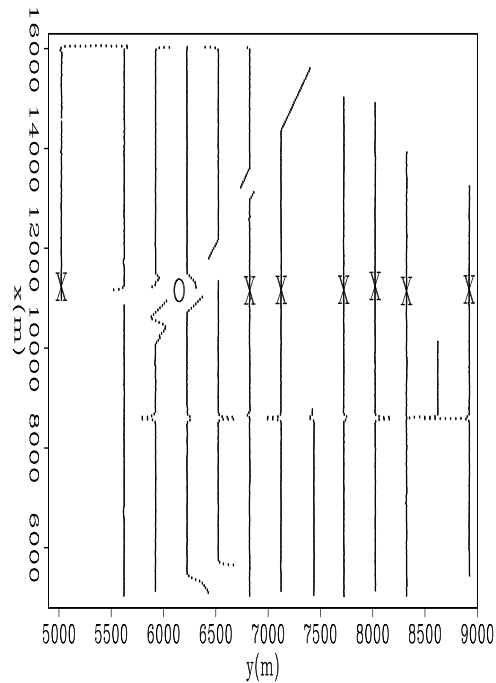
CORRELATION

Following the theory from Wapenaar et al. (2004), the basic principle of time processing passively recorded seismic records to synthesize the kinematics of conventional shot gathers dictates that individual transmission wavefields should be correlated before summing the results. Because windowing individual arrivals from the 79 Gbytes of data associated with the hydrophone measurement is impractical, it is impossible to honor equation 1.3 precisely by correlating traces within time windows when only a single event is active. However, I make the assumption that processing data in 12 second sections is a reasonable approximation to the requirement. The available data were divided into three sections for processing: one 13.5 hours, two 9 hours, and one 4.5 hours in duration. Sequential 12 second records from all the stations were correlated and the results stacked over the duration of the (hours long) section of

available data. Thus

$$\tilde{R}(\mathbf{x}_r, \mathbf{x}_s, \omega) = \sum_{\xi} T(\mathbf{x}_r, \omega, \xi) T^*(\mathbf{x}_s, \omega, \xi) \tag{3.1}$$

was calculated where ξ is the set of sequential 12 second windows of data.

Figure 3.7 shows correlations from one of the 9 hour sections of data. A trace from the NW corner was used to correlate all the other traces in the array. If the processing was perfect, the figure would show something equivalent to an off-end areal shot-gather. Energy to 60 Hz energy was used in the correlations, with cosine tapers beginning at 3 and 57 Hz. Unfortunately, the image, and the other similar ones, do not reveal anything of interest. Figure 3.8 plots the power spectrum of the sum of 100 traces from the gather in Figure 3.7. Versions of the data bandpassed around and between the peaks in the plot of the spectrum made no appreciable increase in interpretability.

Spectral whitening by division of a trace in the frequency domain with its power produces remarkable results when compared to simple correlation. I produced gathers using the relation

$$\tilde{R}(\mathbf{x}_r, \mathbf{x}_s, \omega) = \sum_{\xi} \frac{T(\mathbf{x}_r, \omega, \xi) T^*(\mathbf{x}_s, \omega, \xi)}{|T(\mathbf{x}_r, \omega, \xi) \, T^*(\mathbf{x}_s, \omega, \xi)|} \, , \tag{3.2}$$

where the denominator is simply the complex absolute value of the numerator. The algorithm was very stable and showed no need to smooth the denominator as is common practice in similar deconvolutional efforts. In fact, smoothing the denominator across the trace axis produced large amounts of hard zero results. This can only be explained by large amplitude, uncorrelated energy on neighboring traces. Dividing a weak sample by a very strong neighbor will produce very small values. Smoothing across the frequency axis mixes early- and late-time energy, and therefore is not appropriate.

Figure 3.9 shows the same gather as in Figure 3.7 using the spectral whitening of equation 3.2. A very clear event is now apparent. Similar features can be seen upon close inspection of the correlated result after viewing the deconvolved gather. The various features of the obvious event in the gather reveal important features of the presumed source energy. The event is coherent across the entire 50 km^2 of the array and has reasonable bandwidth. Versions of
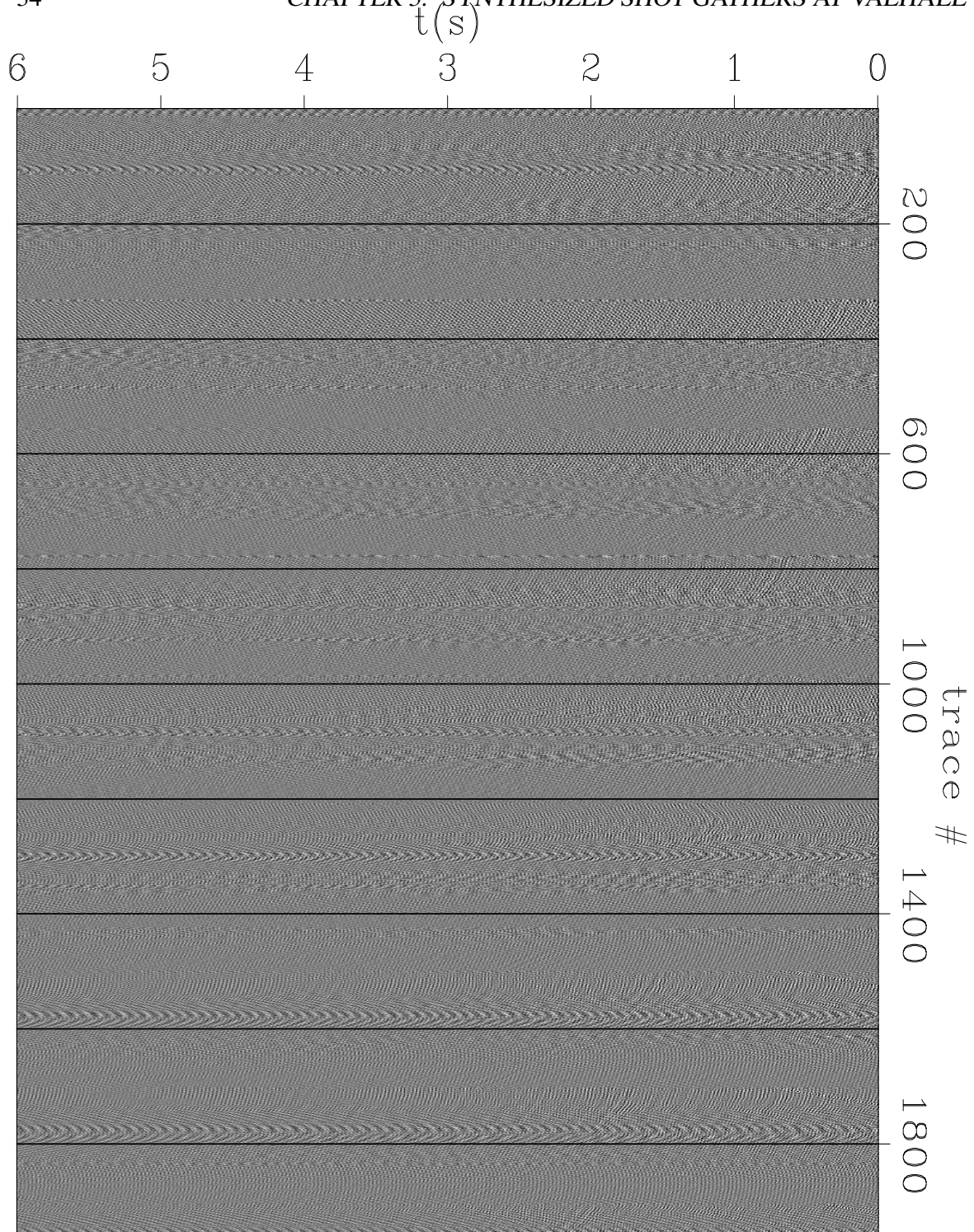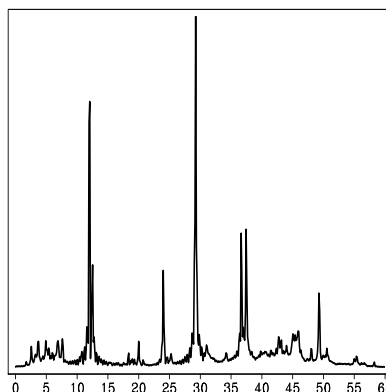
Figure 3.7: Correlated shot-gather with master trace from the left side of the panel. 0-60 Hz energy was used for the correlation performed in the frequency domain. TimeShot-cor [ER]

Figure 3.8: Power spectrum of the gathers in Figure 3.7. TimeShot-spec [ER]

data were bandpassed across the intervals between and surrounding the energy maximums in power spectrum, Figure 3.8, of the correlated gathers. No coherent energy was present at all through 8 Hz. Five versions from 9-40 Hz showed only minor differences, with the strong event equally well represented. After 42 Hz, the signal begins to diminish markedly, and no coherent energy is present at all after 52 Hz.

Events revealed by deconvolution processing, equation 3.2, in Figure 3.9 are hyperbolic. This indicates the event could be caused by any solution to the Eikonal equation. Possible solutions include the direct arrival of a buried source, or the reflection of a point source from a plane. No line contains the top of the hyperbola. This means that the source of the energy is not contained within the area of the array. The moveout of the event is monotonic to the South and East (in field coordinates). This indicates that the source must be to the North and West. Moveout is curve-linear in the inline direction, but very close to linear in the cross-line direction (white line over data in Figure 3.9). The line connecting the minimum traveltimes of the event on each cable deviates from straight only at the far right side. Notice from the map in Figure 3.6 that cross-line offset there is twice the normal spacing. This indicates that the array is close to the asymptote of the hyperbolic event in the cross-line direction, and nearer the hyperbola top in the inline direction. These two observations indicate that the source is much closer to the array in the inline direction than the perpendicular direction. Given the true geographical orientation of the array, Figure 3.2, the source of energy should be somewhere S. West of the location toward the English coast.

The fact that the event does not contain the center of the hyperbola, proves that processing
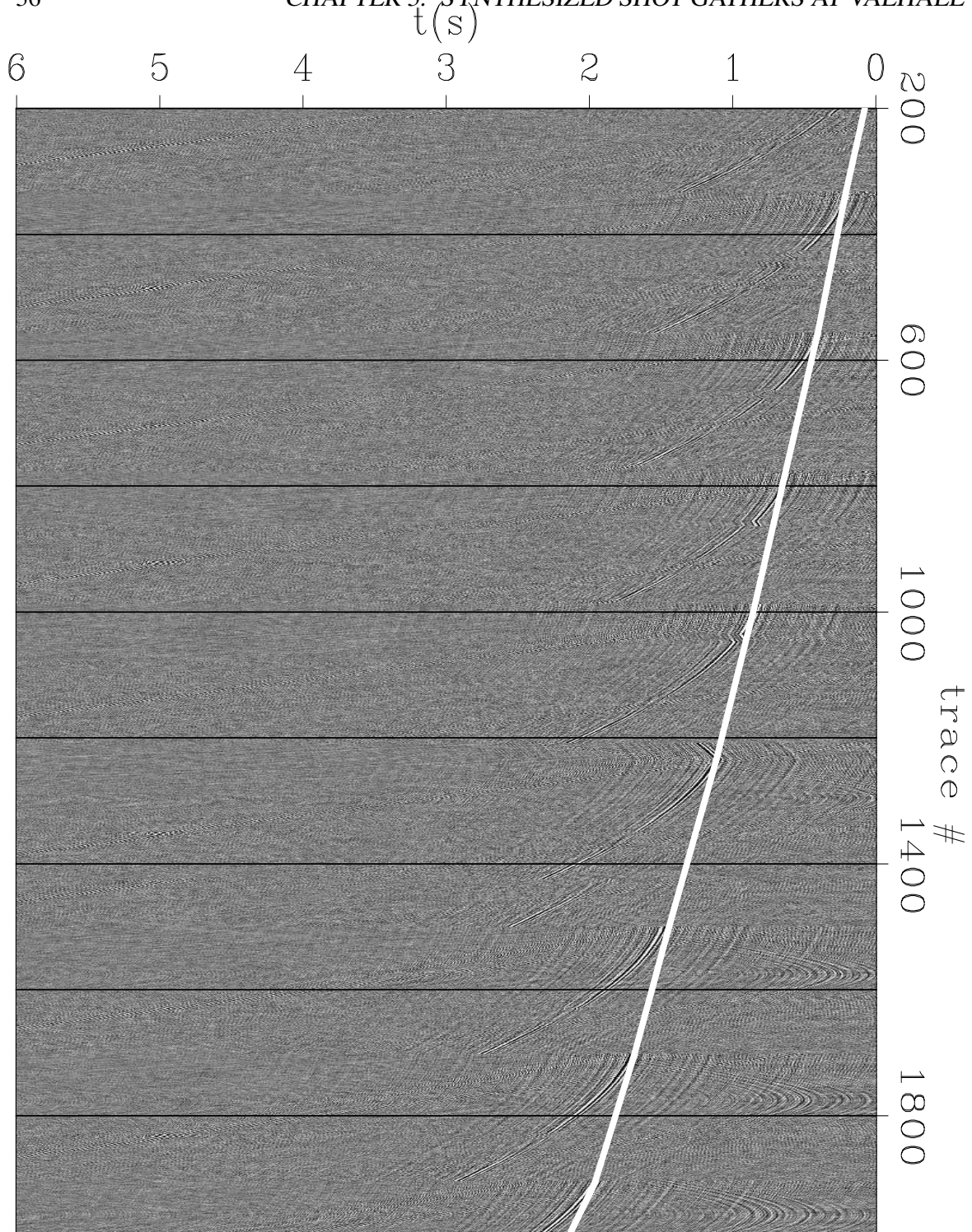
Figure 3.9: Trace-by-trace spectral whitening applied to shot-gather in Figure 3.7. Line connecting minimum traveltimes almost perfectly straight. TimeShot-dec [ER]

by correlation to produce conventional shot-gathers has failed. Specifically the requirement to sum the correlations from many shots in equation 1.3 has not been honored. If only a single impulsive shot is captured by a passively recorded array, correlation simply shifts the event up such that it arrives at $t = 0$ at the master trace used to create the gather. The panels in Figure 3.10 show four lines (OBC cables) from the array from the same correlated shot-gather in Figure 3.9. The panels were processed from the four time intervals into which the raw data volume was divided (the third panel is a subset of the gather shown in Figure 3.9). The lines are the first two and last two complete receiver lines from Figure 3.6. Each trace was correlated with the first trace in the panel. Thus the event on the autocorrelated trace is at time zero. Some of the character of the panels changes, but the presence and kinematics of the strong event is the same. The last panel was computed with data recorded almost one year after the data used in the first three panels.

Figure 3.11 shows four synthesized shot-gathers using data shown in the previous section investigating the character of the raw recordings. The master trace used for correlation in all panels was the first trace on North end of line 2 (trace 119). All panels use the deconvolutional variant of correlation described by equation 3.2 and show the Eastern half of the array. Only 12 s of passive data was used to produce all the results in the figure. The first panel is the first three seconds of the causal lags using the data shown in Figure 3.4, which showed the ringing noise-train. No strong events are present, but there are some faint ringing hyperbolas that show water velocity.

Correlation subtracts the time to the master trace from all the other traces in the gather. The beginning of the noise-train in Figure 3.4 on the first trace (number 119), is at about 3.5 s. The top of the noise train on the last receiver line (trace number 2150) is at about 2 s. The complicated coda evident in the raw data should collapse to a simple wavelet during correlation. Therefore, we expect the top of the energy on the last receiver line to lie at about -1.5 s. The second panel of Figure 3.11 contains the acausal lags associated with the correlations presented to its left. The water velocity event overlain on the traces have a clear minimum on the last hyperbola at -1.5 s, which is exactly what is to be expected for the single-source scenario just described. The hyperbola under the picks ring to the bottom ($t = 0$) of the panel. Given the long (7 s) coda of the noise train in the input, these are probably correlation
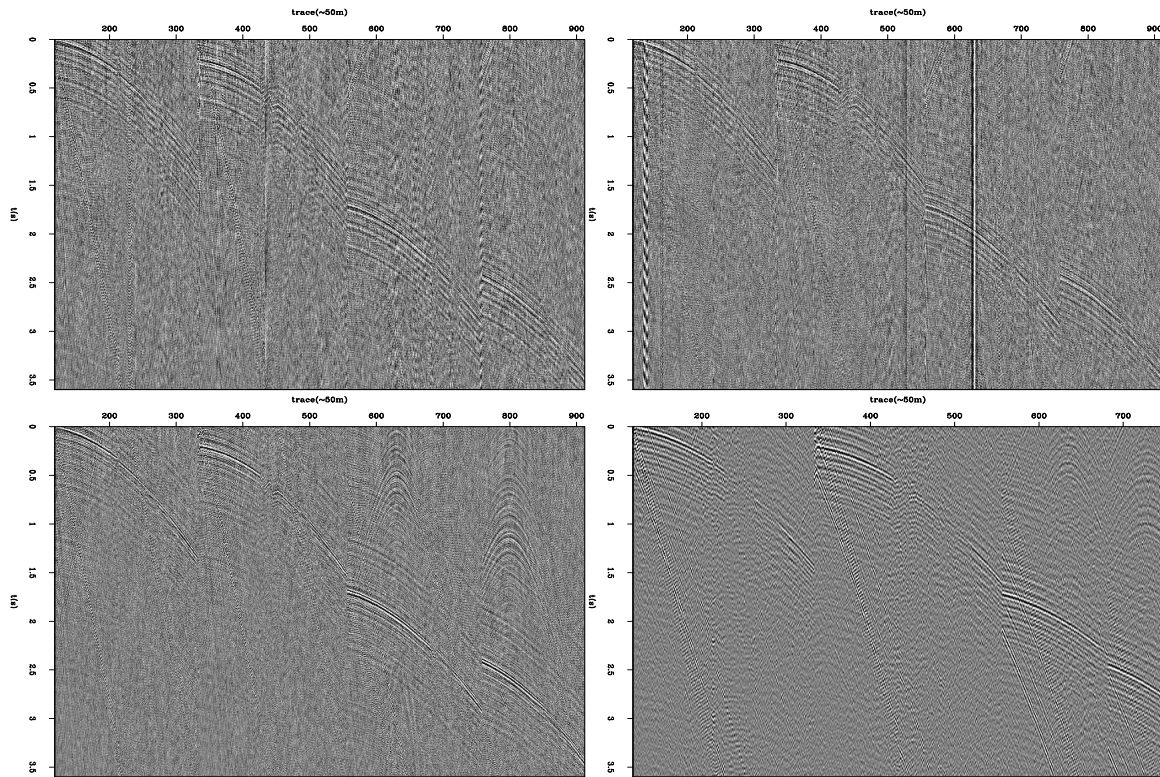
Figure 3.10: Ubiquitous strong event in correlated and deconvolved shot-gathers from data recorded 9 hours, 25 hours, and 11 months after that used to produce the first panel. The third panel is a subset of the gather shown in Figure 3.9. TimeShot-event [CR]

side-lobes. These continue past $t = 0$ and on to the panel to the left.

The bottom row of Figure 3.11 are the deconvolved correlated gathers using trace 119 as the master trace, and the raw data from Figures 3.3 & Figure 3.5 respectively. In each case only 12 s of data was correlated. Faint hyperbolas are discernible in the early time of the right side of the right panel. The most obvious energy however is in the events similar to the strong arrivals in Figure 3.9 and Figure 3.10.

Correlation masks some of the characteristics of the source energy. The source could be impulsive or a long sweep containing many frequency components like a vibrator source. The latter seems less likely. The source could be a single strong event, or many similar repeated events with low amplitude. Given the presence of the event during all the time intervals in
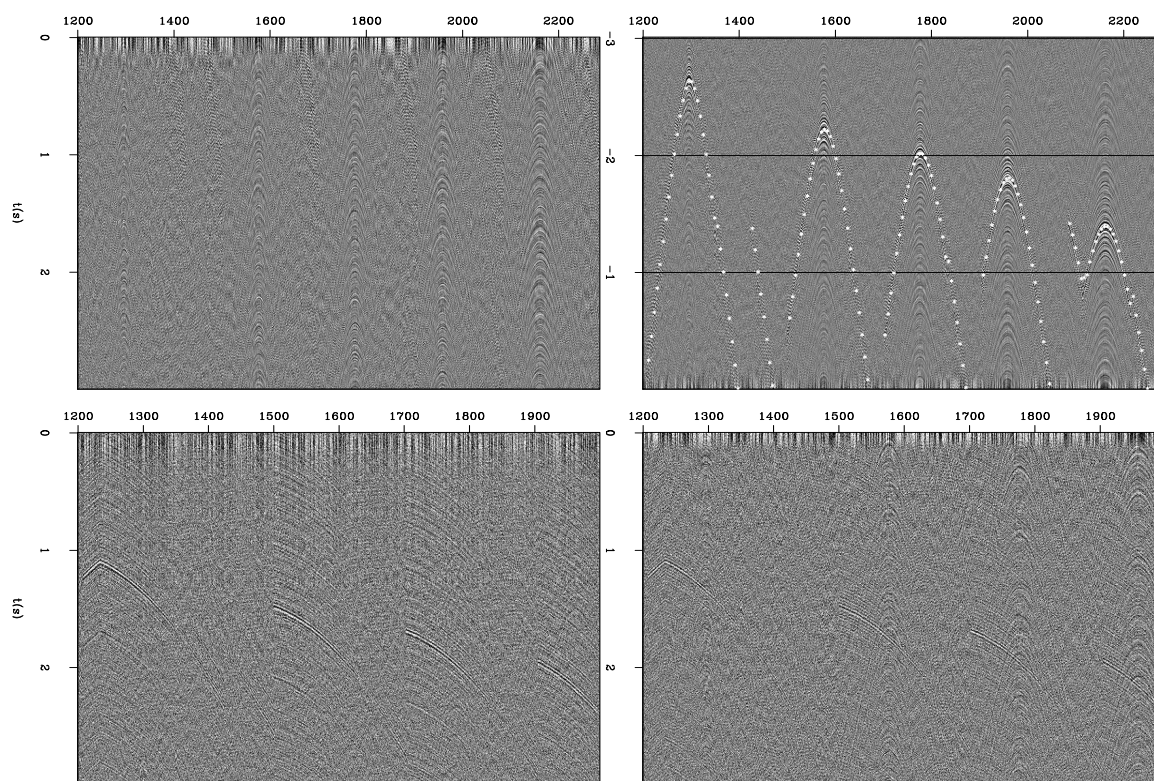
Figure 3.11: Correlated gathers from the Eastern side of the array. Top row: Early causal and acausal lags of a correlated gather using data in Figure 3.4. Picks on right panel have same (1450 m/s) velocity and origin as those delineating the top and bottom boundaries of the raw data. Bottom row: Early causal lags of correlated gathers using data in Figure 3.3 and Figure 3.5. TimeShot-evtest [ER]

Figure 3.10 (spanning 1 year), I favor the latter explanation. I believe the energy is localized in space, but seemingly ubiquitous in time. The fact that the event is clear after correlating 12 s or 9 hours of raw data means that the source must be fairly powerful.

Figure 3.12 is the frequency-wavenumber power spectrum of the raw data from a line of receivers. The frequency axis has been windowed to 23 Hz to avoid some aliasing of steep linear events which have water velocity slope. Two energy trains are observable on either side of zero on the wavenumber axis that could be subsurface signal. The right panel in the figure is an example of the mute applied in the Fourier domain to identify this energy. One hundred minutes of raw passive data was preprocessed in this manner before correlation.

Figure 3.12: Frequency-wavenumber power spectrum of on OBC cable calculated from raw passive data. Right panel shows mute applied to focus on energy near wavenumber equal zero. TimeShot-fkpow [ER]
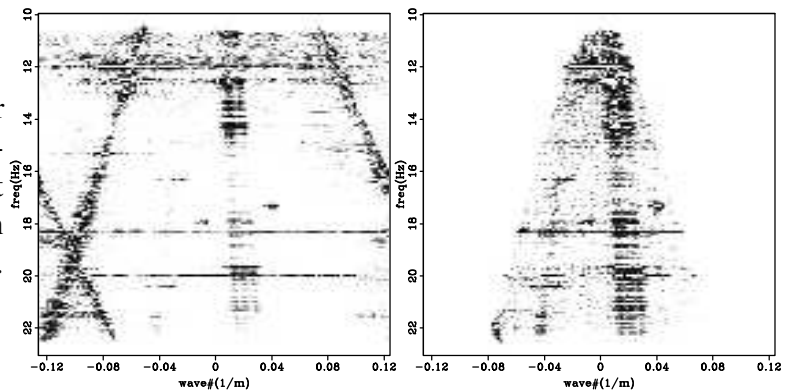
Figure 3.13 shows a shot gather produced from the data preprocessed with a Fourier domain mute (right panel Figure 3.12). A single strong event is apparent. It is aliased and has much lower frequency content that the event in Figure 3.9. It is however the same event. This proves that the kinematics of the distant event have sufficient curvature in the inline direction that a FK-mute of slow linear events does not remove the energy.
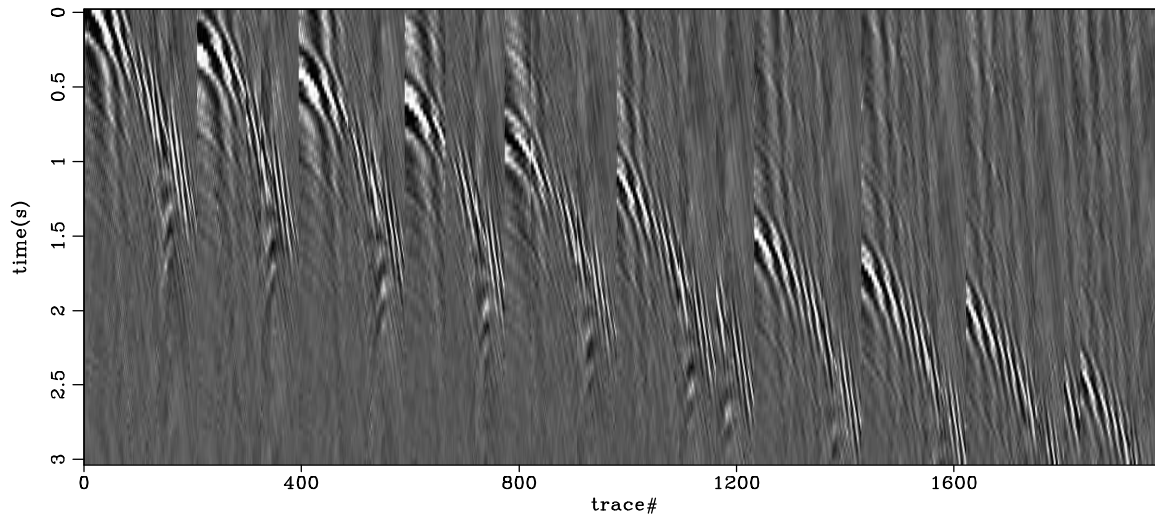
Figure 3.13: Correlated shot gather synthesized from passive data after preprocessing with the Fourier domain mute shown in Figure 3.12. TimeShot-fkgath [ER]
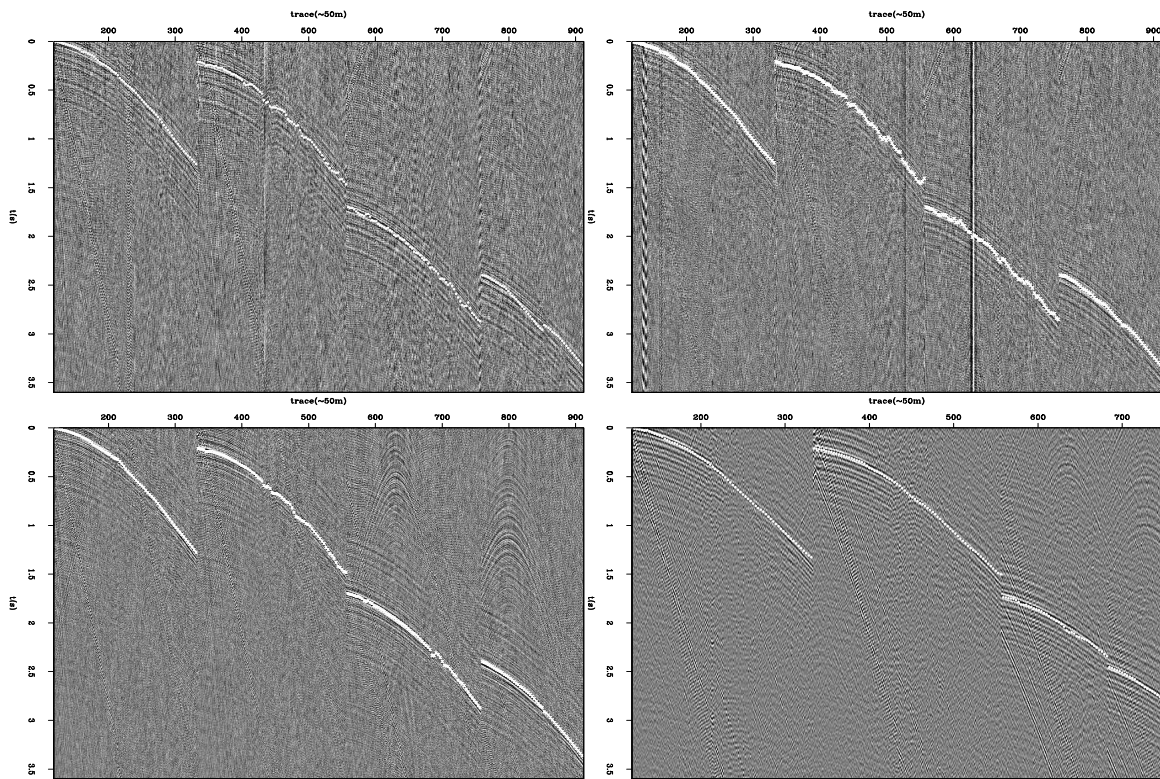
Figure 3.14: Auto-picked time values along event in correlated gathers. Picks are used as data when inverting for the source location. TimeShot-picks [CR]

SOURCE IDENTIFICATION

Figure 3.14 shows the same correlated gathers as Figure 3.10. Superimposed on the data are time picks auto-picked by amplitude. The picks are not very accurate due to the effort required to maintain continuity. These picks, and the receiver locations, are the data used to invert for the source location that caused these events. The forward model used for the inversion was the kinematics of a reflection from an arbitrary plane. Inversions for a direct arrival from depth were not successful.

The equation describing the traveltime of a planar reflection in shot-receiver coordinates

is

$$v^2 t^2 = \mathbf{x}^2 + 4z^2 + 4z\mathbf{x}\cos\alpha\sin\phi \, , \tag{3.3}$$

where $\mathbf{x}$ is the horizontal distance from the source to the receiver, z is the depth to the reflector under the source location, α is the dip direction of the plane, and ϕ is the dip of the plane. At the limit of $z = 0$, the operator solves for a co-planar direct arrival. Correlating the data with a single master trace subtracts the time to the source trace from each trace in a gather. Therefore, the forward modeling operator used in the inversion solves for the time to every receiver, and then subtracts the time, t_s, to the master trace used for correlation. This calculates the time differences produced in a correlated panel. Subtracting t_s in the forward modeling operator greatly diminishes the operator's sensitivity to the terms on the RHS involving depth. Also, large horizontal distance will result in the first term, $\mathbf{x}^2$, to dominate the others. The inversion technique used was a micro-genetic algorithm (Krishnakumar, 1989). The model space was defined as the ensemble: Average velocity through which the rays have passed, areal location of the source, and orientation angles and depth of the reflection plane. The best results were obtained using the l^2 norm to evaluate the power of the data fitness function.

The panels in Figure 3.15 show the effectiveness and quality of the inversion results. The genetic algorithm uses a seed number that determines the characteristics of the first population as well as how the future generations change. I used fifteen random seeds to begin inversions with the synthesized shot gathers from each of the four subdivisions of the passive data. Figure 3.15 is a representative sample of the inversion results from the four volumes of synthesized shot gathers. The panels in the figure show the convergence of the algorithm in total (panel 1), and the convergence of the model parameters velocity, distance East, and distance North (respectively) to specific values. In each case the vertical axis is the fitness value that the inversion is minimizing. The relative width of the data cloud at the minimum fitness value for the 3 model parameters shown reflects the model precision of the inversion for this data. Velocity and distance North are very well constrained, but distance East has a wide region of minimum energy. Not shown are the plots for the model parameters associated with the planar reflection: depth, dip, and azimuth. These parameters were completely within the null space of the inversion.

Convergence

Velocity dependence
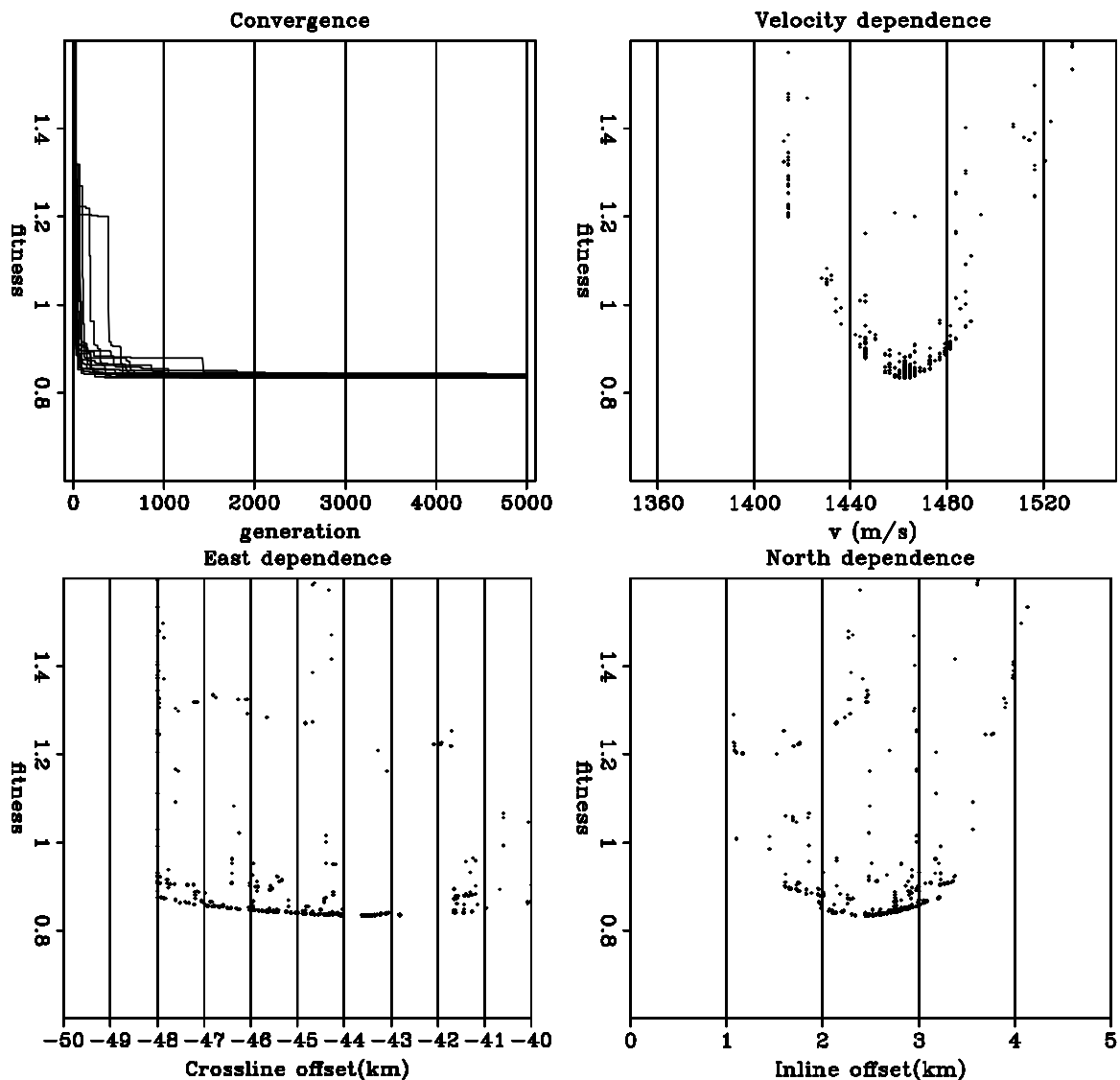
East dependence

North dependence

Figure 3.15: Residual energy of the inversion to locate the source of the energy causing the event in Figure 3.14. As the inversion iterates, overall residual energy decreases, and the standard deviation of the model parameters diminishes for well constrained members.
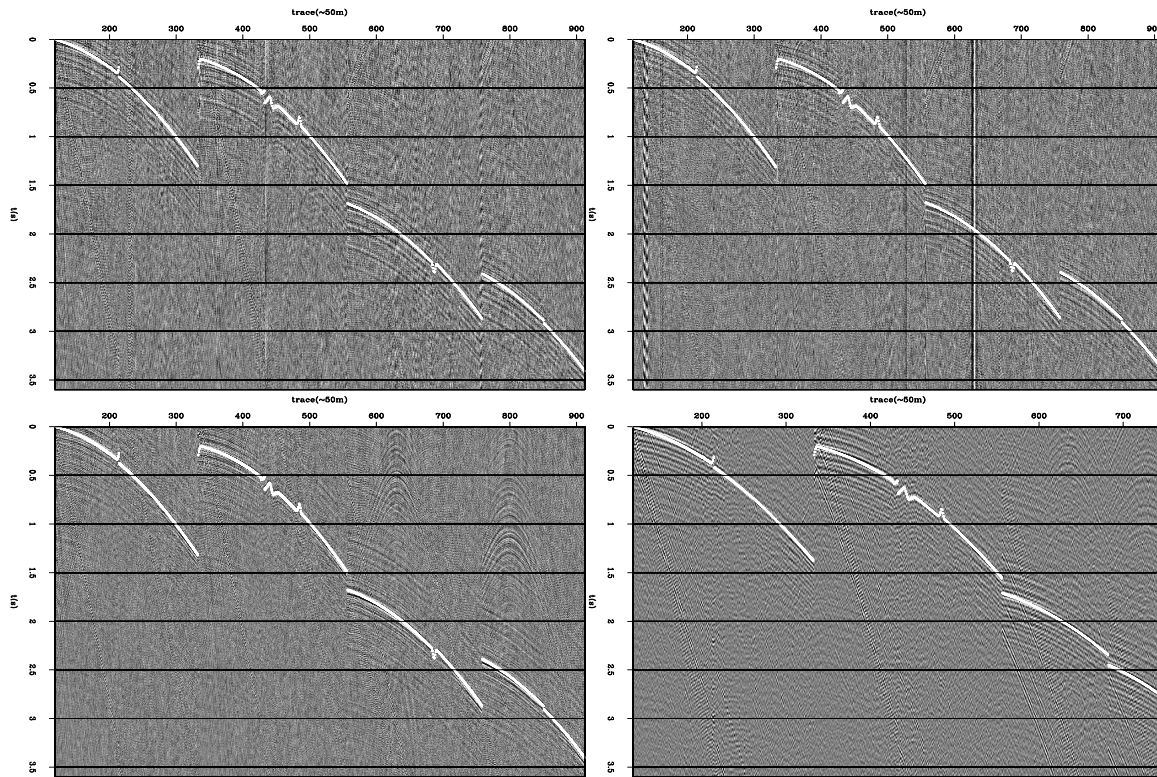TimeShot-inv [CR]

Figure 3.16: The forward modeled time picks from 15 inversions are plotted over each panel in Figure 3.10. The accuracy of the times is very faithful to the irregularities in the array layout in contrast to the auto-picked data input to the inversion, Figure 3.14. TimeShot-mods [CR]

Figure 3.16 shows the forward modeled time picks from the inversions. The best model parameters from all 15 inversions are forward modeled and plotted. The points are very precise in picking even minor deviations from regularity in the array geometry and are much better at describing the event than the auto-picks in Figure 3.14. The quality difference between the picks that went into the inversion and the result account for the residual energy in the first panel of Figure 3.15. The inversion was not able to match the bad data because it was required to honor equation 3.3. Therefore, the result is not able to converge to zero.

I sected values for map coordinates and velocity at the location of the minimum from all 15 inversions using graphs such as those shown in Figure 3.15. The best velocity for the three data sets from 2004 was $\sim$ 1455 m/s, and 1440 m/s for the data from 2005 which are appropriate for compressional waves in the water column. The (x, y) locations measured as

distance from the NW corner of the array in kilometers were: (2.5,-44.0), (2.0,-40.0), (2.4,-43.5), (2.75,-43.5) in the rotated local coordinate system shown in Figure 3.6. Because the inversion returned very reliable water velocities in all cases, the lack of sensitivity for the subsurface parameters is mitigated. If the event must be in the water column, the sea surface and floor are the only potential reflectors. Since the receivers are on the sea floor, the most simple explanation of these results is that the source is also at the sea floor and reflects from the sea surface once before arriving at the array. However, at such large horizontal offset, and shallow water column, the forward model cannot distinguish between this model and the direct arrival from the surface to the sea floor. Conversely, the source could be at the array, travel roughly horizontally, and reflect back from a nearly vertical object. The latter situation is not plausible, nor is the direct arrival associated with this model recorded by the array.

Using the velocity and location values calculated by the inversion, time picks at every receiver station were forward modeled for use as data in an exhaustive search over location variables. The gray scale in Figure 3.17 shows the inverse of the total data misfit normalized by the maximum value. The same line is plotted on the panels for reference (8° from the horizontal axis). All data volumes resolve the source location at very similar locations with very similar trends in precision. A request to the Norwegian Petroleum Directorate indicates that there were no active seismic vessels acquiring data in this region on or around February 15, 2004. The similar location derived from the data collected in 2005 indicates that surface seismic acquisition is likely not the cause of this feature.

Figure 3.18 is a detailed map of the Norwegian and British oil infrastructure around the Valhall reservoir. Also plotted are the footprint of the receiver array (gray rectangle), and the location of the trace used as the virtual source function (push-pin icon G118) for the gathers presented above. Superimposed as a flat image (and thereby introducing small projection errors) is a map of the British production facilities in the area. The black line from the receiver array ends in the SW at the locus of the energy provided by the inversions described above. The location of the source for the events in Figure 3.10 is exactly over the Ardmore field in British controlled section of the North Sea. The British Department of Trade and Industry [3] describes the Ardmore development. It differs from the Valhall operation in several important

---

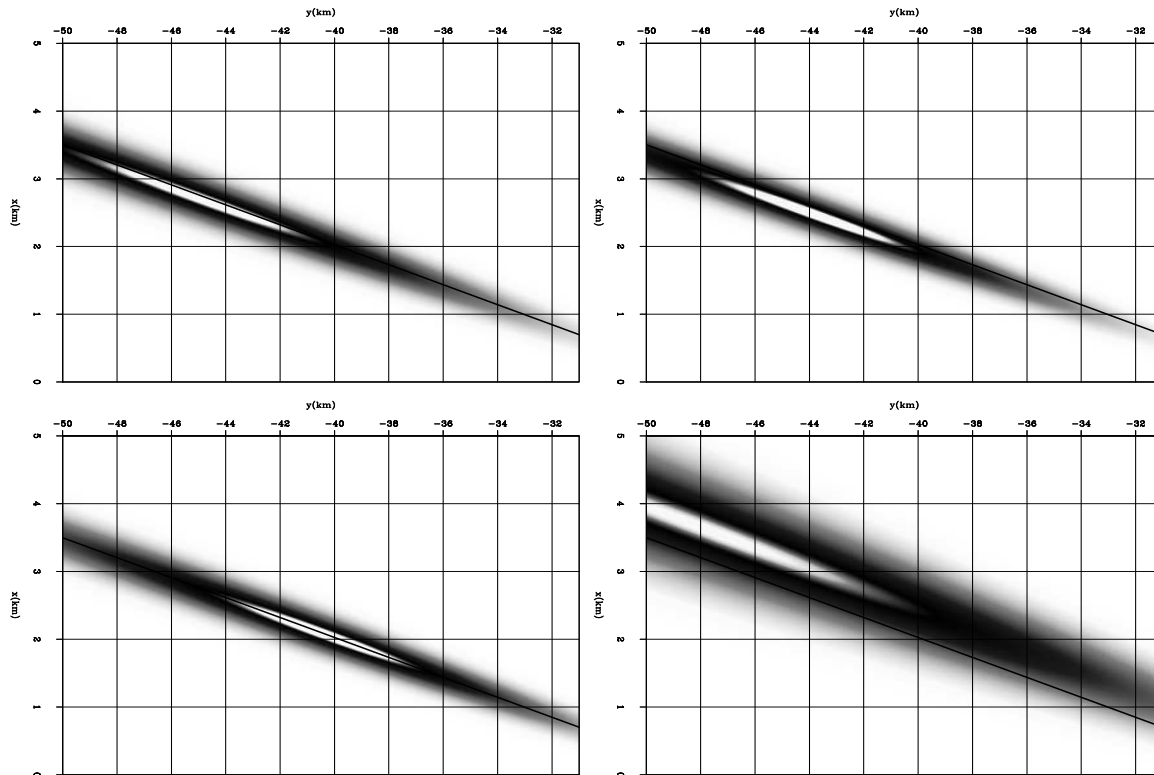[3] www.og.dti.gov.uk/environment/permits/TuscanArdmore.htm

Figure 3.17: Sensitivity to horizontal location for the inversions using the four periods of passive data. TimeShot-xy [CR]

Figure 3.18: A small section of the British infrastructure map with the Norwegian infrastructure and the location of the energy source inverted in Figure 3.17. The inverted locus of the energy in the gathers (the SW terminus of the black line) corresponds exactly to the location of the Ardmore development. TimeShot-mapa [NR]
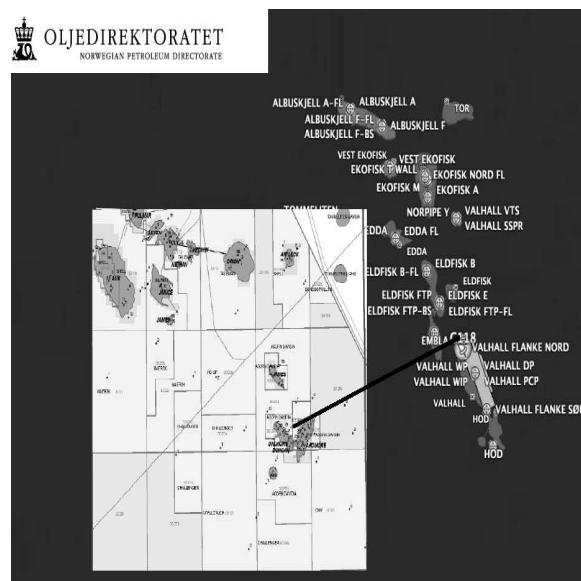
ways. The Ardmore facility has no permanent structures, nor is it connected to an export pipeline. Instead, two vessels are permanently moored over the field: One contains production facilities, the other is used to store liquids for periodic transport to shore. This operation has no way to handle produced gas and must therefore flare gaseous production to the atmosphere.

CONCLUSIONS

The energy responsible for the features in the raw data, Figures 3.4 - 3.5, are likely related as they share a locus at the production facilities within the array and are both characterized by water velocity. I suspect that the noise-train in Figure 3.4 is a collection of aliased repeating events similar that found in Figure 3.5. Also, the gathers generated with data showing the diffuse noise-train, Figure 3.4, have almost identical kinematics to the raw data showing a crisp water-velocity event, Figure 3.5. It is possible that some activity on the platforms could sometimes be impulsive in nature or more drawn out. In either case, the correlated gathers do not show convincing evidence that the energy penetrates the subsurface after the direct arrival is recorded. Forward modeling a direct arrival to the receiver stations shows that water velocity events are aliased close to the source with this array. This analysis of the raw data also shows that the predominance of observable energy shares a single source location.

Correlating the raw data to produce synthesized shot gathers was always performed on 12 second records. These individual results were summed from 1-2500 times to generate some approximation to active seismic shot gathers in accordance with the theory of interferometric imaging. The gathers produced with a single 12 second record, Figure 3.11, or 15 hours of data correlated in 12 second sections, Figure 3.10, show an identical strong event with velocity 1450 m/s arriving from the Ardmore field 40 km to WSW.

It seems counter-factual to attribute the source of this energy to a distant production operation when major production, drilling, and injection activities within the array have not generated events that match the power and bandwidth of this most obvious arrival. Gas flaring at Ardmore is the most significant difference between the two operations. I believe that the gas flare is

Chapter 4

Time sections from Valhall

SUMMARY

NMO-stack, using a velocity model generated with active seismic surveys, was performed on synthesized shot gathers produced from the Valhall passive seismic array. Inline sections show strong events with minor deviation from horizontal. The events are completely discontinuous across the cross-line direction and uninterpretable on time slices. While some events can be tracked between sections, some events appear/disappear completely. While aliasing could be the cause, the apparent dips of these events are geologically unreasonable.

I performed various muting and modeling tests using the kinematics of the energy from the production facility 40 km away from the array. Shot gathers were modeled, windowed, and muted before NMO-stack processing to compare to original results. These exercises show that the direct arrival and ringing correlation side-lobes are a coherent noise source that gives rise to the events that are immediately apparent with this simple processing sequence. The most interpretable volumes of stacked data are produced by muting to the time of the direct arrival. However, such volumes do not yeild obvious and convincing subsurface information. Two artifacts within the results may be indicative of the presence of real subsurface reflections that this processing methodology has not revealed.

INTRODUCTION

After synthesizing shot gathers from the passive data collected by the Valhall subsea array, I applied normal moveout (NMO) and stacked the volume across offset. A similar processing flow with the same data was described by Hohl and Mateeva (2006) using a single 2D section of raw data. Several coherent, roughly flat events are imaged in all time windows over which the raw data were subdivided for processing.

Of paramount concern with this workflow is the presence of the strong direct arrival from 40 km away from the array investigated in Chapter 3. With a single strong source dominating the data, so far away from the array, the synthesized gathers did not reveal obvious subsurface information that conform to the kinematics of surface data. Therefore, the event violates the kinematic assumptions inherent in both the CMP coordinate system and NMO. As such, this event is strong coherent noise in this processing strategy. I employed two main methodologies to investigate whether the events in the stacked sections are actual subsurface reflections or the coherent stack of the distant production noise: Constant water velocity NMO, and various modeling and muting experiments using the inversion results described in the previous chapter. I generated comparative volumes of NMO-stacked results with these strategies to investigate the effects of the distant noise source in the stacked domain.

NMO-STACK TIME SECTIONS

Figure 4.1 shows strong coherent events in inline sections after NMO-stack using a velocity model provided by BP. The shot gathers were synthesized with the spectral whitening correlation strategy explained in the previous chapter. The top images were generated using 9 hours of raw data from February 14, 2004. The bottom images used 5 hours of data 25 hours later. The two volumes were processed identically. The left columns are sections extracted from the 3D cube at the central CMP in the cross-line direction and bandpassed from 10 to 25 Hz. The sections on the right column were extracted from the next cross-line CMP, +150 m. The dissimilarity between the results produced with data differentiated only by when they were recorded leads immediately to the conclusion that the events seen in the sections are artifacts

rather than subsurface reflections.

The shallow geology under the array consists of very flat, continuous layers (Figure 6.1). These NMO-stack results have no continuity on time-slice panels, and the events change radically in character and timing between adjacent cross-line CMP locations. Some events (dis)appear with no continuity between cross-line locations. Other events are correlatable between the two sections, though move approximately 0.25 s. Assuming that the events are aliased in the cross-line direction, this is an apparent velocity of 600 m/s. The strength of the events is also suspicious as the shot-gathers shown in the previous chapter do not contain obvious refelctions from geologic events at these times.

Figure 4.2 used the data from the beginning of the recording period, top row of Figure 4.1, and a constant water velocity for the NMO correction. The two panels correspond to the same cross-line sections from the center of the array. Coherent events are present, but have troubling characteristics again. There is now some consistency of events in the cross-line direction. Figure 4.3 is a time slice through the same data cube at 0.4 s. The panel shows some uniformity of polarity, which is indicative of flat events. The parallel events at early time in Figure 4.2 are all at integer multiples of 0.048 s, which is the time through the 70 m water column. Since NMO was performed with water velocity, multiples are flattened in this volume. It is impossible to determine if these events are multiples or correlation side-lobes in this case.

Figure 4.2 does show that NMO with water velocity drastically changes the events that coherently stack compared to Figure 4.1. The sections are obviously different. Despite the fact that the synthesized shot gathers did not capture the apex of the hyperbola due to the distant noise source, the water velocity NMO was close enough to destroy almost all the odd events in Figure 4.1.

The inversion described in the previous chapter provides very accurate prediction of the kinematics of the strong event in the correlated gathers. Figure 4.4 shows three correlated shot gathers from the Valhall array. The first is a modeled gather synthesized by convolving a wavelet with the forward modeled time picks for the distant production noise described in Chapter 3. The first trace is the zero-offset trace. The second gather was synthesized from
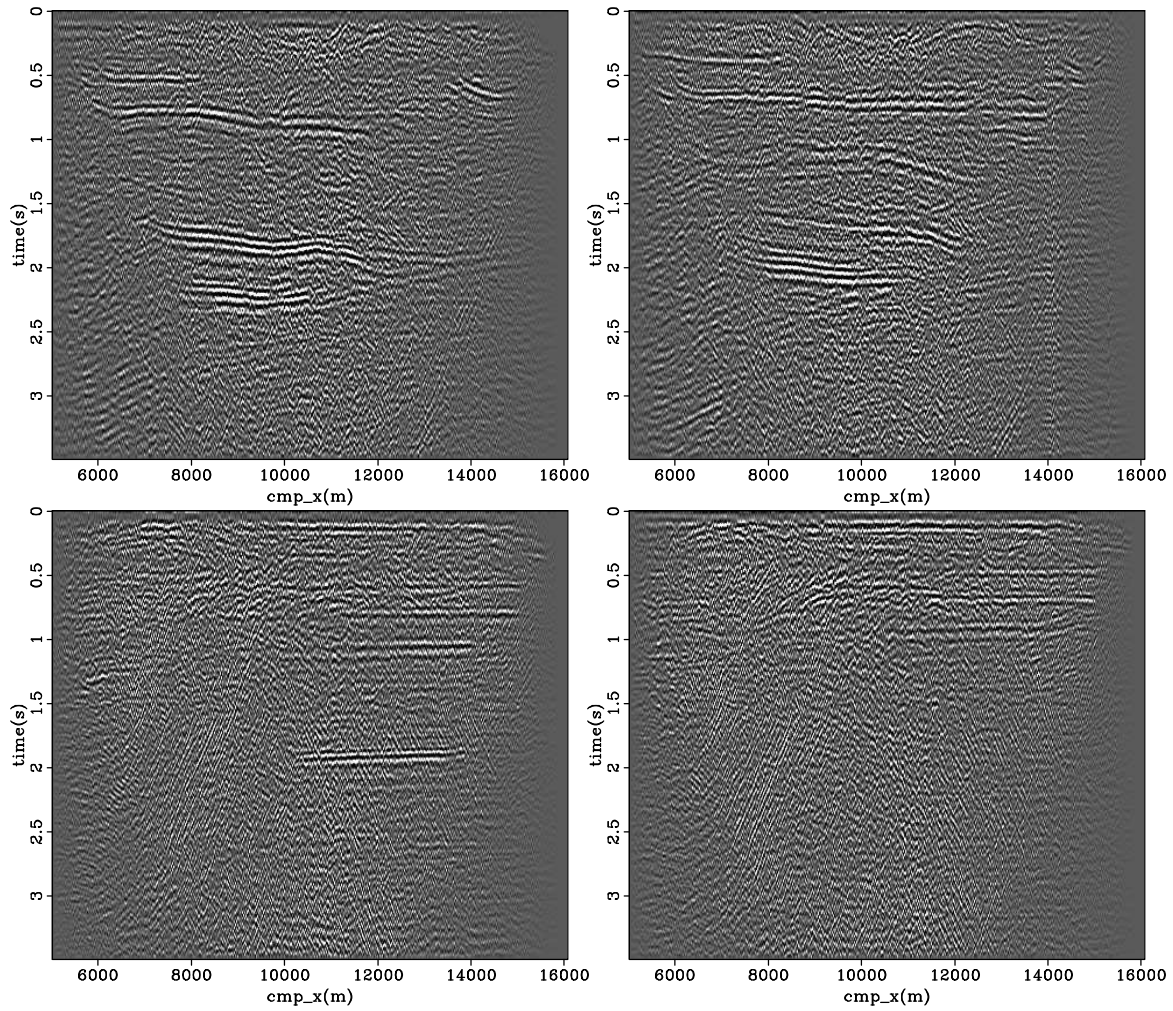
Figure 4.1: NMO-stacked inline sections from the center (left) of the Valhall array, and (right) the neighboring cross-line CMP (+150 m). Top and bottom rows were processed identically using data collected 25 hours apart in February, 2004. TimeCMP-events [NR]
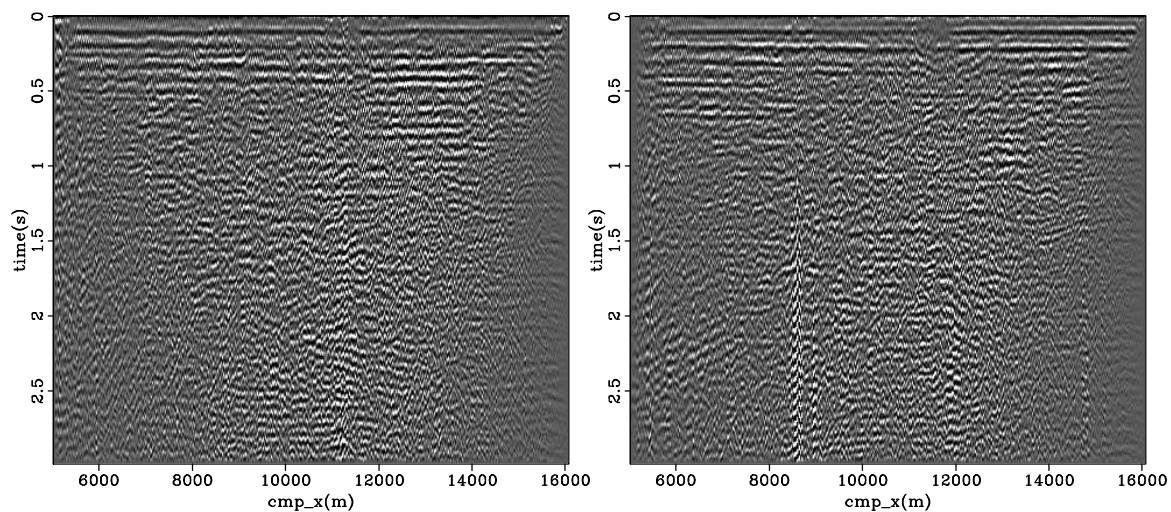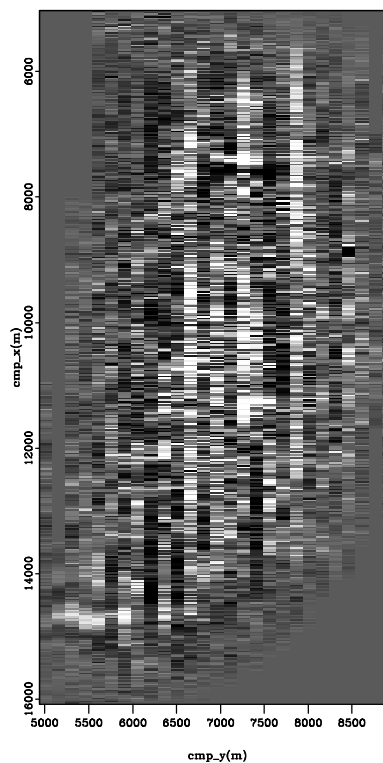
Figure 4.2: Inline sections corresponding to the data on the left of Figure 4.1 using constant water velocity for NMO. TimeCMP-water [NR]

Figure 4.3: Time slice, 0.4 s, through the NMO-stacked cube using water velocity. Some continuity exists across the panel, while individual lines in the CMP-X direction show stong coherence. TimeCMP-waterslice [NR]

passive data and muted to 0.32 s after the arrival time of the event in the top panel. Notice that there are significant correlation side-lobes that remain in the section. The bottom panel shows the application of a short time window around the arrival of the noise event to the synthesized gather in the center panel.

Figure 4.5 shows inline time sections exploring the relationship of the direct arrival from the distant production facilities to the events observed in Figure 4.1. The top-left panel is the same as Figure 4.1. The top-right panel used the muted data from the top of Figure 4.4. The bottom-left panel used the data from the bottom of Figure 4.4. The final section is the result of processing the forward modeled shot gathers. The difference between the first two panels is nearly impossible to find. Inspecting the bottom two panels from the bottom up, there are many similar features, that share characteristics: location, length, dip, and curvature. The steeply dipping events on the right section are not present in the left, but almost all other main features are.

In contrast to the results above, Figure 4.6 compares the muted shot gather generated with the first 9 hours of passive data (top Figure 4.4) with one, bottom panel, from data collected 25 hours later. In contrast to the top panel, the data from a day later do not suffer as much from the ringing correlation side-lobes. The mute leaves a small amount of co-parallel energy, but this time does a much better job at removing most of it. Figure 4.7 compares the NMO-stacked sections created with this second volume of data with and without the mute to the direct arrival. Most of the strong flat events have been eliminated.

The right panel of Figure 4.7 shows a stacked section of data from which the bulk of co-herent noise has been removed before NMO-stack processing. This volume then represents the best candidate for finding subsurface structure. Notice a roughly sinusoidal series of iden-tifiable events, on the right panel, across the section between 0.0-0.2 s. This event will be described further in the next figure.

Figure 4.8 shows two neighboring slices through the 3D volume of the stacked data pro-duced with the muted shot gathers. The volume has been bandpassed from 8-15 Hz. The time slice in the right panel was extracted through the curving events near time zero mentioned above. There is a roughly circular feature on the time slice identifiable as a mostly white ring
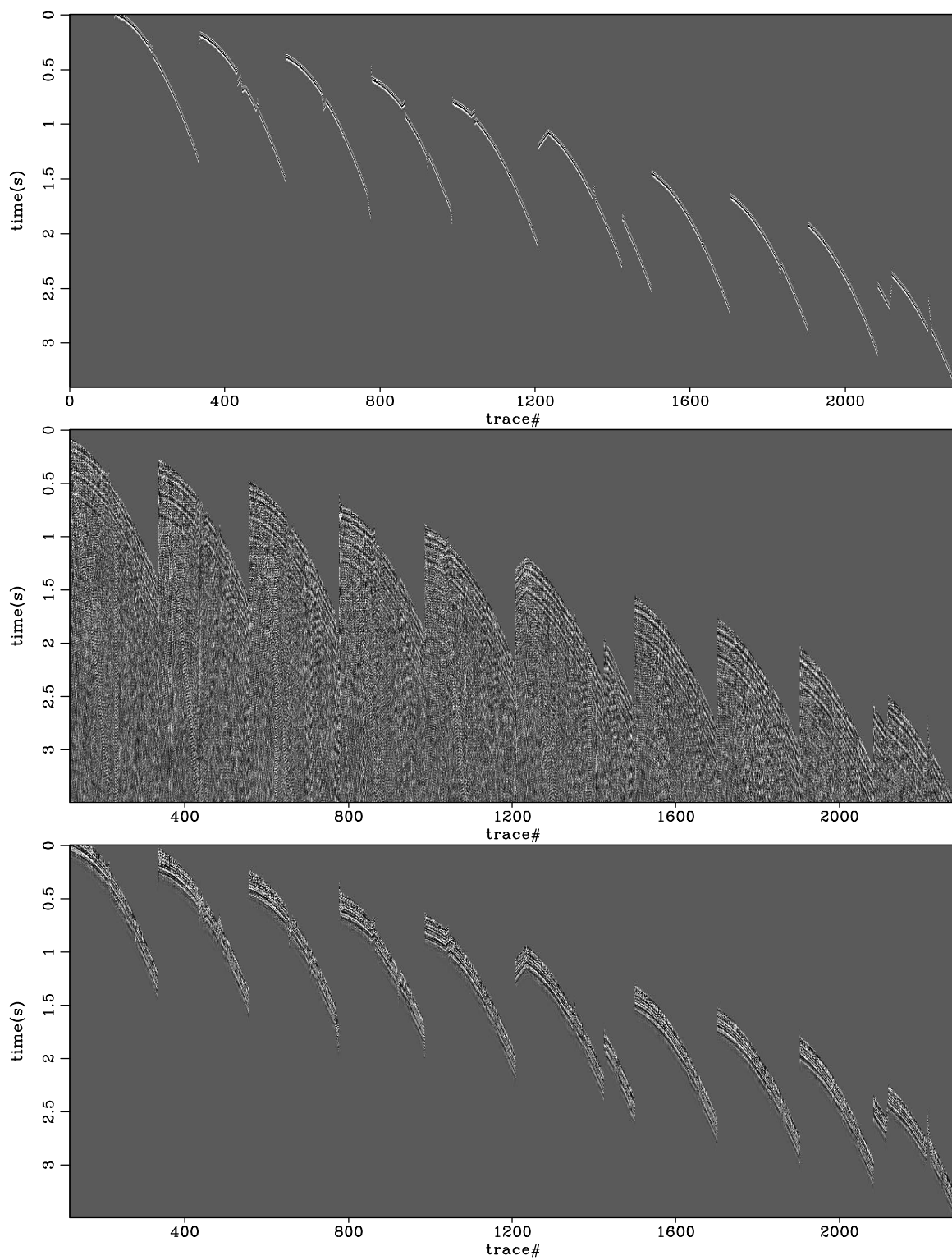
Figure 4.4: Two shot gathers used to test the origin of the events in Figure 4.1. Top panel is muted to the kinematics of the distant noise source. Bottom panel is a narrow window around the modeled event. TimeCMP-modldat [NR]
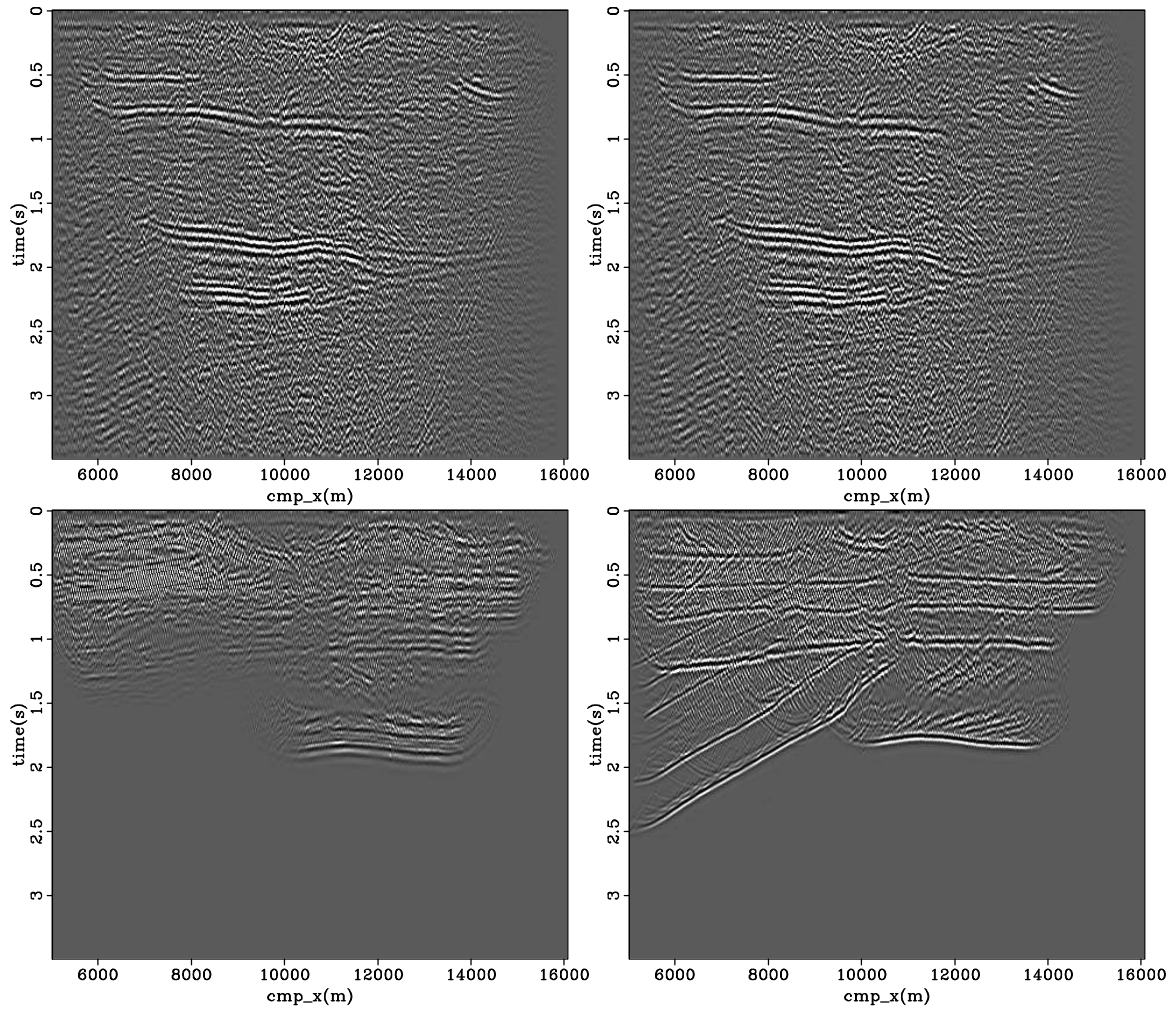
Figure 4.5: Left to right: Inline NMO-stacked sections from the center of the array. Top-left panel processed all data available after correlation. Top-right panel processed the correlations after application of a mute to the modeled direct arrival. Bottom-left panel used data windowed around the direct arrival. Bottom-right panel used a modeled data volume containing only the direct arrival. TimeCMP-evnttest [NR]
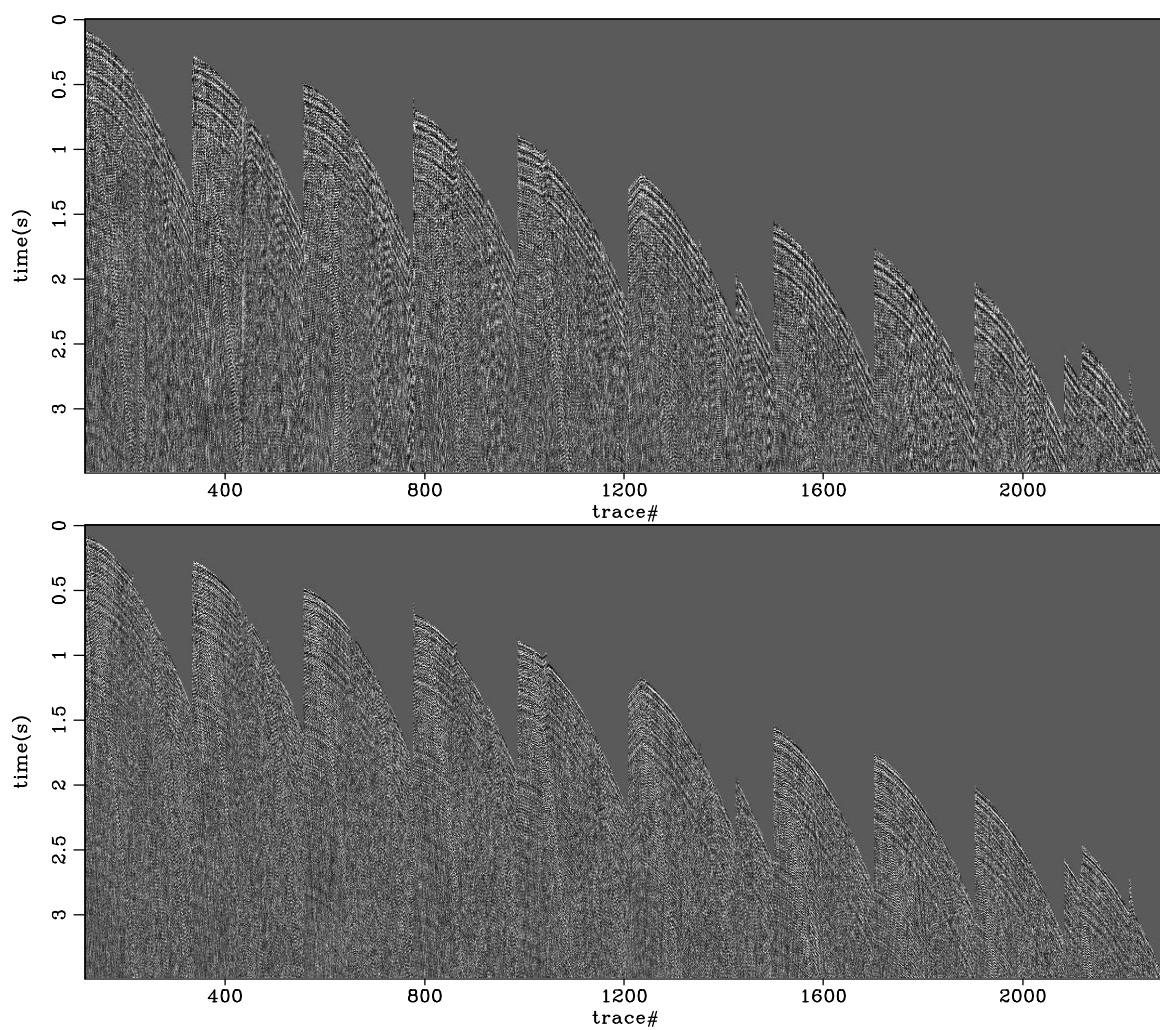
Figure 4.6: Top:Muted shot gather corresponding to the first 9 hours of data (Figure 4.4 top). Bottom: Muted shot gather generated with 5 hours of data 25 hours after those used to create the top panel. TimeCMP-modldat2 [NR]
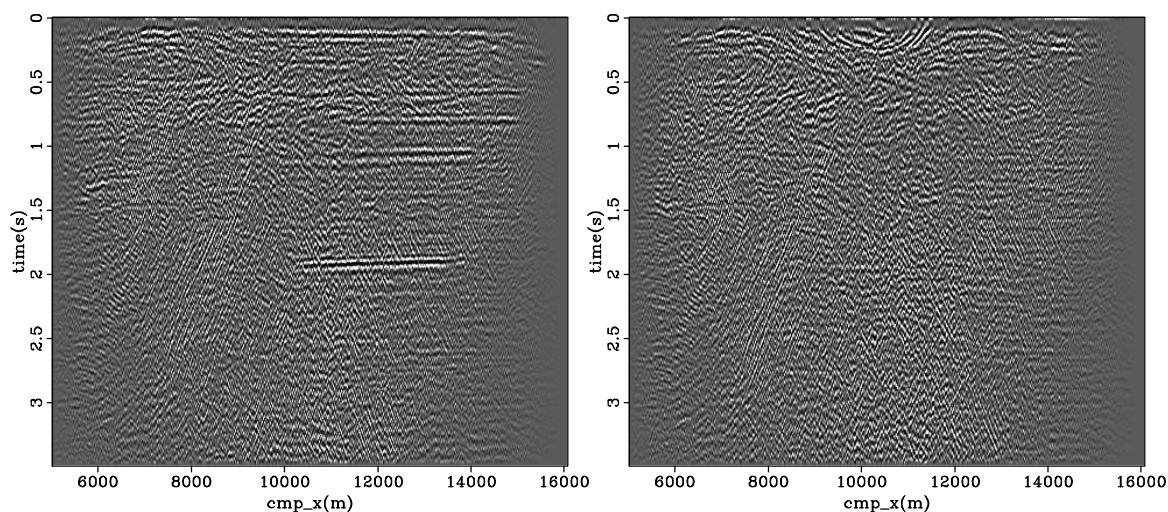
Figure 4.7: NMO-stacked sections from center cross-line location generated from data a day after those used for Figures 4.4 and 4.5. Muting down to the direct arrival (Figure 4.6) successfully eliminates all coherent events. TimeCMP-evnttest2 [NR]

with a dark center. In movies progressing down the time axis, these events are expanding, circular rings emanating from the location of the surface production facilities.

There is little continuity in the cross-line direction throughout the bulk of the stacked volume. The events that approach the edge of the array are suspicious since these CMP locations do not have the complete coverage needed to build reflection gathers as discussed in Chapter 2. However, there are a few events that could be real subsurface structure. There is a fairly continuous event that can be interpreted across the first panel at just less than 2 s. The event is interpretable on the next section as well, and there is a black-white-black event on the cross-line panel where the section line cuts through 2 s. This seems like the most likely event in the volume to be a genuine reflector. There is also a coherent reflection between 6000-8000 m just beneath the time line on the right panel that is stationary in both sections. At approximately the same time on the right edge of the section, this event is again present. This event could be associated with the production noise, or a real subsurface reflection that is interrupted in the middle of the array. Better results are presented in Chapter 6 that compare well with images produced with active seismic surveys and allow more convincing interpetation of the subsurface.
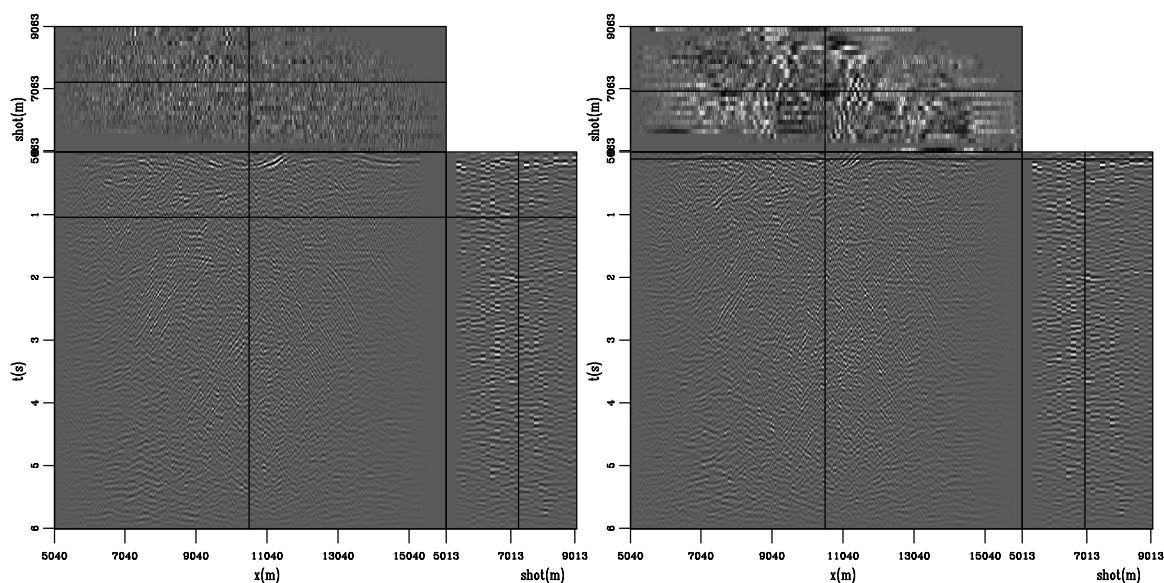
Figure 4.8: Best 3D time section processed from 5 hours of passive data in February, 2004. Preprocessing included correlation, deconvolution, mute, NMO, and stack. TimeCMP-Fourth.mute [NR]

Figure 4.9 uses the synthetic transmission functions from Chapter 2 to analyze the effects of the distant production noise on passive data. The first panel is a transmission wavefield due to a source below the center of the model domain with the addition of an event (0.75-0.9 s across the wavefield) modeled from the analysis of the farfield noise in Chapter 3. The second and third panels present NMO-stacked sections using correlations of transmission wavefields due to sources below the left edge and center of the model domain which included the farfield direct arrival.

The steeply dipping events in the NMO-stacked sections in Figure 4.9 are readily identified in the sections produced with real data above (e.g. Figures 4.7 and 4.8). Also note the concave-up features that connect to the steep events at the surface. These noise events are large hyperbolas, centered over the location of the subsurface source, that have been folded down at the free surface. These events are not present in sections produced without the addition of the farfield noise event. Sections produced with only the farfield noise event contain only a strong event at $t = 0$ s. The steep events are caused by the convolution of the farfield noise event with the subsurface reflections. Therefore, the presence of the bowl-shaped structures in Figure 4.8
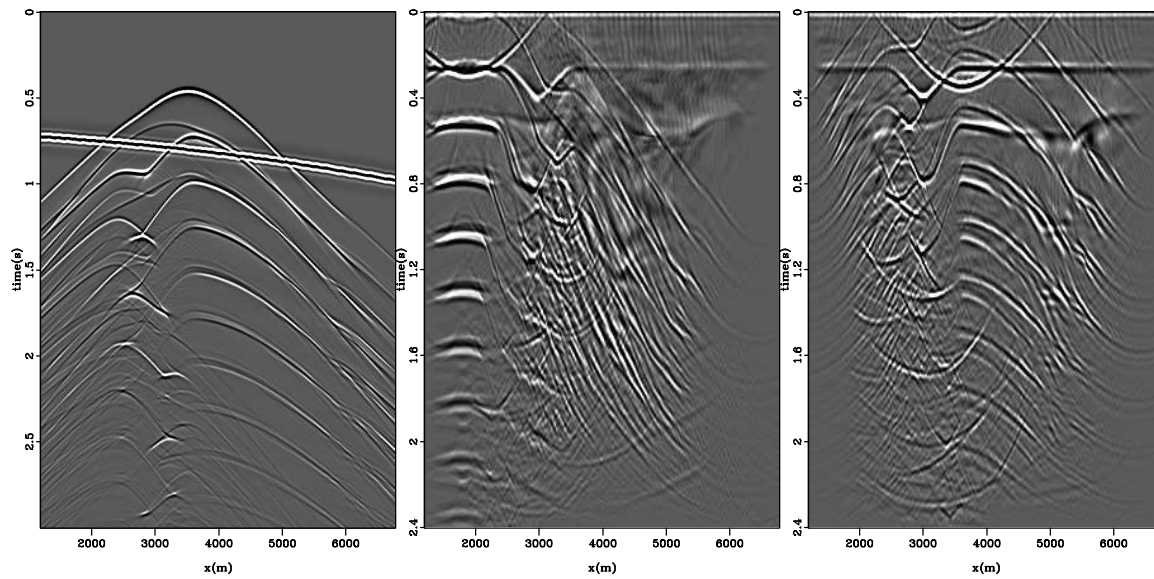
Figure 4.9: Left: Transmission wavefield from the modeled data presented in Chapter 2 due to a source below the center of the domain. Center: Depth image produced from a transmission wavefield due to a source below the left edge of the domain. Right: Depth image produced from the transmission wavefield in the left panel. TimeCMP-synNMO [NR]

and the steep events in Figure 4.7 are indicative that there is real subsurface information within the passive data that was not well imaged with this processing strategy (NMO-stack).

DISCUSSION AND CONCLUSIONS

NMO-stack of the whitened, correlated passive data from the Valhall OBC array produces strong events on inline sections. Other authors have interpreted these as geologic information after processing only 2D subsets of the data (Hohl and Mateeva, 2006) spanning a limited acquisition duration. Very good imaging from many active surveys has been carried out at the location that shows the shallow geology to be very flat and uniform in both map directions. The time sections shown above however have several unbelievable characteristics that lead me to conclude that these are coherent noise rather than genuine subsurface information. Of paramount importance, the same section produced with the same processing flow using different windows of the available data (separated by 25 hours) have dramatically different

events.

Because the kinematics of the direct arrival in the shot gathers are well constrained, it was possible to mute, model, and window around the event before NMO-stack. Using the gathers generated with the first section of data, the muted volume was barely distinguishable from the result produced with the full data. However, the section produced with a short time window around the event was very comparable to that produced by forward modeling the event. The final observation that links these two facts was provided by performing similar experiments with data from 25 hours later. The NMO-stacked section produced with the muted gathers had no evidence of the events seen previously. Comparing the muted shot gathers produced from the two data subsets showed that there is significant ringing (correlation side-lobes) present in the gathers produced with the data from early in the recording effort.

Using a constant (water) velocity for NMO processing completely eliminates the strong events present when using the correct velocity model. Instead, several orders of flat multiples are present in this volume relating the 70 m of water above the geophone array. The OBC cables are buried approximately 1 m beneath the sea floor. Thus, within the first 2 samples (10 m depth), the interval velocity model has values in excess of 1600 m/s which is much too fast to correctly NMO the direct arrival that arrives at 1450 m/s. This experiment again confirms that the strong events produced by NMO processing the correlated gathers are coherent noise rather than reflections from the subsurface.

These experiments lead me to conclude that the strong events revealed by sorting the gathers to CMP coordinates, NMO, and stack are artifices of the coherent noise event generated by a production facility 40 km away from the Valhall array. While there may be subsurface information in the time domain gathers, it is weak and distributed amongst the volume to make it very hard to interpret in both the shot domain and the CMP domain. Muting the synthesized shot gathers before NMO-stack to remove the distant production noise source, produced a much simplified data volume. A second noise source was then identified in the volume as expanding rings emanating from the production facilities contained within the array. These features correspond to the local water-velocity events picked on the raw and correlated data sections in the previous chapter. The best stacked volume of data has several intriguing features that are indicative of subsurface reflectors, but do not provide convincing results that the

passive data recordings can illuminate the subsurface at Valhall.