PASSIVE SEISMIC IMAGING

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF GEOPHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Bradley Artman
March 2007

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Biondo Biondi) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Simon Klemperer)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Gregory Beroza)

Approved for the University Committee on Graduate Studies.

Abstract

Passive seismic imaging is a sound theoretic idea that has languished in the realm of theory for nearly four decades since the original concept was presented 1968. Only a small number of experimental tests have been fielded in that time, and none provided convincing results. The recent surge in interest in the idea has generated a substantial literature that has proved the validity of the theory beyond the initial conjecture, as well as experimentally at the lab scale and the global scale. This thesis makes a substantive contribution to the field by extending the processing algorithms available for passive seismic data and by analyzing a real 3d data volume from a 2300 station areal array.

The experiments with synthetic transmission wavefields in Chapters 2 and 5 show that investigating subsurface structure with body waves can be accomplished with a limited distribution of subsurface sources directly beneath the sensor array. Sources at shallow depths near the edges of the array are not required. Without them, however, some artifacts remain in synthesized shot gathers. Chapter 5 shows that migration of passive data produces interpretable images of the subsurface even when the assumptions behind the rigorous theory are violated. In extreme cases, such as capturing only a single subsurface source in passively recorded data, synthesized gathers look completely incorrect, but migrating the data produces very good results. The same conclusion applies when processing transmission wavefields due to many sources.

Direct migration of transmission wavefields, introduced herein, is a robust, convenient processing method. Any conventional shot profile migration algorithm that maintains separate up-coming and down-going wavefields can be easily modified to use passive data. The

only change required is to use the data for the initial condition of the source wavefield. Processing passive data by direct migration can provide substantial computational savings. This is especially true if multiple images are produced from the total volume of data collected. Because we are relying on the natural rate of seismicity to supply the energy for investigating the subsurface, much of the total recording time is likely devoid of useful information. Therefore, it is useful to produce quality control images with multiple short windows of data. However, synthesized shot gathers may be required for velocity analysis or addressing common data problems such as noise removal or aliasing.

Over 35 hours of passive data collected at the Valhall development in the North Sea was processed to produce synthesized shot gathers and migrated images. The synthesized shot gathers did not provide any subsurface information. Instead, they were dominated by a direct arrival through the water column from the gas flare at a production facility 40 km away from the array. The normal moveout correction applied to the gathers did not produce any interpretable subsurface information. Importantly, the energy from the gas flare is a coherent noise source that arrives at the array at such a shallow angle that it is confined to the water column and does not refract into the subsurface. As a coherent noise source, the energy manufactures events in results that are not subsurface reflectors. Removing the far-field noise from the synthesized gathers also removes the interpretable events from the NMO-stacked result.

Images produced with direct migration of the Valhall passive data in Chapter 6 contain interpretable subsurface events that seem to correlate with events in active seismic images. The basic syncline structure under the array is well represented in the passive images. However, data modeled after the coherent noise from the distant gas flare produces almost identical images to those produced with field data. This fact drastically discounts any value in the images produced with field data even for events that are highly correlatable between the active and passive images. Direct migration and the migration of synthesized shot gathers from modeled data produces almost identical images. For the Valhall data, however, the two strategies produced markedly different results. The migration of synthesized gathers always produced substantially worse results than direct migration. The reason for this difference is again due to the far-field production noise.

The Valhall Life of Field Seismic array is a unique installation of seismic sensors. The geophones can continuously record data between the active surveys that are performed several times per year. The chief obstacle in continuously processing data from the array is the transportation of the large volumes of data from the offshore installation. Most, or all, of the tele-communications bandwidth currently available at the installation would be used by continuously moving data to shore for processing. For a single component, a single day of raw data is more than 46 Gbytes. The correlation volume, in the frequency domain, produced from the array is similarly sized. However, if a modest computer cluster were available on-site for processing, a small image volume can be transferred easily. The nature of the ambient noisefield at Valhall seems to require correlation processing to remove the far-field energy rather than direct migration, which would save computation demands. Unfortunately, my analysis of the day-long passive records from Valhall do not provide convincing evidence to justify such an experiment.

Passive seismic imaging shares many characteristics with addressing multiples in conventional data. Imaging with subsurface sources is kinematically equivalent to imaging primaries with multiples. These two imaging techniques rely on the correlation of up-coming and down-going events in the data. Predicting multiples for subsequent removal is often an important processing step for active data. Prediction techniques for surface related multiples rely on convolution of the traces in a data volume. I introduce a new, convenient multiple prediction method that performs the convolution during the course of shot profile depth migration of active seismic data. This new multiple prediction is defined in the image domain with space and subsurface-offset dimensions. Image-space surface-related multiple prediction (IS-SRMP) is velocity independent as is the data-domain equivalent (SRMP). The result is the same as that produced by migrating the SRMP volume. Because wavefield extrapolation is such an expensive algorithm, the additional computational cost of multiple prediction during migration is small.

Preface

The markings [ER], [CR], and [NR] included in the figure captions in this thesis denote the extent to which each figure is reproducible by anyone desiring to do so. The electronic version of this thesis[1] provides the original figures and Fortran90 programs written to produce all of the results. Most programs are included locally in the chapter directories. The ShotMig directory contains my split-step Fourier plus interpolation shot-profile depth migration program. I assume you have a UNIX-based workstation with Fortran90, and C compilers, an X-Windows environment, and the Stanford Exploration Project (SEP) software available from our website (make rules, programming libraries, and LaTeX packages).

Reproducibility is a way of organizing computational research that allows both the author and the reader to verify and regenerate results. Reproducibility also facilitates the transfer of knowledge within SEP and between SEP, its sponsors, and the geophysical community at large.

ER denotes Easily Reproducible results of processing described in the paper. I claim that you can reproduce such a figure from the programs, parameters, and makefiles included in the electronic document. The data must either be included in the electronic distribution, be easily available to all researchers (e.g., SEG-EAGE data sets), or be available in the SEP data library[2]. Before the publication of the electronic document, someone other than me tested my claim by destroying and rebuilding all ER figures.

---

[1] http://sepwww.stanford.edu/public/docs/sep128

[2] http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html/

CR denotes Conditional Reproducibility. I certify that the commands are in place to reproduce the figure if certain resources are available. CR results have been tested to make sure that the makefile rules exist and the source codes referenced are provided. Conditional reproducibility applies to figures produced with excessively long or parallel computing processes or when using proprietary data.

M denotes a figure that may be viewed as a movie in the web version of the report. A movie may be either ER or CR.

NR denotes Non-Reproducible figures. SEP discourages authors from flagging their figures as NR except for figures that are used solely for motivation, comparison, or illustration of the theory, such as: cartoons, drawings, scans, or figures taken from other publications.

Our testing is currently limited to LINUX 2.4 (using either the Intel or Portland Group Fortran90 compiler), but the code should be portable to other architectures. Reader's suggestions are welcome. For more information on reproducing SEP's electronic documents, please visit

`http://sepwww.stanford.edu/research/redoc/.`

Acknowledgments

I thank the sponsors and faculty of the Stanford Exploration Project (SEP) for technical and financial support given to me during my studies. My colleagues at SEP provided invaluable expertise, resources, and ideas to facilitate all of my work and make the study of Geophysics engaging and rewarding.

BP provided a data set as yet unique in the world for analysis in this thesis. The timely installation of the acquisition system at the Valhall reservoir in the North Sea provided an excellent opportunity to test the theory of passive seismic imaging at the kilometer scale. BP also gave me the opportunity to work with the data at their Advanced Imaging research group as an intern for a summer. While there, Ken Matson helped flesh out my ideas related to multiple prediction while I waited for computers to process the enormous passive data volume.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 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 in interferometry that has generated a wealth of literature (Wapenaar et al., 2006).

The second reason that passive imaging is not new science, is that interferometry is simply a generalization of the seismic problem formulated by earthquake and active-source exploration seismology. On 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 (Owens and Zandt, 1985; 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; Shapiro et al., 2005). 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.

To date, the literature concerning subsurface structural imaging at the reservoir scale has been largely relegated to theoretic and modeling pursuits due mainly 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.

1.2 Development of the correlation theory

Figure 1.1, simplified for clarity, shows the basic kinematics exploited in processing transmission 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 dashed 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 features 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.

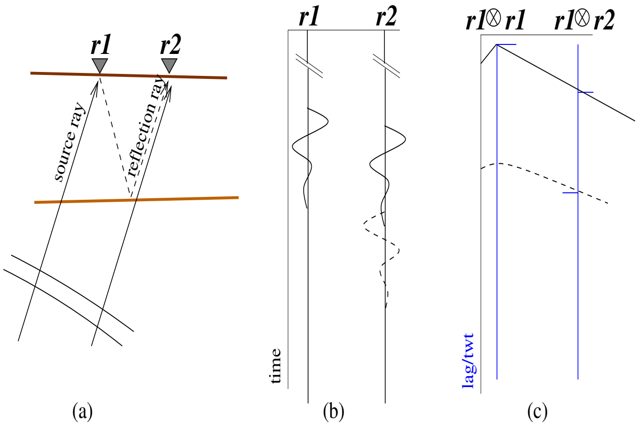


Figure 1.1: (a) Approximately planar arrival with rays showing important propagation paths for passive imaging. (b) Idealized traces from a transmission wavefield. (c) Shot-gather (reflection wavefield) synthesized using trace *r1* as the source. Intro-sketchP [NR]

Choosing trace *r1* as the comparison trace (or source function), panel c depicts the correlation spikes associated with the arrivals in the data (panel b), where $\otimes$ is correlation. A solid line with linear 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.

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

$$2\Re[R(\mathbf{x}_r, \mathbf{x}_s, \omega)] = \delta(\mathbf{x}_s - \mathbf{x}_r) - \int_{\partial D_m} T(\mathbf{x}_r, \xi, \omega) T^*(\mathbf{x}_s, \xi, \omega)\, \mathrm{d}^2\boldsymbol{\xi}\ , \tag{1.1}$$

where the symbol * represents complex conjugation, and subscripts r and s indicate different station locations, $\mathbf{x}$, at which a transmission wavefield, T, is recorded. Only the real part of the correlation, $\Re$, is maintained. After correlation r and s acquire the meaning of receiver and source locations, respectively, associated with an active survey, R. The RHS represents summing correlations of windows of passive data around the occurrence of individual sources, ξ. The symbol ∂D_m represents an arbitrary domain boundary that surrounds the subsurface region of interest on which the sources are located. The delta function, δ, is a negligible artifact of correlation processing. 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, indexed by ξ. If these requirements are met, the sum of the many correlation volumes synthesizes active surface seismic data. 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. 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. However, 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 Shapiro et al. (2005) extends a similar analysis to crustal scale observations.

1.2.1 Direct migration

The basic premise behind imaging with uncontrolled seismic energy is simple. Correlation of data traces calculates the time differences that energy packets require to travel between two locations. In a heterogeneous environment, the various reflections and diffractions that add travel paths beyond the simple direct ray between the receivers provides information about the medium. The entire field of seismic analysis is built around the interpretation of time differences between two locations. Most often, the time differences that are recorded and analyzed are those between the initiation of source energy (earthquakes, explosions, or vibrations) and the arrival of that energy at receiver locations. Correlating traces removes the initial waiting time for energy to arrive at a recording station. The resulting volume keeps the time delays of all the possible ray paths through the medium between the receivers. Simultaneously, complicated wavelet functions are collapsed to a bandlimited impulse.

In fact, the seismology community has been routinely analyzing complicated data by correlation algorithms that correspond to various specific cases of interferometric processing. Global seismologists correlate teleseismic arrivals with windows of data after the first arrival. These traces, and more recently wavefields, are then back-projected along the arrival path of the energy in receiver function analysis (Morozov and Dueker, 2003). Results are then summed from many earthquake sources to investigate the deep structure of the Earth. The cascade of projection and summation developed in receiver-function analysis can be defined as a migration strategy. Migration is a processing concept that is not specific to a particular implementation. It is applied with varying degrees of complexity and accuracy with many different algorithms to move energy in data records laterally and vertically (time or depth) to locations in an Earth model where energy has changed propagation mode or direction due to subsurface discontinuities. Producing images with receiver-function analysis, therefore, can be simplified to a two step algorithm: Correlate, migrate.

In the exploration community, vibrator sources sweep across a frequency spectrum using various source functions that are not impulsive in time. To define time-zero, records are correlated with the known source function to collapse the wavelet coda. The preprocessed shot gathers are then migrated to image the subsurface structure. Imaging with this acquisition strategy can be simplified to a two step processing flow: Correlate, migrate. For marine acquisition, the source bubble associated with underwater sources reflecting from the sea surface makes the real source signature a function of time rather than a simple impulse. Deconvolving the effect is an important step in processing marine data which can be analyzed in terms of a multi-step correlation process. Again, producing a final image of the subsurface can be simplified to the familiar two step algorithm.

Passive seismic imaging with many subsurface sources was presented in equation 1.1 as another two step process: Correlate, sum. The summation required is a sum of the correlations produced by individual unknown sources. The synthesized shot gathers can then be processed (migrated) as a volume of conventional data with many surface sources. The total processing flow is therefore a three step process: Correlate, sum, migrate. The basic components used to process passive data are the same as those described above, with the introduction of a summation.

I have capitalized on the flexibility of a specific migration strategy in this thesis: the wave-equation shot-profile migration algorithm (Claerbout, 1971). The descriptor "wave equation" denotes a method of extrapolating seismic energy through a (accurate) subsurface velocity model with a propagator derived from the kinematic solution to the wave equation.

The physics of wave propagation is the same for impulsive source functions or complicated wavefields that are a function of space and time. Seismic data records provide initial conditions for wavefields extrapolated with with these propagators. Extrapolation moves data away from the acquisition plane to a new datum, by modeling the physics of wave propagation and maintains appropriate spatial and temporal differences between events within the wavefield. After changing the experimental datum, the traces in the extrapolated wavefields can be still be processed in accordance with all appropriate techniques developed for seismic data to produce any desired result. Extrapolation is the first step of wave-equation migration to produce a subsurface image.

Shot profile migration algorithms extrapolate the source wavefields and data wavefields independently, in opposite temporal directions. When the energy within the two wavefields is collocated after propagation, we have located a subsurface discontinuity that gave rise to a reflection in the field experiment. Finding collocated energy is accomplished by calculating time differences between the two wavefields, and keeping only energy at $t = 0$. The zero-lag of the correlation of the traces within the two wavefields provides the desired result. Images from many experiments (shot gathers) are summed for a final result. Therefore, this migration algorithm can be distilled to components: Extrapolate, correlate, sum.

In this thesis I introduce direct migration as a simplified processing algorithm for passive seismic data. Instead of the three step processing flow explained above (correlate, sum, migrate), direct migration uses the correlation and summation within the migration algorithm to process transmission wavefields. If the extrapolation has been performed accurately, correlation will calculate accurate time differences between source events and reflections at a new reference datum. The migration algorithm calculates time differences by correlation at all subsurface depth levels in the model instead of only at the acquisition surface. The resulting volumes due to individual sources are then summed to provide a complete subsurface wavefield. This procedure honors the correlate/sum processing flow, with the addition of an extrapolation, or depth, axis. However, after extrapolating the data, we can discard non-zero lags in the correlations to remove the now unnecessary time axis. In so doing, the final result is a migrated image of the subsurface produced without creating shot gathers by correlation (that would then be ready for migration). This processing flow can be simplified to: Extrapolate, correlate, sum = migrate.

Direct migration of passive data completely skips the generation of synthesized gathers 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 data from a large passive seismic recording campaign.

Shragge et al. (2006) presented images produced with transmission wavefields due to earthquake arrivals and the resulting local reverberations beneath a recording array. Teleseismic arrivals were used as source functions for migrating data windows. The authors capitalize on the relatively simple, planar source functions to separate the up-coming from down-going energy in the transmission wavefields with a deconvolution algorithm. The separated wavefields from each earthquake event were migrated with a shot profile algorithm to produce a subsurface image. Acknowledging the kinematic similarity between correlation and deconvolution, teleseismic migration represents a hybrid of direct migration and migrating correlations.

Imaging with transmission wavefields is the same as imaging primaries with multiples in a marine survey. Multiples are due to strong up-coming events that reflect from the free surface and return to depth as source energy. Therefore, the water bottom reflection fulfills the same role as the direct arrival from a subsurface source. If the water bottom reflection is used as a source function in shot-profile migration (Guitton, 2002), all the first-order multiples are imaged at the subsurface location of primary reflections. That migration is an example of the processing flow: Extrapolate, correlate. If all shot gathers are correlated and summed, the time delay to the water bottom reflection is removed from the records, and we can migrate the resulting data volume (Shan, 2003). That processing flow is an example of: Correlate, sum, migrate. Further similarities to passive seismic imaging within the literature addressing the presence of multiples in seismic data can be immediately drawn from the "Inverse data space" concept (Berkhout and Verschuur, 2006).

1.2.2 Chapter description

Chapter 2 analyzes 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 are 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 the normal moveout correction (NMO) followed by stack. NMO-stack is another example of a simple migration algorithm performed in the time domain.

Chapter 5 presents depth 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. I illustrate the effectiveness of the method with the same 2D synthetic data used previously.

Chapter 6 presents depth migration results from the Valhall passive data set produced by both correlation followed by 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.

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 source and data wavefields. In this case, shot gathers are convolved instead of correlated, 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 a shot profile migration at minimal extra computational cost.

Chapter 2

Correlation processing experiments

2.1 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 vertically propagating 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.

2.2 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. Figure 2.1 shows two representative transmission wavefields and the velocity model with which they were generated. Imaging the subsurface structure with body waves is the goal of this study of the details of processing transmission wavefields. Therefore, I use only modeled transmission wavefields due to sources located directly beneath the model domain. All surface-related and internal multiples are included, and the transmission wavefields have been flux-normalized so we can neglect attenuation (Draganov et al., 2003). 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, axis **x**, spaced 20 m apart.

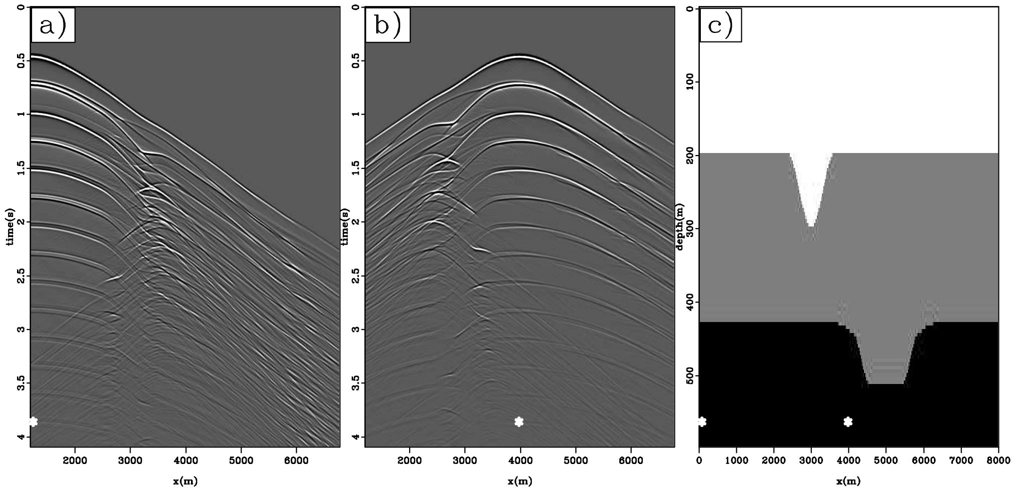


Figure 2.1: Representative transmission wavefields (a,b) from impulsive subsurface sources through a three layer velocity model with two synclines (c). All surface-related and internal multiples are included. Asterisks show subsurface source locations. TimeSynth-data [ER]

Figure 2.1a shows data from a source at the left edge of the model domain, whereas panel b is a transmission wavefield due to a source directly beneath the center of the model domain.

The asterisks at the bottom of the panels indicate the location of the subsurface sources. The first hyperbolas are direct arrivals and the rest are reflections due to the structure within the velocity model. The third event, with minimum travel time 1 s, is the reflection from the second layer in the model. Subsequent events are surface related and internal multiples. Panel c is the velocity model with which the transmission wavefields were modeled. For comparison, Figure 2.2 shows two modeled surface seismic shot gathers (where the "shot" axis is also defined across **x**) at locations in the model from which the forthcoming synthesized active 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.

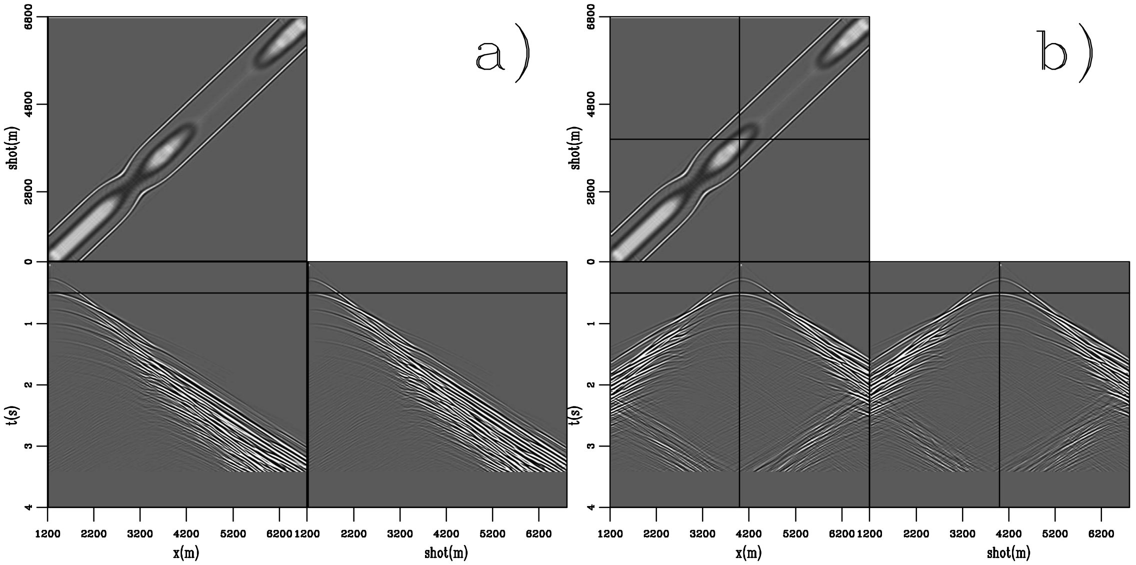


Figure 2.2: Two active reflection seismic shot gathers. These same shot locations, out of the total volume, are used in the forthcoming figures showing synthesized gathers. Events in the right panel dipping down toward the center of the frames are reflections from the edges of the model domain. TimeSynth-actshots [CR]

2.3 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 2.3 shows synthesized active shot gathers (at locations corresponding to previous figures) produced by summing the correlations from subsurface sources across the entire bottom of the model. The front face of panel a is not substantially different from the correlations of a single transmission wavefield in Figure 2.4. This gather, from the left edge of the domain, suffers from the limited distribution of subsurface sources used to synthesize active surface data. The quality of a gather is predicated on up-coming 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 this virtual 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 Figure 2.4a.

Figure 2.3b is a gather from the center of the model domain, which should be the most complete gather in the synthesized volume. Notice that on the right face (time-virtual shot axes), the events develop a slight concave-up moveout toward the edges of the domain rather than approach an asymptote. The inflection points of the events around virtual shot locations $x = 2700$ and 6700 m represent the limits of accurately synthesized shot gathers for this distribution of subsurface sources. 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.

Comparing Figure 2.2 with Figure 2.3 highlights several key differences between surface seismic data and the synthesized gathers. Obvious amplitude differences should be ignored,

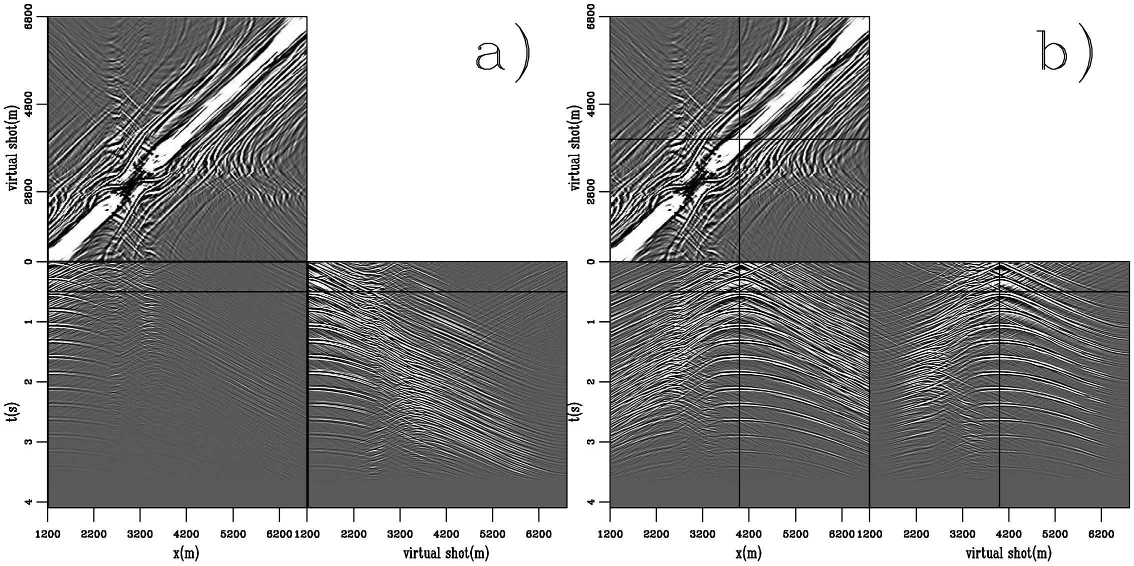


Figure 2.3: Two correlated shot gathers from the volume generated by summing correlations from 225 sources across the bottom of the model. The "virtual shot" axis is built by correlating every trace in the data wavefield with all of the other traces. TimeSynth-allshots [CR]

as this work is focused on recovering only 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 in Figure 2.3 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 2.3 and 2.2 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.

Figure 2.4 shows two slices through a 3D volume of synthesized gathers produced by correlating all traces in a single transmission wavefield (Figure 2.1a). The trace under the vertical black line in the correlated shot gathers is the zero-offset trace in the synthesized gather. It is also the location of the source function, or master trace, from the transmission wavefield for correlation to generate the synthesized gather. This volume is a partial processing of the source summation required to synthesize active data from transmission data. This experiment also represents the situation where a powerful or repeated source from a single subsurface location dominates the ambient noisefield.

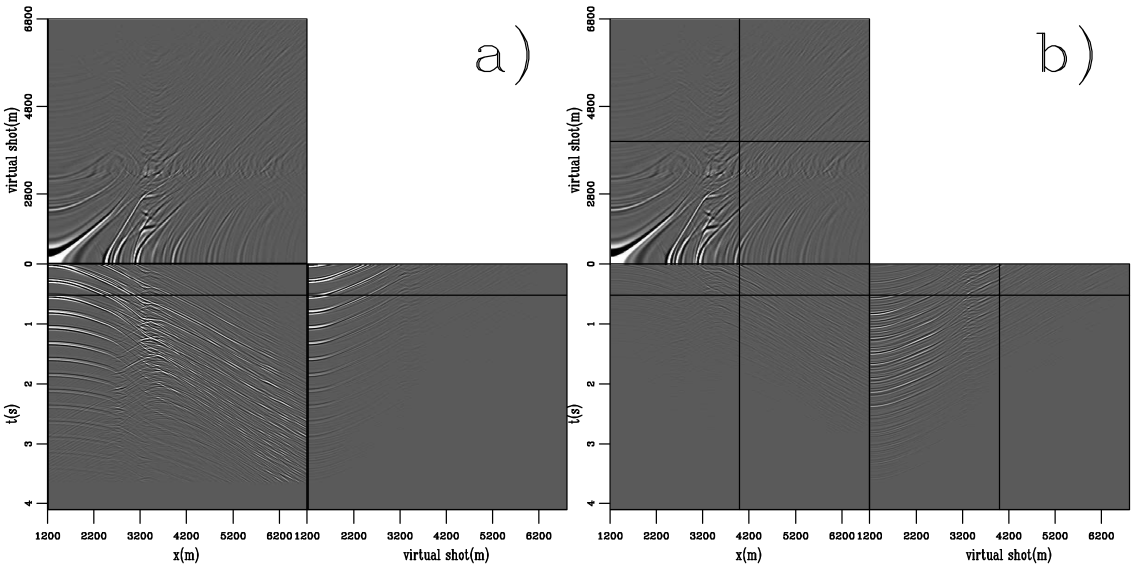


Figure 2.4: Two correlated shot gathers from the volume generated with a single source below the left edge of the model. TimeSynth-shot1200 [CR]

The time line at $t = 0.5$ s is at the zero-offset two-way traveltime to the second reflector. The front face of Figure 2.2a 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 2.1a. 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 of Figure 2.4) crosses the right edge of the gather at about $t = 2$ s. The same event in Figure 2.2a reaches the right edge of the gather at about $t = 3$ s. These observations highlight the inaccuracies incurred by not completing the integral in equation 1.1.

The first event in Figure 2.4a that intersects the x, t axes at (1200 m, 0 s), is the manifestation of the delta function, $\delta(\mathbf{x}, \mathbf{x}_s)$, in equation 1.1. It is the time-shifted direct arrival. 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 front face of Figure 2.4b 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. If only a small number of subsurface sources are captured in a passive seismic recording campaign, it is likely that such synthesized gathers would be discarded.

2.4 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 by processing 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. To model field data, T_f, composed of two single source transmission wavefields, a and b, a bulk time shift τ can introduce time delays between the activation of the sources. Therefore, data modeling multiple sources has the form

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

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 in equation 2.1 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 1.1. 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 ($x = 1200, 3500$ 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 2.3.

The differences in the time sections and time slices in Figures 2.5 and 2.6 show the inaccuracies associated with 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 (upper ovals) that are not present in Figure 2.6. A strong saddle shape within the central oval overlay mimics the energy in the center of the time slice. The lowest oval on the left edge of Figure 2.5b highlights strong half-hyperbolas that are not present in Figure 2.6b. Correlating the sum of the transmission wavefields causes all of the events due to each source to be present in the result instead of inappropriate events canceling due to destructive interference. Both shot positions in Figure 2.5 look very similar, and the noise is symmetric across the zero-offset diagonal in the time-slices. Note the similarities in Figure 2.5 and the correlations of a single subsurface source, Figure 2.4.

As more sources are processed together, such as a storm of small subsurface sources

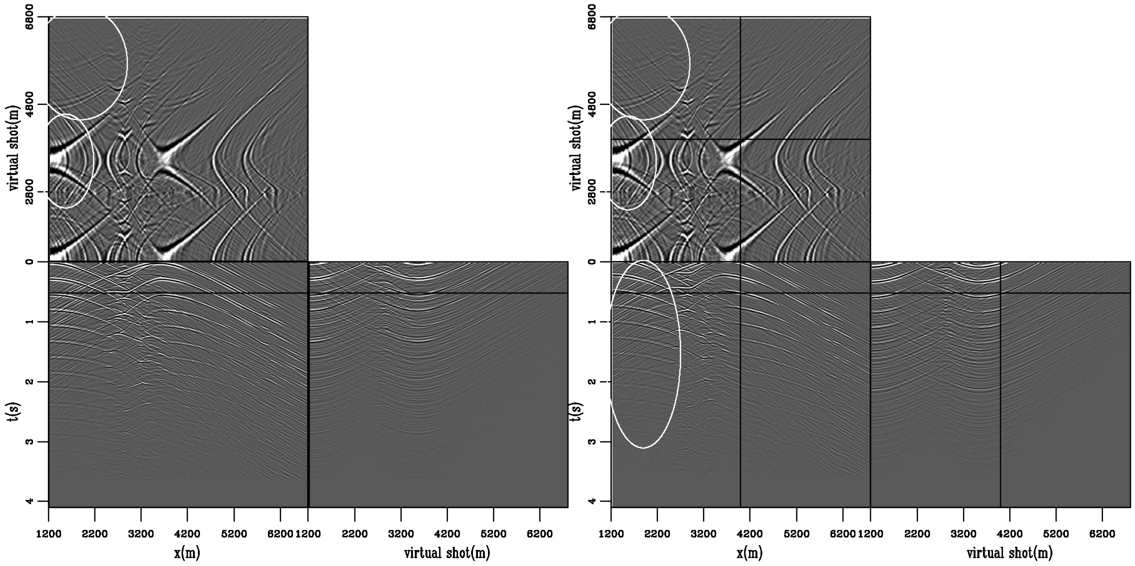


Figure 2.5: Synthesized shot gathers produced by correlating the sum of transmission wavefields due to sources beneath the left edge and middle of the model. Compare artifacts within the ovals to the more correctly processed result in Figure 2.6 TimeSynth-vshot1200.3500 [CR]

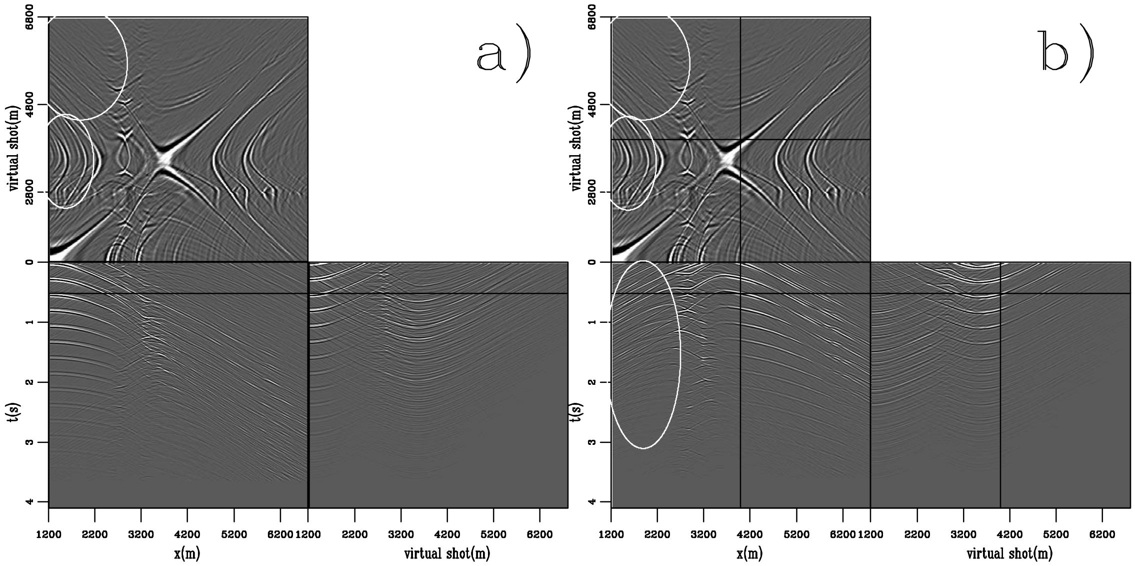


Figure 2.6: Synthesized shot gathers produced by summing the correlations from sources beneath the left edge and middle of the model. Ovals indicate locations of artifacts in the results shown in Figure 2.5. TimeSynth-vshot1200csum3500 [CR]

during natural or induced fracturing, the introduction of cross talk noise can overwhelm the correlations. Figure 2.7a 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 coherent noise that the gathers are uninterpretable.

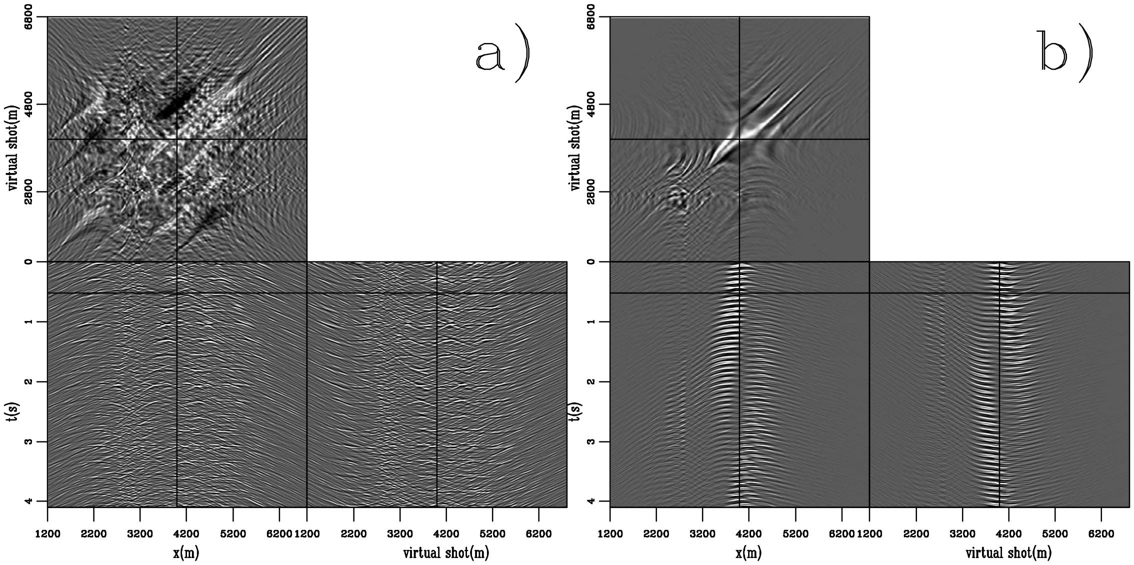


Figure 2.7: Left: Synthesized shot gathers with 25 subsurface sources spaced evenly across the bottom of the model. Right: Synthesized shot gathers with 30 subsurface sources clustered under the center of the model. TimeSynth-manysrc [CR]

Figure 2.7b was produced with 30 shots every 25 m clustered around the center of the model. Clustered subsurface sources may be a common situation experienced with field data when energy from a localized area dominates the energy budget of the ambient noisefield. 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.

2.5 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 transmission wavefields are perfectly correlatable. 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, I convolve the individual transmission wavefields with a random number sequence. 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 associated with correlating their sum 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 reflection is repeated $n_s!/(n_s-2)!$ times. The swarm of events in Figure 2.7 is caused by the repeated inclusion of every event due to each source in the result. The ratio of desirable zero phase terms, I_e^2, 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. 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, especially 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 figure is directly comparable to Figure 2.5 to show the influence of the randomized source functions on the resulting synthesized gathers. The ovals drawn on the figure are the same as those on the comparison figure. The noise identified within the ovals in Figure 2.5 is mostly absent. The half hyperbolas identified by the lowest oval in Figure 2.5 has now moved later in time as indicated by the arrow. These are the time shifted events from the source at the left edge of the model domain seen in panel a. 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. The cross talk introduced by summing the transmission wavefields is at comparatively late time. This could be windowed away since the zero offset two-way traveltime to the deepest reflector in this model is only 0.5 s. Without knowing that two sources were active during the recording interval, these gathers would be incorrectly interpreted.

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 cross talk below 2 s on panel b is not present. The random source wavelet introduces an extended wavelet coda to each event that leads to the events overlapping and an overall ringing character that diminishes the quality of the gathers. Random 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 introduced a ringing coda that can look like events. If the source 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

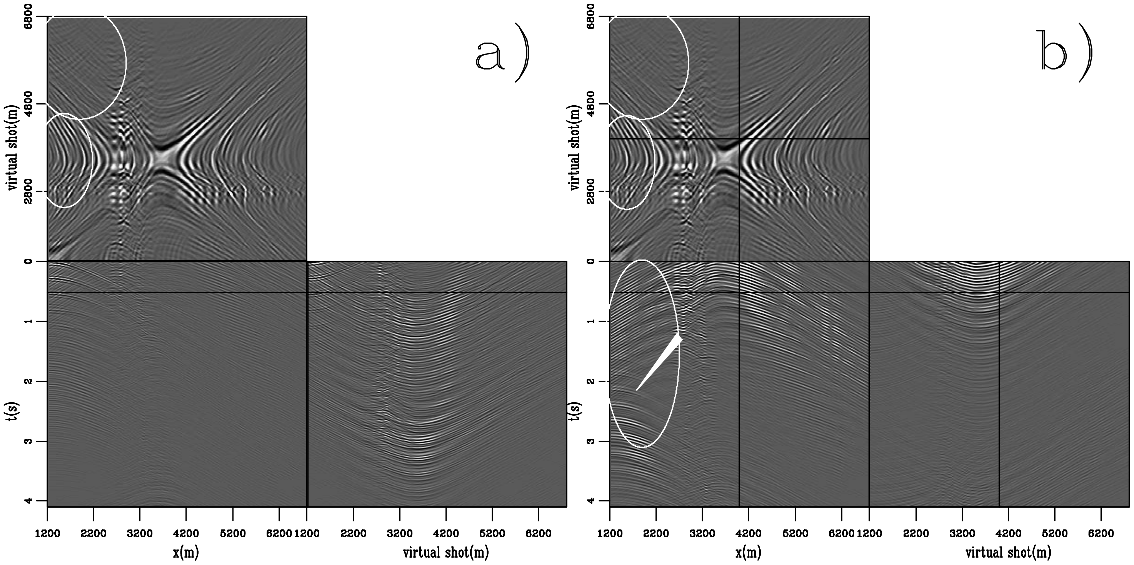


Figure 2.8: Synthesized shot gathers produced by correlating the sum of transmission wavefields due to randomized sources beneath the left edge and middle of the model. TimeSynth-vshot1200.3500.rand [CR]

to 1 s in duration does not allow for much cancellation of the correlation side-lobes.

Figure 2.10a 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 worse than that produced without the randomized wavelet, Figure 2.7a. The synthesized gathers including randomized wavelets contain more events and are not interpretable.

Figure 2.10b 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, Figure 2.7b. However, these events are not subsurface reflections. They are cross talk artifacts that completely mask the subsurface information.

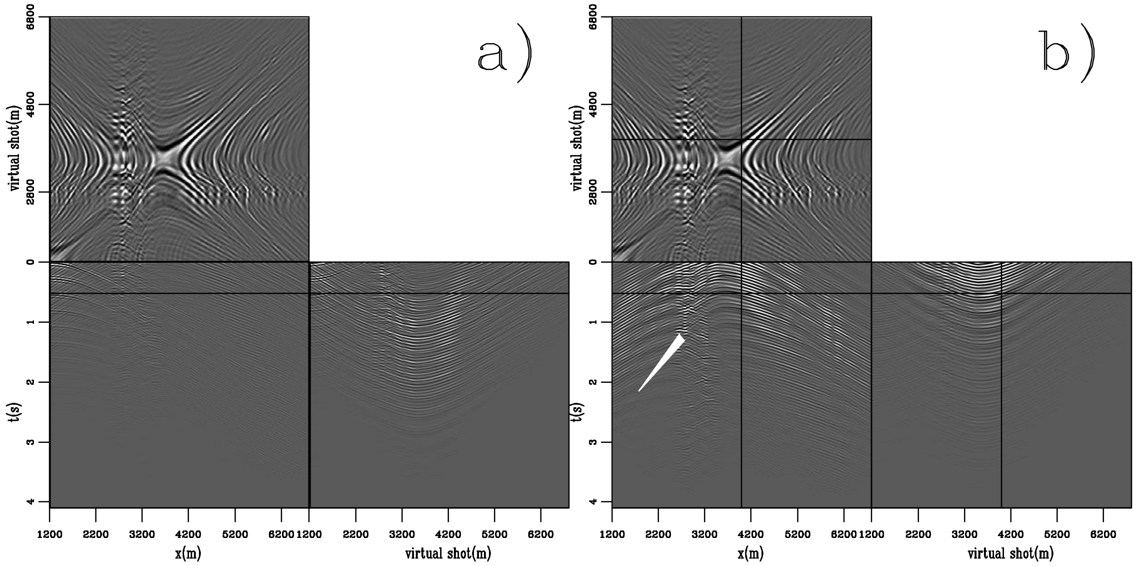


Figure 2.9: Synthesized shot gathers produced by summing the correlations from randomized sources beneath the left edge and middle of the model. TimeSynth-vshot1200csum3500rand [CR]

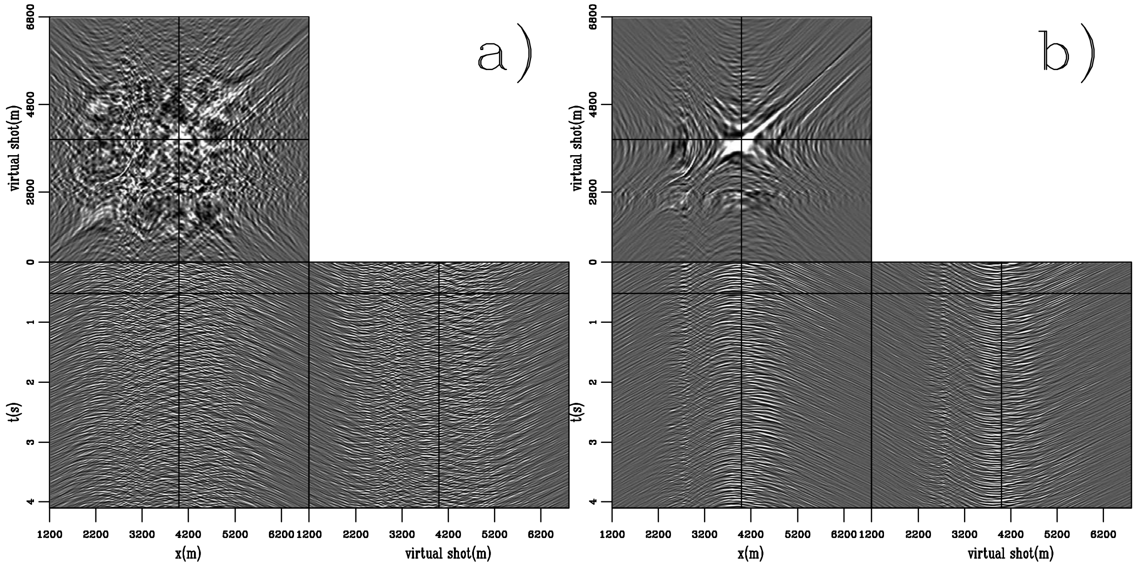


Figure 2.10: Left: Synthesized shot gathers with 25 randomized subsurface sources spaced evenly across the bottom of the model. Right: Synthesized shot gathers with 30 subsurface sources clustered under the center of the model. TimeSynth-manysrcrand [CR]

2.6 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. 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. Upon comparing such correlated gathers to the active gathers they are meant to synthesize, it is likely that these results would be discarded when produced with field data. However, in Chapter 5 it is shown that despite the poor quality, these gathers contribute positively to the image produced by migrating all possible synthesized shot gathers. The subsurface information in these synthesized gathers is due to the high-order multiples included in these models. If high-order multiples are not recorded in field data, these gathers would not contain useful information.

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 clustered or well distributed subsurface sources correlated simultaneously, the resulting volumes 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 provide very substantial damping of the cross talk. 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 from Valhall

3.1 Summary

BP made available passive recordings from an array of 2300 four component receiver stations at the Valhall reservoir in the Norwegian North Sea from February 2004 and January 2005. I used the hydrophone records for this analysis. The bulk of the raw records do not contain obvious 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 the 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. Dominating the gathers 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.

3.2 Introduction to the Valhall array

BP, with partners Hess, Shell, and Total, provided approximately 43 hours of continuous passive seismic records from the permanent sensor installation above the Valhall reservoir in the Norwegian North Sea. Figure 6.1 is a map provided by the Norwegian Oil Directorate which shows the production infrastructure between the coasts of Britain and Norway with the many developments in the region. It is 630 km from the Southern tip of the Norway shown and the city of Teeside in Britain. The Valhall development is roughly in the center of the line connecting those two geographic points.

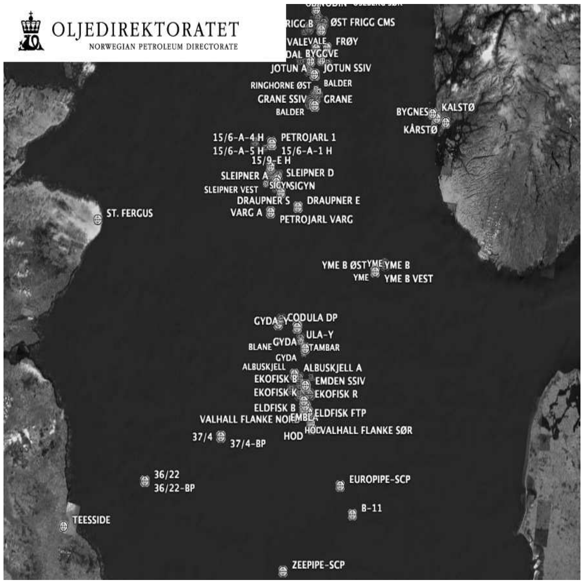


Figure 3.1: British, Norwegian, and Danish coasts surrounding the North Sea oil development area. The Valhall reservoir is in the South-central area of the map and only 20 km from U.K.-controlled waters to the West, and Danish waters to the South. TimeShot-areamap [NR]

There are five surface structures (roughly over the center of the reservoir), and two wellhead 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 to facilitate time-lapse imaging of the subsurface for production monitoring. The $\sim 9,200$ sensors are permanently installed 1 m below the sea floor, in roughly 70 m of water. The array forms a roughly 4x11 km rectangle. Nominal inline spacing of receiver stations is 50 m, while the crossline separation is 300 m.

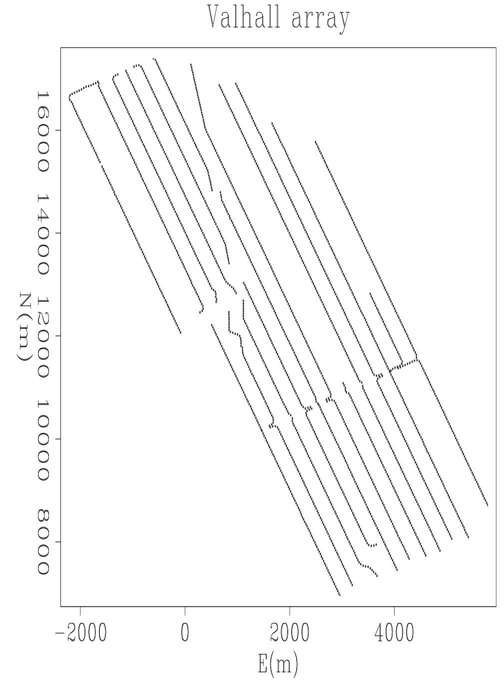


Figure 3.2: Valhall permanent subsea geophone array. Each station measures 4 components (4C) of the seismic wavefield: Pressure with a piezoelectric sensor, and three orthogonal velocities of ground displacement with coil-spring geophones. TimeShot-actg [NR]

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 the instruments are telemetered. Also noticeable is the deviation of the ends of cables on the North side of lines 5-7 (E=1000 m) 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 electric power with gas turbines, flow through the 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 U.K. 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 withdrawal during production. In its first 15 years of production (to 1997), the sea-floor subsided 3.5 m at the crest of the reservoir structure (Gebara et al., 2000).

---

[1] www.npd.no/engelsk/

3.3 Passive seismic recording

The passive recordings donated by BP are from 29 hours February 15, 2004 and 14 hours January 19, 2005. Of the former, some 7 hours was lost due to a failed disk-drive. The most likely teleseismic arrival recordable at the array, as determined by query of the National Earthquake Information Center event catalog[2], arrived approximately 5 hours after the last records available. Of the 2300 4C stations, this work uses the hydrophone measurement from each location. The data are sampled at 0.004 s, and written as contiguous files between 12-20 s long. Figure 3.3 shows 12 s of data from a single receiver. The ten traces are contiguous 1.2 s sections from the total 12 s record. Therefore, both the vertical and horizontal axes are time in seconds. Figure 3.4 shows an example of the character of most of the data. The traces are balanced against each other, but are otherwise raw recordings. Scanning through much of the raw data, I was not able to identify clear events. Most of the raw data look like incoherent noise, but I found several interesting features that will be explored below.

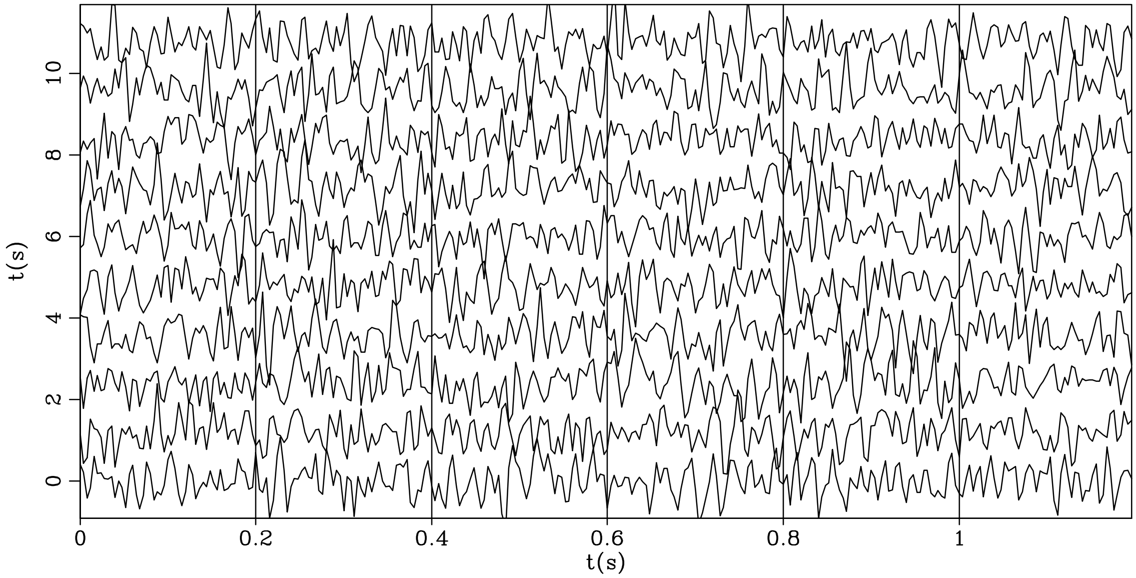


Figure 3.3: Ten contiguous 1.2 s sections of data from a single raw record from the array. TimeShot-wiggle [ER]

---

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

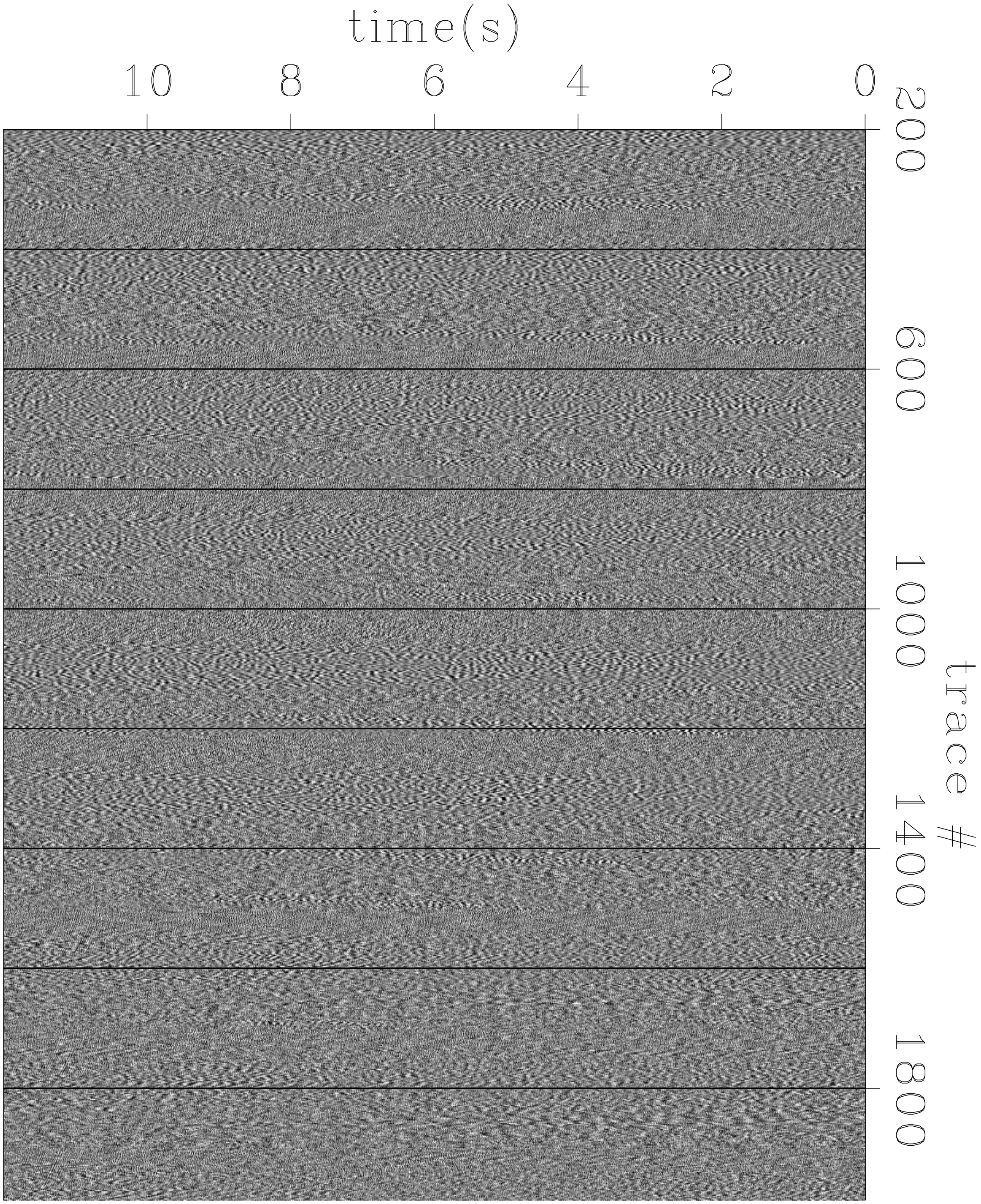


Figure 3.4: Twelve seconds of representative passive hydrophone data from February 16, 2004. Trace number increases from the top left to bottom right corners of the array shown in Figure 3.2. TimeShot-raw1 [ER]

Figure 3.5 shows a conspicuous noise-train in a section of raw data. The figure shows all of the traces within the array, progressing (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 3.6 shows a 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 half. The minimum traveltimes of the hyperbolas in the data decrease to the West 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 3.6 show the kinematics of a a co-planar direct arrival from the large 'O' in Figure 3.7 (explained below) 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 3.7 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. A line connecting the X's has no inline offset variation and includes the area around $y = 6250$ m, which is the location of the platforms (presumably near the

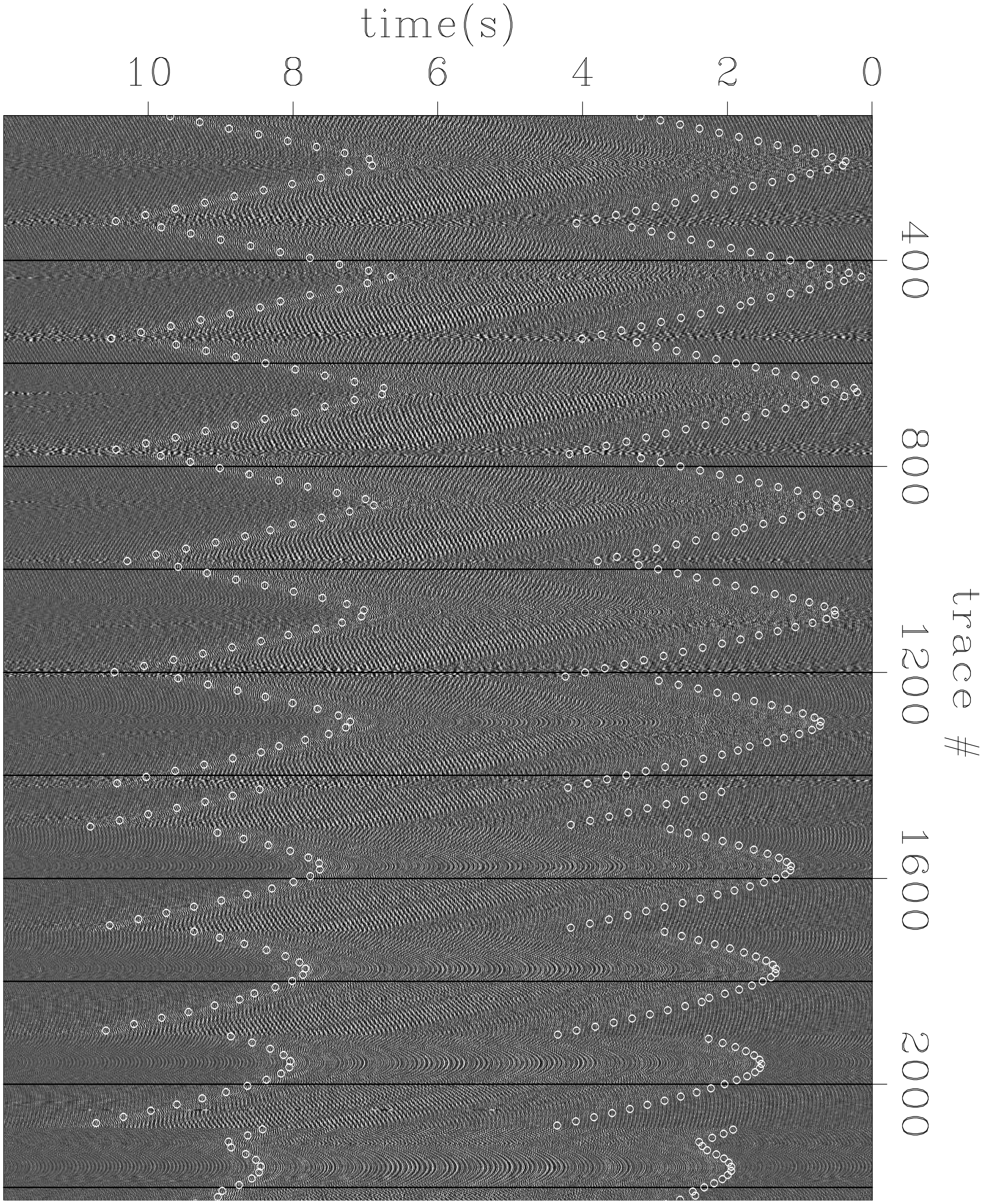


Figure 3.5: Raw data with a powerful noise train envelope. Ten lines of the array are apparent as the hyperbolas are repeated and shifted over crossline offset. Circles are co-planar direct arrivals from the center of the array traveling at 1450 m/s. TimeShot-raw2 [ER]

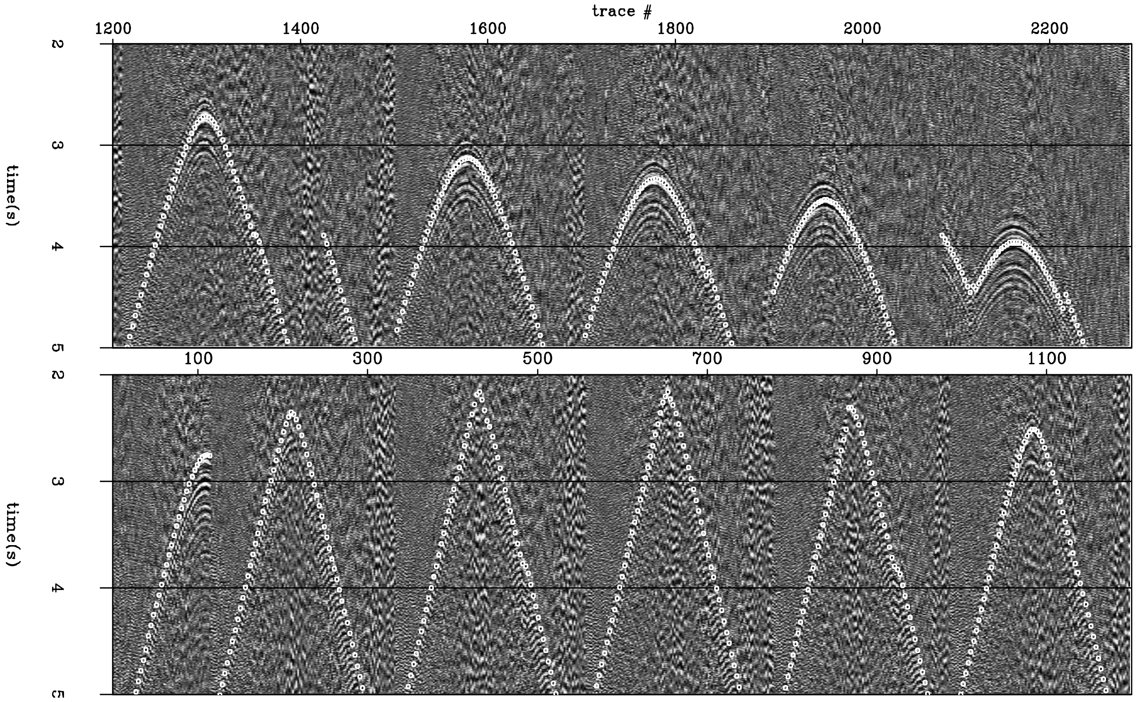


Figure 3.6: Raw data with a clear hyperbolic event annotated with the time calculated for a coplanar direct arrival from the array center traveling at 1450 m/s are. TimeShot-raw3ev [ER]

large ’O’ symbol). This figure provided the inline coordinate for the overlaid event plotted in Figures 3.5 & 3.6. The crossline coordinate was deduced by matching the picks to the data windows. The four cables with circled receiver stations indicate a subset of the array used for later figures.

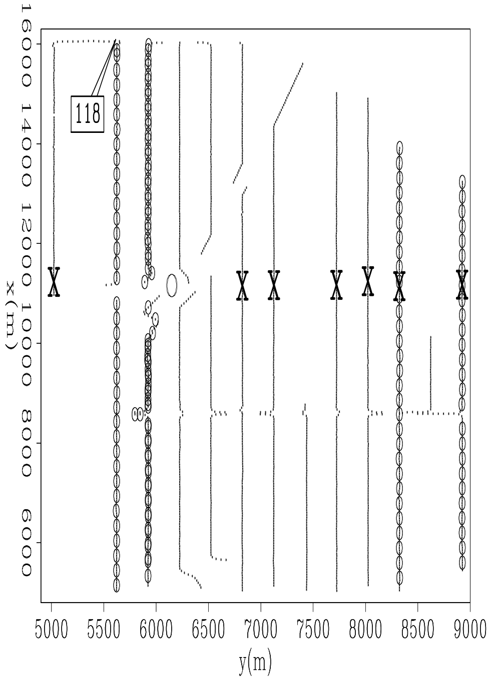


Figure 3.7: Receiver array in field coordinates. Marked with ’X’ are the locations of traces with minimum traveltimes picked from the hyperbolas in Figure 3.6. Surface facilities are roughly at the large *O*. Trace 118 is the virtual source location for the synthesized gathers in Figures 3.8-3.16, 3.18 and the origin in the black line in Figure 3.20. Cables marked with circles indicate receiver locations used in Figures 3.12-3.16,3.18. TimeShot-minloc [ER]

3.4 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 +36 hours of passive data is impractical (because of the large volume and the lack of identifiable events), it is impossible to honor equation 1.1 precisely by correlating traces within time windows when only a single event is active. 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 s records within these sections were correlated and the results stacked over the duration of the (hours long) section of the total 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 index of sequential 12 second windows of data.

Figure 3.8 shows correlations from one of the 9 hour sections of data. A trace from the NW corner (receiver 118 in Figure 3.7) 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 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 3.9 plots the power spectrum of the sum of 100 traces from the gather in Figure 3.8. The power spectrum of the gathers produced from all four subsections of the raw data were very similar. Versions of the gathers (produced with all subsections of the raw data) bandpassed around and between the peaks in the spectrum (Figure 3.9) showed no 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.

Figure 3.10 shows the same gather as in Figure 3.8 using the spectral whitening of equation 3.2. A very clear event is now apparent. Similar features can be seen upon close

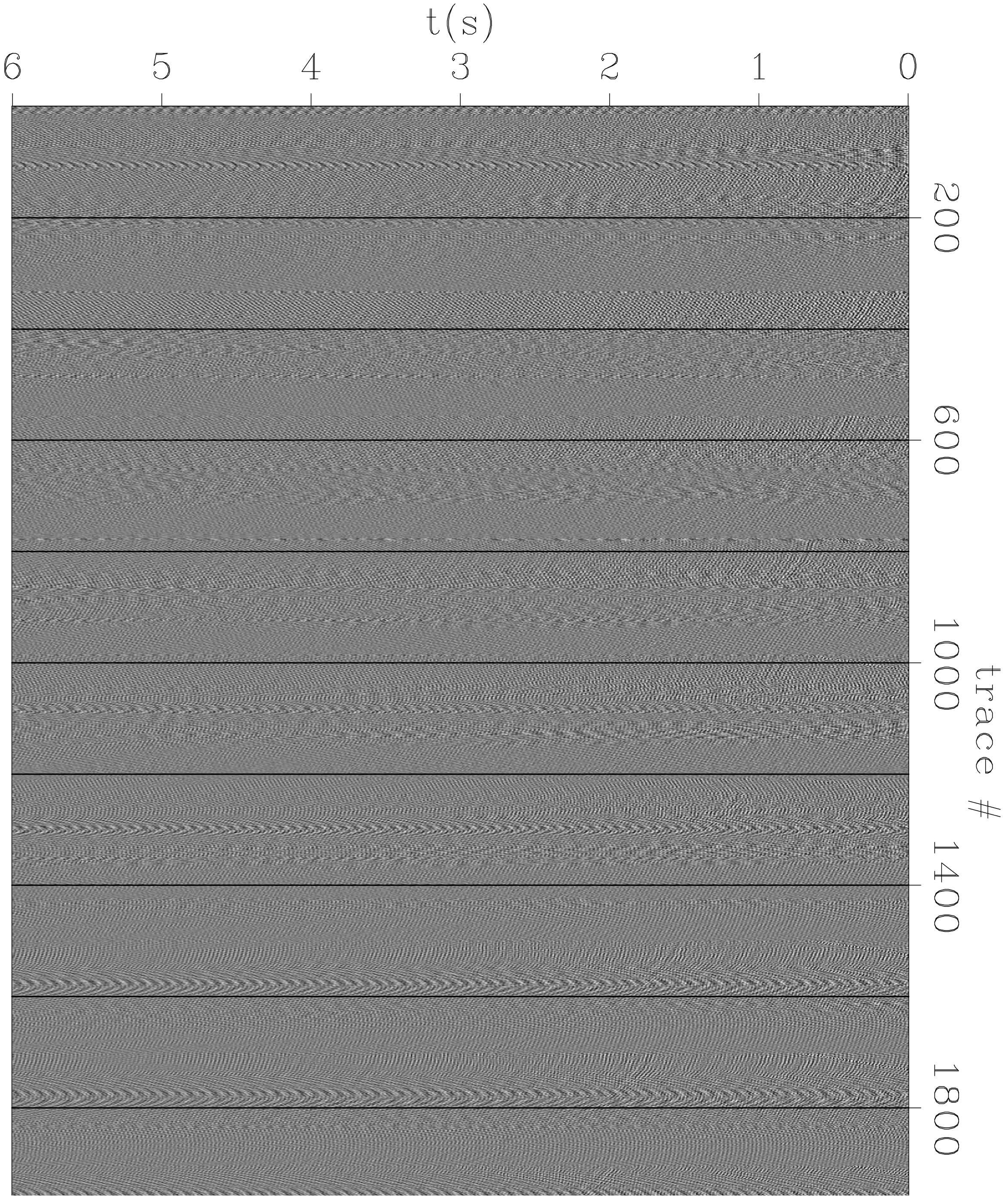


Figure 3.8: Correlated shot-gather with master trace from the left side of the panel. 0-60 Hz energy was used for the correlation performed in the frequency domain. TimeShot-cor [CR]

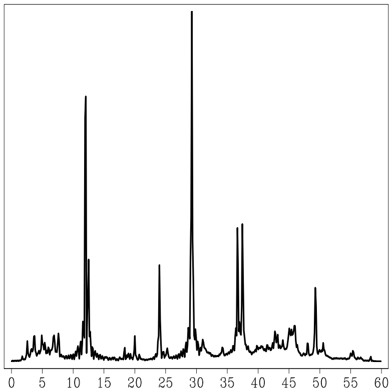


Figure 3.9: Power spectrum of the gathers in Figure 3.8. TimeShot-spec [CR]

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. The deconvolved gathers, produced with all subsections of the raw data, were bandpassed across the intervals between and surrounding the energy maximums in power spectrum, Figure 3.9, 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. Above 42 Hz, the signal begins to diminish markedly, and no coherent energy is present at all above 52 Hz.

Events revealed by deconvolution processing, equation 3.2, in Figure 3.10 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. Therefore, 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 curvi-linear in the inline direction, but very close to linear in the crossline direction (white line over data in Figure 3.10). 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 3.7 that crossline offset there is twice the normal spacing. These observations indicate that the array is close to the asymptote of the hyperbolic event in the crossline direction, and nearer the hyperbola top in the inline direction. Therefore, the

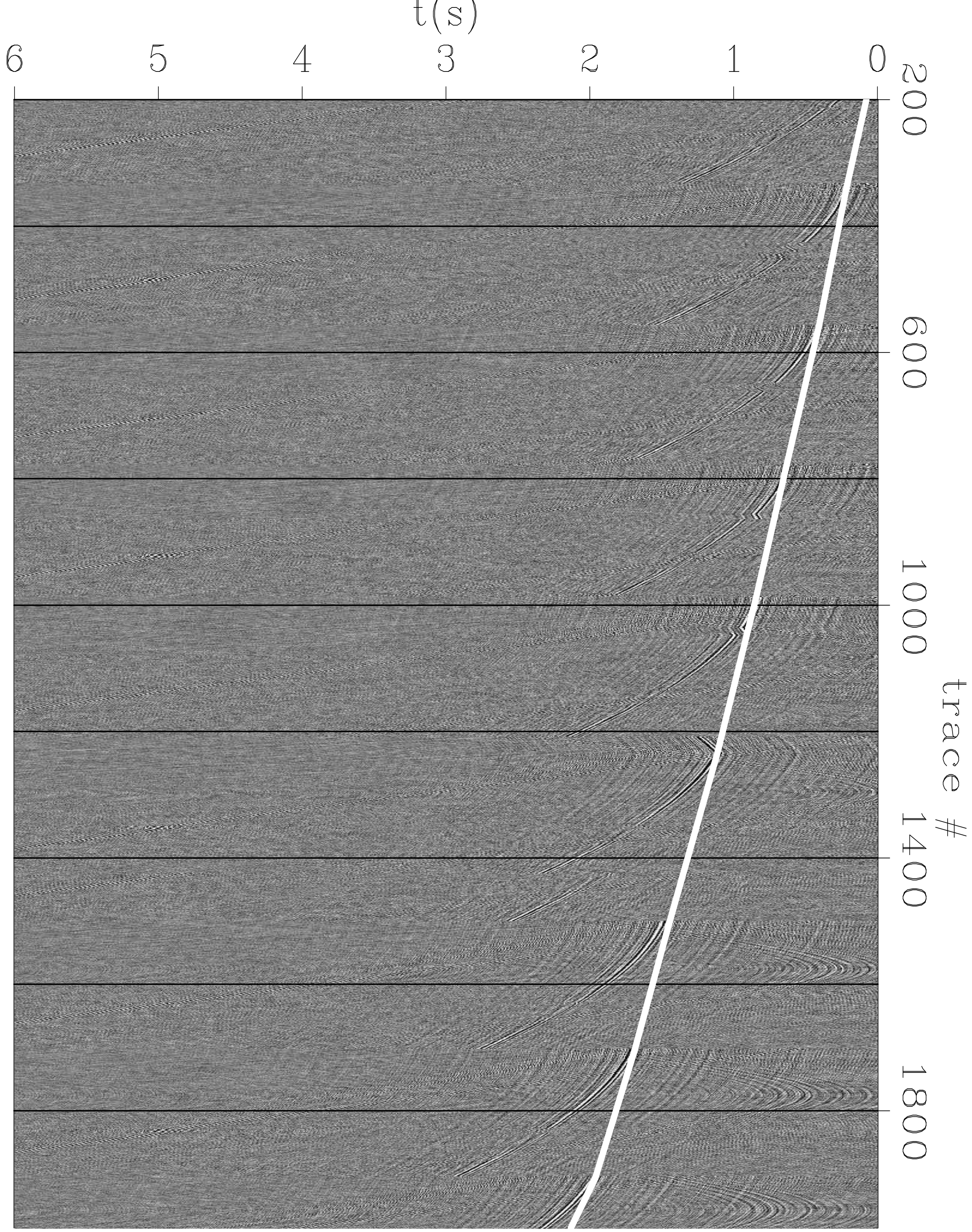


Figure 3.10: Trace-by-trace spectral whitening applied to shot-gather in Figure 3.8. Line connecting minimum traveltimes almost perfectly straight. TimeShot-dec [CR]

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 array toward the British coast. The fact that the event does not contain the center of the hyperbola, proves that processing by correlation to produce conventional shot-gathers has failed. Specifically the requirement to sum the correlations from many shots in equation 1.1 has not been honored.

Figure 3.11 shows two constant crossline offset sections extracted from a volume of deconvolved synthesized shot gathers. The crossline offset is 2400 m between the two OBC lines with bold receiver locations in panel a. Panel b was produced by correlating traces from West to East (bottom to top of panel a), while panel c shows reciprocal correlations. Active seismic data are symmetric for these two sections. Therefore, events in the synthesized sections must be present in both panels to be geologic reflections. Alternatively, panels b and c are mathematically equivalent to the causal and acausal parts of the correlations performed in a single direction. The causal and acausal parts of the correlations for a successfully synthesized volume of shot gathers are also symmetric. The lack of identifiable events in panel b, and the concomitant lack of symmetry between the panels, indicates that the raw data processed by correlation do not fulfill the source distribution requirements of the theory of interferometry, equation 1.1. The strong event below 1 s in panel b is event discussed in Figure 3.10 on which the bold white line is overlain. These sections, and similar ones exploring a full range of crossline offset, do not reveal geologically feasible events.

If only a single impulsive subsurface source is captured by a passively recording 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 3.12 all show four lines (receiver locations marked with circles in Figure 3.7) from the same virtual shot location as Figure 3.10. The panels were produced with data from the four time intervals into which the raw data volume was divided. Panel c is a subset of the gather shown in Figure 3.10. 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.

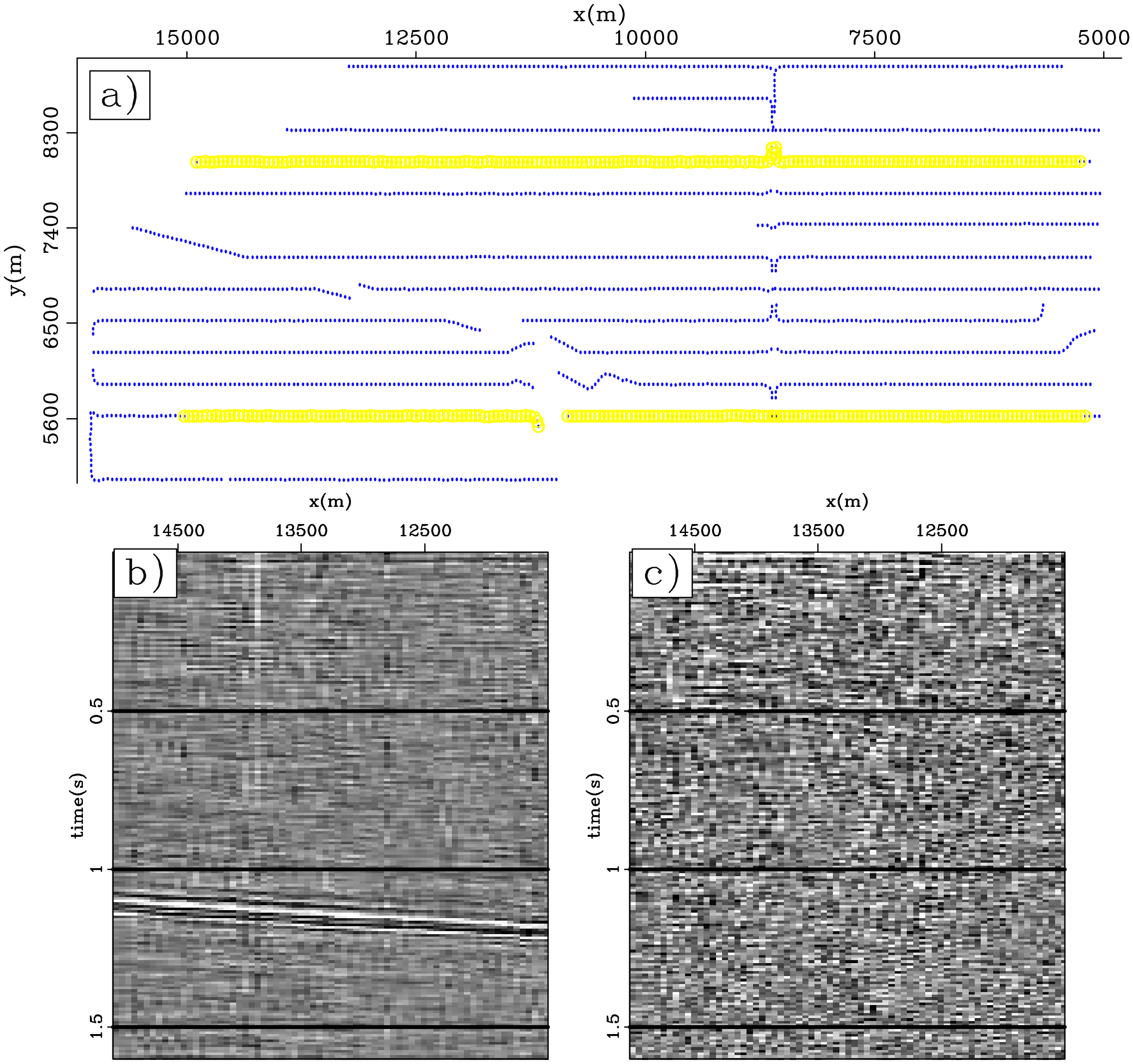


Figure 3.11: Reciprocal constant offset sections produced by correlating traces in the two bold OBC lines in panel a. TimeShot-conoff [NR]

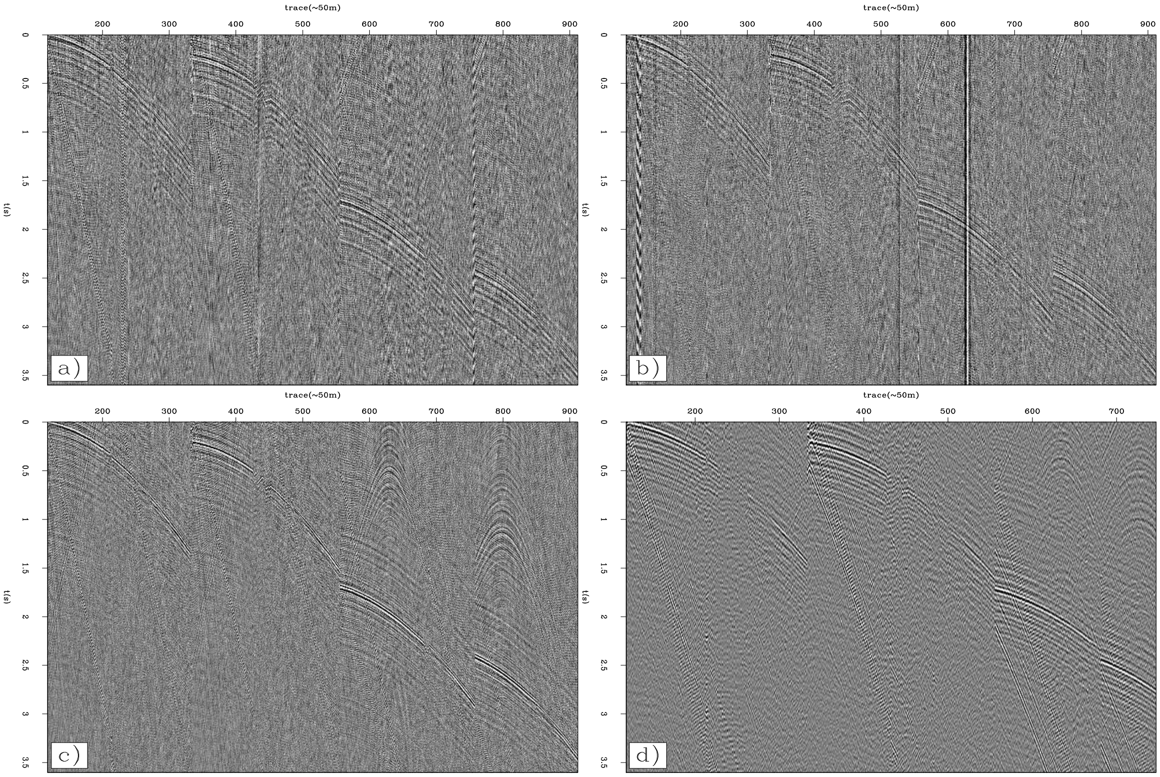


Figure 3.12: Ubiquitous strong event in deconvolved shot-gathers from data recorded 9 hours (b), 25 hours (c), and 11 months (d) after that used to produce the panel a. The third panel is a subset of the gather shown in Figure 3.10. TimeShot-event [CR]

Figure 3.13 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 118). 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. Panel a shows the first three seconds of the causal correlation lags using the data shown in Figure 3.5, which showed the ringing noise-train. No strong events are present, but there are some faint ringing hyperbolas described by water velocity.

Correlation, to first order, subtracts the initial time to the direct arrival in the master trace from all the other traces in the gather. The beginning of the noise-train in Figure 3.5 on the first trace (number 118), 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 after correlation. Figure 3.13b contains the acausal lags associated with the correlations presented in panel a. The water velocity event overlain on the traces has 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 overlays rings 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 side lobes. The side lobes continue past $t = 0$ and on to panel a.

The bottom row of Figure 3.13 are the deconvolved correlated gathers using trace 118 as the master trace, and the raw data from Figures 3.4 & Figure 3.6 respectively. In each case only 12 s of data was correlated. Faint hyperbolas are discernible in the early time of the right side of panel d. The faint hyperbolas have the same kinematic description as the water velocity events in panel b and Figure 3.6. The most obvious energy however is in the events similar to the strong arrivals in Figure 3.10 and Figure 3.12.

Correlation masks some of the characteristics of the source energy. The source could be impulsive or a long sweep of 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

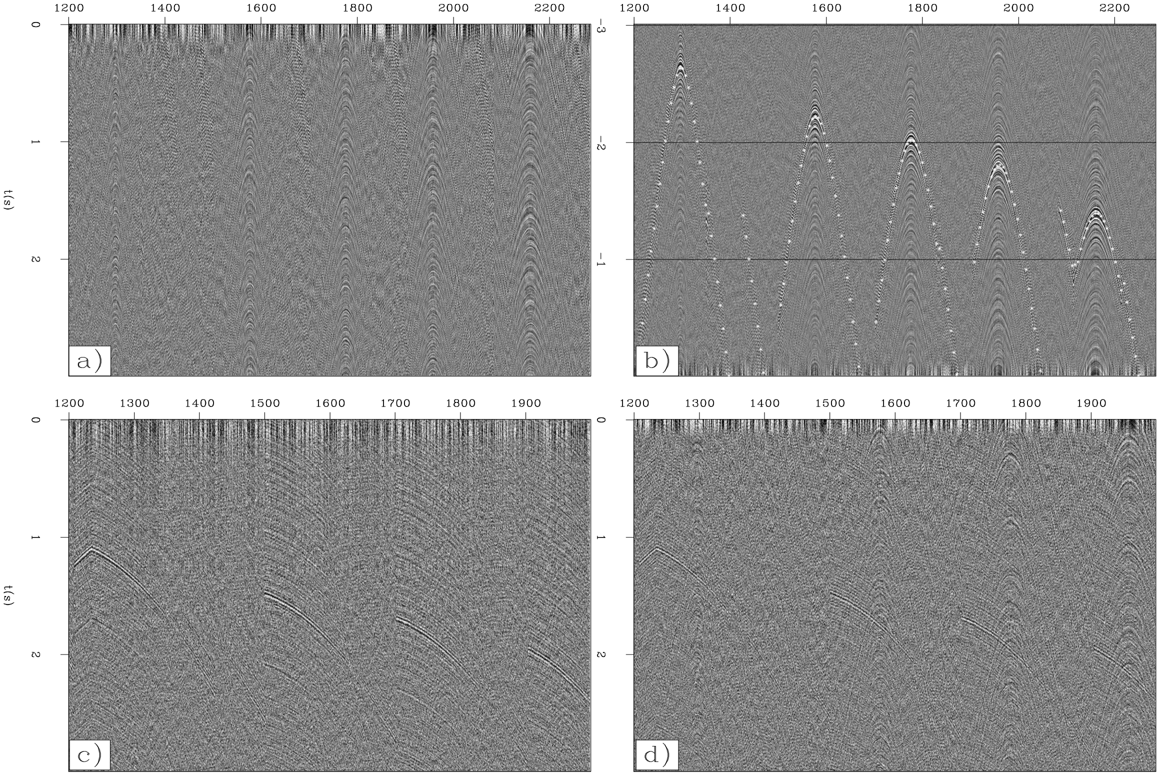


Figure 3.13: Correlated gathers from the Eastern side of the array. Top row: Early causal (a) and acausal lags (b) of a correlated gather using the data in Figure 3.5. Overlays on panel b have the same velocity (1450 m/s) and origin as those delineating the top and bottom boundaries of the noise train in the raw data. Bottom row: Early causal lags of correlated gathers using 12 s of raw data. Panel c used data in Figure 3.4 and panel d used data in Figure 3.6. TimeShot-evtest [CR]

Figure 3.12 (spanning 1 year), I favor the latter explanation. I believe the energy is localized in space, and 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. The extended noise-train in the data shown in Figure 3.5 used to synthesize the gathers shown in Figure 3.13a,b indicates that over short time intervals, local cultural noise can overpower the temporally stable event revealed in Figures 3.12a-d. However, the local production noise generated only water velocity in the correlated gathers. Therefore, energy from the production facility did not provide subsurface information.

Figure 3.14 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 Frequency domain wrap-around of steep linear events characterized by water velocity slope. Two energy trains are observable on either side of zero on the wavenumber axis that could be subsurface signal. Panel b has been muted to preserve the energy near zero wavenumber and discard the strong planar water velocity energy. This Fourier domain mute was applied to 100 minutes of raw passive data before correlation to investigate the steeply dipping (fast) energy near the center of the panels.

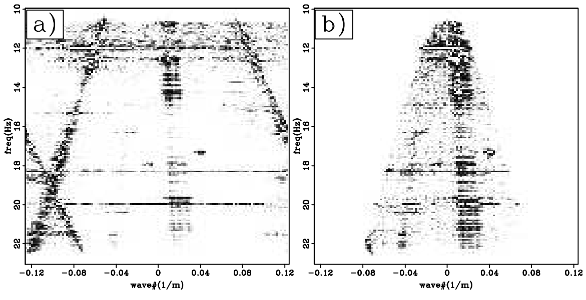


Figure 3.14: Frequency-wavenumber power spectrum of on OBC cable calculated from raw passive data. Panel b shows mute applied to focus on energy near zero wavenumber. TimeShot-fkpow [ER]

Figure 3.15 shows a shot gather produced from the data preprocessed with the Fourier domain mute in Figure 3.14b. A single strong event is apparent. It is aliased and has much lower frequency content that the event in Figure 3.10. 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.

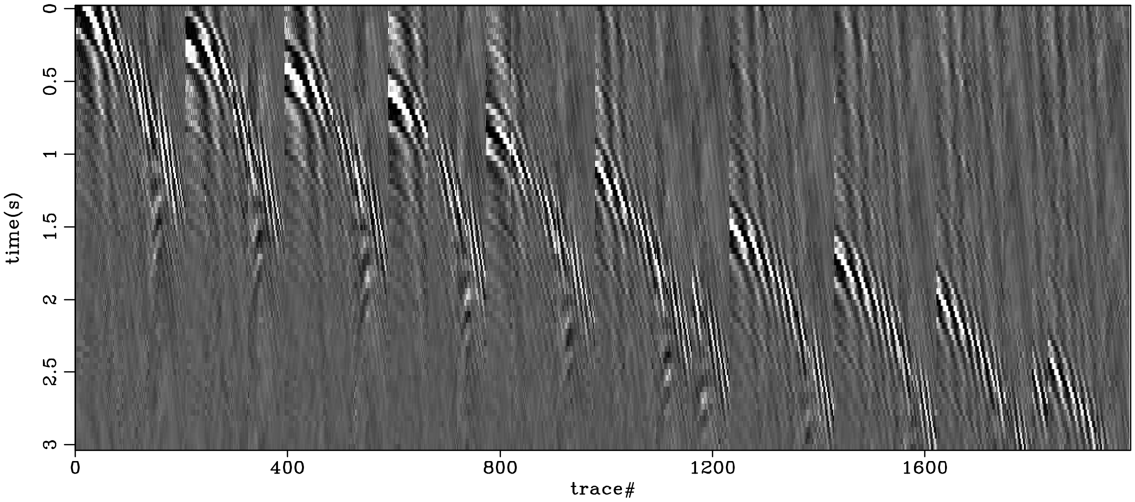


Figure 3.15: Correlated shot gather synthesized from passive data after preprocessing with the Fourier domain mute shown in Figure 3.14. TimeShot-fkgath [ER]

3.5 Source identification

Figure 3.16 shows the same correlated gathers as Figure 3.12. 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

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

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 data with a single master trace subtracts the time from the source to the master trace from

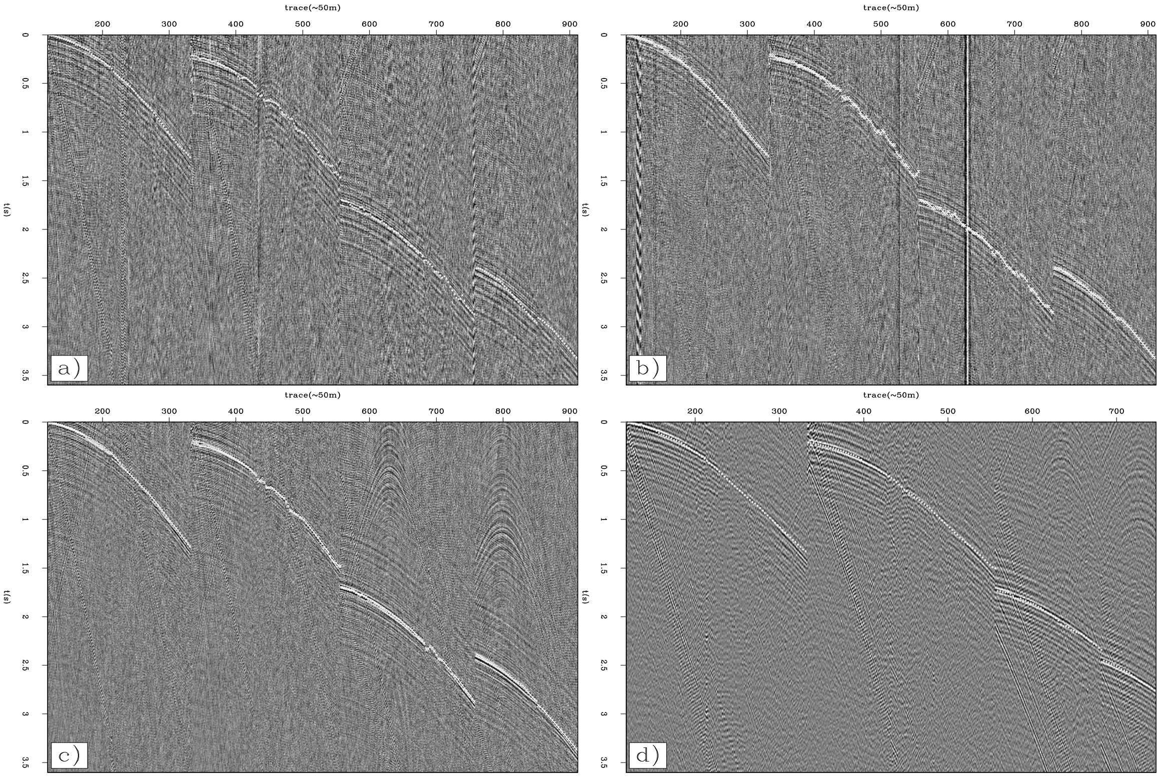


Figure 3.16: Auto-picked time values along event in correlated gathers. Picks are used as data to invert for the source location. TimeShot-picks [CR]

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 cause 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 3.17 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 seed numbers to begin inversions with the synthesized shot gathers from each of the four subdivisions of the passive data. Figure 3.17 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 inversion. The total convergence of the inversion as a function of generation number is displayed in panel a. The convergence of the model parameters velocity, distance East, and distance North (respectively) to specific values is shown in panels b-d. 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 reflector: depth, dip, and azimuth. These parameters were completely within the null space of the inversion.

Figure 3.18 shows the forward modeled time picks from the inversions over a synthesized shot gather. 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 3.16. The difference in quality between the picks that went into the inversion and the

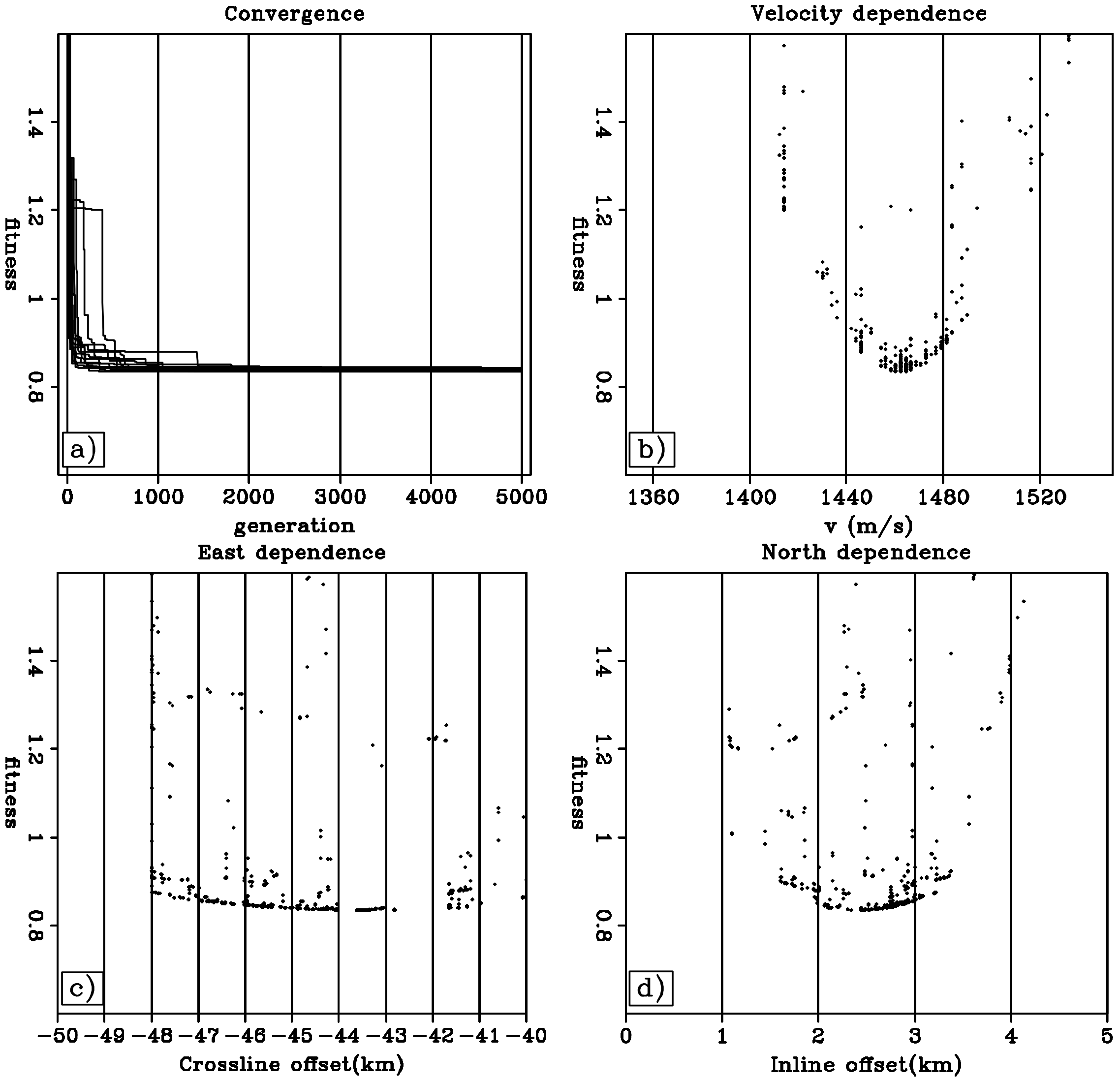


Figure 3.17: Residual energy of the inversion to locate the source of the energy causing the event in Figure 3.16. As the inversion iterates, overall residual energy decreases, and the standard deviation of the model parameters diminishes for well constrained members. TimeShot-inv [CR]

result accounts for the residual energy in Figure 3.17a. The inversion was not able to match the bad data (imprecise picks) because it was required to honor equation 3.3. Therefore, the result is not able to converge to zero.

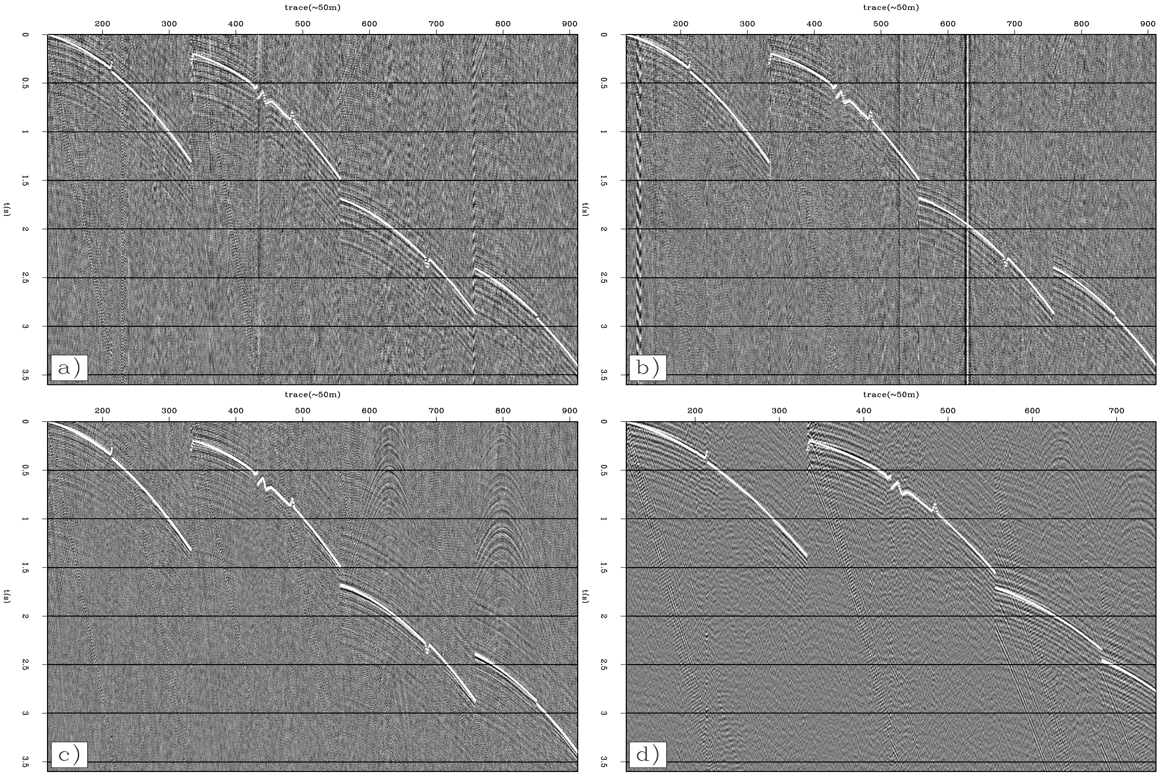


Figure 3.18: The forward modeled time picks from 15 inversions are plotted over each panel in Figure 3.12. The accuracy of the times is very faithful to the irregularities in the array layout in contrast to the auto-picked data input to the inversion, Figure 3.16. TimeShot-mods [CR]

I selected values for map coordinates and velocity at the location of the minimum from all 15 inversions using graphs such as those shown in Figure 3.17b-d. 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 3.7. 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. From these results, I interpret the energy as arriving in a direct path through the sea to the array. The data do not contain sufficient information to distinguish between sea floor and sea surface source locations.

I used the mathematic description of the direct arrival from the distant source to design a stacking operator. Summing along the trajectory of the arrival at all times in the raw data returns a measure of the energy in the data associated with this event. I also summed the energy in the raw data volumes associated with zero ray parameter. The horizontal plane wave sum was consistently at least two orders of magnitude less than the value returned by the stacking operator designed with equation 3.3 and the parameters calculated by inversion. Similar partitions of energy were observed in the summations using the synthesized gathers instead of raw data.

Using the velocity and location values calculated by the inversion, I forward modeled time picks at every receiver station for use as data in an exhaustive search over location variables. Figure 3.19 shows the results of the search using the four time sections of passive recordings. The gray scale shows the inverse of the total data misfit normalized by the maximum value. The same line is plotted on the panels for reference (8^o from the horizontal axis). All data volumes resolve the source location at similar locations with 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 3.20 is a detailed map of the Norwegian and British oil infrastructure around

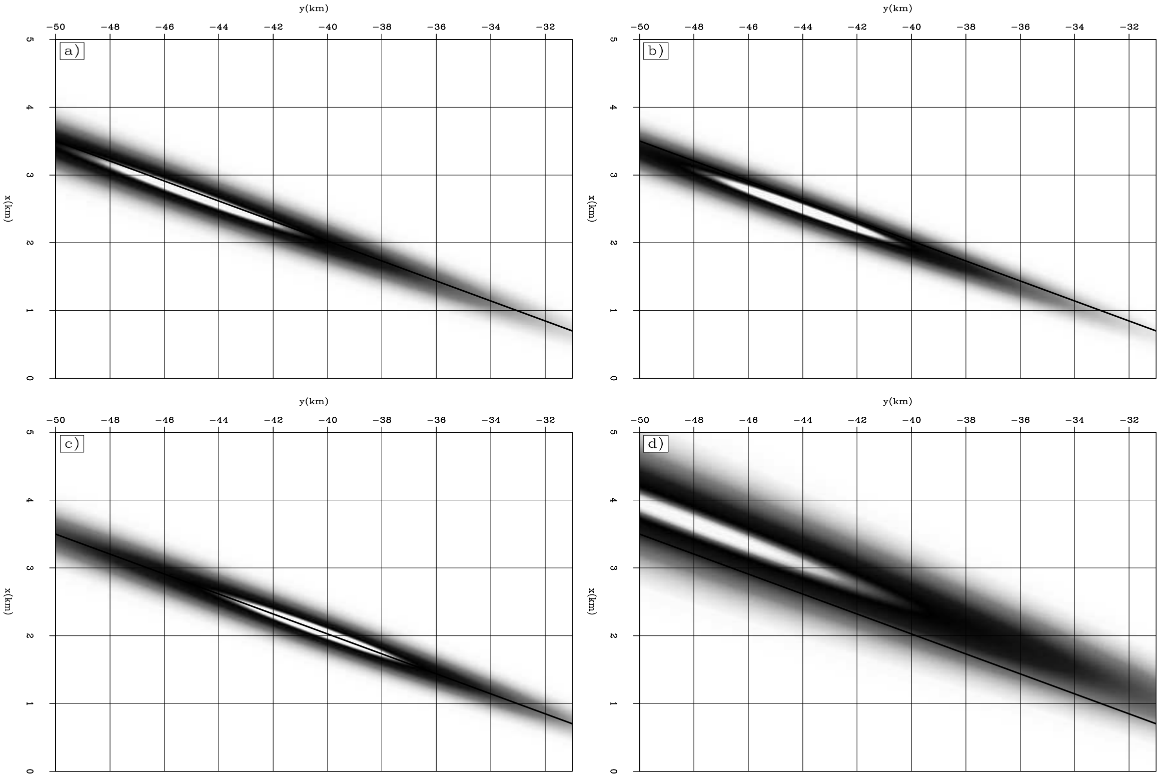


Figure 3.19: Sensitivity to horizontal location for the inversions using the four periods of passive data. TimeShot-xy [CR]

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 calculated by the inversions described above. The location of the source for the dominant events in the synthesized gathers 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 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 all produced gas.

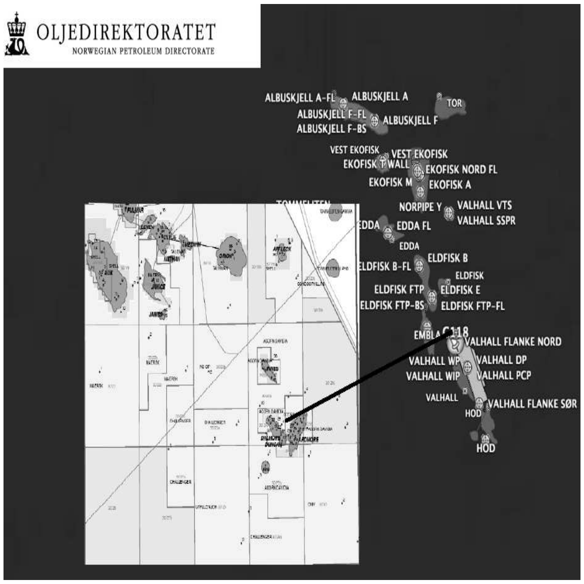


Figure 3.20: A small section of the British infrastructure map, the Norwegian infrastructure, and the location of the energy source inverted in Figure 3.19. The inverted locus of the energy in the gathers (the SW terminus of the black line) corresponds exactly to the location of the Ardmore development. TimeShot-mapa [NR]

---

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

3.6 Conclusions

The energy responsible for the features in the raw data, Figures 3.5 - 3.6, 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 3.5 is a collection of aliased repeating events similar that found in Figure 3.6. Also, the gathers generated with data showing the diffuse noise-train, Figure 3.5, have almost identical kinematics to the raw data showing a crisp water-velocity event, Figure 3.6. 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. This analysis also shows that the predominance of observable energy in the raw data 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 3.13, or 15 hours of data correlated in 12 second sections, Figure 3.12, 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 production 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 the energy source for the strong off-array event revealed by whitened correlation processing. Synthesized shot gathers produced with short, 12 s, time windows of raw data show very different results depending on the character of the raw data used. Figure 3.13 shows that when there are high levels of local energy generated by the Valhall production facilities, synthesized gathers produced with short time windows do not produce the far-field production event. Otherwise, the far-field production energy dominates the correlation result using raw data windows as short as 12 s and as long as 9 hours.

The synthesized gathers produced with 36 hours of passive recordings do not provided subsurface information. The gathers are dominated by a cultural noise source. The energy within this noise is several orders of magnitude stronger than anything else in the data. The strength and continued presence of this energy makes it very difficult to remove or ignore. Several subtraction methods, including inversion, were not successful in removing the energy.

Chapter 4

Time sections from Valhall

4.1 Summary

I performed NMO-stack, using a velocity model generated with active seismic surveys, 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 crossline direction and uninterpretable on time slices. While some events can be tracked between crossline sections, others 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 muted, windowed, and modeled before NMO-stack processing to compare to results produced with the synthesized gathers. These tests 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 yield obvious and convincing subsurface information.

4.2 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 (in inline sections) events are revealed in the resulting data volumes produced with all the various 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. The coherent noise event violates the kinematic assumptions inherent in both the CMP coordinate system and NMO. 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.

4.3 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 crossline direction and bandpassed from 10 to 25 Hz. The sections on the right column were extracted from the next crossline 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 (active seismic image is presented in Figure 6.3). These NMO-stack results have no continuity on time-slice panels, and the events change radically in character and timing between adjacent crossline CMP locations. Some events (dis)appear with no continuity between crossline locations. Other events are correlatable between the two sections, though move approximately 0.2 s. Assuming that the events are aliased in the crossline direction, this is an apparent velocity of 725 m/s. The definition of the CMP coordinate system requires CMP spacings half of the receiver spacing (for this acquisition geometry). However, a direct arrival through the water column violates the two-way travel path assumed in the CMP coordinate transform. The appropriate source receiver distance for a direct arrival is not 1/2 the receiver spacing, but still the full distance. Therefore, 300 m/0.2 s=1450 m/s is an appropriate water-velocity arrival consistent with the distant noise event identified in the previous chapter. Because this event is not a subsurface reflection, the CMP coordinate transform does not map the event correctly into the CMP domain. The strength of the events is also suspicious as the shot-gathers shown in the previous chapter do not contain obvious reflections from geologic events at these times.

Figure 4.2 used the same data used to produce top row of Figure 4.1 and a constant water velocity for the NMO correction. The two panels are the same crossline sections from the center of the array in Figure 4.1. Coherent events are present, but have troubling characteristics again. There is now some consistency of events in the crossline 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

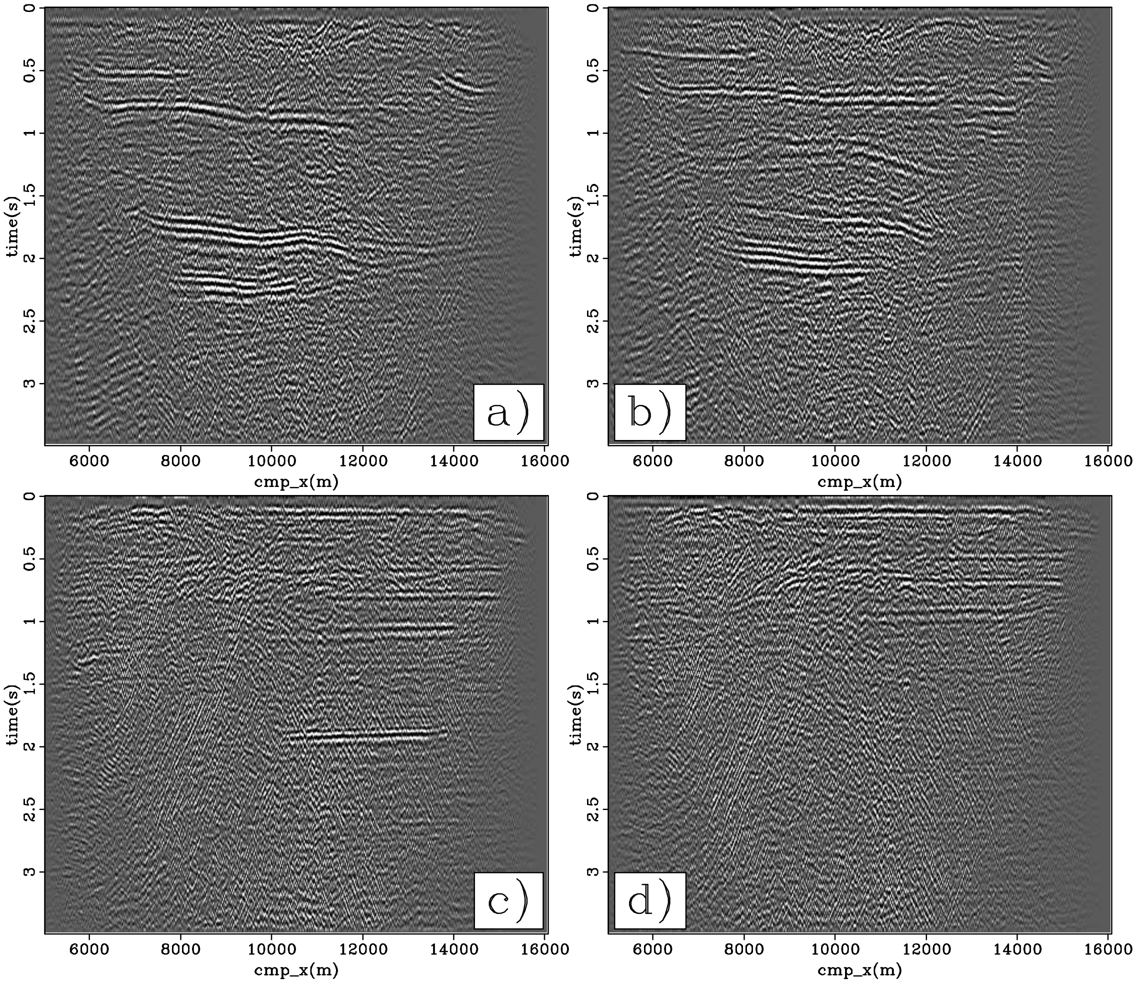


Figure 4.1: NMO-stacked inline sections from the center (left) of the Valhall array, and (right) the neighboring crossline CMP (+150 m). Top and bottom rows were processed identically using data collected 25 hours apart in February, 2004. TimeCMP-events [NR]

due to the distant noise source, the water velocity NMO was close enough to destroy almost all of the odd events in Figure 4.1.

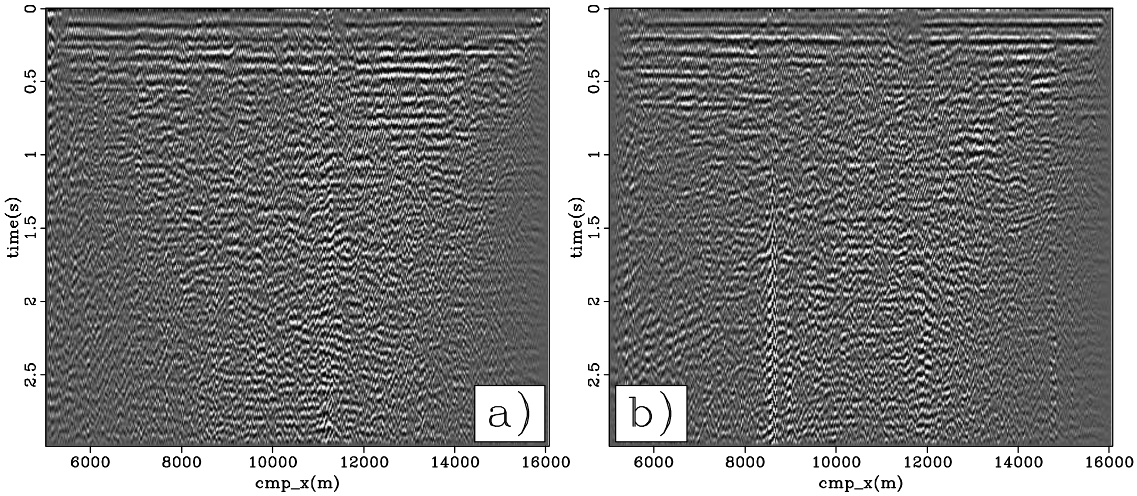


Figure 4.2: Inline sections corresponding to the data on the left of Figure 4.1 using constant water velocity for NMO. TimeCMP-water [NR]

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. Panel a is a modeled gather produced 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. Panel b was synthesized from passive data and muted to 0.128 s after the arrival time of the event in the top panel. Notice that there are significant correlation side-lobes remaining in the section. Panel c was produced by keeping only a short time window around the arrival of the noise event in the synthesized gather.

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. Panel a is the same as Figure 4.1a. Panel b was produced with the muted data from Figure 4.4b. Panel c used the data in Figure 4.4c. The final section, d, is the result of processing the forward modeled shot gathers, Figure 4.4a. Differences between Figure 4.5a and b are nearly impossible to find. Inspecting panels c and d from the bottom up, there are many similar features, that share characteristics: location, length, dip, and curvature. The steeply dipping events in panel c

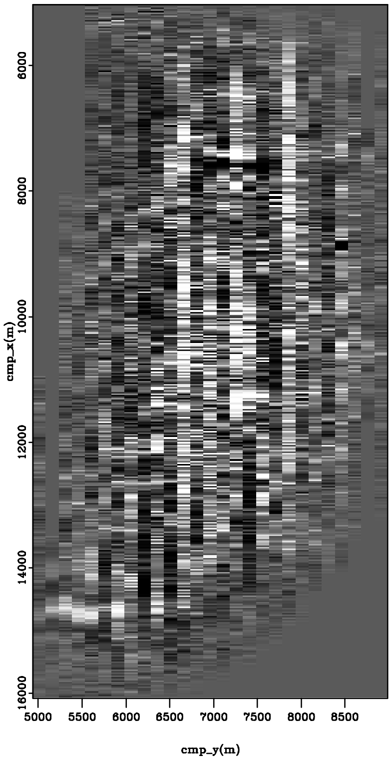


Figure 4.3: Time slice, 0.4 s, through the NMO-stacked cube using water velocity. Some continuity exists across the panel, while individual lines in the CMP-X direction show strong coherence. TimeCMP-waterslice [NR]

are not present in d, but almost all other main features are identifiable. The inclusion of a small amount of data around the direct arrival, and the instability of the wavelet associated with the distant source energy differentiates the results in panels c and d. Correlation side lobes in the real data add repeated events and more noise to the section at $t = 0.6$ s.

Figure 4.6 compares the muted shot gather generated with the first 9 hours of passive data (Figure 4.4b) 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. Panel a is the same as Figure 4.1c and panel b was produced with the muted volume of synthesized gathers. Most of the strong flat events have been eliminated in panel b.

Figure 4.7b shows a stacked section of data from which the bulk of coherent noise has

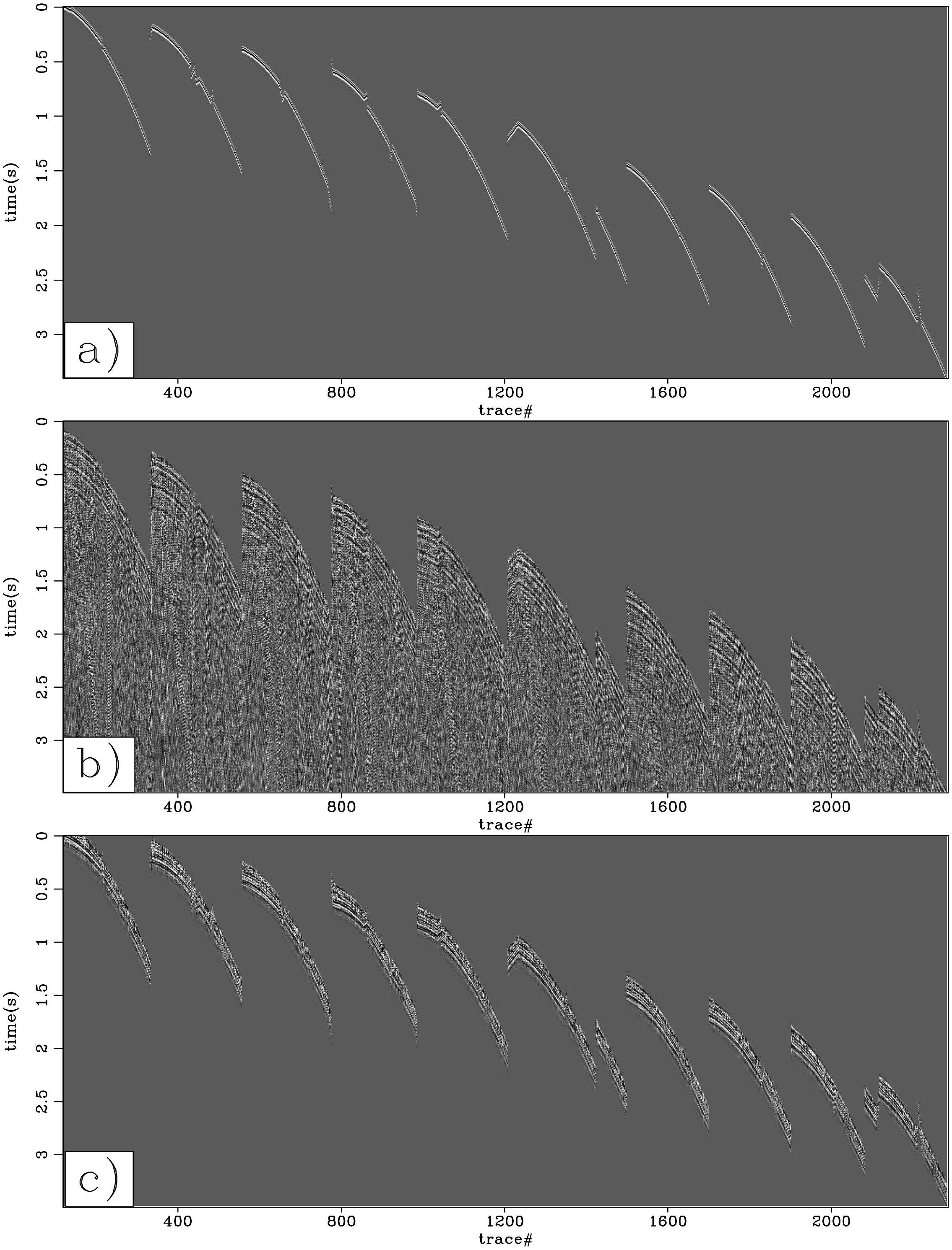


Figure 4.4: Shot gathers used to test the origin of the events in Figure 4.1. Panel a is a modeled shot gather using the inversion results presented in Chapter 3. Panel b is muted to the kinematics of the distant noise source. Panel c is a short window of data around the modeled event. TimeCMP-modldat [NR]

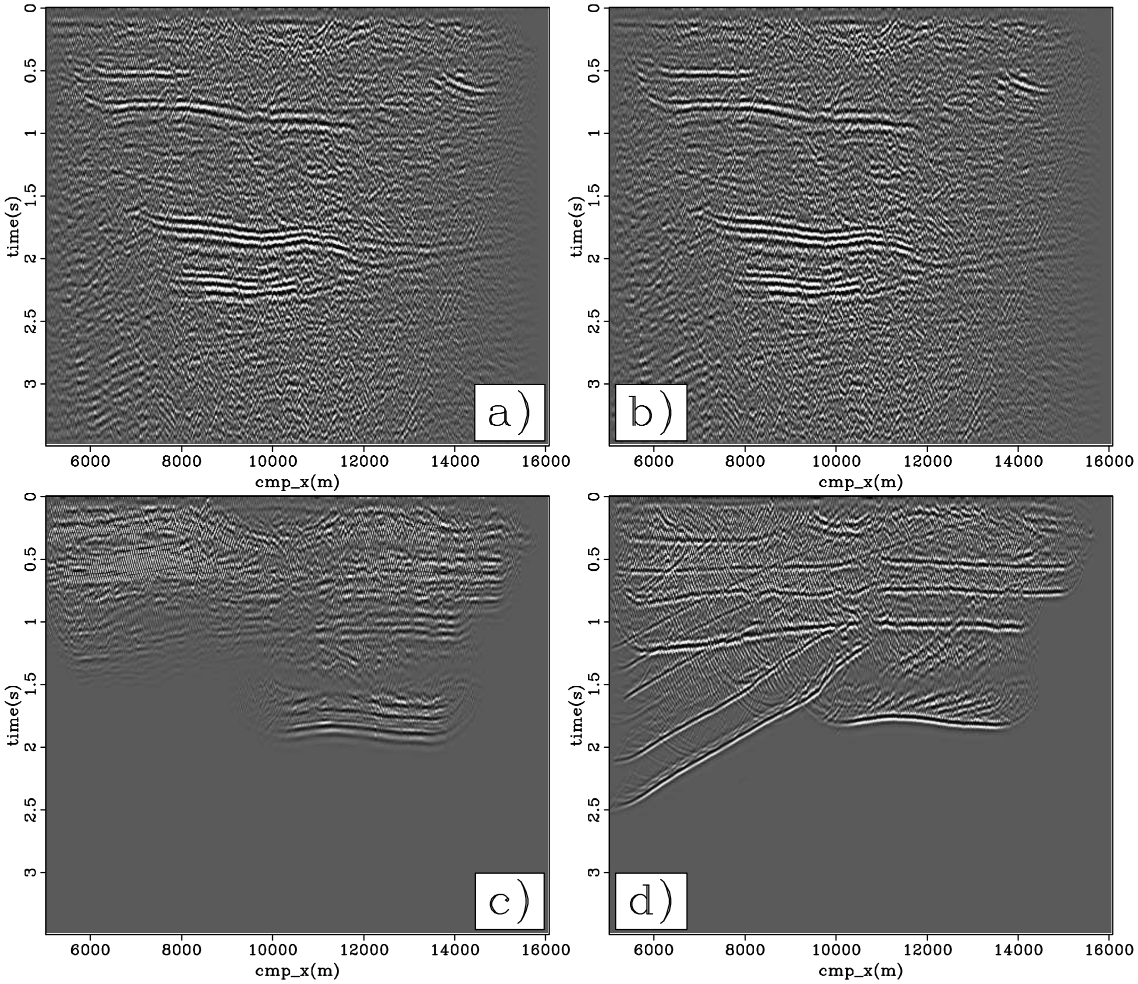


Figure 4.5: Inline NMO-stacked sections from the center of the array. a) Section produced with used all data available after correlation. b) Section produced with correlations after application of a mute to the modeled direct arrival. c) Section produced with data windowed around the direct arrival. d) Section produced with a modeled data volume containing only the direct arrival. TimeCMP-evnttest [NR]

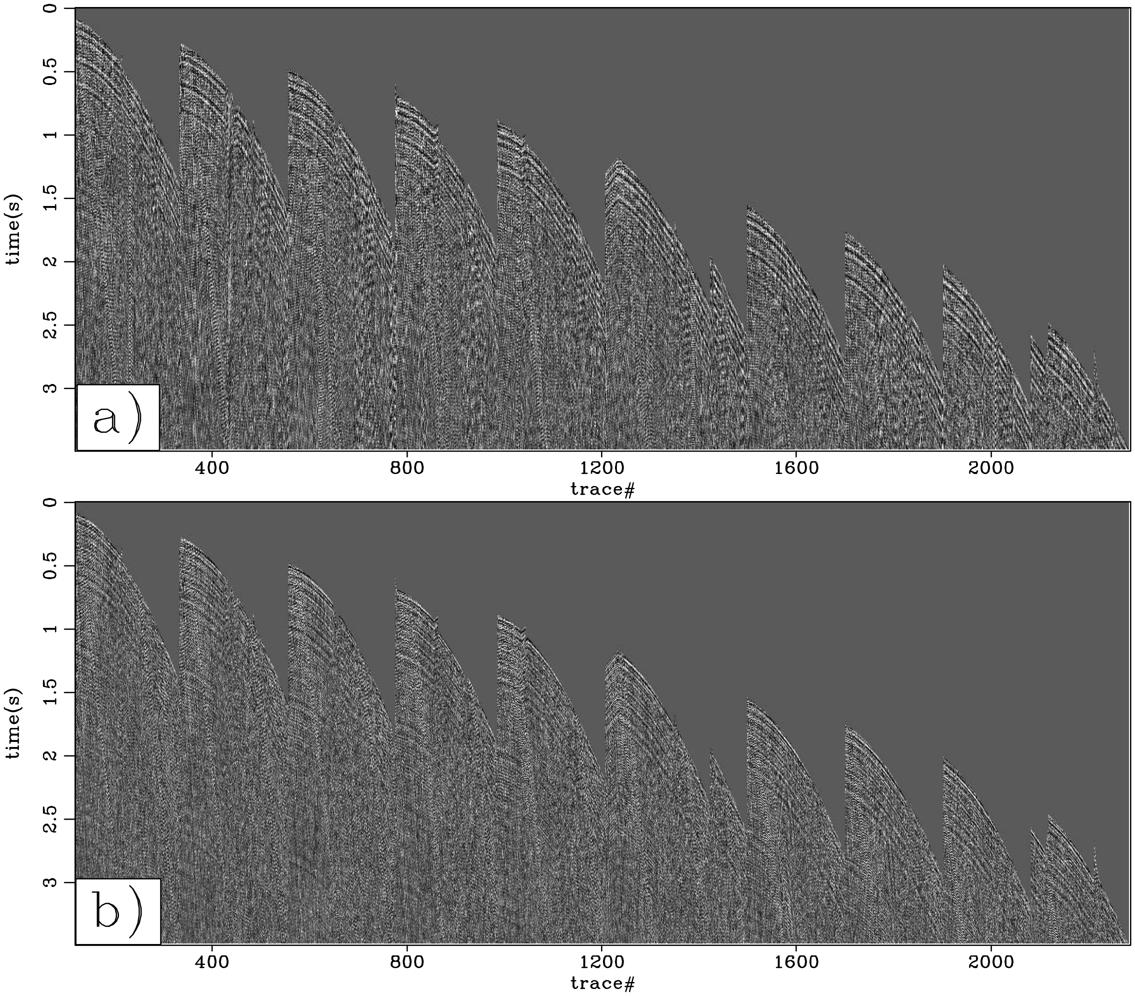


Figure 4.6: Top:Muted shot gather corresponding to the first 9 hours of data (Figure 4.4 top). Bottom: Muted shot gather generated with 5 hours of data 25 hours after those used to create the top panel. TimeCMP-modldat2 [NR]

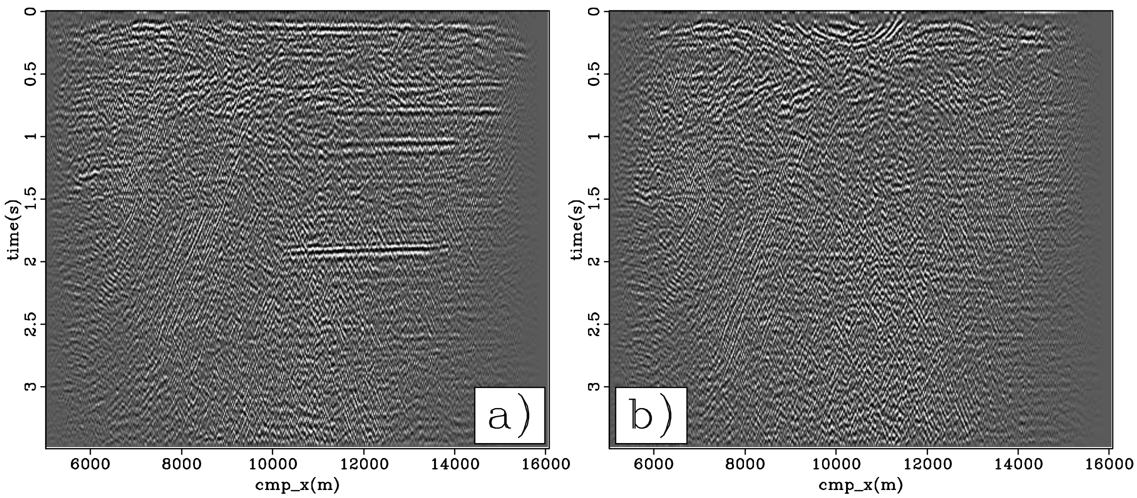


Figure 4.7: NMO-stacked sections from center crossline location generated from data a day after those used for Figures 4.4 and 4.5. Muting down to the direct arrival (Figure 4.6) successfully eliminates all coherent events. TimeCMP-evnttest2 [NR]

been removed before NMO-stack processing. This volume then represents the best candidate for finding subsurface structure. Notice a roughly sinusoidal series of identifiable events across the section between 0.0-0.2 s including the bowl-shaped events in the center of the section. Similar events are present in Figure 4.5.

Figure 4.8 shows two neighboring slices through the 3D volume of the stacked data produced with the muted shot gathers. The volume has been bandpassed from 8-15 Hz. The bottom panels show the same sections as those above, but produced with a gain function to enhance the events at in the middle of the time axis. The time slices in panels b and d were extracted through the curving events near time zero within the ovals in panels a and c. Within the circle on the time slices, there is a roughly circular feature identifiable as a mostly white ring with a dark center. In movies progressing down the time axis, these events are shrinking, circular rings centered around the location of the surface production facilities. This location suggests that the events are associated with production noise at the production facilities.

There is little continuity in the crossline direction throughout the bulk of the stacked volume. The events that approach the edge of the array are suspicious since these CMP

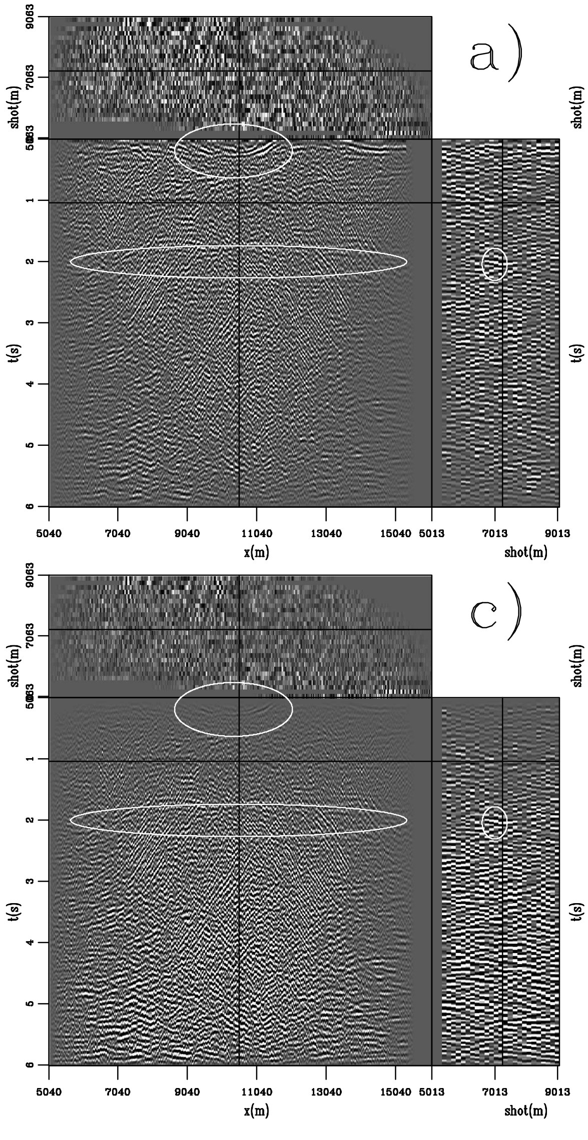


Figure 4.8: Best 3D time sections processed from 5 hours of passive data in February, 2004. Processing flow: Correlation, deconvolution, mute, NMO, and stack. TimeCMP-Fourth.mute [NR]

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 front frames of all panels at just less than 2 s (lower ovals). The event is interpretable between the adjacent inline sections (left and right columns) and corresponds to a black-white-black event on the crossline panels (circled) where the section line cuts through 2 s. This seems like the most likely event in the volume to be a genuine reflector.

A second coherent reflection in Figure 4.8 is between 6000-8000 m just beneath the time line on the front frames of panels b and d. The event is broken and somewhat discontinuous on the crossline axis, but could be interpreted with marginal confidence. At approximately the same time on the right edge of the sections, this event is again present. This event could be associated with production noise (bowl shape described above), or a real subsurface reflection that is interrupted in the middle of the array.

Because the distant production noise is localized in space, the direction of its arrival can be used to avoid correlating traces along the azimuth of the arrival. In so doing, it might be possible to avoid the overpowering influence of the production noise when synthesizing shot gathers. Figure 4.9 is an NMO-stack image produced by selecting from the total correlation trace pairs toward the location of the source energy. In this way, the strong event is in the acausal part of the correlations (and discarded), and geologic reflectors might be imaged away from its influence. However, even though the arrival is not kept in the correlation output, its effects are still in the calculation. Figure 4.9 does not show any coherent reflections. The swooping down to the left character in the inline panel mimics the effects of distant production noise in other results. The crossline panel shows no consistency.

Figure 4.10 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 far-field noise in Chapter 3. The second and third panels present NMO-stacked sections produced from the correlations of single transmission wavefields. The result in panel b was produced with a transmission wavefield due to a subsurface source below $x = 1800$ m and the modeled direct

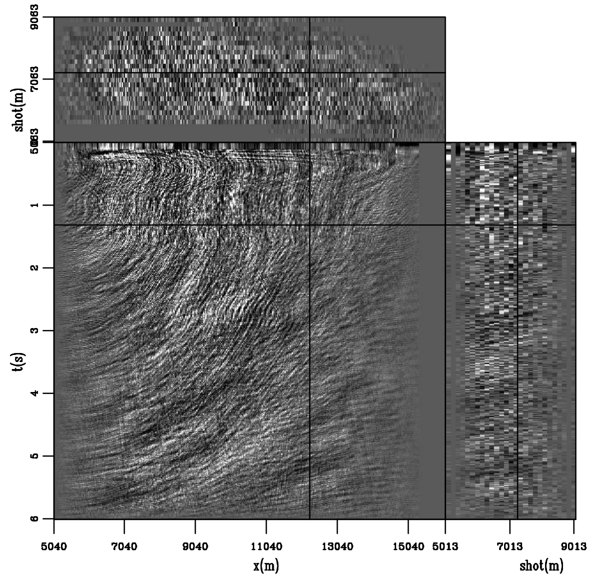


Figure 4.9: NMO-stack image created with only source-receiver trace combinations resulting in negative offset values. By correlating toward the source location, the strong arrival in the synthesized shot gathers is discarded with the acausal correlation lags. TimeCMP-oneside [NR]

arrival. The result in panel c was produced with the wavefield in panel a.

The steeply dipping events in the NMO-stacked sections in Figure 4.10 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 far-field noise event. Sections produced with only the far-field noise event contain only a strong event at $t = 0$ s. The steep events are caused by the convolution of the far-field noise event with events in the transmission wavefield. Therefore, the bowl-shaped structures in Figure 4.8 are due to the presence of the distant noise source and a local energy source in the transmission wavefield. The energy is centered at the production facilities. Obvious water velocity events sourced at the Valhall production facilities were presented and analyzed in Chapter 3. I interpret these events as the convolution of the far-field energy with local cultural noise produced during the synthesis of shot gathers by correlation of multiple direct arrivals.

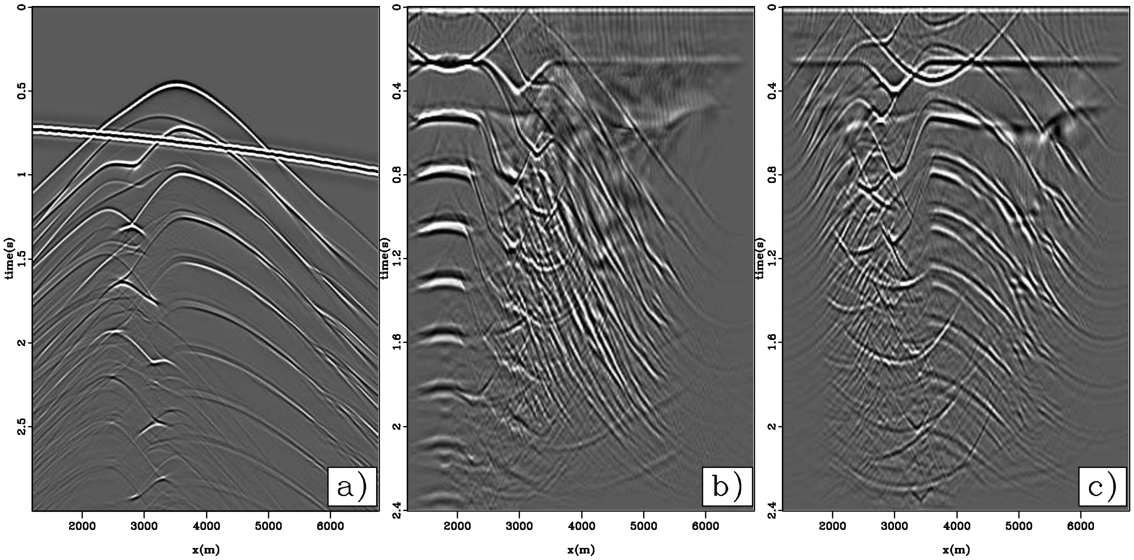


Figure 4.10: a) Transmission wavefield from the modeled data presented in Chapter 2 due to a source below the center of the domain. b) NMO-stack image produced from a transmission wavefield due to a source below the left edge of the domain. c) Stacked time image produced from the transmission wavefield in panel a. TimeCMP-synNMO [NR]

4.4 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 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. I performed similar experiments with data from 25 hours later. The NMO-stacked section produced with the muted gathers from that data 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 related to 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 due to coherent noise arriving through the water column 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 making 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. These experiments do not provide convincing results that the passive data recordings can illuminate the subsurface at Valhall.

Chapter 5

Depth imaging passive seismic data

5.1 Summary

Migrating the synthesized shot gathers produced in Chapter 2 produces interpretable images even when the synthesized gathers look noisy or completely incorrect. Incomplete distribution of subsurface sources surrounding the domain of interest and simultaneous processing of multiple subsurface sources are the most likely weaknesses in processing data from a passive recording campaign. These problems cause coherent noise events in the synthesized gathers that do not conform to the kinematics of the active reflection seismic experiment. Migration minimizes the effect of the noise and correctly images subsurface structure even when events in the gathers do not appear to conform to the conventional experiment. For this reason, I believe migration should be a near mandatory processing step for transmission wavefields. Direct migration of raw passive data produces accurate subsurface images without first synthesizing shot gathers. The images produced by direct migration are almost indistinguishable from those produced by migrating the synthesized shot gathers with the synthetic tests presented in Chapter 2. Under most scenarios, direct migration saves substantial computational cost. Any migration strategy that maintains separate up-coming and down-going wavefields and employs a correlation based imaging condition to combine them at depth can be employed to migrate transmission wavefields.

5.2 Introduction

A 5D volume of seismic data is a very large and complicated volume to interpret, that doesn't recapitulate humans' 3D concept of space. Offset, moveout, traveltime, interference, and high dimensionality make interpreting the subsurface in the data domain a challenging task. For passive data composed of transmission wavefields due to unknown sources with incomplete subsurface distribution, the gathers synthesized via correlation may be impossible to interpret correctly. Under scenarios expected to be realistic with field data, shown in Chapter 2, small violations of the assumptions behind correlation processing of transmission wavefields can quickly make the synthesized gathers uninterpretable. This chapter investigates the extent to which the artifacts in the synthesized gathers affect the image space after migration.

Migration is a powerful data processing tool to collapse dimensionality and sum redundant information as a function of offset to subsurface location. Shot-profile depth migration is a robust algorithm to process seismic data and return an interpretable subsurface volume. Fortuitously, migration can also reduce coherent noise events in the data volume that do not comply with the kinematics of simple reflections. Images still suffer from coherent noise, but it is often more recognizable as implausible, and easily ignored or removed. Interpreting a dense 3D image volume is also much easier than a sparse 5D data volume.

In Chapter 2, cross-correlation of transmission wavefields was analyzed in terms of its efficacy in synthesizing shot gathers from a conventional surface seismic active survey. In this chapter, I present depth images produced with shot-profile migration using the correlation volumes shown previously. The best synthesized volume of data was produced by summing the correlations from the transmission wavefields from 225 impulsive sources across the bottom of the model domain. This volume still contains many artifacts. Progressively poorer correlation volumes were produced by using very few of the transmission wavefields and by correlating the sum of transmission wavefields. The first section of this chapter presents the images produced by migrating those synthesized gathers.

Next, I develop the processing of passive seismic data as a migration operation: Direct migration. Artman and Shragge (2003) introduced the applicability of direct migration for transmission wavefields. Artman et al. (2004) provided the mathematical justification for zero-phase source functions. Shragge et al. (2006) showed results for the special case of imaging with teleseismic events. Direct migration of transmission wavefields requires an imaging algorithm composed of wavefield extrapolation and a correlation based imaging condition (Artman, 2006). Shot-profile wave-equation depth migration (Claerbout, 1971) fulfills these requirements. Images produced by direct migration are presented to show the validity of the new method.

5.3 Migration of synthesized gathers

The first two panels in Figure 5.1 are two shot gathers from the volume of synthesized gathers produced by summing the correlations calculated from all 225 transmission wavefields as described in Chapter 2. The virtual shot locations for the gathers are indicated by vertical lines. The right panel is the zero subsurface-offset image produced by shot-profile migration of the entire volume from which the gathers were extracted. All images produced in this chapter are zero subsurface-offset images. The delta function in equation 1.1 results in a strong flat event at $z = 0$ m in the image. To diminish that event and boost the amplitude of the second reflector, the image was processed with a depth-dependent gain function. All images in this chapter are presented with this same post-processing step. The event that partially obscures the second reflection is the first order multiple between the surface and the first reflector. Despite the imperfections in the synthesized gathers compared to active surface seismic gathers, the resulting image is very clear. Because most of the noise events in the synthesized gather do not conform to the kinematics of the reflection experiment from which the migration algorithm is designed, their effects are minimal in the image domain.

The left column of Figure 5.2 (panels a, c, e) are correlated gathers from a single source beneath the left edge of the model domain whose generation was explained in Chapter 2. The right column (panels b, d, f) shows active data modeled at the same (virtual) shot locations. The correlated gathers suffer many problems that make them obviously poor reproductions of the surface data modeled to the right. The gathers are a partial result of the summation over subsurface source required by the development of correlation processing to synthesize

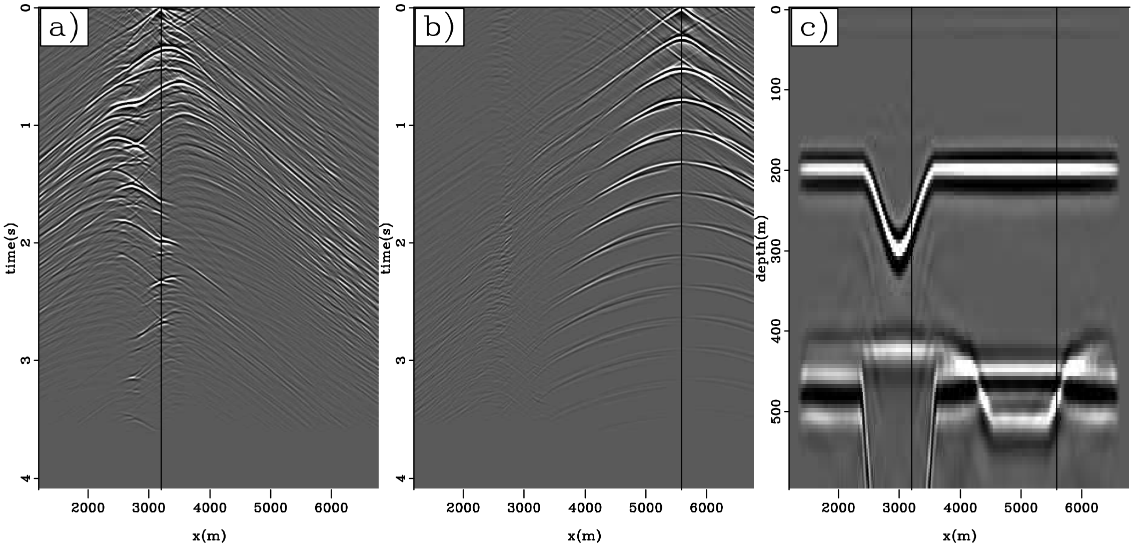


Figure 5.1: Representative gathers synthesized by summing correlations from all modeled transmission wavefields (zero-offset trace under black line) and the depth image produced with shot profile migration of all such gathers. DepthSyn-mig.all [CR]

active surface data. Produced with only a single transmission wavefield, the synthesized gathers are not just poor quality, but produce completely wrong kinematics in many places. The synthesized gather with virtual shot location directly above the subsurface source is the closest to correctly reproducing an active gather. At far offset the events are incorrect, but the near-offset traces in this synthesized shot gather are accurate. Synthesized gathers with virtual source locations further away from the horizontal coordinate of the subsurface source become less like the shot gathers which one is trying to synthesize via correlation. Most of the energy in panels c and e is from the high order multiples in the transmission wavefield. There is seemingly no correspondence between synthesized gathers and the modeled active gathers presented in panels d and f. Most of the events in these synthesized gathers dip to the right. There is little to no flat or left-dipping energy in the synthesized gathers. Such results would likely be discarded if produced from field data since they do not obviously conform to active seismic shot gathers.

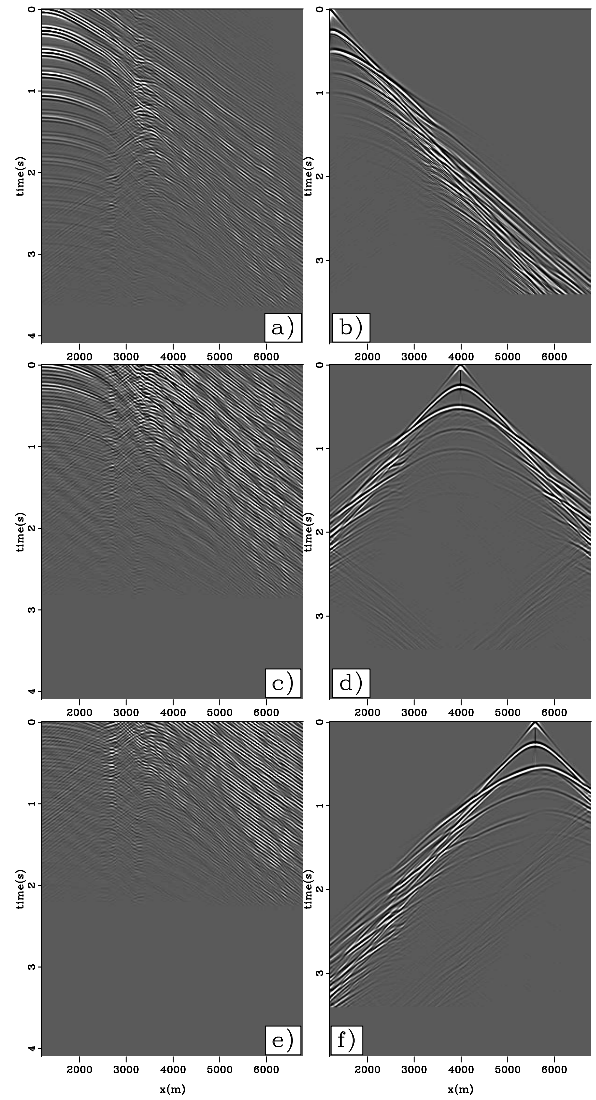


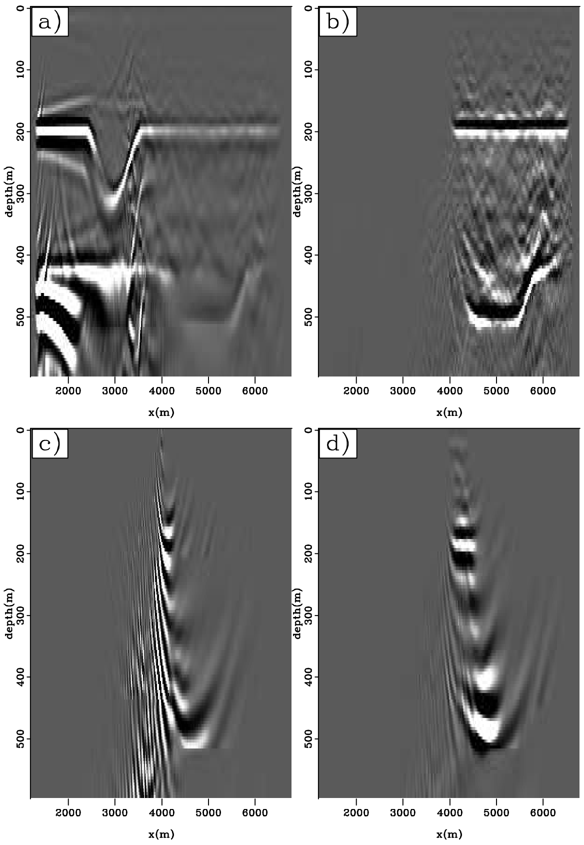
Figure 5.2: Left column: Synthesized gathers from left and middle of synthesized data volume produced with a single transmission wavefield due to a source below the left edge of the model domain. Right column: active surface data at the same (virtual) shot locations. DepthSyn-dat1200 [CR]

Figure 5.3 shows images produced from the volume of synthesized gathers using a transmission wavefield due to a single source beneath the left edge of the model (left column Figure 5.2). Panel a is the migration of all gathers produced by correlation, which images the subsurface structure well above the noise level. Structure on the right side of the model and down-right dipping structure is diffuse and weak. However, the entire domain is correctly imaged. Panel b was produced by migrating the synthesized shot gathers from only the right side of the model (such as Figure 5.2c,e). Without the strong, well-imaged left side of the domain, the gain for the figure is substantially boosted which shows more of the background noise. However, the noise is not coherent, and the power of events is substantially above the noise level. The negative (black) part of the wavelet below the reflector is lost, but the positive (white) peak is in the same location ($z = 200$ m) as in panel a. Despite the correlated gathers on this half of the domain looking completely incorrect (bottom Figure 5.2), using them as data for migration has correctly imaged the subsurface structure.

Zero subsurface-offset images produced with a single shot gather from active or passive surveys image flat reflections across a very limited lateral distance. Figure 5.3c shows the image resulting from the migration of a single synthesized shot gather with virtual source location $x = 4000$ m. It is impossible to interpret structure and noise in this image. Panel d is the migration of 20 synthesized shot gathers, beginning with virtual source location $x = 4000$ m. Destructive interference of the noise within the image produced with each individual shot gather begins to build the two subsurface reflectors seen above in panels a,b. These images show that although the synthesized gathers with virtual source locations at large horizontal distance from the subsurface source look completely incorrect (Figure 5.2c,e) they do contain appropriate events to image the subsurface structure during migration.

Figure 5.4 presents a suite of migrations using the correlation volumes produced in the previous chapter generated to explore the deleterious effects of processing two transmission wavefields simultaneously. The synthesized gathers presented previously were obscured by the superposition of two correlation results. The gathers were not uninterpretable, but without knowing that two sources were contained within the transmission wavefield, it is likely that the subsurface would be misinterpreted. The structure revealed in these depth

Figure 5.3: a) Image produced by migrating all synthesized gathers from a single transmission wavefield due to an impulsive source at the left edge of the domain. b) Image produce by migrating right half of the total volume of synthesized gathers. c) Migration of synthesized gather with virtual source location $x = 4000$ m. d) Migration of 20 synthesized gathers from virtual shot locations $x = 4000 - 4380$ m. DepthSyn-mig.shot1200 [CR]



images however is very easy to accurately interpret and noise is easily identified as such. The top row are images produced by migrating the correlations of summed transmission wavefields. The left panel used impulsive wavefields, and the right panel used randomized wavelets. The same is true for the bottom row of images, which show the migration of summed correlation volumes from the two wavefields. The bottom row thus represents a correctly processed partial result (independent correlation).

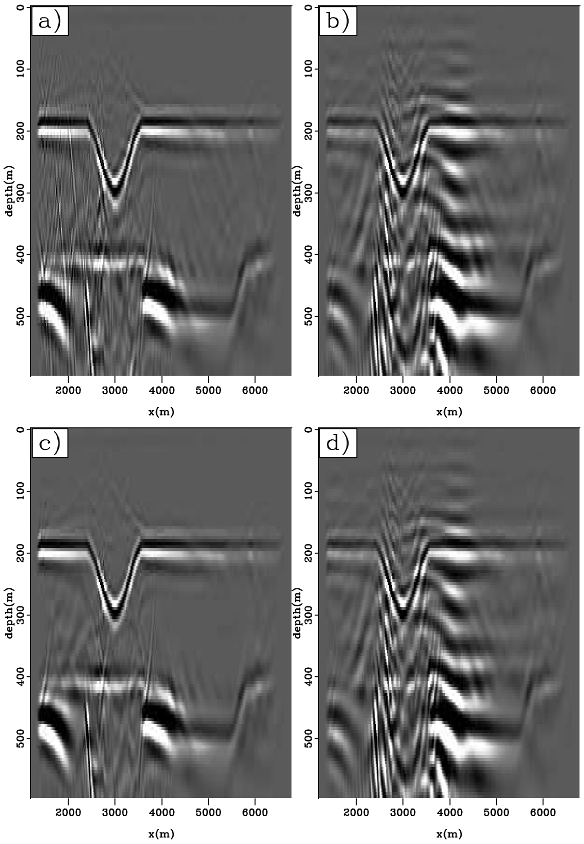


Figure 5.4: Images obtained by migrating correlation volumes due to two sources (located below the left edge and center of model domain). Left column used impulsive sources, while source functions were randomized for the right column. Top row was produced with summed transmission wavefields, and the bottom row by summing the correlations of single transmission wavefields. DepthSyn-mig.shot1200.3500 [CR]

Despite the fact that the model was illuminated with only two subsurface sources, all the structure is well imaged in Figure 5.4. Much of the noise in panel a is eliminated in panel c by correctly processing the transmission wavefields one at a time. The images in the right column are almost identical. In fact, migrating the correlations from the transmission wavefield modeled due to a single source below the middle of the model domain produces a similar image. Figure 2.8 shows that the random wavelet used for the source at the left edge of the model also contains a nearly 2 s time delay. Therefore, this energy is not included in the image at such shallow depths.

Figure 5.5 shows images produced by migrating the correlations produced with wavefields that contain more than two subsurface sources. Both panels were produced with the synthesized correlations from the sum of 25 transmission wavefields due to sources evenly distributed across the bottom of the model domain. The sources used to generate panel a were all impulsive. The image in panel b was produced by first adding a random phase shift and wavelet coda to each transmission wavefield before summing. The subsurface structure is fully illuminated, though steeply-dipping noise-trains are seen due to the sparse sampling of subsurface sources and the correlation interference of all the individual wavefields. Similar features are in visible in Figure 5.4a. The addition of randomized wavelets in panel b does not substantially differentiate the result from the image in panel a. The noise events are less structured and the second reflector is marginally less continuous in panel b.

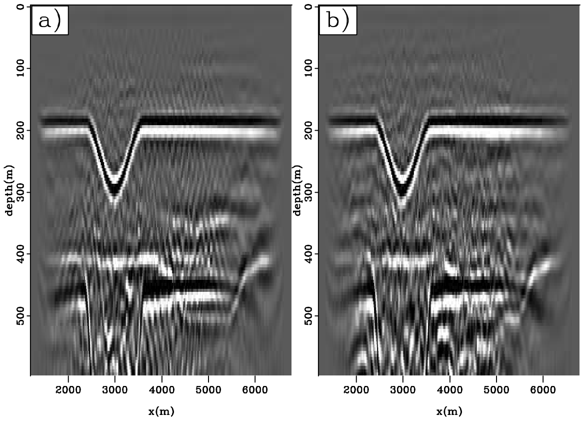


Figure 5.5: Images produced by migrating synthesized gathers of a summed transmission wavefield due to 25 sources across the bottom of the model domain. Panel a used impulsive wavelets, while Panel b was produced after introducing randomized source functions. DepthSyn-mig.shotsmall1 [CR]

Figure 5.6 was produced by migrating the synthesized shot gathers from a single transmission wavefield due to 30 subsurface sources clustered around the center of the model domain. These images have less artifacts than the previous images, but do not illuminate the structure as well due to the limited lateral distribution of subsurface sources. Since the subsurface sources used were so close together, the events are highly correlatable across the space-axis near the flat hyperbola tops in the raw data. This leads to the strong, low frequency noise in the center of the images. A low-cut filter can remove this energy to provide better imaging of the second reflector, but the filter cut-off must be carefully chosen since the frequency content of the noise events is only marginally below that of the real events.

The randomized wavelets introduced to the input for the image in panel b removes much of the low frequency noise in the center of panel a. The noise is less structured, and the signal to noise ratio of the result is improved.

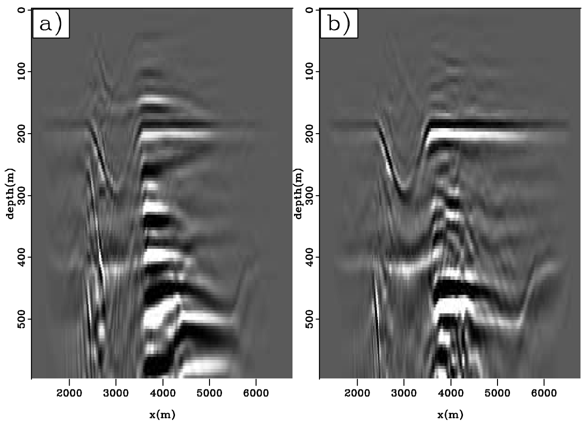


Figure 5.6: Images produced by migrating synthesized gathers of summed transmission wavefields from 30 sources clustered around the center of of the model domain. Panel a used impulsive wavelets, while Panel b was produced after introducing randomized source functions. DepthSyn-mig.shotsmall2 [CR]

5.4 Direct migration

5.4.1 Theory

Shot-profile migration uses one-way extrapolators to independently extrapolate up-going, R, and down-going, S, energy through the subsurface velocity model. R, extrapolated acausally, is a single shot gather. S, extrapolated causally, is a wavefield modeled to mimic the source used in the experiment. I will define the causal extrapolation operator, E^+, and the acausal operator, E^-. These one-way operators can not extrapolate evanescent or overturning waves, but this limitation does not adversely affect imaging simple structure with mild dip. In this thesis, I use the split-step Fourier plus interpolation (SSF-PI) (Kessinger, 1992) extrapolator for all migration operations.

Wavefields are extrapolated to progressively deeper levels, z, by recursive application of $E^{\pm}$

$$\begin{aligned} R_{z+1}(\mathbf{x};\mathbf{x}_s,\omega) &= E_z^-(\mathbf{x},\omega)R_z(\mathbf{x};\mathbf{x}_s,\omega) && (5.1)\\ S_{z+1}(\mathbf{x};\mathbf{x}_s,\omega) &= E_z^+(\mathbf{x},\omega)S_z(\mathbf{x};\mathbf{x}_s,\omega). && (5.2) \end{aligned}$$

With two wavefields at each depth, an imaging condition is invoked to extract energy for the image, i, that is collocated within the two wavefields. The zero-offset imaging condition for shot-profile migration is defined as the zero lag of the cross-correlation of the two wavefields at every depth

$$i_z(\mathbf{x}) = \sum_{\mathbf{x}_s} \sum_{\omega} R_z(\mathbf{x};\mathbf{x}_s,\omega) S_z^*(\mathbf{x};\mathbf{x}_s,\omega)\,. \tag{5.3}$$

The sum over frequency extracts the zero lag of the correlation. The sum over shots, $\mathbf{x}_s$, stacks the overlapping images from all the individual shot gathers. The important implication of the shot profile strategy is that the causal and acausal extrapolation operators are applied to independent (down-going and up-coming energy) components of the seismic experiment. This is the reason that two one-way operators can be used in place of two-way operators to recapitulate the field experiment in the computer.

Direct migration capitalizes on the correlation in the imaging condition, equation 5.3, to satisfy the one in equation 1.1 after extrapolating (or redatuming) the wavefield to levels progressively further away from the acquisition surface. The imaging condition performs the trace comparison required to synthesize active data from transmission data as explained in Chapter 1. Performing this operation in the image space simply moves the correlation processing into the subsurface from the acquisition plane. This result is proved in terms of a redatuming experiment in Artman et al. (2004), and independently derived in Borcea et al. (2003).

A simple framework within which to view correlation processing of transmission wavefields is phase-encoded migration (Romero et al., 2000; Biondi, 2006). 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 (Liu et al., 2006). 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 at the source location and zero elsewhere, $\delta(\mathbf{x}-\mathbf{x}_s,\omega)$. After convolving the impulses with encoding functions the source wavefield, $\hat{S}$, required for simultaneous migration is

$$\hat{S}_{z=0}(\mathbf{x},\omega) = \sum_{\mathbf{x}_s} A(\mathbf{x};\mathbf{x}_s,\omega)\delta(\mathbf{x}-\mathbf{x}_s,\omega)\ . \tag{5.4}$$

The five dimensional data volume of shot gathers, $R(\mathbf{x},\mathbf{x}_s,\omega)$, is similarly convolved with encoding functions and summed to reduce dimensionality

$$\hat{R}_{z=0}(\mathbf{x},\omega) = \sum_{\mathbf{x}_s} A(\mathbf{x};\mathbf{x}_s,\omega)R_{z=0}(\mathbf{x};\mathbf{x}_s,\omega)\ . \tag{5.5}$$

The image is produced at each depth level z with the same correlational imaging condition as equation 5.3, but does not require the sum over $\mathbf{x}_s$, and the encoded wavefields $\hat{S}$, $\hat{R}$ are used in place of S, R. 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 j encoded migrations with different function sets, A_j. 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 undo the convolution in equation 5.5. For a single source the correlation of $\hat{R}$ will reproduce the kinematics of R. The autocorrelation of A becomes a scalar amplitude and R^2 has the same kinematics as the reflection data R. Unfortunately, after the summation in the equation, this is not possible. To avoid this problem, we must correlate wavefields before source summation. The rigorous development of the correlation theory developed in the literature, equation 1.1, also relies on processing data due to a single source, ξ. The individual transmission wavefields indexed by ξ play the same role as a suite of encoded gathers $\hat{R}_j$ produced with functions A_j. Viewing a suite of transmission wavefields as phase-encoded shot gathers provides an simple explanation for need to sum the results of the correlations rather than correlate summed wavefields.

While phase-encoding can provide a framework for the analysis of passive seismic records, the particular methodology was developed as a migration algorithm. Instead of convolving conventional gathers with a synthetic encoding function for migration, a transmission wavefield is data convolved with a subsurface source function. In this case, the Earth provides the encoding function and the data is recorded by the geophone array as $\hat{R}$. Recognizing the similarities between a transmission wavefield and a phase-encoded wavefield naturally leads to direct migration of transmission wavefields.

Transmission wavefields are defined as the superposition of the up-coming and down-going seismic energy convolved with some unknown function: $T = A(S+R)$. Therefore, all the boundary conditions for the wavefields needed to perform a migration, $\hat{R}$ and $\hat{S}$, are recorded directly in the data. Propagating a transmission wavefield acausally will correctly extrapolate the up-coming energy contained therein during migration. Similarly, propagating a transmission wavefield causally will correctly extrapolate the down-going energy. Any spatial/temporal energy source (such as an array of Vibroseis trucks) can be used in a seismic experiment as long as the source characteristics are modeled correctly for the wavefield S used to migrate the data. For direct migration, we do not need to model the source wavefield because we record it.

The imaging condition for direct migration is a hybrid of the standard form presented in equation 5.3, and the one appropriate for phase-encoded gathers. Because the transmission wavefield is parameterized as a function of individual subsurface source, ξ, the direct migration imaging condition retains a summation over the source term. Therefore the imaging condition for direct migration is

$$i_z(\mathbf{x}) = \sum_{\xi} E_z^-(\mathbf{x},\omega) T_{z-1}(\mathbf{x},\omega,\xi) \left[E_z^+(\mathbf{x},\omega) T_{z-1}(\mathbf{x},\omega,\xi) \right]^* , \tag{5.6}$$

which in practice is an identical operation to the conventional algorithm. The only difference

$$\begin{tikzcd}
 & \text{shot profile migration} & & & & \text{transmission imaging} & \\
R_{z=0}(\mathbf{x};\mathbf{x}_s,\omega) \arrow[d, "E^-"] & \otimes & S_{z=0}(\mathbf{x};\mathbf{x}_s,\omega) \arrow[d, "E^+"] & = & T_{z=0}(\mathbf{x};\xi,\omega) \arrow[d, "E^-"] & \otimes & T_{z=0}(\mathbf{x};\xi,\omega) \arrow[d, "E^+"] \\
R_{z=1}(\mathbf{x};\mathbf{x}_s,\omega) & \otimes & S_{z=1}(\mathbf{x};\mathbf{x}_s,\omega) & = & T^-_{z=1}(\mathbf{x};\xi,\omega) & \otimes & T^+_{z=1}(\mathbf{x};\xi,\omega)
\end{tikzcd}$$

Table 5.1: Equivalence of shot-profile migration of reflection data and direct migration of transmission wavefields, $T(\mathbf{x},\xi,t)$. $\mathbf{x}_s$ has a similar meaning to ξ. $\sum_{\mathbf{x}_s,\xi}$ and $\sum_\omega$ produces the image $i_z(\mathbf{x})$ for both methods. Only first and second levels of the recursive process are depicted.

between a direct migration of passive data and a conventional shot profile migration is the initial condition of the down-going wavefield. Instead of modeling a source wavefield, the conjugate of the data wavefield is used for the down-going wavefield in the migration.

Table 5.1 pictorially demonstrates how direct migration of transmission wavefields fits into the framework of shot-profile migration to produce the 0^{th} and 1^{st} depth levels of the zero-offset image. The correlation in the imaging condition takes the place of the correlation in equation 1.1 to synthesize shot gathers. The summations over shot locations and frequency in table 5.1 are omitted to reduce complexity. Note however, that the sum over subsurface source locations in equation 5.6 takes the place of the integral over subsurface sources in equation 1.1. Also, after the first extrapolation step, using the two different propagation operators, the two transmission wavefields are no longer identical (and can be redefined R and S). This is noted with superscripts on the T wavefields at depth. Synthesizing shot gathers by correlation followed by migration is represented by movement from the top-right of the table to the conventional data domain on the left, and then down via conventional migration.

The principle difference between the phase-encoded sum of all shots and and a transmission wavefield, is that the transmission wavefield is defined as a function of subsurface source, ξ, while equations 5.4 and 5.5 are defined as functions of surface source locations $\mathbf{x}_s$. However, whether the sources are initiated at the surface or anywhere else is unimportant as long as the source wavefield used for migration mimics the spatial and temporal characteristics of the source energy that gives rise to subsequent reflections from the subsurface. The appropriate source energy is the direct arrival from the subsurface. Therefore, ξ is merely a wavefield index rather than a vector location containing the required information to model $S(\mathbf{x},\mathbf{x}_s,\omega) = \delta(\mathbf{x}-\mathbf{x}_s,\omega)$. Summing the images produced with each transmission wavefield due to all available sources ξ builds the complete subsurface image and cancels artifacts associated with incomplete illumination.

The source wavefield, $\hat{S}$, required to migrate phase encoded gathers is the encoding function with which the gathers were convolved. Transmission wavefields contain that function and all the events represented by $\hat{R}$, which introduces noise to individual images. The extraction of only the zero lag of the correlation at the imaging condition discards energy in the two wavefields that is not collocated. This includes much of the energy that has been extrapolated in the wrong direction (since the same data is used initially for both S and R wavefields). However, some cross talk will be introduced since this approach strictly violates the phase-encoded migration strategy. cross talk events in direct migration images are progressively eliminated by summing results from many individual transmission wavefields. This same cancellation happens while summing correlation volumes when synthesizing shot gathers at the acquisition surface or summing images produced with a suite of encoding functions. If the number of subsurface sources recorded is small, some cross talk artifacts will remain in the final image. If the subsurface sources do not completely surround the volume of interest, artifacts in the image are due to incomplete illumination.

5.4.2 Multiple subsurface sources

When considering the case of direct migration of transmission wavefields due to multiple sources, a summation over ξ is implied in field data used as data on the right column of Table 5.1. This can happen when subsurface sources are not well separated in time, or are too weak to be apparent in the raw data to facilitate separation by time windowing. If we assume that we can not separate passive records into individual transmission wavefields, field data, T_f, is most likely a partial sum over subsurface source

$$T_f(\mathbf{x},\omega) = \sum_{\xi}^{n} T(\mathbf{x},\xi,\omega)\,, \tag{5.7}$$

where n is an unknown number of sources that were active during a time window of the passive data. Processing relatively short time windows independently will tend to minimize n, though can not address the possibility that multiple sources may be firing simultaneously as can be expected if the ambient sonic energy is caused by cultural noise or reservoir compaction, for example.

Source summation is explicitly included in the conventional migration flow shown on the left column of Table 5.2. The implicit sum over subsurface sources for the transmission wavefields on the right column is indicated by not parameterizing the wavefield as a function of ξ. After summing across the source axis of a conventional reflection seismic data volume, the appropriate source wavefield, the sum of all the impulsive down-going wavefields, is a horizontal plane wave. The direct migration of summed transmission wavefields is very analogous to the horizontal plane-wave migration strategy, though if the number of subsurface sources is small, a complete plane wave will probably not be manufactured. Further, unless the sources fire simultaneously, they will not construct a horizontal plane wave no matter how many are included in the transmission wavefield. In this case, the combined arrivals of the multiple sources forms an unknown wavefront, or series of wavefronts, that is the source function to be used as the down-going source energy. For this reason, I label the approach wavefront imaging.

Figure 5.7 shows direct arrivals from subsurface sources that will reflect from the free surface to provide the down-going energy to investigate the subsurface. These are the wavefront source functions to which I am referring with the name “wavefront imaging.” Panel a is a single direct arrival from a source below the center of the model domain. Panel b is the sum of direct arrivals from 30 subsurface sources clustered over 600 m beneath the center of the domain. Arbitrary time delays have been added to each source. For clarity, no random wavelet coda was added to the arrivals. Panel c shows the wavefront sum of 25 subsurface

$$\begin{array}{ccccccc}
\multicolumn{3}{c}{\text{plane-wave migration}} & & \multicolumn{3}{c}{\text{wavefront imaging}} \\
\hline
\left(\sum_{\mathbf{x}_s^n} R_{z=0}(\mathbf{x};\mathbf{x}_s,\omega)\right) & \otimes & \left(\sum_{\mathbf{x}_s^n} S_{z=0}(\mathbf{x};\mathbf{x}_s,\omega)\right) & = & T_{z=0}(\mathbf{x},\omega) & \otimes & T_{z=0}(\mathbf{x},\omega) \\
| & & | & & | & & | \\
E^- & & E^+ & & E^- & & E^+ \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
R_{z=1}(\mathbf{x},\omega) & \otimes & S_{z=1}(\mathbf{x},\omega) & = & T^-_{z=1}(\mathbf{x},\omega) & \otimes & T^+_{z=1}(\mathbf{x},\omega)
\end{array}$$

Table 5.2: Equivalence between direct migration of passive field data and simultaneous migration of many shots in a reflection survey. Only first and second levels of the iterative process are depicted. $\sum_\omega$ produces the image i_z for both methods.

sources evenly distributed across the entire bottom of the model domain. As more subsurface sources explode close together in time, destructive interference progressively simplifies the resulting superposition to generate a plane wave at the limit. The direct arrival is actually only a part of the appropriate source wavefield required to image the subsurface. It is the principle component, but multiple reflections image subsurface reflectors as down-going energy for the next higher-order multiple in the transmission wavefield during migration. This is why the bad looking synthesized gathers in Figure 5.2 were able to image the subsurface correctly in Figure 5.3. The importance of the multiple reflections in the source wavefield will also be discussed in the interpretation of the images in Figure 5.8.

5.4.3 Synthetic tests

The modeled transmission wavefields used in previous examples are again used below to produce images with the direct migration strategy presented above. In addition to direct migration, images are also produced using the direct arrival from subsurface sources for the down-going source wavefield and horizontal plane-wave migration. Some figures from above are presented again next to their direct migration counterpart to facilitate comparisons between the two approaches.

Figure 5.8 shows images produced with several processing strategies all beginning with the transmission wavefield due to a single impulsive source at the left edge of the model

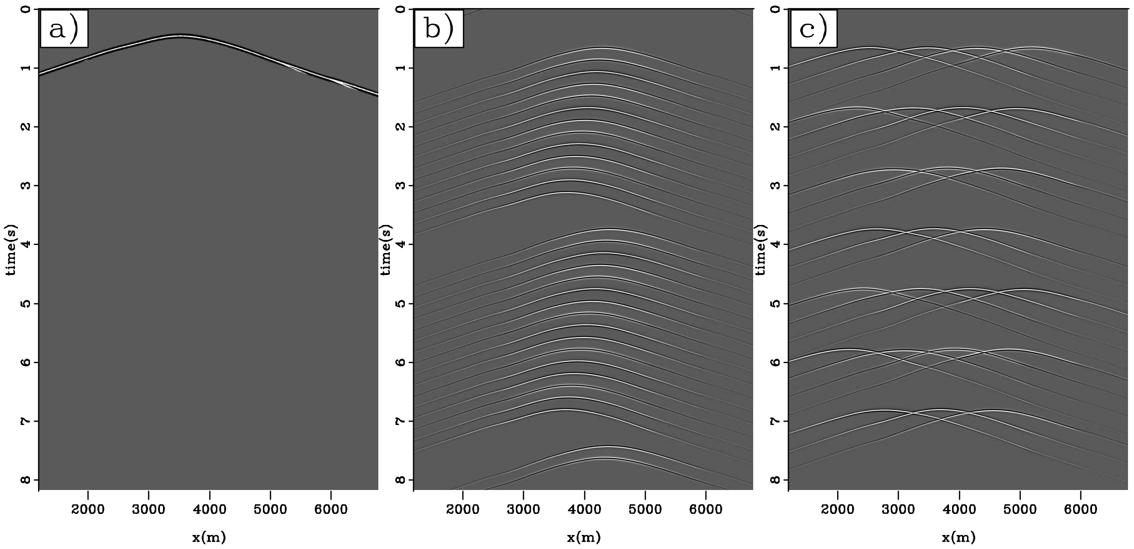


Figure 5.7: a) Direct arrival of energy from a source below the center of the model domain. b) Wavefronts due to the direct arrivals from 30 subsurface sources spanning 600 m around the center of the domain. c) Direct arrivals of 25 subsurface sources evenly distributed across the bottom of the domain. Arbitrary time delays were added to each direct arrival wavefront in panels b) and c). DepthSyn-directsum [CR]

domain. Panel a is the direct migration result. As can be expected, events are much stronger on the left edge of the model above the subsurface source. The entire domain is imaged fairly well, and the imperfections are easily identified. Because the right side of the domain is so weak however, it is probable that this structure would be below the noise threshold in a field experiment with weak, unidentifiable sources.

Figure 5.8b is the same image shown in Figure 5.3a (migration of synthesized gathers) for comparison to the direct migration result, panel a. Panel b shows a slightly different image than the direct migration to its left. The images were gained identically. There is marginally less cross-hatched noise on the right side of the image in panel b, but the reflectors are not as strong as in panel a on the right side of the domain. The migration of correlations, b, enjoys a slightly higher frequency content. All events, artifacts, and edge effects in the imaging results are identifiable in both images. Both images are incomplete results since only a single transmission wavefield was processed. The image produced by direct migration of 280 traces is over two orders of magnitude less expensive to produce

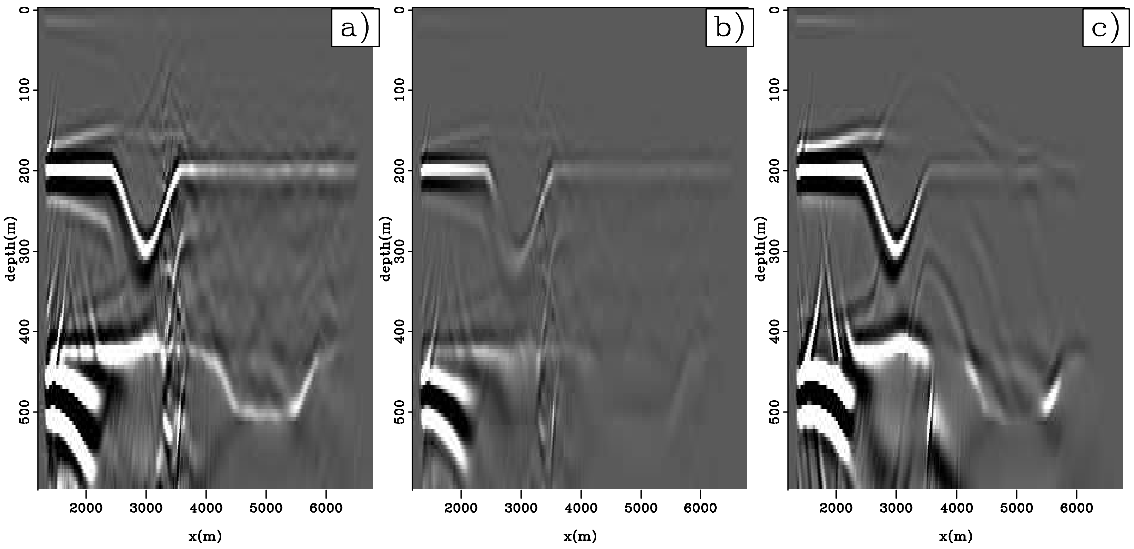


Figure 5.8: a) Direct migration of a transmission wavefield due to an impulsive subsurface source at the left edge of the model domain. b) Migration of synthesized gathers produced with the same transmission wavefield. c) Image produced with the transmission wavefield as data and only the direct arrival as the source. DepthSyn-dmig1200 [CR]

than migrating the $280^2 = 78,400$ traces in the volume of synthesized gathers. This cost difference does not include the cost of calculating the correlations required to compute the synthesized gathers, which is almost the same as the cost of a single direct migration.

Figure 5.8c was produced using the same transmission wavefield due to a single subsurface source below the left edge of the domain used for input in panels a and b. This image was created using the total transmission wavefield for the the up-going (data) wavefield. The down-going (source) wavefield was only the direct arrival from the subsurface source. This scenario is directly analogous to imaging with teleseismic arrivals (Shragge et al., 2006). The image has less cross talk noise than panel a. However, the right edge of the model and the second reflector are not imaged as well. Recall that the synthesized gathers on the right edge of the domain produced with this transmission wavefield, Figure 5.2c, contain energy from only high-order multiples. Therefore, the appropriate source wavefield for migrating a single transmission wavefield is not simply the direct arrival, but should also include primary and multiple events. Using the entire transmission wavefield as a source function produces cross talk in an image due to a single transmission wavefield. However, using only the direct arrival as a source function does not fully utilize all of the events in the transmission wavefield.

Figure 5.9 shows images produced with the various migration strategies available when processing a transmission wavefield due to two subsurface sources, one at the left edge and one below the center of the model domain. Panel a shows the direct migration of the sum of the two impulsive transmission wavefields. The result in panel b includes randomized source functions with the same summed transmission wavefield. The two transmission wavefields were imaged independently, by direct migration, and summed to produce the result in panel c. Panel d is the result of migrating the correlations produced with the summed transmission wavefield using randomized wavefields (Figure 5.4d). The sum of direct migrations, panel c, is clearly the best image, as dictated by theory. The main problem in the image is partial illumination. There is very little difference between the upper two images (with and without randomized wavelets). The comparison between panels b and d shows that there is very little difference between the direct migration, b, and the migration of the synthesized correlations, d. There is a small phase roll between the two images, but the two methods provide roughly equivalent imaging capability. The left side of the second reflector is a little more interpretable in b.

Figure 5.10 presents images produced with various processing strategies, but always using all 225 transmission wavefields from sources across the entire bottom of the model domain. Panel a was produced by direct migration of the sum of all 225 modeled wavefields. In this case, the minimum traveltime of the direct arrival for each transmission wavefield is constant, and therefore the summation builds a wavefront very close to a horizontal plane wave. The data for direct migration used to produce the image in panel b and long image in panel e included the addition of randomized source functions. Panel e is the same as panel b with an extended depth axis to show more of the deep artifacts. The second reflector stands out more clearly, and the multiple has been weakened in comparison to panel a. Some coherent cross talk events have been introduced however. There is an event at $z = 300$ m, a mirrored event below the flat-bottomed syncline, and many strange events in the deep image. The deep events are due to Fourier-domain wrap-around on the time axis. These

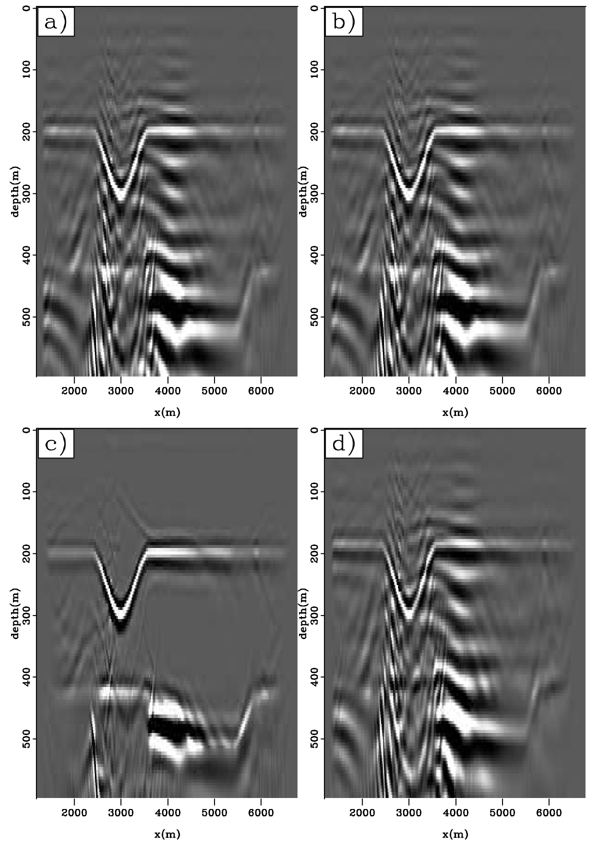


Figure 5.9: a,b) Direct migration of two summed transmission wavefields without and with randomized source functions. c) Sum of images produced by direct migration of individual transmission wavefields. d) Migration of synthesized gathers produced from the summed transmission wavefield. DepthSyn-dmig1200.3500 [CR]

can be simply removed by truncating the depth axis below the deepest reflector of interest similarly to truncating the correlation lag axis after the two-way traveltime to the deepest reflector of interest when synthesizing shot gathers.

Figure 5.10c shows the sum of the direct migrations of individual transmission wavefields. This result is very high quality and conforms to the requirement of independently processing each transmission wavefield. Panel d shows the image produced by horizontal plane-wave migration of the gathers individually correlated (Figures 5.1a,b). This image suffers from steeply-dipping noise events, and the reflectors are weaker compared to panels b and c despite all images being gained to the same level. The amplitudes in panels a and d are close to the same level. The source summations applied to the data used for these images discards almost all energy but vertically propagating plane waves. This limitation on the spectrum of plane wave constituents diminishes the quality of the imaging results.

Figure 5.11 presents comparisons between direct migration and the migration of synthesized gathers using modeled transmission wavefields below the entire domain. The top row of images was produced by independent processing of wavefields due to impulsive sources. The bottom row used a single transmission wavefield as input, which was produced by summing all the modeled wavefields due to randomized source functions. The left column images are direct migration results, and the right column images are migrations of synthesized shot gathers. Panel b is a reproduction of Figure 5.1c. The direct migration result in panel a is nearly indistinguishable from the migration of synthesized gathers in panel b. There are some minor amplitude differences and the migration of synthesized gathers, b, seems to have a slightly broader wavelet. This minor difference could be caused by creating a source wavelet for use in the down-going wavefield, S, with slightly different characteristics than the wavelet resulting from the correlations. I used a standard Ricker wavelet for all migrations conforming to the conventional shot profile migration. The number of modeled transmission wavefields used to produce image a was almost the same as the number of virtual shots (and receiver) locations to produce image b. Thus, the cost of the two migrations is roughly equivalent. However, synthesizing the gathers used as input for the migration to produce panel b requires substantial preprocessing that is not required for direct migration.

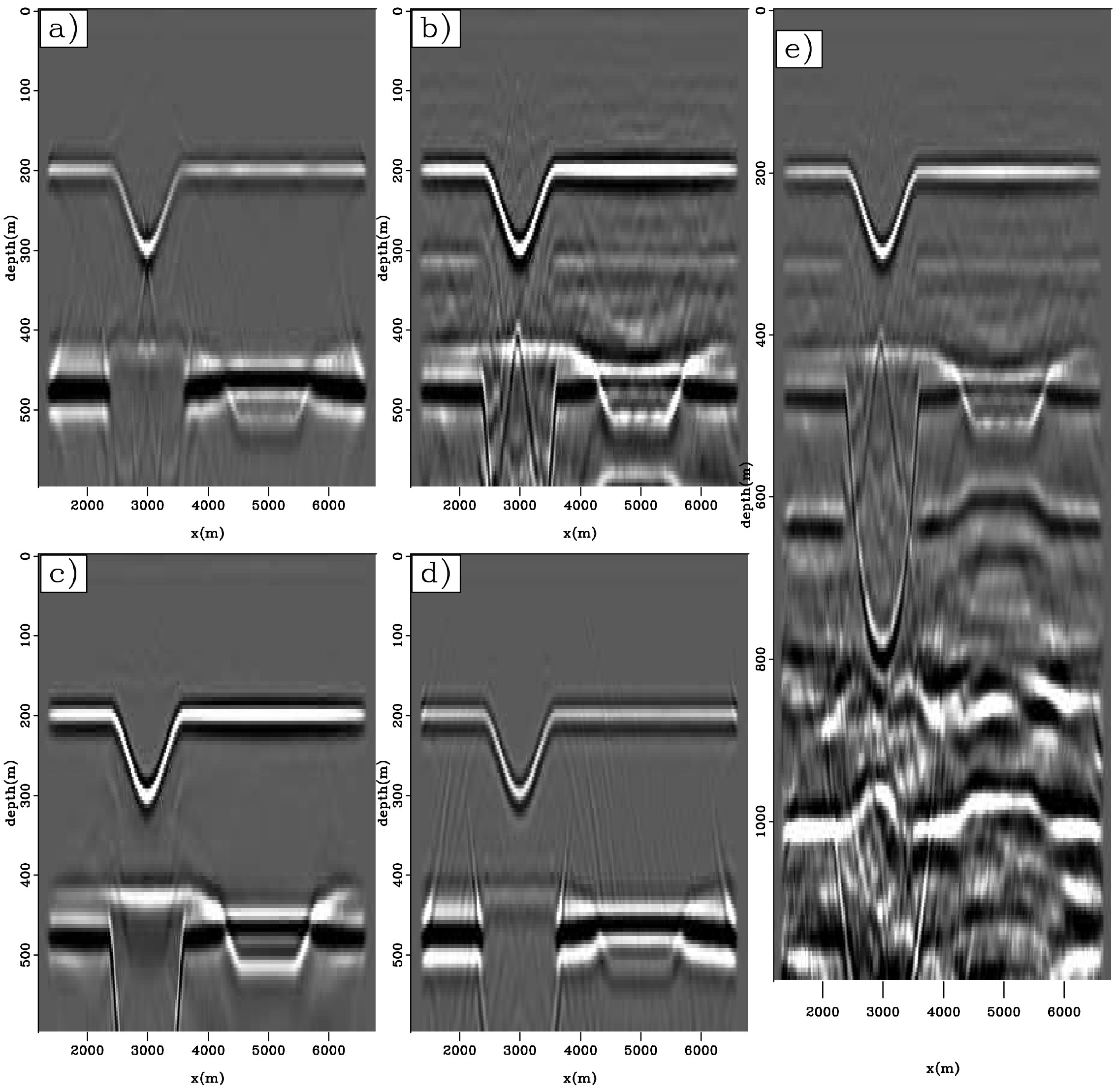


Figure 5.10: a) Direct migration of the transmission wavefield due to all 225 subsurface sources. b,e) Direct migration of summed transmission wavefields including randomized wavelets. c) Sum of images produced by direct migration of individual transmission wavefields. d) Horizontal plane-wave migration of synthesized gathers produced by correlating individual transmission wavefields. DepthSyn-dmigall [CR]

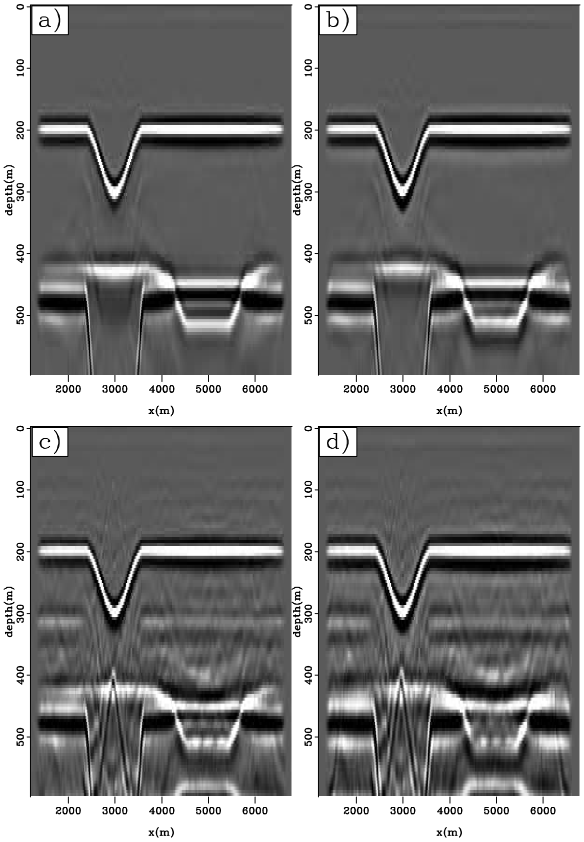


Figure 5.11: Left column shows direct migrations, while right column shows migrations of synthesized gathers. Top row processed impulsive transmission wavefields independently, while the bottom row used a single transmission wavefield generated by summing the transmission wavefields due to randomized subsurface sources distributed densely below the entire model domain. DepthSyn-allcomp [CR]

Figures 5.11c and d compare direct migration (left) to the migration of synthesized gathers (right) using a single transmission wavefield due to subsurface sources below the entire domain. These experiments represent processing the passive data from an entire data collection campaign simultaneously. The image in panel d took 123 minutes to produce, while the direct migration image took only 0.41 minutes: a cost ratio of 300. Again, the images are almost indistinguishable. All the main features in the images are directly comparable, though the events in the migration of correlations look to be of marginally lower frequency content.

Minimizing the number of subsurface sources active within a single transmission wavefield reduces the amount of cross talk introduced to the final image. The rows of Figure 5.11 compare images produced by processing 225 transmission wavefields due to single subsurface sources to those produced by processing a single transmission wavefield due to 225 subsurface sources. Figure 5.12 presents images produced with a scenario between those two end-member cases. Figure 5.12a is a transmission wavefield due to five subsurface sources. The 225 available transmission wavefields were summed in groups of five to generate 45 similar transmission wavefields. The five subsurface sources in each summed wavefield were all spaced 1125 m apart. All transmission wavefields were convolved with random source functions before summation. Panel b was produced by direct migration of the 45 summed transmission wavefields. For comparison, panel c is the result of summing the images produced by direct migration of all 225 individual transmission wavefields with randomized source functions. There are some weak steeply-dipping noise events crossing the bottom of the second syncline in b. The similarity between the images suggests that the assumption that transmission wavefields due to single sources must be processed independently is more restrictive than it needs be in practice. The phase-encoding development of processing passive seismic data provides the justification for this observation. The wavelet functions applied to the subsurface sources before summation are meant to model the differentiation of individual sources that can be expected in a real experiment. These natural encoding functions decorrelate the events within the transmission wavefields such that the cross talk induced by combining the wavefields is dampened. Spatial separation of the subsurface sources is also an effective natural encoding function.

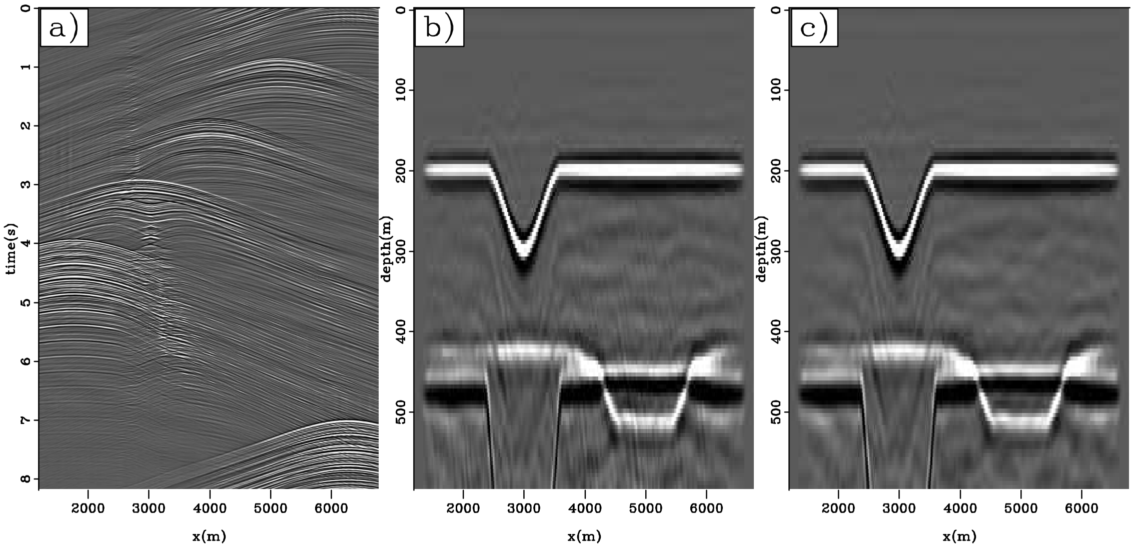


Figure 5.12: a) Transmission wavefield due to five subsurface sources with randomized wavelet functions. b) Sum of 45 images produced with direct migration of wavefields due to five subsurface sources with randomized wavelets. c) Sum of 225 images produced with direct migration of wavefields due to single subsurface sources with randomized wavelets. DepthSyn-migsums [CR]

Figure 5.13 shows images produced with the sum of transmission wavefields due to sources sparsely, but evenly distributed across the bottom of the model domain. The left column used wavefields with impulsive source functions, while the right column introduced randomness to the length and wavelet, and time delays to the source functions. The top row, a and b, was produced by direct migration: Source and data wavefields were identical. The bottom row, c and d, was produced using the sum of all direct arrivals as the source wavefield and the same data wavefield as used to create the images above them. Both migration strategies incur the same computational cost, as they differ only in the initial conditions of the source wavefield. The top row used the data itself as the source wavefield, while the bottom row results required knowledge of the up-coming source energy to model an appropriate source wavefield. Without the multiples in the source wavefield however, the bottom row of images do not image the structure as well as the images produced by direct migration above them.

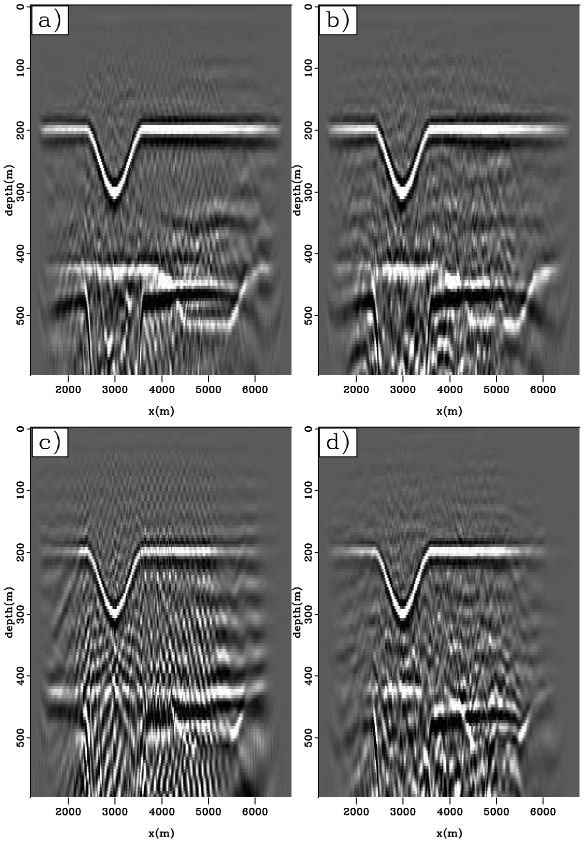


Figure 5.13: Images produced from a transmission wavefield due 30 source clustered around the center of the bottom of the model domain. Top row produced by direct migration. Bottom row produced using the direct arrivals for the source wavefield in the migration. Wavefields used for results in the left column had impulsive wavelets, while those used for the right column had randomized wavelets. DepthSyn-dmigsmall [CR]

Figure 5.14 shows comparisons between direct migration and the migration of synthesized gathers. It reproduces the images from Figure 5.5 and the top row of Figure 5.15. The transmission wavefield used for this example was the sum of 25 modeled transmission wavefields due to sources distributed evenly beneath the model domain. The left column used data due to impulsive wavelets, while the right column used data due to randomized wavelets. The top row were produced by migrating the correlations. The bottom row was produced by direct migration. The two rows of images are largely comparable. There is a minor phase roll between the images produced with the two different processing algorithms. The second reflector, is more continuous and less effected by noise in the direct migration images, panels c and d.

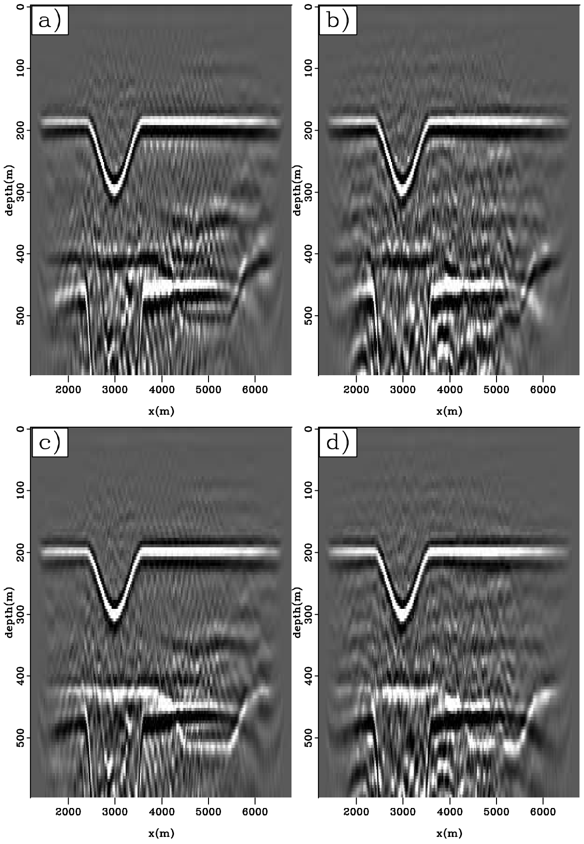


Figure 5.14: Comparison of processing strategies to image a transmission wavefield due to 25 sources distributed evenly across the bottom of the domain. Top: Migration of correlations. Bottom: Direct migration of the summed wavefield. Wavefields used for results in the left column had impulsive wavelets, while those used for the right column had randomized wavelets. DepthSyn-smallcomp [CR]

Figure 5.15 shows images produced with the sum of transmission wavefields due to 30 subsurface sources clustered around the center of the model domain. This experiment represents the situation in field data were the predominance of subsurface source energy is localized in a small area, such as might be expected in the case of induced fracturing. The left column used wavefields with impulsive source functions, while the right column introduced randomness to the length and wavelet, and time delays to the source functions. The top row, a and b, was produced by direct migration, and the bottom row was produced using the summed direct arrivals as the source wavefield for the migration. Panel b shows a better result, with the introduction of randomized source function, than panel a. There is less energy in the noise, and the structure is better illuminated at the edges of the model domain. Panel b shows a result with a modeled data volume that can readily be expected to provide insight into real situations encountered with field data recordings.

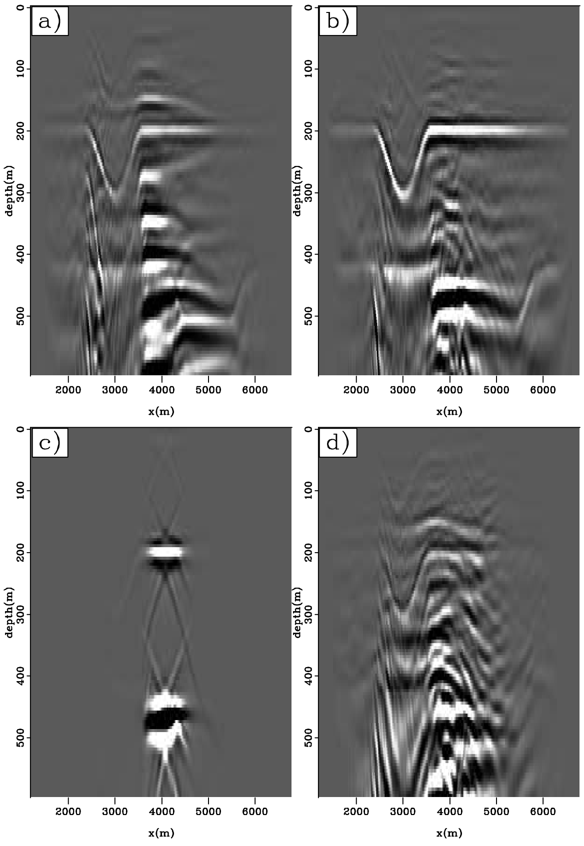


Figure 5.15: Images using summed transmission wavefields due to 30 sources clustered around the center of the model domain. Top row, left to right: Direct migration with impulsive sources and randomized sources. Bottom row, left to right: Horizontal plane-wave migration using synthesized gathers without and with randomized source functions. DepthSyn-dmigsmallsmall [CR]

Figure 5.15, in contrast to Figure 5.13 above, shows a much more marked difference between the images in the two rows. For the case of a localized cluster of subsurface sources, using the direct arrivals as the source wavefield for migration (c and d) has produced much worse images. When all the subsurface sources contain the same wavelet function and are perfectly in phase, the result using the summed direct arrivals as the source function for migration, c, images a tiny fraction of the model domain. The randomized source functions used to produce panel d move the individual transmission wavefields up and down the time axis such that the destructive interference leading to the poor result in panel c is not as severe. Direct migration (panels a,b) provides the highest quality images with passive data due to a localized cluster of subsurface sources as compared to using only the direct arrivals as the source function for migration.

Figure 5.16 shows comparisons between direct migration and the migration of synthesized gathers. It reproduces the images from Figure 5.6 and the top row of Figure 5.15. The summed transmission wavefield used for these images was the result of 30 subsurface sources below the center of the model domain. The results in the left column used data due to impulsive wavelets and those in the right column used randomized wavelets. These images roughly reinforce the observations made in analysis of Figure 5.14. The migration of correlations (a,b) are marginally poorer than the direct migration results (c,d). The ringing noise in the center of panel a is not as problematic in panel b. The most realistic data scenario with respect to potential field data (randomized wavelets) results in little noticeable difference in the quality of the images produced with the two imaging methodologies. As noted above in reference to Figure 5.8, however, the direct migration result was 300 times less expensive to compute.

5.5 Discussion

When passive data does not complete the requirements of the rigorous development of seismic interferometry, equation 1.1, the synthesized gathers may be misinterpreted or uninterpretable. The results may even be misleading enough that they are discarded. For the case of imaging the subsurface with only a single subsurface source, Figure 5.3 showed that synthesized gathers that seem completely devoid of reflections corresponding to the kinematics of active seismic shot gathers still contribute significantly to imaging the subsurface via migration. The synthesized gathers with virtual shot locations at far horizontal distance from the subsurface source were completely devoid of interpretable events and do not provide a

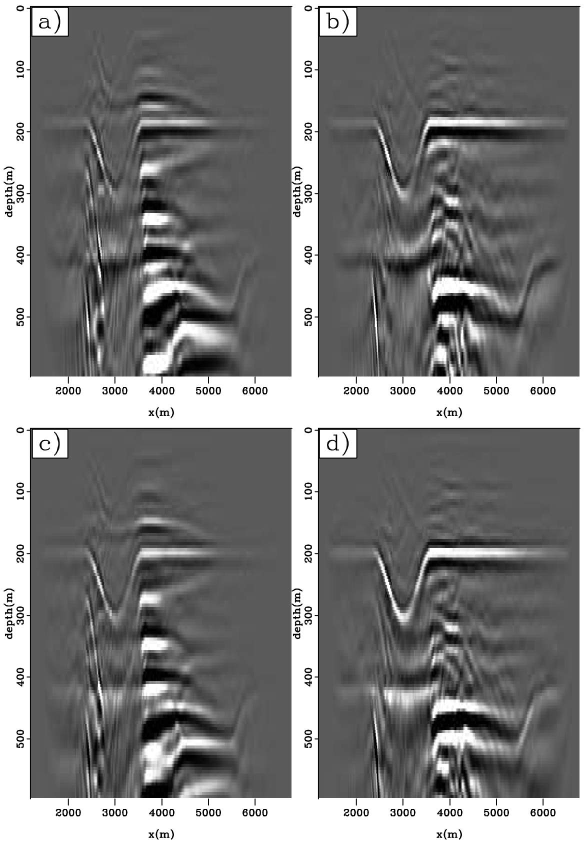


Figure 5.16: Comparison of processing strategies to image a transmission wavefield due to 25 sources distributed evenly across the bottom of the domain. Top: Migration of correlations. Bottom: Direct migration of the wavefield. Left column summed wavefields as modeled, while right column convolved each wavefield with a random wavelet before summing. DepthSyn-smallsmallcomp [CR]

useful data volume with which to assess the quality of the partial result.

The cost of synthesizing shot gathers compared to the cost of direct migration of a single transmission wavefield is almost the same for the modeled data presented above. For the same computational cost, the rows of Figure 5.17 show the two results possible when processing a single transmission wavefield. The left column shows a single synthesized shot gather, where the virtual source location is indicated by the vertical line. The right column shows direct migration results. The top row are results obtained by processing a transmission wavefield which contained subsurface sources clustered around the center of the model domain. The subsurface sources were evenly distributed beneath the domain in the transmission wavefield used to produce the bottom panels.

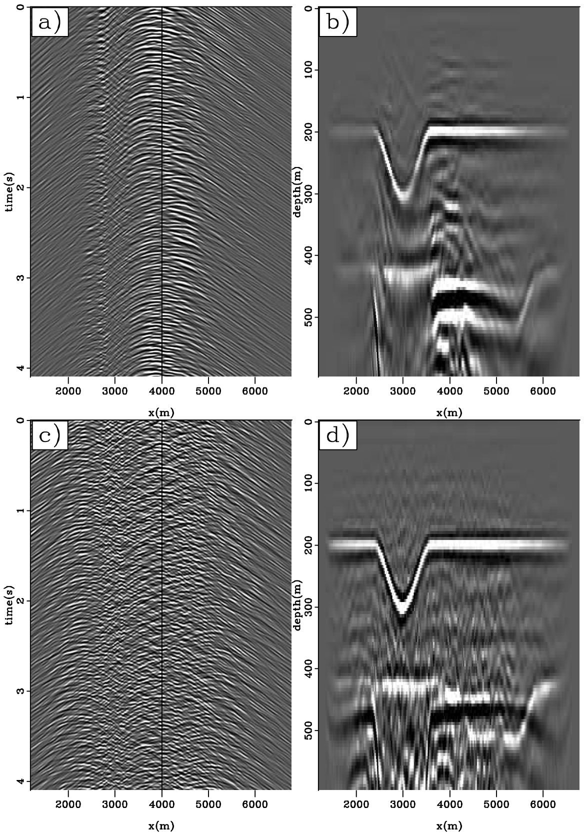


Figure 5.17: Synthesized shot gathers (left) and direct migration images (right) generated at the same computational cost using transmission wavefields due to a cluster of subsurface sources below the center of the domain (top) and a even distribution of sources below the entire domain (right). DepthSyn-disccomp [CR]

The synthesized gathers due to multiple subsurface sources in Figures 5.17a,c are uninterpretable. For this scenario, the images in panels b,d (and Figures 5.11, 5.14, and 5.16) show that migration produces very good results despite the poor quality of the synthesized gathers. Images produced by direct migration or migrating the correlated gathers are very similar in quality. When immersed in the lower signal-to-noise regime likely for field data, the subtle differences would likely be hard to find. However, because the synthesized gathers in Figure 5.17 are not useful, creating the synthesized gathers with these transmission wavefields is a useless intermediate processing step that still requires an expensive migration to produce images like those next to them produced by direct migration. If the synthesized shot gathers provide a useful domain in which to address noise or aliasing problems, this expense could be justified. However, it may be impossible to know what subsets of the energy in the gathers contribute to the image positively or negatively. Direct migration images of time windows from a passive recording campaign are cost effective and easily interpretable intermediate results that can be produced for quality control purposes for transmission wavefields containing either a single subsurface source or multiple sources.

The ratio of the cost to migrate a single transmission wavefield to that of preparing synthesized gathers from the data depends, most significantly, on the number of traces in the transmission wavefield and the number of depth samples in the image space. Figure 5.18 shows the expense ratio as function of the number of geophones, g, and the number of time samples, t, in the transmission wavefield and the number of depth levels, z, in the image domain. For a single transmission wavefield, the curves show the ratio of the cost to synthesize gathers, C_{syn}, over the cost of direct migration, C_{dm}

$$\frac{C_{syn}}{C_{dm}} = \frac{(gt\log t + g^2 t\log t) + 2tg^2}{gt\log t + Mztg\log g} \ . \tag{5.8}$$

The terms in parenthesis are forward and inverse Fourier transform costs. The last term in the numerator is the cost of correlating the traces. The first term in the denominator is another Fourier transform over time, and the second term is the migration cost, which is proportional to a Fourier transform over the trace variable g. M is a cost multiplier representing migration expense, which is usually dominated by the accuracy of the extrapolation operator used and the calculation of common image gathers. The migration cost multiplier, M, for my implementation of SSF-PI shot profile depth migration to produce zero subsurface-offset images is about 5.

The three groups of curves in Figure 5.18 are associated with the size of the image domain $z = (60, 120, 700)$. The cost of migration with a given number of receivers, C_{dm}, is dominated by the size of the image space. Therefore, the lowest group of curves used $z = 700$ for the number of depth samples in the image space. The synthetics presented above used 120 depth samples for migration. The images using the Valhall data I present in the next chapter use about 700 depth samples. The three curves within each group correspond to the number of time samples within the input wavefields, $t = 1024, 2048, 8192$. The dotted line is the ratio when the transmission wavefield contains 8192 time samples, and the solid line is due to 1024 time samples. The cost of creating the gathers with a given number of receivers, C_{syn}, is dominated by the number of time samples in the wavefield. Therefore, the highest curve in each group is associated with the largest number of time samples. Since migration is the most expensive operation, the number of depth samples in the image space has the most drastic effect on the cost ratio. It is interesting that for both the synthetic and the Valhall data, the combination of the number of receivers in the data and the number of depth samples in the image happen to result in a cost ratio ~ 1.

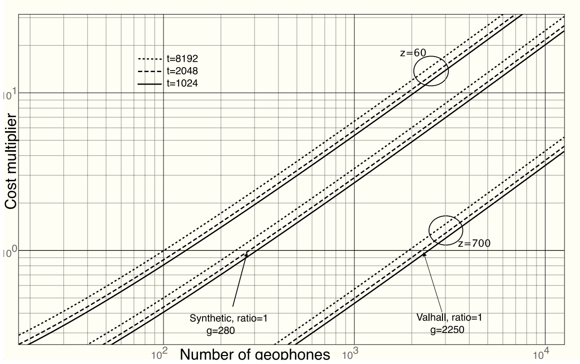


Figure 5.18: Cost ratio of processing a single transmission wavefield by synthesizing shot gathers or imaging with direct migration. Cost ratio values greater than 10^0 indicate that it is more expensive to synthesize shot gathers. DepthSyn-cost2 [NR]

When passive data is collected over a long period of time, the best way to process the volume is to work with short time windows from the total data set. This provides more reasonably sized data volumes to process, and will minimize the cross talk in the results associated with processing transmission wavefields due to multiple sources. When strong identifiable events are present, such as teleseismic arrivals or local explosions, it may be relatively easy to collect a small number of time windows from the total volume of data. Otherwise, the passive seismologist must window the data into arbitrary sections to minimize the number of subsurface sources captured within an individual transmission wavefield. The length and number of such time windows from a long recording campaign that can be efficiently processed can be determined by analyzing the relative computational costs incurred by processing the data with different strategies.

Figure 5.19 shows the computational cost ratio of migrating synthesized shot gathers compared to direct migration as a function of the number of time windows processed. Assuming time windows 15 s long, 10^5 windows is 17.36 days. The two curving lines represent processing transmission wavefields with 2800 (dashed) and 280 (solid) traces. These numbers correspond roughly to the size of the Valhall data and the synthetic used herein. The ordinate of the graph is the ratio of the cost to create and migrate synthesized gathers, C_{msyn} over the cost of direct migration, C_{dm},

$$\frac{C_{msyn}}{C_{dm}} = \frac{w(gt\log t + 2tg^2) + g^2 t \log t + Mztg^2 \log g}{w(gt\log t + Mztg\log g)}\,, \tag{5.9}$$

where the independent variable w is the number of time windows to process. This cost ratio, as compared to the previous, contains a migration cost in the numerator. However, this is a fixed cost that is not a function of w since the synthesized gathers can be migrated once after they have been summed. Likewise, the inverse Fourier transform of the synthesized gathers need only be performed one time. The entire denominator, the cost of direct migration, is multiplied by the number of time windows, w. This equation simplifies to approximately g/w for small numbers of time windows since the migration dominates the operation count to produce an image and the variables t and z appear on both sides of the ratio. The thin solid lines plot g/w for the two transmission wavefield sizes, $g = 280, 2800$. These asymptotic lines cross $C_{msyn}/C_{dm} = 1$ when $g = w$. The departure of the cost curves from the asymptote is due to the $2tg^2$ cost of correlation in the numerator of equation 5.9. For an array the size of Valhall (upper solid line) it is less expensive to process with direct migration until the number of time windows processed is approximately 10^4. Processing 15 s time windows, more than 40 hours of data need be available before it is more cost effective to produce a subsurface image by first synthesizing shot gathers.

The curves in Figure 5.19 both approach a horizontal asymptote. The solid lower curve,

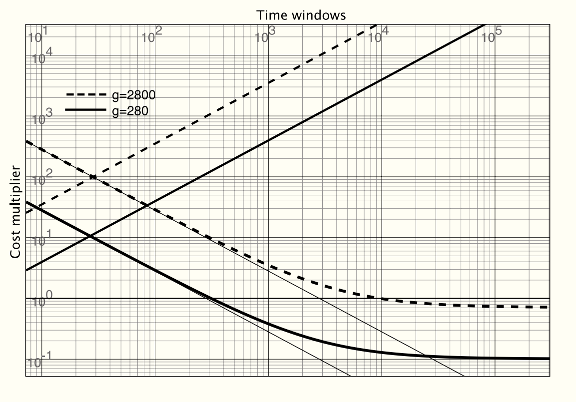


Figure 5.19: Cost ratio of processing many transmission wavefields by synthesizing shot gathers or imaging with direct migration. Cost ratio values greater than 10^0 indicate that it is more expensive to create and migrate synthesized shot gathers. DepthSyn-cost1 [NR]

calculated with a 280 receiver array, reaches an asymptote of 0.1. However, processing data from such a small array takes so little time with modern computers, that either processing selection is tractable. The dashed upper curve was calculated assuming a number of receiver stations similar to the Vahall array. The horizontal asymptote for the large array is $C_{msyn}/C_{dm} = 0.75$. This is likely not a severe computational performance cost if synthesized gathers are never made. This also eliminates storage of the g^2 traces in a correlation volume and produces continuous intermediate images that can be evaluated for quality control purposes. A parallel curve with horizontal asymptote of 10^0 can be computed by assuming 4000 sensors in a hypothetical array.

These cost functions assume that all time windows of data are equally important. If that is not the case, it would be beneficial to produce image volumes more often than one time to determine if the passive data recorded were useful. Producing such intermediate volumes quickly increases the computational burden of migrating synthesized gathers. Producing a few gathers with each time window does not have significant impact on the cost ratio. However, a second migration of the gathers approximately doubles the number of time windows that can be processed less expensively by direct migration. No matter how many intermediate migrations are produced, the cost ratio approaches the same horizontal asymptote.

5.6 Conclusions

Migration is a powerful operator that substantially reduces the complexity and size of the volume of data to be analyzed and often diminishes noise in the input data volume, or transforms it to an obvious character. Transmission wavefields are inherently imperfect data volumes compared to those produced with active seismic experiments. Migration greatly simplifies the complexity and volume of data to interpret, provided a reasonably accurate velocity model is available. While this could be a potential disadvantage, the expense related with mounting a passive recording campaign will, for the near future, limit passive acquisition to fairly well understood areas where a good velocity model is likely available. If not, a reasonable initial model can be constructed and then refined through iterative methods. While the discussion above was limited to zero subsurface-offset, both migration methods (direct migration and conventional migration of synthesized gathers) can produce images with complete common-image gathers as a function of subsurface offset (Rickett and Sava, 2002; Shragge et al., 2006) by generalizing the imaging condition, equation 5.3. Doing so introduces the possibility of residual migration to correct the image for errors in the migration velocity model (Biondi and Symes, 2004).

Using transmission wavefields modeled only across the bottom of a subsurface domain, the rigorous equation to synthesize the conventional active seismic data volume is not obeyed. Even though coherent noise events remain in the synthesized gathers, the migrated images do not suffer substantially from their presence. When multiple sources are captured in a transmission wavefield during the recording of a passive seismic experiment, the cross talk associated with correlation can lead to misinterpreting the shot gathers. The same is true when only a single subsurface source in captured in the transmission wavefields. This misinterpretation of the subsurface is minimized, though not mitigated completely, by migration.

Correlation removes a round trip from depth to the surface. Therefore, the direct arrival is the appropriate areal/temporal source function to images the the first-order multiples. Similarly, the first-order multiples correctly image the subsurface structure as the source energy for second-order multiples, and so on. If high-order multiples are recorded in the transmission wavefield, these events are important contributors to imaging reflectors at distant horizontal location from the subsurface source. If the sources are weak, and attenuation of the energy does not allow recording high-order multiple reflections from the reflectors at depth, synthesized shot gathers and direct migration images will not image structure at large horizontal distance from the subsurface source.

Rather than framing the correlation processing of transmission wavefields as an interferometric calculation of Green's functions, I believe that analyzing the use of transmission data as an imaging problem provides an intuitive and simple framework: direct migration. The experiments presented above show that processing transmission wavefields as an imaging problem within the framework of shot profile depth migration is a valid strategy akin to phase-encoded migration or a repeated datuming exercise. The images created by correlation followed by migration are very similar to those created by direct migration. Correlation and extrapolation are commutable operators, and we are free to choose a strategy most convenient for the task, or data, at hand. If passive data are available from an area with no knowledge of the subsurface velocity field, it may be necessary to correlate the records before migration for velocity analysis. Synthesized gathers may also provide a convenient domain to identify and remove coherent noise. However, even if the synthesized gathers seem to show little or no correct information about the subsurface, migrating the data may provide surprisingly good results.

If a velocity model is available, and migrating the data is a foregone decision, direct migration can save substantial computational cost. Because depth migration is an expensive operation, minimizing the volume of input data drastically effects the total cost to process a data set. The cost comparisons discussed above show that if the raw data can be windowed into fewer transmission wavefields than approximately 4 times the number of deployed receivers, it is always less expensive to migrate the raw data. Of course these ratios depend on the parameters explained above and the break-even point will decrease by increasing migration costs. These cost comparisons are true however for computing only one inverse Fourier transform and migration of the g^2 traces in the volume synthesized gathers. If intermediate volumes of gathers are viewed in the time domain and/or migrated, the number of time windows necessary to make direct migration more expensive increases almost linearly with the number of intermediate results produced.

For a passive experiment of the size of the Valhall array, the added computational burden of processing more time windows by direct migration than the calculated break-even point reaches an asymptotic limit of 1.3 times the cost of migrating synthesized gathers. Direct migration of data from an array with 4000 channels is always less expensive or equivalent to migrating synthesized gathers with these parameters. Direct migration of data from larger arrays is always the less expensive option.

With current computer technology, a single transmission wavefield collected with a large array takes approximately one hour to image with direct migration. Assuming that transmission wavefields 15 s long are windowed to processes independently, a single day of passive records results in 5760 time windows. Therefore, a cluster of 240 computers can process passive data in real time. The images presented above showed that direct migrations of transmission wavefields due to multiples sources are still interpretable, with often minor imperfections. Therefore, the number of windows that can be summed before direct migration can likely be increased by several times more to further reduce the computational cost. Summing four transmission wavefields before processing reduces the number of processors to $240/4 = 60$ to produce images with minimal cross talk artifacts associated with imaging data due to multiple sources.

The cost ratios presented above implicitly assume that all windowed transmission wavefields add equal contributions to the final image. For field data this is probably not the case, as much of the records may be useless background/instrument noise between episodes providing useful data. Summing data windows containing only background noise will decrease the signal to noise ratio of the product. However, it is unrealistic to view 5760 results per day for quality control purposes. Therefore, a summation is required either before or after imaging. Therefore, I view decreasing computation costs by some level of summation prior to processing to be a reasonable compromise. Hopefully, balancing the local rate of seismicity and the available personnel and computer resources can be facilitated with the analysis presented above.

Chapter 6

Depth imaging Valhall passive data

6.1 Summary

BP made available 43 hours of passive data from the Valhall Life of Field Seismic array of 2300 four component receiver stations over the reservoir in the Norwegian North Sea from February 2004 and January 2005. The quality and quantity of receivers associated with the acquisition, and the extent and uniqueness of processing algorithms I have applied to the data represent an industry first in the study of the passive seismic experiment. Unfortunately, the success of a passive seismic imaging experiment is predicated on the strength and availability of the local ambient noisefield. The records donated from the Valhall array do not contain sufficient body wave energy to image the Earth beneath the array. Instead, the records contain a strong coherent noise source that can lead to a false-positive interpretation of the success of the experiment.

6.2 Introduction

The main hindrance to the development of passive seismology since the original conjecture presented in Claerbout (1968) has been the availability of quality data sets. Cole (1995) presented work with a 13x13 station array in the foothills behind Stanford University. This effort did not manage to image subsurface structure due, I believe, to short time records, too few receiver stations, and a poor location for the experiment. Since then, large dense arrays have not been deployed to test the practicality of the theory. Artman (2006) showed marginally convincing results with a 2x2.25 m, 8x9 geophone array on low attenuation beach sand. Once again, too few receivers and a small amount of data lead to questionable results. That experiment was also marginalized by imperfect planning. Draganov et al. (2006) showed positive results synthesizing shot gathers with 17 traces in a linear array on land. The correlations show two likely reflections in a low signal to noise regime.

Figure 6.1 is a map provided by the Norwegian Oil Directorate which shows the production infrastructure between the coasts of Britain and Norway with the many developments in the region. It is 630 km from the Southern tip of the Norway shown and the city of Teeside in Britain. The Valhall development is roughly in the center of the line connecting those two geographic points. BP, with partners Hess, Shell, and Total, provided approximately 43 hours of continuous passive seismic records from the permanent sensor installation above the Valhall reservoir, which is located in the Norwegian controlled water of the North Sea.

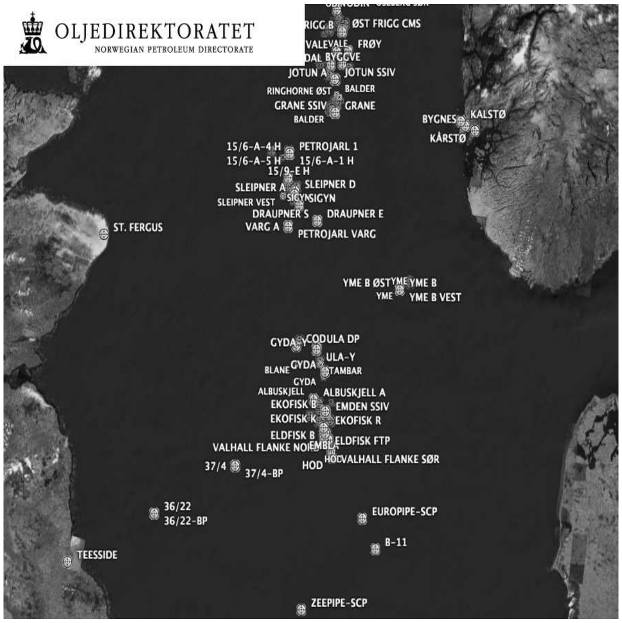


Figure 6.1: British, Norwegian, and Danish coasts surrounding the North Sea oil development area. The Valhall reservoir is in the South-central area of the map and only 20 km from U.K.-controlled waters to the West, and Danish waters to the South. DepthVal-areamap [NR]

The passive records are from 29 hours February 15, 2004, and 14 hours January 19,

2005. Of the former, some 7 hours was lost due to a failed disk-drive. The most likely teleseismic arrival recordable at the array, as determined by query of the National Earthquake Information Center event catalog[1], arrived approximately 5 hours after the last records I have. The data are sampled at 0.004 ms, and written as contiguous files between 12-20 s long. Chapter 3 presents the analysis of the raw data and synthesized shot gathers generated by correlating each trace with every other trace in the array.

The sub-sea array of ocean bottom cable (OBC) seismic sensors, Figure 6.2, at Valhall was installed in 2003 to facilitate time-lapse imaging of the subsurface for production monitoring. There are 13 lines of OBC 4-component (4C) sensors in the array. Of the 2300 4C stations, this work uses the hydrophone measurement from each location. This is mostly due to the fact that is currently unknown how to separate (with matrix rotations) the energy components on the arbitrarily orientated 3C geophones without any knowledge of the seismic source. The 9,200 sensors are permanently installed 1 m below the sea floor in about 70 m of water. The array forms a roughly 4x11 km rectangle. Nominal inline spacing of receiver stations is 50 m, while the crossline separation is 300 m. There are five surface structures (roughly over the center of the reservoir), and two wellhead 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.

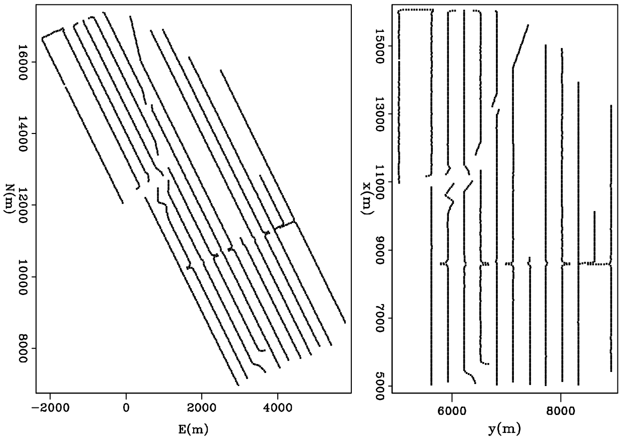


Figure 6.2: Valhall receiver array in North-East UTM coordinates and rotated field coordinates. Coordinates x, y are inline and crossline coordinates used in the images below. DepthVal-maps [NR]

---

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

Approximately half-way down the North axis of Figure 6.2, the cables on the Western side of the array are truncated where the production facilities are located to which the instruments are telemetered. Also noticeable is the deviation of the ends of cables on the North side of lines 5-7 (E=1000 m) 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 electric power with gas turbines, flow through the export lines, and the continual presence of multiple modes of transportation[2]. Valhall is located in a heavily developed area, within kilometers of developments in the British 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 withdrawal during production. In its first 15 years of production (to 1997), the sea-floor subsided 3.5 m at the crest of the reservoir structure (Gebara et al., 2000).

To date, there have been numerous towed marine surveys, multi-component OBC surveys, and extensive well log analysis of the reservoir. That extensive body of work has generated an accurate subsurface velocity model that was donated by BP along with the passive seismic data. To evaluate the results of my work with the passive data, BP also provided an image produced with active surface seismic data using the same velocity model. BP's image was generated with a shot profile depth migration algorithm that is very similar to my own migration implementation.

Synthesizing shot gathers from transmission wavefields produces a volume of data with the assumption that an impulsive virtual source for each gather is located at the receiver location used as a master trace for correlation. Therefore the virtual source axis and the receiver axis are identically sampled. The active surveys performed to target fine-scale reservoir changes over time use more than an order of magnitude more shot locations than the 2300 receivers in the array. To more accurately represent the best possible results I can achieve with purely passive data, the data from an active survey were subsampled before migration along the shot axis to provide a volume analogous to synthesized shot gathers

---

[2] www.npd.no/engelsk/

from passive data. The active data used to generate the control image for my experiments only used shot gathers collected from source locations directly over a receiver location.

Figure 6.3 shows two cuts through the image provided by BP produced with the active data. The near surface reflectors have very little structure. Note the migration diffraction noise on the crossline panel due to the large source and receiver spacing in that direction. The oil-filled chalk reservoir is below 2300 m and spans much of the image domain. The oil has a high gas-oil ratio which has lead to the accumulation of low-saturation gas in the central sedimentary column above the reservoir. This low-velocity zone above the reservoir leads to imaging problems that can be partially addressed with converted-wave seismic experiments, which was part of the justification for installing the OBC array.

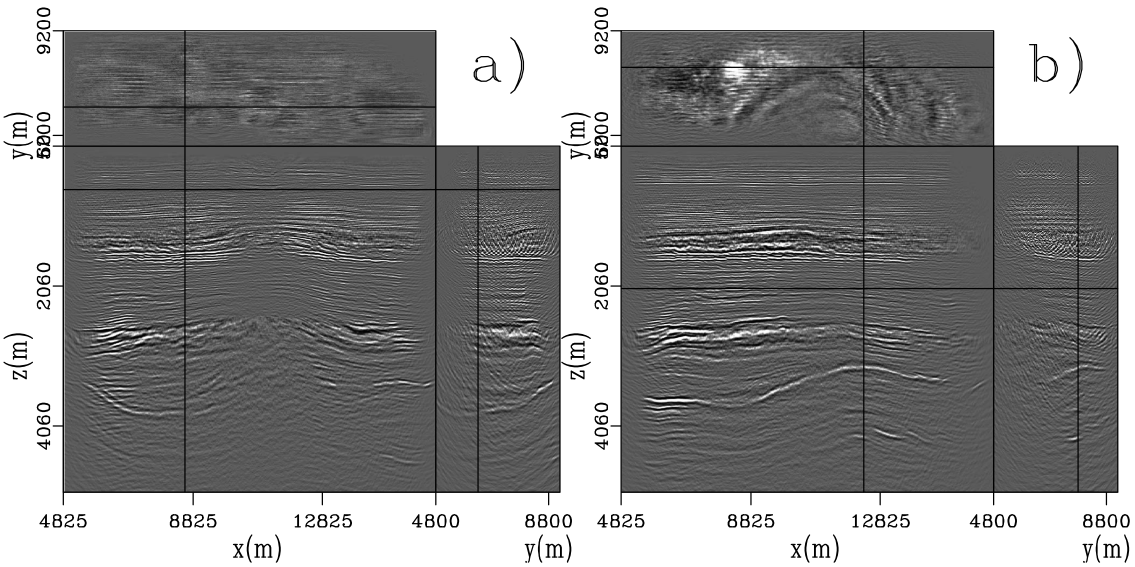


Figure 6.3: Control image produced with active seismic data using only shots located directly above a receiver location. DepthVal-Vactive [NR]

A section of the interval velocity model provided by BP is displayed in Figure 6.4. Merged with the velocity model is the image obtained by a zero-offset plane-wave migration using several horizontal plane waves as data. The first plane wave was placed at $t = 0.4$ s and the rest were spaced every 2 s after that. A perfect location to test passive seismic experiments would contain substantial geologic structure within a $v(z)$ medium. The velocity model at Valhall has a large slow column below the center of the array that causes the flat plane waves to image as a large anticline after depth $z = 1000$ m. Unfortunately, the subsurface structure and the velocity model are highly correlatable spatially.

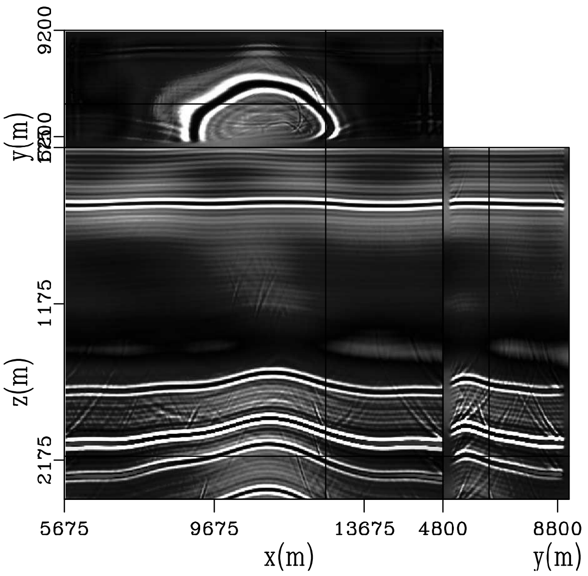


Figure 6.4: Interval velocity model merged with an image produced by migrating 5 horizontal plane waves. DepthVal-vplane [NR]

Figure 6.5 shows two control images produced to test the validity of images produced with the passive data. Both images were produced with the direct migration strategy. The left image was produced by migrating data replaced entirely with random noise traces. Notice the receiver pattern visible on the depth slices in the figure. The right image was produced by migrating real passive data with random noise (scaled between ±1 km) added to the header coordinates of the receiver locations. The appearance of coherent structure in either image would indicate that the convolution of randomness with the velocity model during migration could lead to believable events. Neither image indicates that the migration algorithm itself produces artifacts that mimic Earth structure. Therefore, I assume that the events in the various images presented below using volumes of real passive data are not simply artifacts of the processing algorithm (depth migration).

Figure 6.6 contains two time sections of data modeled to produce more control images. These data are modeled after likely coherent noise in the data. Only panel a is a series of horizontal plane waves. Analysis of the synthesized shot gathers in Chapters 3 and 4 showed that far-field noise from the gas flare at the Ardmore production facility is the dominant energy source in the passive records. Panel b is a series of repeated events modeled after

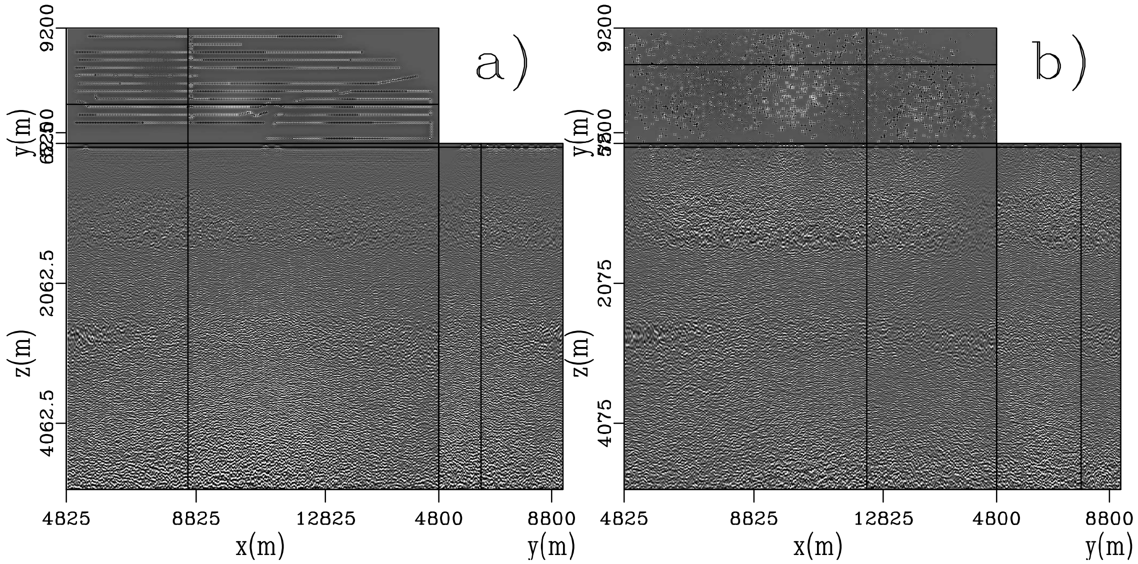


Figure 6.5: Control images produced with: a) random noise as data, b) 100 minutes of passive field data with random receiver locations. DepthVal-random [NR]

that noise using the kinematics of a direct arrival from 40 km to the SSW from the array as described in Chapter 3. The time axes of the two input data volumes, Figure 6.6, were sampled identically and the repeated events were produced by convolution with the same random series.

Figure 6.7 shows the images produced by migrating the data shown in the figure above. I produced panels a and b by direct migration using the respective data panels in Figure 6.6. Both modeled data sets produce image volumes that look realistic. Some events in these images even correspond to events in the active data. Such similarities will be explored below. Panel a, produced with horizontal plane waves, has lower frequency content than panel b. Panel a also shows less continuity in the crossline direction. Both images faithfully reproduce flat events in the shallow section and the anticline at depth. Panel b looks like a very realistic and interpretable image.

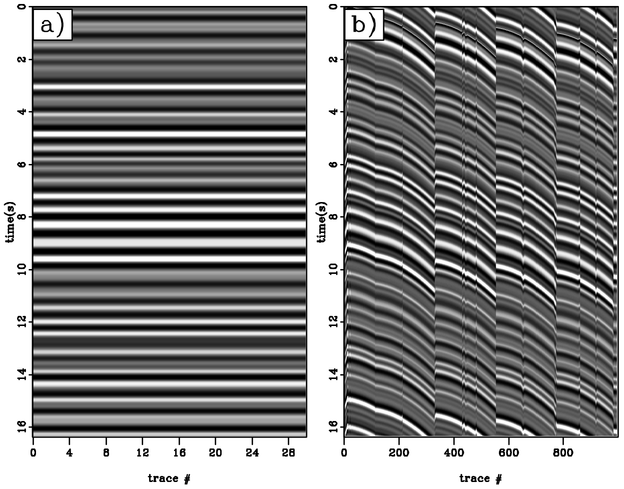


Figure 6.6: Modeled data to produce control images. a) Repeated horizontal plane waves. b) Repeated far-field noise arrival. DepthVal-manydat [NR]

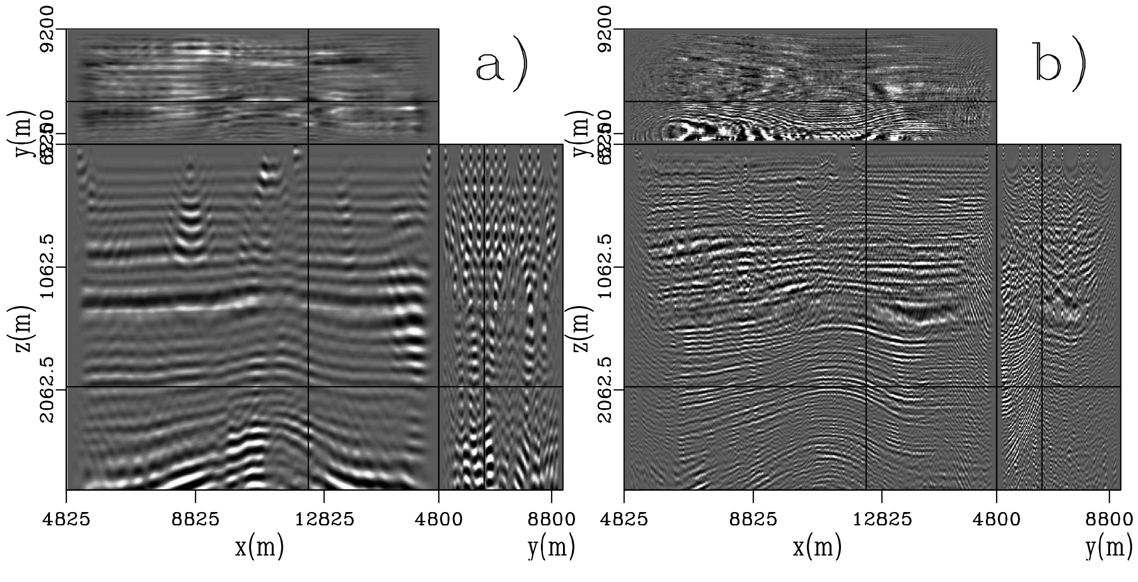


Figure 6.7: Images produced with likely coherent noise sources as data. Panel a was produced with repeated horizontal plane waves, Figure 6.6a. Panel b was produced with repeated far-field production noise, Figure 6.6b. DepthVal-manyimg [NR]

6.3 Passive imaging results

Following the theory from Wapenaar et al. (2004), the principal assumption behind processing passively recorded seismic records to produce results equivalent to those produced with the conventional active seismic experiment is that individual transmission wavefields should be processed before summing the results. Because windowing individual arrivals from the +36 hours of passive data is impractical (because of the large volume and the lack of identifiable events) it is impossible to identify and process individual transmission wavefields when only a single event is active. I make the assumption that processing data in 12 s sections is a reasonable approximation to the requirement, which corresponds to the approach adopted to synthesize shot gathers presented in Chapter 3. Thus images, i, produced by direct migration were extrapolated and imaged according to

$$i_z(\mathbf{x}) = \sum_{\xi} E^{+} T_{z-1}(\mathbf{x},\omega,\xi)\, [E^{-} T_{z-1}(\mathbf{x},\omega,\xi)]^{*} \,, \tag{6.1}$$

where ξ is the set of sequential 12 s windows of data and $E^{\pm}$ are the mixed domain the split-step Fourier plus interpolation (SSF-PI) (Kessinger, 1992) extrapolators used in the migration code. The development of direct migration and examples with synthetic data are presented in Chapter 5. I also migrated the synthesized shot gathers presented in Chapter 3 with the conventional shot profile migration algorithm. I used the same 12 s windowing strategy to parameterize individual wavefields for correlation processing. Two preprocessing steps were applied to all data before migration. Spectral whitening was performed to balance the frequency content of the traces by

$$\tilde{T}(\mathbf{x},\omega,\xi) = \frac{T(\mathbf{x},\omega,\xi)}{T(\mathbf{x},\omega,\xi)\, T^{*}(\mathbf{x},\omega,\xi)} \,. \tag{6.2}$$

Frequencies between 8-30 Hz were used for migration, with a cosine taper applied at the cutoff frequencies. In the frequency-wavenumber domain, I applied a tapered mute to remove energy slower than 1500 m/s.

The center row of Figure 6.8 shows sections through the 3D images produced by direct migration of 100 minutes of data. This image is one of the best results from those produced by direct migration. There is low frequency noise through the center of panels c and d that obscures possible events at these depth levels. When imaging with ambient noise, a noisy image is to be expected. The columns of the figure all show the same sections extracted from the three image volumes. The top row shows the image produced with active data. The bottom row is the control image shown in Figure 6.7a produced with the horizontal plane waves in Figure 6.6a. The criteria for evaluating success of imaging subsurface structure with passive data with this figure is that: Panels c,d share characteristics of panels a,b and are dissimilar to panels e,f.

The upper-most oval overlays in panels a, c and e highlight an event in the crossline panel that is continuous and interpretable in the upper two images. The same area in panel e is not continuous, but could possibly be present. The sparse crossline sampling of the array leads to migration diffraction noise throughout the image, though it is particularly bad at shallow depths. Within and directly above the upper ovals in panels b and d there are flat, broken events. Though not well imaged due to acquisition sampling, the events are interpretable with some confidence. The passive data image has lower frequency content. While the events in and around the upper oval in panel f could be interpreted across the gaps in the image, the panel lacks realism that is present in the image above it. The circles on the lower left corners of each panel highlight the flat the events away from the slow velocity column in the center of the domain. The plane waves in Figure 6.4 also exhibit this characteristic at the model edges which have little lateral velocity change.

Figure 6.9 shows the same direct migration result as above, but the bottom row of control images are those produced with the repeated kinematics of the far-field noise source in Figure 6.6b. The upper row of images is again the result of migrating the active surface seismic data. All three rows of images look remarkably similar. The central anticline in the active image is weak compared to the flanks due to the presence of the low velocity gaseous anomaly above that area. The passive image has events in this region with similar amplitude compared to the flanks of the anticline. The ovals on the crossline frames of panels a, c, and e contain an almost identical event. This event can not be interpreted unequivocally in the passive image given the presence of an exact counterpart in panel e.

The event in the upper oval of Figure 6.9d is not as well represented in the control image

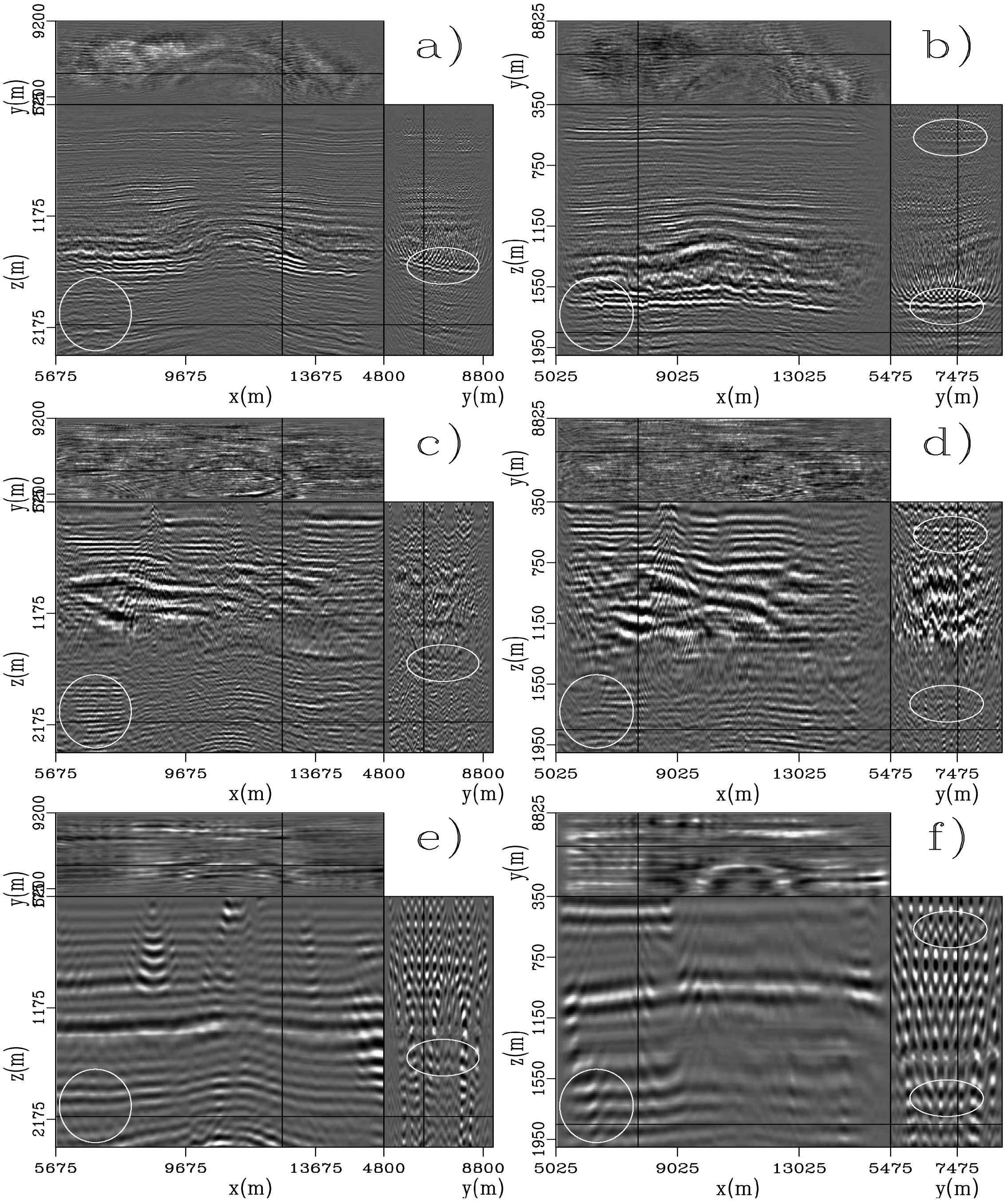


Figure 6.8: a,b) Image produced with active data. b,c) Image produced by direct migration of 100 minutes of passive data. e,f) Image produced with modeled data in Figure 6.6a. DepthVal-first200.2 [NR]

below it. The event in the center of the upper oval in panel d and the shallowest of the two events in the corresponding oval in panel b are both centered at $z = 537$ m. The same events in the inline section of b show a brighter amplitude than surrounding events. The same claim is marginally supported in the passive image. However, these events are interpretable with only marginal confidence in panel d. The structure exhibited in the deepest levels of the passive image shows some promise. The circles in the lower left corners of the active image and passive image highlight events with no dip. The control images in the bottom row have significant dip in these areas.

Figure 6.10 shows two versions of a synthesized gather extracted from a volume created with 9 hours of passive data recordings. Only a 2 s and a few traces were extracted for the panels. The gathers are the summation of 12 s correlation volumes. Panel a is the product of the correlation and spectral whitening described in Chapter 3. Panel b is the same gather, with a short time mute applied to mask the direct arrival from the gas flare 40 km away from the array. In Chapter 4, this volume of synthesized gathers was identified as containing the least contamination due to ringing correlation side lobes. Therefore, the mute applied before migration removed the far-field noise in this volume of data most successfully. For this reason, I assume that this data volume has the most promise of revealing subsurface structure that is not otherwise masked by the dominant seismic energy in the records. These two volumes of synthesized shot gathers are used to produce the images in the next figure.

Figure 6.11 shows images produced by migrating the synthesized shot gathers in Figure 6.10. Panel a was produced by migrating the synthesized gathers as produced. Panel b is the result of migrating the muted gathers. Panel a does not exhibit the detail in the images in Figure 6.9. A few flat events are visible on the left side of the inline section, but the section is mostly devoid of realistic detail. Panel b does not contain the events seen in panel a. Therefore, these events must be artifacts of the far-field direct arrival. Most of the image in panel b does not contain interpretable events, and resembles the random control images presented in Figure 6.5. The bottom of panel b has a few events that correspond to the anticline structure in the active image. However, it is unlikely that deep events are well imaged without correspondingly good imaging of shallow events close to the array.

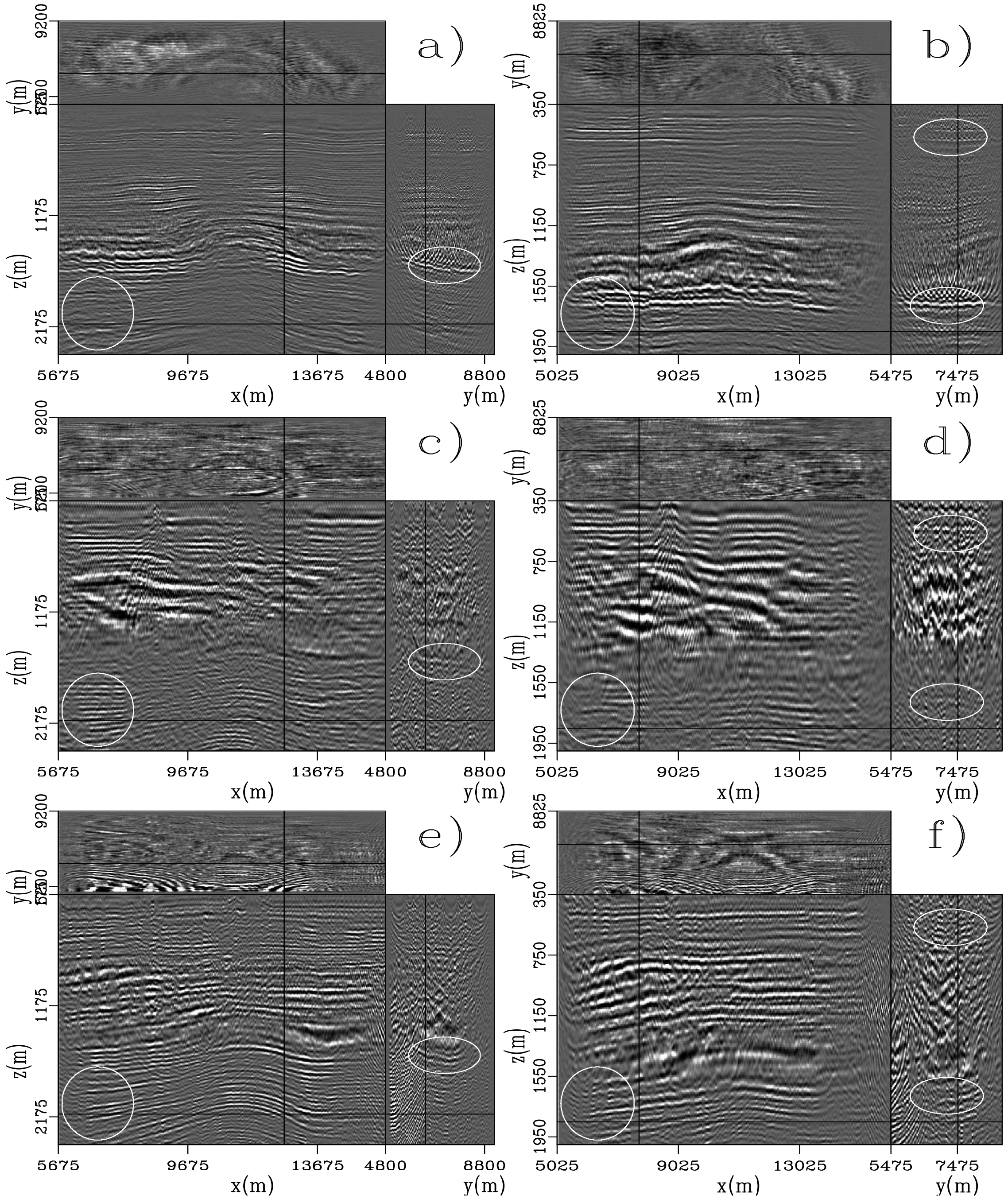


Figure 6.9: a,b) Image produced with active data. b,c) Image produced by direct migration of 100 minutes of passive data. e,f) Image produced with modeled data in Figure 6.6b. DepthVal-first200 [NR]

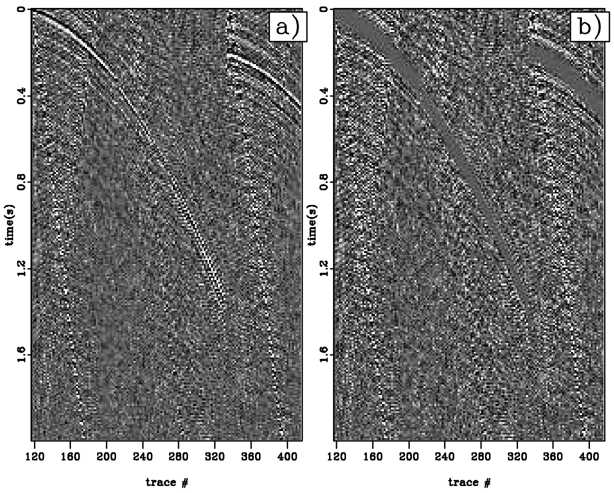


Figure 6.10: a) Synthesized shot gather using 9 hours of passive data. b) Direct arrival from production noise windowed away. DepthVal-noeventdat [NR]

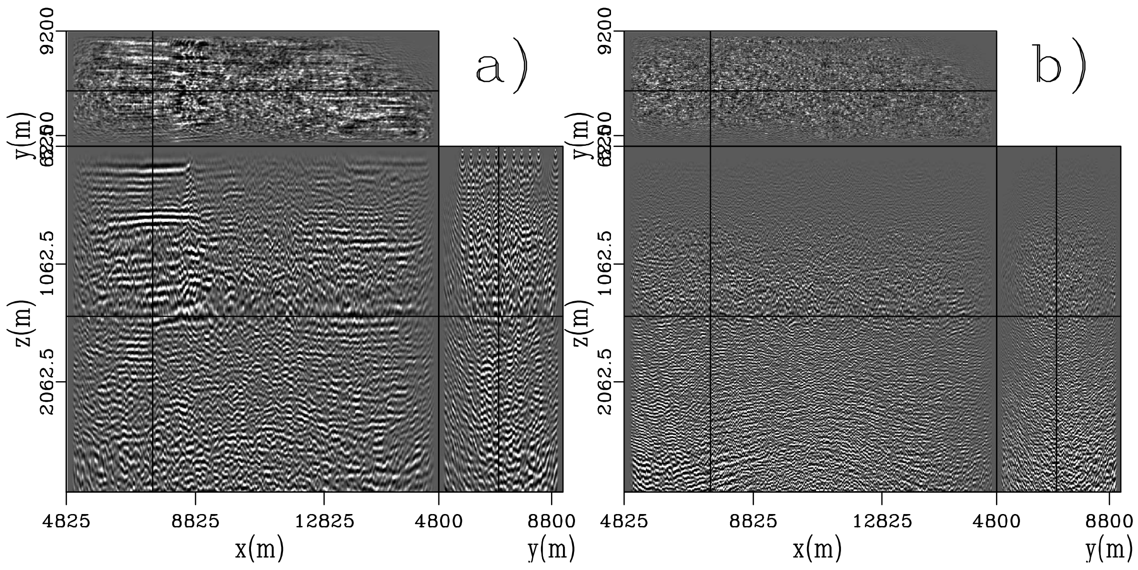


Figure 6.11: a) Migration of synthesized gathers produced with 9 hours of passive data. b) Migration of the same gathers after muting a short time window around the direct arrival from the Ardmore production facility. DepthVal-noevent [NR]

6.4 Conclusions

Direct migration of the Valhall passive data generates deceptively interpretable images. However, imaging with modeled data designed to mimic the dominant sonic energy in the records identified in Chapter 3 produces results with all of the main features found in results using field data. In some sections of the images generated with field data, direct migration results imaged events with dips corresponding to the active image and contrary to the image produced with modeled data. However, this is weak evidence that the direct migration results are reliable images of the subsurface.

The ambient noisefield in the region of the Valhall subsea seismic array is dominated by an energy source that does not contain travel paths through the subsurface. The energy arrives as such a shallow angle that it does not investigate the subsurface within the aperture of the array. The energy is a coherent noise source that overwhelms the passive seismic experiment. The fact that the energy is continuous in time and unidentifiable in raw data, makes it especially problematic for the direct migration strategy. The modeled data in Figure 6.6b is an accurate visualization of the field data if the source wavelet was more compact. Because the event is present at all times, events in the down-going source wavefield and up-coming receiver wavefield are collocated at every depth in the image space during migration. Because there is almost always correlatable energy in the two wavefields, events are manufactured at nearly every depth level. The slow velocity column in the center of the subsurface domain delays the energy within it to produce structure in the direct migration images in the same manner seen in Figure 6.4.

After correlating the raw data, the coherent noise energy is localized near time zero in the synthesized shot gathers. Therefore, migrating the correlations produces a very different imaging result than direct migration. Only a few events are present in these images at shallow depths. However, the analysis of the synthesized gathers in Chapters 3 and 4 showed that there are no coherent geologic events in the synthesized shot gathers at early time. Because the far-field noise is a coherent event after correlation, it is possible to remove most of the energy from the synthesized gathers. Migrating the data after preprocessing with a mute function designed to remove the noise event results in an image that looks like the migration of random noise. The events present in the shallow section of the migration of the synthesized gathers are removed by removing the coherent noise source in the data.

In Chapter 5, synthetic transmission wavefields were processed with both direct migration and by migrating synthesized shot gathers. The results produced with these two work flows were almost indistinguishable. The results produced with the two methods using field data were drastically different. The presence of the far-field noise in the field data is the cause of this difference. The noise is continuously present in the raw data used for direct migration, but temporally localized in the correlations. Neither processing strategy with the Valhall passive seismic field data showed capability of imaging subsurface structure. The ambient noise field in the vicinity of the array does not contain body waves carrying subsurface information strong enough to overcome the predominant noise source in the area.

Chapter 7

Image space multiple prediction

7.1 summary

A very important aspect of removing multiples from seismic data is accurate prediction of their kinematics. I cast the multiple prediction problem as an operation in the image space parallel to the conventional surface-related multiple prediction methodology. Though developed in the image domain, the technique shares the data-driven strengths of data-domain surface-related multiple elimination (SRME) by being independent of the Earth (velocity) model. Also, the data are used to predict the kinematics of the multiples exactly so that a Radon transform need not be designed to separate the two types of events. The cost of the prediction is approximately the same as that of data-space methods, though it can be computed during the course of migration. The additional cost is not significant compared to that incurred by shot-profile migration.

Image-space multiple predictions are generated by convolving the traces in each shot-gather at every depth level during the course of a shot-profile migration. The prediction in the image domain is equivalent to that produced by migrating the data-space convolutional prediction. Adaptive subtraction of the prediction from the image is required. Subtraction in the image domain, however, provides the advantages of focused energy in a smaller domain since extrapolation removes some of the imperfections of the input data.

7.2 Introduction

Removing multiples from seismic data is often an imperative to producing interpretable images of the subsurface. Multiple attenuation has a rich history in the geophysical literature ranging from methods that predict the multiples via convolution of the data (Anstey and Newman, 1966) to methods that use a differential characteristic between primaries and multiples as a discriminator for separating the two types of events (Hampson, 1986; Weglein, 1999).

Kinematic prediction of multiple reflections by convolution leads to amplitude and bandwidth inconsistencies. Therefore, direct subtraction of such predictions is not possible. Tsai (1985) suggests modeling the waveform of the multiples to subtract events at times calculated by convolution from the data. The surface-related multiple elimination (SRME) method (Verschuur et al., 1992; Berkhout and Verschuur, 1997) convolves traces within shot-gathers to predict multiples (surface-related multiple prediction, SRMP) followed by an iterative subtraction scheme to eliminate them from the data. Alternatively, after predicting multiples, via convolution or filter-based methods, Guitton (2005) uses a pattern-based subtraction technique that resembles the match filter application described in Biersteker (2001).

While the above references all operate in the data domain, authors have also suggested removing multiples in the image space, after migration (Sava and Guitton, 2003, 2005; Alvarez et al., 2004). These authors capitalize on kinematic differentiation of the primaries and multiples and separate the two via a Radon transform tailored to residual moveout or the analytic expression of over-migrated multiples, respectively. There are several motivators for attacking multiples in the image space. First, the image space is usually much smaller than the data space. Second, given a reasonably accurate velocity model, the kinematics of the image domain are simplified. Appropriately migrated primary events have little to no residual moveout, and multiples, migrated with velocities that are too fast, have predictable moveout in both angle- and offset-domain common-image gathers. Alternatively, one could migrate the data-domain SRMP volume to produce an image-space prediction of the multiples.

I extend the SRMP approach to the image domain through the commutability of wavefield extrapolation and convolution to produce a multiple prediction in the image domain without needing to migrate two data volumes. My approach is directly analogous to SRMP, though the prediction is calculated during the course of a shot-profile migration. The added cost of image-space surface-related multiple prediction (IS-SRMP) is only a second imaging condition. Because extrapolation dominates the cost of shot-profile migration, IS-SRMP does not significantly increase its cost.

Muijs (2005) develops an analogous strategy to IS-SRMP. That work requires substantial preprocessing, as its input is a previous migration result. Furthermore, the imaging procedure uses a multi-dimensional Fourier domain (transformed over source position) deconvolutional imaging condition that is iteratively performed over frequency and position that would seem very computationally burdensome. Generating a multiple model in the image domain during survey-sinking migrations was also discussed in Malcolm et al. (2006). Because the up-coming and down-going energy are not explicitly separate, as in shot-profile algorithms, the authors require extra work that is not necessary for this method. That extra work effectively doubles the cost and memory requirements of the (DSR) migration. In contrast to our method, it implicitly acts as a layer-stripping methodology that is also capable of predicting intrabed multiples.

The multiple prediction produced in the image space with this technique is mathematically equivalent to migrating the multiple prediction produced by SRMP convolutions. The predicted multiples are then removed from the data via adaptive subtraction or pattern-matching techniques. The need for adaptive subtraction is two-fold. First, convolution squares the wavelet in the data which reduces the effective bandwidth, squares the amplitude, and destroys the polarity of the output. Therefore, even if a true-amplitude migration algorithm is available, the multiple prediction will not share the amplitude characteristics of the multiples in the migrated image. Second, adaptive subtraction can account for imperfections including kinematic errors and the presence of residual primaries or higher-order multiples in the prediction.

As wave-equation migration becomes more of a commodity, iterative migration as part of estimating the Earth's velocity has become more common. The presence of multiples in the data makes velocity analysis more difficult since events with conflicting velocity properties are present at the same time or depth. For this reason, there will always be a need to remove multiples in the data domain. However, if a shot-profile migration strategy is planned, it will be shown that simultaneously producing a multiple prediction adds no significant cost, and does not require the presence of multiples in the data being migrated. Therefore, the prediction can be generated even after perfect elimination as a comparative volume for interpreters to use as an example of events that are not geologic.

In this chapter, I develop the image-space surface-related multiple prediction (IS-SRMP) technique by combining the SRMP approach with a wave-equation shot-profile depth migration algorithm. Guitton (2005) shows that pattern-based and adaptive subtraction of multiple models work better when higher dimensionalities can be exploited by the subtraction algorithm. The various imaging conditions I present can all produce the image as a function of subsurface offset (and from there reflection angle) to facilitate better subtraction. A simple flat-layer synthetic and the Sigsbee2B synthetic data set are used to show the efficacy of the prediction and its subtraction from the migrated image. Multiple prediction and subtraction is also presented with a Gulf of Mexico data set acquired by Western-Geco in the Mississippi Canyon lease area.

7.3 Theory

The image space is the output of migration, which I produce with a shot-profile depth migration algorithm. Shot-profile wave-equation depth migration (Claerbout, 1971) is the cascade of two component operations: extrapolation and imaging. Extrapolation is carried out with an anticausal wave-equation operator applied to the up-going wavefield, U, and a causal operator applied to the down-going wavefield, D. U is a shot record with traces placed along a wavefield axis $\mathbf{x}$. D is a zero valued wavefield, also defined along the $\mathbf{x}$-axis, where a source function is placed at the location of the shot being migrated, $\mathbf{x}_s$. The wavefields are recursively extrapolated to all depths z using one-way mixed-domain solutions to the wave-equation

$$U_{z+1}(\mathbf{x};\mathbf{x}_s,\omega) = E^{+}(\mathbf{x},\omega)U_z(\mathbf{x};\mathbf{x}_s,\omega) \tag{7.1}$$

and

$$D_{z+1}(\mathbf{x};\mathbf{x}_s,\omega) = E^{-}(\mathbf{x},\omega)D_z(\mathbf{x};\mathbf{x}_s,\omega)\,. \tag{7.2}$$

In this work, the form of the extrapolator E used to propagate the wavefields is the split-step Fourier plus interpolation (SSF-PI) (Kessinger, 1992) with multiple reference velocities, though the degree of complexity of the operator does not change the discussion herein. The importance of these equations is that the operator that extrapolates a wavefield from one depth level to the next is commutable with convolution. Using the simple, though less accurate, bulk phase-shift operator of the form $\mathrm{e}^{i\omega s\Delta z}$ for E, makes this commutativity more obvious. The extrapolation can also be viewed as repeatedly redatuming the shot gathers into the image space.

The correlation based multi-offset imaging condition for shot-profile migration is (Rickett and Sava, 2002)

$$i_z(\mathbf{x},\mathbf{h}) = \sum_{\mathbf{x}_s}\sum_{\omega} U_z(\mathbf{x}+\mathbf{h};\mathbf{x}_s,\omega)D_z^{*}(\mathbf{x}-\mathbf{h};\mathbf{x}_s,\omega) \tag{7.3}$$

where the $*$ represents conjugation, and $\mathbf{h}$ is subsurface offset. Extraction of the zero lag of the correlations, by summation over ω, combines the energy in the two wavefields that is collocated at each depth level. Overlapping acquisition patches from the individual shots are stacked by the sum over $\mathbf{x}_s$.

Acknowledging the approximation of convolving raw data traces in lieu of only primaries, the prediction of multiples in the data space (SRMP) can be written in the Fourier domain (Berkhout and Verschuur, equation 13f, 1997)

$$M(\mathbf{x}_g;\mathbf{x}_s,\omega) = \sum_{\mathbf{x}_a} R(\mathbf{x}_g;\mathbf{x}_a,\omega)R(\mathbf{x}_a;\mathbf{x}_s,\omega)\,, \tag{7.4}$$

where R is the data-space volume of shot-gathers defined at geophone and source locations on the acquisition surface. Equation 7.4 is a trace-by-trace operation to produce the multiple prediction with any geophone-source, $(\mathbf{x}_g, \mathbf{x}_s)$, combination by convolving each trace of every shot gather with all the others followed by summing over the convolution index $\mathbf{x}_a$. Note however, the similarity of the SRMP equation to the imaging condition of shot-profile migration, equation 7.3.

Wave-equation extrapolation is performed on wavefields where data and source-functions are used as initial conditions to propagate energy into the subsurface. To begin, traces at locations $\mathbf{x}_g$ are inserted into a zero-valued wavefield defined along the axis $\mathbf{x}$. Although data-space SRMP is a trace-by-trace operation, equation 7.4 can be redefined in terms of the wavefield $U(\mathbf{x}, \mathbf{x}_s)$. With null-traces at locations $\mathbf{x}$ where no data were collected, a multiple prediction can also be written

$$M_{z=0}(\mathbf{x}; \mathbf{x}_s, \omega) = \sum_{\mathbf{x}_a} U_{z=0}(\mathbf{x}; \mathbf{x}_a, \omega) U_{z=0}(\mathbf{x}_a; \mathbf{x}_s, \omega) \,, \tag{7.5}$$

where the resultant volume has been regularized along $\mathbf{x}$ by adding zero-traces, and I have added the specification that the operation is being performed at the recording surface $z = 0$.

Using the principle of reciprocity between the receiver and source locations (first and second arguments of the wavefields respectively), the multiple prediction becomes

$$M_{z=0}(\mathbf{x}; \mathbf{x}_s, \omega) = \sum_{\mathbf{x}_a} U_{z=0}(\mathbf{x}; \mathbf{x}_a, \omega) U_{z=0}(\mathbf{x}_s; \mathbf{x}_a, \omega) \,. \tag{7.6}$$

Here, the subscript s on the RHS represents any different receiver location (since it is the first argument of the wavefield), and the dummy index $\mathbf{x}_a$ is recognized as a summation over source location. Therefore, using arbitrary index subscripts c, d and restoring the significance of source location to subscript s

$$M_{z=0}(\mathbf{x}_c; \mathbf{x}_d, \omega) = \sum_{\mathbf{x}_s} U_{z=0}(\mathbf{x}_c; \mathbf{x}_s, \omega) U_{z=0}(\mathbf{x}_d; \mathbf{x}_s, \omega) \,. \tag{7.7}$$

Finally, I define the dummy indices c, d in terms of physically significant variables location and half-offset, $\mathbf{x}_c = \mathbf{x} + \mathbf{h}$ and $\mathbf{x}_d = \mathbf{x} - \mathbf{h}$, such that

$$M_{z=0}(\mathbf{x}, \mathbf{h}, \omega) = \sum_{\mathbf{x}_s} U_{z=0}(\mathbf{x} + \mathbf{h}; \mathbf{x}_s, \omega) U_{z=0}(\mathbf{x} - \mathbf{h}; \mathbf{x}_s, \omega) \; . \tag{7.8}$$

Thus reconfigured, equation 7.8 is now of parallel construction to the shot-profile imaging condition, equation 7.3, lacking only the summation over frequency.

Extrapolation by $E(\mathbf{x}, \omega)$, in equation 7.1, simply redatums the shot-gather U, after which the conventional SRMP volume is calculated by convolution. Image-space SRMP (IS-SRMP) is the application of a second imaging condition evaluated at each subsurface depth level during the migration that images only multiples. It is the chain of multiple prediction (convolution) and zero-time extraction (summation over frequency). The image-space multiple prediction, as a function of sub-surface offset, is therefore

$$m_z(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{x}_s} \sum_{\omega} U_z(\mathbf{x} + \mathbf{h}; \mathbf{x}_s, \omega) U_z(\mathbf{x} - \mathbf{h}; \mathbf{x}_s, \omega) \; . \tag{7.9}$$

There are two important ramifications associated with the equation for predicting multiples with the imaging condition above. The first is that this operation is intrinsically a shot-domain manipulation of the data. After sorting to midpoint-offset coordinates, the source and receiver coordinates are mixed in such a way as to make IS-SRMP more difficult for survey-sinking style migration algorithms (Malcolm et al., 2006). Second, because reciprocity was invoked to derive equation 7.9, off-end (marine) acquisition geometries will need to have split-spread gathers generated via reciprocity. The split-spread gathers will include the ray-paths from multiples that emerge in front of the receiver spread (boat) which need to be included in the shot-gathers to predict all possible multiple events.

Further understanding of the IS-SRMP imaging condition for those familiar with shot-profile migration algorithms can be elicited by defining the down-going wavefield in equation 2 as $D \equiv U$. Therefore, equations 1 - 3 become

$$U_{z+1}(\mathbf{x};\mathbf{x}_s,t) \quad = \quad E^{-}(\mathbf{x},t)U_z(\mathbf{x};\mathbf{x}_s,t) \tag{7.10}$$

and

$$\hat{U}_{z+1}(\mathbf{x};\mathbf{x}_s,t) \quad = \quad E^{+}(\mathbf{x},t)\hat{U}_z(\mathbf{x};\mathbf{x}_s,t) \tag{7.11}$$

with the the imaging condition

$$i_z(\mathbf{x},\mathbf{h}) = \sum_{\mathbf{x}_s}\sum_{\omega} U_z(\mathbf{x}+\mathbf{h};\mathbf{x}_s,\omega)\hat{U}_z^*(\mathbf{x}-\mathbf{h};\mathbf{x}_s,\omega)\,, \tag{7.12}$$

where ˆ denotes that after extrapolation in different directions, the wavefields are no longer identical. Because the conjugation of D in the imaging condition of equation 7.3 can be commuted with the causality of the extrapolation operator, it is not necessary to extrapolate U in two different directions. Instead, the second extrapolation above can be ignored and the imaging condition becomes convolutional rather than correlational. In this case, equation 7.12 is exactly the multiple prediction, equation 7.9.

Cast in this manner, the migration shows similarity with reverse-time migration (Baysal et al., 1983) and using multiples to migrate primaries (Shan and Guitton, 2004). The difference is that the data used as source function is not first time-reversed. IS-SRMP uses the shot gathers as both an areal source function and data to image multiples. Conversely, time-reversing the data will use the primaries as areal source functions (delayed from time zero) to image the subsurface with the multiple reflections. Using both an impulsive source function at $\mathbf{x}_s$ and time-reversed primaries as a source function would image Earth structure with the primaries and the multiples, but still include the multiples in the image.

7.3.1 Analytic example

In 1D, let a trace be represented in the Fourier domain by the expression

$$R(\omega) = \mathrm{e}^{-i\phi_w} + \mathrm{e}^{-i\phi_e} - \mathrm{e}^{-i2\phi_w} - 2\mathrm{e}^{-i(\phi_w+\phi_e)} - \mathrm{e}^{-i2\phi_e} = P + M_1. \tag{7.13}$$

The trace has primary reflections, P, at phase delays, ϕ, representing a water-bottom and a subsurface event (subscripts w, e). Also included are first order multiples, M_1, which are the water-bottom multiple at $2\phi_w$, two peg-leg multiples at $\phi_w + \phi_e$, and the event multiple at $2\phi_e$. SRMP dictates the autoconvolution of P to derive M_1, which is clearly true. Including the events M_1 in the autoconvolution will predict the higher-order multiples as well.

Extrapolating trace R to a deeper depth applies a common phase shift, say $\mathrm{e}^{-i\phi_z}$, to all terms in equation 7.13. The trace then becomes

$$R_z(\omega) = e^{-i\phi_z}(\mathrm{e}^{-i\phi_w} + \mathrm{e}^{-i\phi_e} + \mathrm{e}^{-i2\phi_w} + 2\mathrm{e}^{-i(\phi_w+\phi_e)} + \mathrm{e}^{-i2\phi_e}) = e^{-i\phi_z} P + e^{-i\phi_z} M_1. \tag{7.14}$$

This equation shows that the extrapolation of data without multiple subtraction produces the superposition of the redatumed primaries and the redatumed multiples. The extraction of the zero-time lag in the imaging condition of migration states that energy in the wavefield should be mapped into the image domain only when the extrapolation phasor ϕ_z, is equal to the time delay of the event in the data. Thus the water bottom primary is imaged when $\phi_z = \phi_w$, the water-bottom multiple is imaged when $\phi_z = 2\phi_w$, etc. Whether the data R is first separated into its constituent parts, P and M_i, or not, it can be seen where in the image domain the various events in the above example will be placed. However, by first squaring the trace, implementing data-space SRMP, the water-bottom primary is mapped into the image domain when $\phi_z = 2\phi_w$. This is the same phasor that maps the recorded water-bottom multiple into the image domain when migrating data contaminated with multiples.

Figure 7.1 is a cartoon depicting the generation of a 1D image-space surface-related multiple prediction without first convolving the gathers. The scenario drawn is for the simple case of a water-bottom reflector and its multiple. The trace U denotes a zero-offset recording from a shot gather, and D is the modeled source function used for shot-profile migration. The trace I represents the result of the conventional imaging condition, while M is the multiple prediction in the image space generated by autoconvolution. The superscript $*$ denotes conjugation. Both imaging conditions extract energy only at $t = 0$.

The first panel, $z = 0$, shows the initial conditions of U and D and the fact that both $I|_{t=0}$ and $M|_{t=0}$ are devoid of events. Since correlation subtracts the time to energy in the source trace D, currently zero, from the receiver trace, U, the image trace I is the same as U. The time for events on trace UU are doubled from the initial condition, which results in the multiple on trace U mapping below the time interval in the cartoon.

The second panel depicts the situation when the wavefields have been extrapolated to the depth of the water column. Now energy in D and U is collocated and the primary is imaged during correlation, UD^*. The multiple prediction is still zero valued. The last panel depicts the wavefields and imaging results after the wavefields have been propagated to twice the depth of the water column. Now the source energy is collocated with the multiple, and the correlation results in a negative polarity event. Simultaneously, the primary is now at $t = 0$ which maps energy into the IS-SRMP volume, M. However, even though the kinematics of the multiple in UU are correct, the event has the wrong polarity. Also notice that the multiple is now at the original time of the primary, which sets up prediction of second-order multiples from the convolution of the first-order multiples. Importantly, even if the velocity used to calculate the extrapolation phasors is incorrect, both I and M will share the same error. Also, U consisting of only primaries will correctly predict the location of multiples even if no multiples exist in the data.

7.4 Examples

7.4.1 Simple synthetic

The left panel of Figure 7.2 is a synthetic shot gather with two reflections and three multiples. The velocity to the first flat layer at depth $z = 400$ m is 1500 m/s. The velocity is 2500

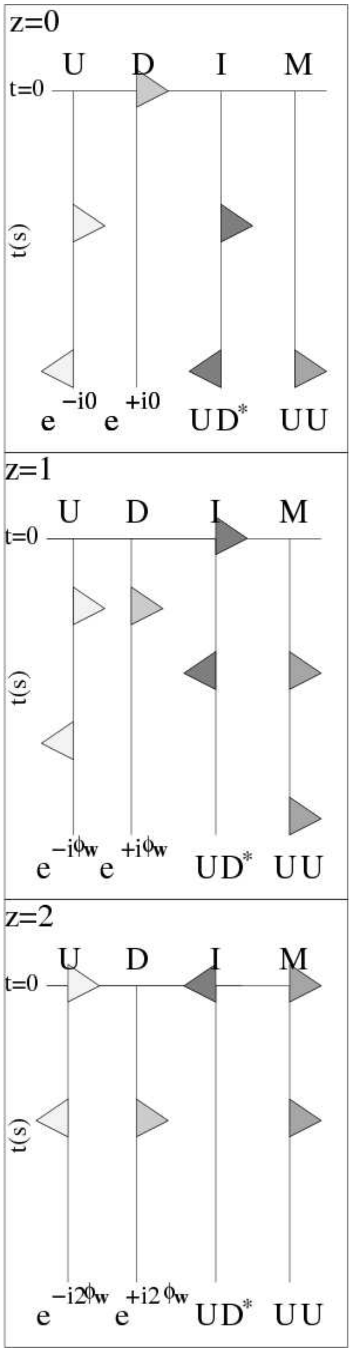


Figure 7.1: One-dimensional example of IS-SRMP during shot-profile migration at three extrapolation levels, $z = 0, 1, 2$. Trace U represents data with a primary and multiple event. Trace D is the modeled impulsive source wavefield. Trace I is the conventional image (UD^*). Trace M is the multiple prediction (UU). Final image volumes for each depth level are produced by extracting the values of I and M at $t = 0$. DepthMult-sketch [NR]

m/s to the second flat layer at depth $z = 1200$ m. The traveltimes were computed analytically and convolved with a wavelet. The white events are primaries, and the multiples have opposite polarity. The three multiple events are the simple multiples to both events, and the asymmetric peg-leg. The intrabed multiples between the two layers were not included since they are not predicted by SRMP. Receiver and source spacing was 20 m. The right panel is a subsurface-offset common-image gather produced by shot-profile migration of 350 shot-gathers identical to the data shown. Because the layers are flat, every location **x**, away from the edges, is the same. Only $1/6^{th}$ of the surface offsets are required to capture the moveout of the multiples in the subsurface-offset common image gathers after migration.

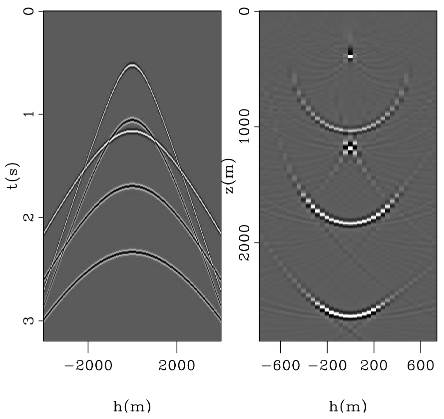
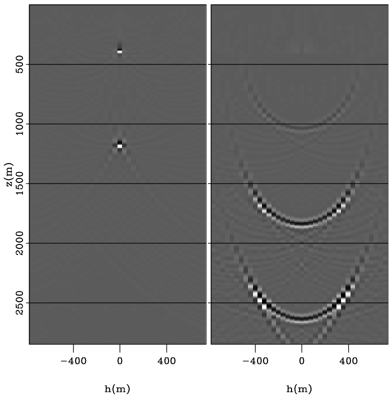


Figure 7.2: Left: Synthetic data with two primaries and three multiples. White events are primaries. Two simple and a peg-leg multiple were modeled. Right: Subsurface-offset common-image gather produced by shot-profile migration. DepthMult-syndat [NR]

The two panels in Figure 7.3 are offset-domain common-image gathers produced by migrating only the primaries in the data above. The first panel shows energy tightly focused at zero offset for the two reflectors. The second panel is the multiple prediction. Three clear events are imaged with concave-up moveout. The middle event is the peg-leg multiple and the others are the simple multiples to the two reflectors. At zero-offset, the events are analytically calculated to arrive at depths $z = 1066, 1866$ and 2666 m, respectively.

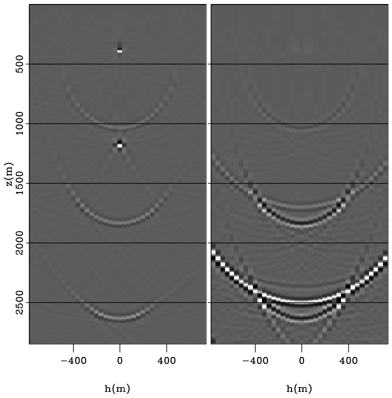
The two panels in Figure 7.4 are offset-domain common-image gathers produced by migrating all the events in Figure 7.2. Autoconvolution predicts multiples from primaries. Convolution of primaries with the first-order multiples produces second-order multiples. Autoconvolution of first-order multiples predicts third-order multiples. The left panel now has the two focused primary events and the three multiples predicted by IS-SRMP in Figure 7.3. These multiples are again predicted in the IS-SRMP gather on the right of this figure.

Figure 7.3: Offset-domain common-image gathers from migrating only the primaries in Figure 2. Left: Conventional imaging condition. Right: IS-SRMP. DepthMult-synoffprim [NR]



However, higher order multiples are now included as well. The broad event at $z = 1733$ m is the multiple that makes three trips through the water column. The faint event at $z = 2400$ m is the double water-column multiple. The last new event is at $z = 2533$ m. This multiple has two trips in the water column and one trip from the surface to the second reflector.

Figure 7.4: Offset-domain common-image gathers from migrating all events in Figure 2. Left: Conventional imaging condition. Right: IS-SRMP with higher order multiples. DepthMult-synoffni [NR]



7.4.2 Sigsbee2B synthetic

The Sigsbee2B data-set was designed to model strong surface-related multiples from an offshore acquisition. Two datasets were generated with a 2D finite difference algorithm: One with the perfectly reflecting free surface, and one without[1]. Therefore, the direct subtraction of the two data volumes yields a nearly perfect multiple model without the need for SRMP.

---

[1] http://www.delphi.tudelft.nl/SMAART/S2Breadme.htm

There are slight differences in the source and receiver ghost effects between the two data sets, so their subtraction retains faint residuals of primary energy. Though the data were modeled with an off-end acquisition strategy, split-spread gathers were computed via reciprocity for all of the examples below. All of the images herein were produced with four reference velocities in a SSF-PI shot-profile migration code.

Figure 7.5 shows zero subsurface-offset images of the Sigsbee2B data set. The top panel, generated with the conventional imaging condition, contains primaries and multiples. The bottom panel has multiples and some migration artifacts. The worst artifacts in the prediction are above the first multiple, especially near steep salt flanks. This noise can be easily muted before subtraction. The IS-SRMP image was produced using the data with primaries only. The data was modeled such that the details of the bottom left corner have simple kinematic differences between primaries and multiples: Events dipping up-right are multiples, and those dipping down-right are primaries.

Figure 7.6 shows the bottom third of the image produced with the Sigsbee2B data sets. The top image used the data modeled without the reflecting free surface and contains only primaries. The second image migrated the data with the free-surface and contains multiples. The third panel is the image produced by migrating the difference between the two modeled data volumes (only multiples). The complex multiples in this deep section quickly overwhelm the primary events and could easily be mistaken for primaries in some instances. Notice the faint primaries in the top left corner and the basement reflector due to the imperfect subtraction. The bottom panel is a zoomed-in version of the multiple prediction in Figure 7.5. No residual primaries are present, but some edge effects are visible at $z = 6000$ m on the left side. The bottom two panels are effectively identical which demonstrates the commutability of SRMP convolutions and shot-profile migration.

7.4.3 Gulf of Mexico data

The left panel of Figure 7.7 is a shot gather from a data set acquired in 1997 by WesternGeco in the Mississippi Canyon lease area of the Gulf of Mexico. Split-spread coverage was generated via reciprocity. The data set contains 1096 shots and recorded to 10 s. The source

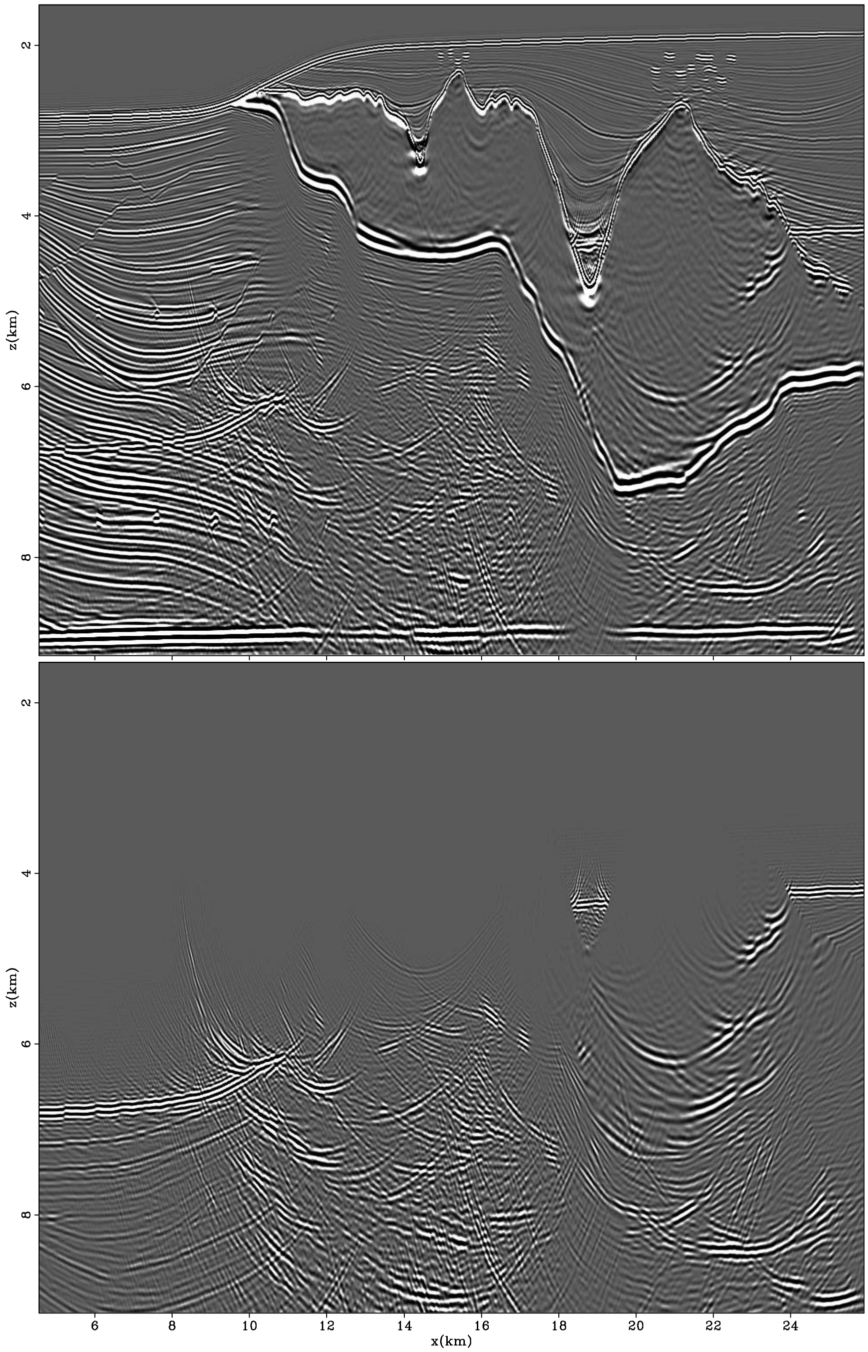


Figure 7.5: Migration of the Sigsbee2B data containing multiples and the IS-SRMP result using data containing only primaries. DepthMult-sig1 [NR]

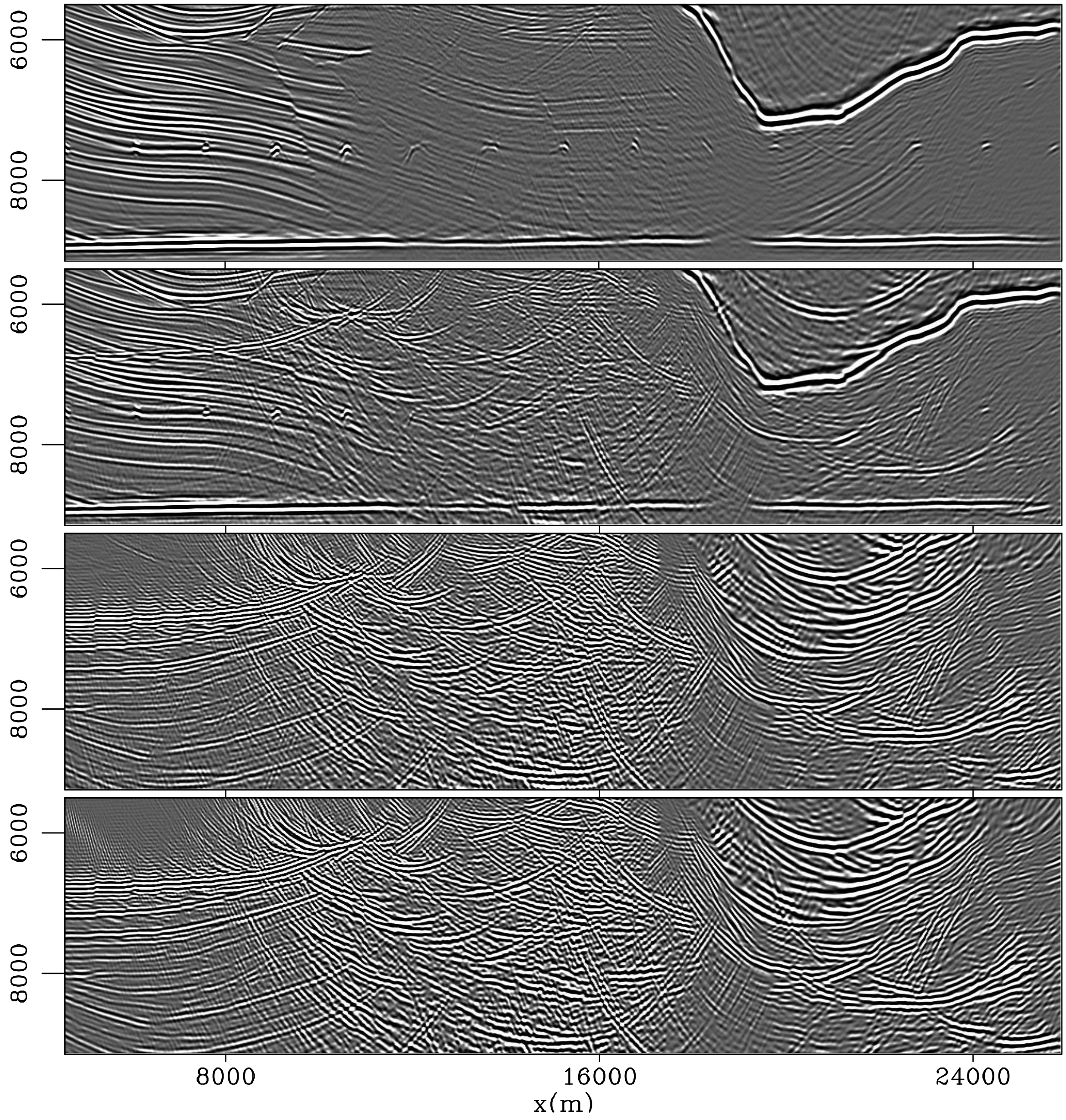


Figure 7.6: Images of the bottom third of Sigsbee2B modeled data. Top to bottom: Migration of primaries only. Migration of multiples and primaries. Migration of multiples only. IS-SRMP image using data containing only primaries. DepthMult-mult [NR]

and receiver sampling was 27 m. The right panel is a migrated angle gather from the same location. The raw data contains 367 traces, while the angle gather contains only 60 traces. Primaries are flat, and multiples are characteristically concave-down. Before migration, the data were regularized, and bandpassed from 3-65 Hz. The data were migrated with the same SSF-PI shot-profile migration algorithm with four reference velocities. To produce only the zero-subsurface offset of the pre-stack migration, the cost of the imaging condition, including the multiple prediction, was 1.4% of the cost of extrapolation. Calculating 60 subsurface offsets approximately equalizes the costs of extrapolation (with four reference velocities) and imaging (including multiple prediction and conventional image generation). This number of offsets is about $1/6^{th}$ of the traces in the split-spread gather, and more than sufficient to capture the unfocused multiple energy at all depths in the subsurface-offset gathers.

Figure 7.8 shows the resulting migration and image-space multiple prediction. The multiples have the opposite polarity as expected. The quantity of multiple events in the prediction is overwhelming. Several events in the deep section of the image that have correspondence in the multiple prediction, look very much like primaries. The event that runs off the image at $z,x = 6,2$ km could very easily be misinterpreted as a primary.

7.5 Adaptive Subtraction

After predicting the kinematics of the multiples, they must be adaptively subtracted from the image due to the amplitude and bandwidth problems associated with convolving (squaring) the traces. Another potential problem with generating a multiple prediction by convolving data instead of only the primaries is the prediction of higher-order multiples, even if they were not recorded in the time interval of the original data. This was noted above in the IS-SRMP predictions with the flat-layer synthetic, Figure 7.4. The adaptive subtraction must be able to ignore events in the multiple prediction that are completely absent from the data. Also, the multiple prediction has more artifacts than the image. This is especially true at shallow depths before the first multiple. These artifacts must not be introduced to the estimate of the primaries when removing the multiples.

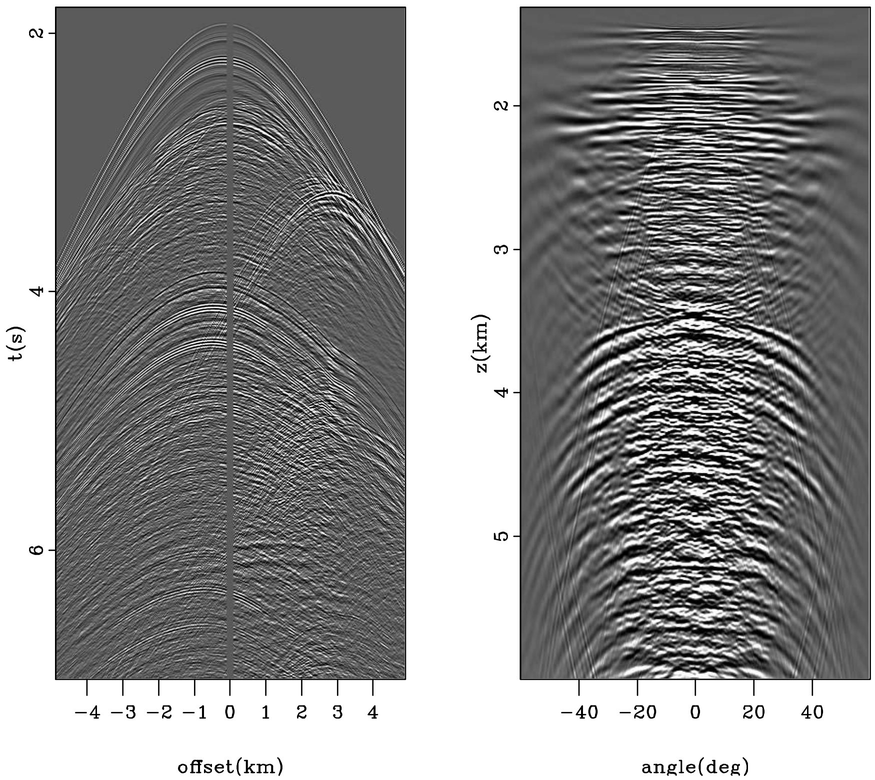


Figure 7.7: Left: Shot gather from a Gulf of Mexico data set. Near-offset traces are null. Split-spread gathers were created via reciprocity. Right: Subsurface-offset gather from the migrated image at the same location. Subsurface gather contains $1/6^{th}$ of the number of traces in the shot gather. DepthMult-gulfp [NR]

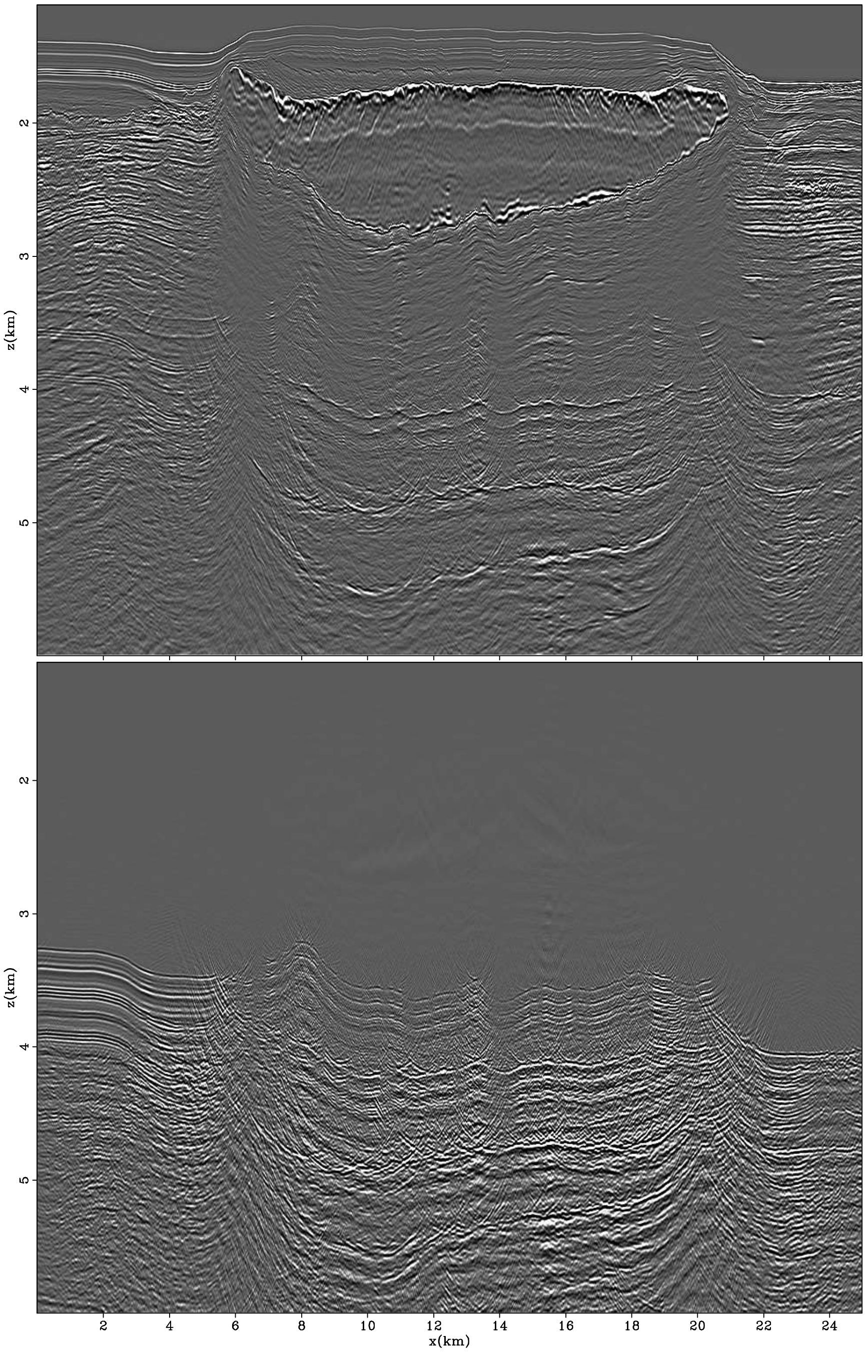


Figure 7.8: Mississippi Canyon, Gulf of Mexico zero-offset image and the IS-SRMP computed during the course of shot-profile migration of 1096 shots. DepthMult-gulf [NR]

In this section I present subtraction results from the previous predictions using adaptive subtraction. Adaptive subtraction proceeds by inverting for match filters connecting two similar objects, convolving the filters with its object, and subtracting this result from the second object. This can be a complicated and subtle art that I make no claims of performing at the optimum level. All subtraction results presented operated in two dimensions, and could likely be improved by incorporating more dimensions in the process (Guitton, 2005).

I pose the subtraction of the predicted multiples from the data as the following linear inversion problem:

$$\begin{aligned} \mathbf{Mf} &\approx \mathbf{d} && (7.15) \\ \epsilon\mathbf{Af} &\approx \mathbf{0}, && (7.16) \end{aligned}$$

where $\mathbf{M}$ is the convolutional matrix of the multiples (a matrix whose columns contain shifted versions of a vector of the multiple prediction, $\mathbf{m}$), and $\mathbf{f}$ is a bank of non-stationary filters acting on patches (Claerbout and Fomel, 2002) of the data vector, $\mathbf{d}$, which contains primaries and multiples. The matrix $\mathbf{A}$ is a regularization operator (in this case a Laplacian) and ϵ controls the strength of regularization. The result of this linear inversion is a multiple model, $\mathbf{Mf}$, that matches the amplitude and wavelet of the data.

7.5.1 Simple synthetic

Figure 7.9 details the steps in the subtraction process for an angle-domain common-image gather from the flat-layer synthetic. The first panel is the conventional image. The second panel is the IS-SRMP gather. The last panel is the matched version of the IS-SRMP gather, $\mathbf{Mf}$, ready to subtract from the image gather. Some energy from the water-bottom primary has leaked into the matched multiple prediction. Where the first multiple crosses the second primary, the filters have difficulty separating the two events. The higher-order multiples at the bottom of the initial prediction have been removed.

Figure 7.10 contains the final subtraction results. The first panel is the subtraction of the first and third panels in the previous figure, $\mathbf{d} - \mathbf{Mf}$. This is a reasonable result, but

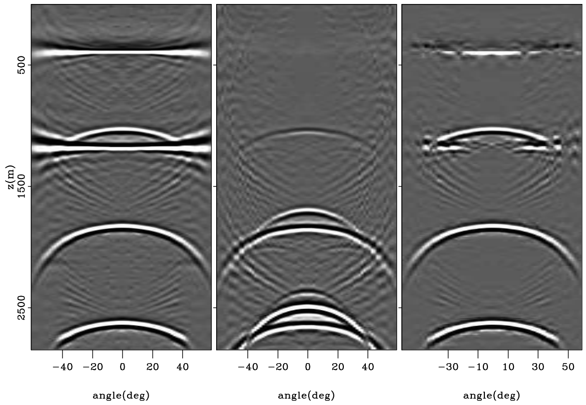


Figure 7.9: Angle-domain image gathers from a flat-layer model. Left to right: Conventional image, multiple prediction, prediction matched to data. DepthMult-m.ni [NR]

suffers from some artifacts. The last panel is the original image gather, **d**. The improved subtraction result in the center is produced with a second application of the the match-filter technique described in equation 7.15. For this application, the convolutional matrix **M** is made from the subtraction result (first panel), and the data vector **d** is once again the conventional image gather. The result shown is therefore the subtraction result matched to the input gather. This two-stage matching methodology is a partial application of the much more thorough approach presented in Alvarez and Guitton (2006).

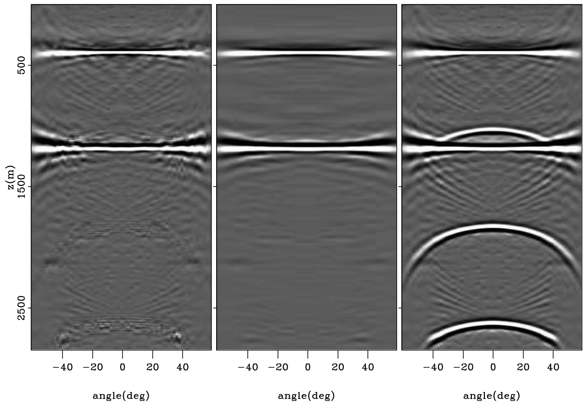


Figure 7.10: Angle-domain image gathers from a flat-layer model. Left to right: Subtraction of matched multiple prediction from data, subtraction matched to data, original data. DepthMult-p.ni [NR]

Missing data, either source or receiver positions, results in incomplete multiple prediction for SRMP. However, extrapolation spreads energy across such gaps after a few propagation steps. This is commonly referred to as wave-front healing. Collecting near-offset traces is always difficult in the field. Figures 7.11 and 7.12 were produced to test the IS-SRMP algorithm when data does not contain near-offset information. Ten null-traces surrounding the source were substituted in the gather shown in Figure 7.2, which corresponds to a gap of 200 m. There is some dimming near zero angle in the primary events, and a small deviation in the continuity of the curvature of the predicted multiples.

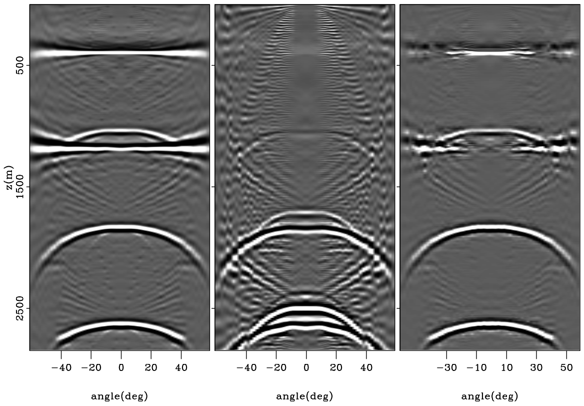


Figure 7.11: Angle-domain image gathers from a flat-layer model with missing near offsets. Left to right: Conventional image, multiple prediction, prediction matched to data. DepthMult-m.0n [NR]

In the interest of direct comparison, the adaptive subtraction parameters were kept constant for both Figures 7.10 and 7.12. The faint energy in the center panel of Figure 7.12 above the second reflector indicates that the same parameters for the adaptive subtraction are not as appropriate for this image volume. However, the remainder of multiple energy is easily removed from the final result when the subtraction is better tuned.

7.5.2 Sigsbee2B

The top panel in Figure 7.13 shows the bottom third of the multiple-contaminated image produced with the Sigsbee2B synthetic data. The bottom image was produced with the data

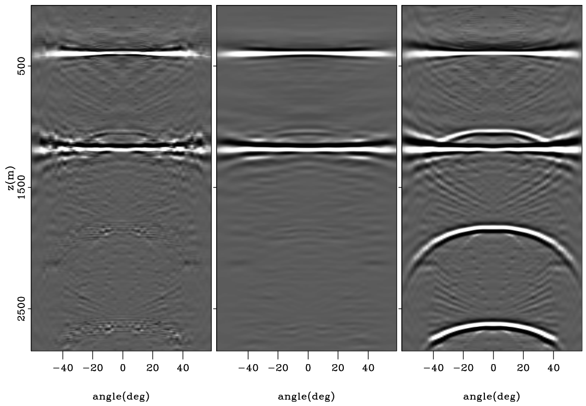


Figure 7.12: Angle-domain image gathers from a flat-layer model with missing near offsets. Left to right: Subtraction of matched multiple prediction from data, subtraction matched to data, original data. DepthMult-p.0n [NR]

volume without multiples. The center image is the result of predicting the multiples during shot-profile migration of data containing multiples and removing them from the top image. The point diffractors masked in the top image are much better realized in the subtraction result. The multiples in the sedimentary section before $x = 10$ km are almost perfectly removed except for some up-right dipping energy at $x = 7$ km emerging from the basement reflector. Several events between $x = 10 - 19$ km are much more continuous and interpretable.

7.5.3 Gulf of Mexico data

The top panel in Figure 7.14 shows the bottom half of the multiple contaminated image produced with the Gulf of Mexico data. The center panel is the multiple prediction generated during shot-profile migration. The bottom panel is the subtraction result. The subtraction still contains some residual multiples associated with the 3D nature of the salt body. It is possible that the prediction could improve with more diligent adaptive subtraction since the prediction captures much of the detail of these multiples. Many diffracted multiples associated with the sharp tips of the salt have been removed around $x = 7$ and 20 km. Primaries at $z = 4.5$ km on the left edge of the image have been brought out and may

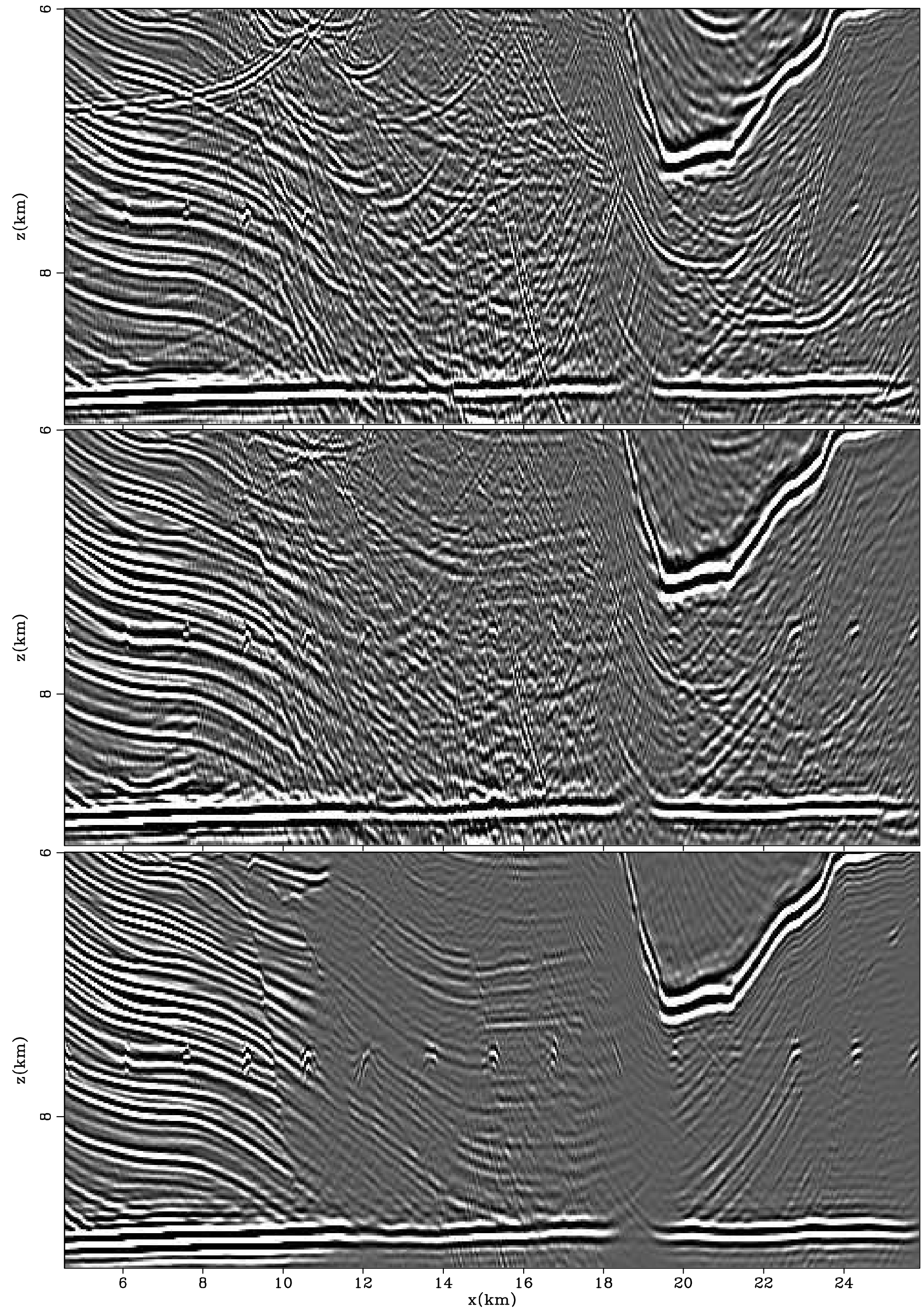


Figure 7.13: Bottom third of the zero-offset image produced with the Sigsbee2B synthetic. Top to bottom: Image with primaries and multiples, subtraction result, image produced from data without multiples. DepthMult-sigsub [NR]

suggest an anticlynal structure. Most of the migration diffraction energy associated with the rugose salt surfaces has also been attenuated.

7.6 Discussion

In many geophysical applications, improved results are generated by performing operations in the pre-stack domain. While more expensive, operating in higher dimensions can often produce better images. Migration performs a sum over frequency, shot position, and (possibly) reflection angle to produce an image volume. I suggest that the amplitude and bandwidth normalization performed in adaptive subtraction can be implemented at the imaging condition using deconvolutional variants (Lee et al., 1991; Guitton et al., 2006). To normalize the frequency content and collapse the wavelet, the imaging condition in equation 7.3 becomes

$$i_z(\mathbf{x},\mathbf{h}) = \sum_{\mathbf{x}_s}\sum_{\omega} \frac{U_z(\mathbf{x}+\mathbf{h};\mathbf{x}_s,\omega)D_z^*(\mathbf{x}-\mathbf{h};\mathbf{x}_s,\omega)}{\ll D_z(\mathbf{x}-\mathbf{h};\mathbf{x}_s,\omega)D_z^*(\mathbf{x}-\mathbf{h};\mathbf{x}_s,\omega) \gg}, \tag{7.17}$$

where the denominator is smoothed across horizontal coordinates (angle braces) for stability. Similarly, the IS-SRMP, equation 7.9, becomes

$$m_z(\mathbf{x},\mathbf{h}) = \sum_{\mathbf{x}_s}\sum_{\omega} \frac{U_z(\mathbf{x}+\mathbf{h};\mathbf{x}_s,\omega)U_z(\mathbf{x}-\mathbf{h};\mathbf{x}_s,\omega)}{\ll U_z(\mathbf{x}-\mathbf{h};\mathbf{x}_s,\omega)U_z^*(\mathbf{x}-\mathbf{h};\mathbf{x}_s,\omega) \gg}. \tag{7.18}$$

Our experiments to date have not shown this approach particularly effective. There are minor improvements in both zero-offset images, but adaptive subtraction of the multiple prediction is still required. The stability of the multi-offset images is very sensitive to the smoothing parameters selected. The deconvolutional variants are substantially more expensive, and have not, so far, proven to be worth the computational burden.

While IS-SRMP, by definition, produces only surface-related multiples, the technique could be manipulated to address strong multiple generators in the subsurface. A layer-stripping type approach derived in parallel to the DSR multiple prediction (Malcolm et al., 2006) could probably be implemented at the cost of extrapolating a third (in addition to the

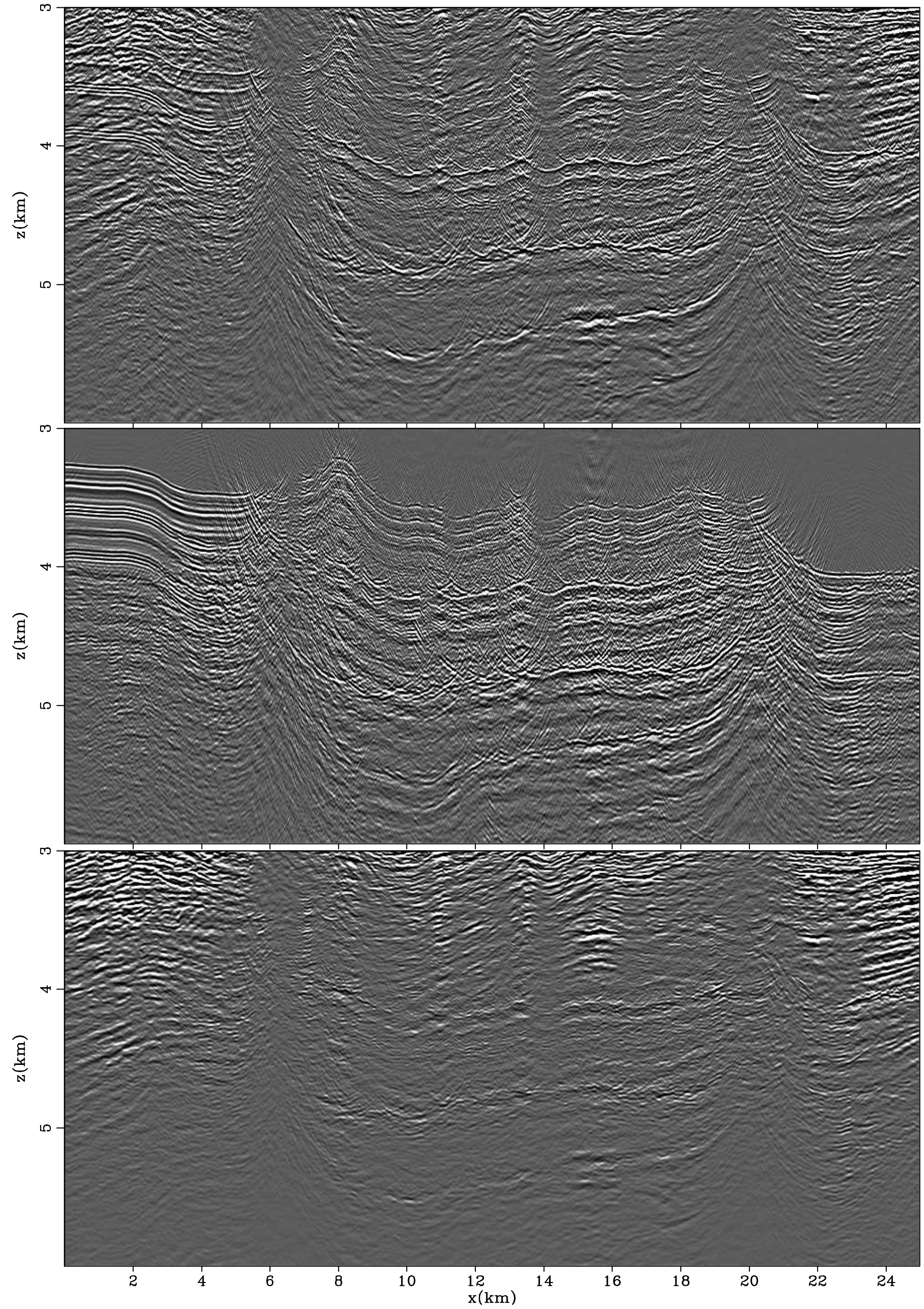


Figure 7.14: Bottom half of the zero-offset image produced with the Mississippi Canyon, Gulf of Mexico, data provided by Western-Geco. Top to bottom: Image with primaries and multiples, image-space multiple prediction, subtraction result. DepthMult-gsub [NR]

up-coming and down-going) wavefield.

7.7 Conclusion

Image space SRMP produces a multiple prediction by convolving the data with itself at every subsurface depth level during the course of a shot-profile migration. The result is the same as migrating the conventional data-space multiple prediction (SRMP). This method is most convenient with shot-profile migration strategies since the convolution operation must operate on distinct gathers rather than the combinations thereof produced by resorting to CMP coordinates. The simplicity of this approach can immediately be leveraged to generate the image-space multiple prediction directly from any shot-profile migration program. The method is immediately applicable to 3D, and non-zero subsurface-offset and angle. Any migration algorithm that maintains separate up-coming and down-going wavefields and uses a combinatory imaging condition (e.g. plane-wave and reverse-time migrations) can be easily modified to produce an IS-SRMP volume.

Given a reasonably accurate velocity model for migration, it is only necessary to compute $O(10)$ subsurface offsets. This results in many fewer traces involved in calculating the multiple prediction than the $O(1000)$ offsets collected at the surface. This savings will be reduced however by the need to convolve the traces at every depth level, $O(100)$, of the image space rather than just at the surface. Whatever the balance of floating-point operations for a particular survey, the convenience of being able to generate the multiple prediction during another required processing step can save file manipulation, sorting, and overhead costs. Furthermore, this technique can also be implemented in a target oriented fashion simply by not calculating the multiple prediction at shallow depths where it is not required.

Importantly, split-spread gathers must be pre-computed via reciprocity for data collected with off-end acquisition geometries. Off-end gathers will not contain (nor therefore predict) emerging rays which pierce the acquisition surface in front of the boat. This may increase the size of the computational domain used for propagating each individual shot-record. The cost increase by performing two imaging conditions is not severe, as the cost of calculating an imaging condition with in-line offsets is usually a fraction of the cost of a shot-profile migration. Therefore, whenever a shot-profile migration is being performed, it may be advantageous to generate the IS-SRMP even if a data-space elimination effort has already been performed, especially when the adaptive subtraction leaves some multiple energy in the final result.

The quality of the multiple prediction produced in the image space with this technique is independent of the accuracy of the velocity model used during the migration. The multiple prediction, propagated with the same velocity model, must be kinematically accurate with the location of any multiples in the migrated image. Though both the image and the multiple prediction may be mis-migrated, events in the volumes always correspond when constructed with the same extrapolation operators and velocity model.

Due to the squaring of the wavelet when convolving the data, the multiple prediction can not be directly subtracted from the data or the image. The adaptive subtraction results above show high quality estimations of the primaries for both the Sigsbee2B synthetic and the Gulf of Mexico data examples. Substantial improvement to the estimation of primaries was achieved by re-using the match filter inversion process to match the subtraction result to the original image. This second iteration mitigates some of the residual energy from the subtraction, restores some of the primary energy inadvertently removed during subtraction, and attenuates some of the artifacts associated with migration and the transform from subsurface offset to angle domain common-image gathers.

7.8 Acknowledgments

This chapter was written for publication in Geophysics (accepted 2006) with two coauthors. Gabriel Alvarez provided expertise with the adaptive subtraction efforts. Ken Matson provided valuable insight into the world of multiple prediction (with which I have little to no experience). During the review process for publication, reviewers C.P.A Wapenaar, D.J. Verschuur, and Paul Sava all made insightful comments to drafts of the manuscript that substantially improved its quality.

Chapter 8

Conclusions

Passive seismology is a very broad field that basically recapitulates the entire field of seismology with unconventional sources. Correlating traces from a receiver array calculates the time differences recorded in an active survey. The central caveat to successfully synthesizing active data from passive data is that wavefields from many sources from a wide distribution of subsurface locations are available. If so, the synthesized shot gathers are, at best, reproductions of a data volume that has been collected and analyzed extensively for decades. As such, all of the same limitations, inconveniences, and problems are included in the processed data volume. In practice, utilizing the ambient sound field of the Earth will lead to synthesized data that have even more difficulties than active data.

Preprocessing passive data is a challenging task. At its simplest, short time windows can be extracted when a seismologist is sure that useful energy is investigating the subsurface. After events are identified, there are a range of options that lead to the standard processing strategies seen in global seismology and the exploration industry. If identifiable events are not present in the raw data, the only preprocessing strategies available are those based on known quantities such are the array geometry and sensor recording bandwidth. After synthesizing shot gathers, a host of signal enhancing algorithms are likely applicable to the data. However, if a complete distribution of subsurface sources was not captured to synthesize the gathers, it may be impossible to correctly identify signal and noise components in the gathers.

For subsurface imaging exploiting body waves, a distribution of subsurface sources directly beneath the recording array is sufficient to synthesize interpretable shot gathers. The results have artifacts that are not present in active seismic data, but the bulk of the events in the volume are accurate for further analysis with conventional processing algorithms. When only a small number of subsurface sources are captured, or when multiple subsurface sources are captured in the same time window, the correlation results can be completely uninterpretable. There is still ample subsurface information contained within the data volume however that can be uncovered by further processing with a migration algorithm. Migrating passive data due to single or summed sources produces stable, interpretable images. The image domain provides a convenient, smaller sized volume of information from which to judge if the data collected contain informative reflection events or just noise.

Direct migration of passive seismic data is an alternate processing strategy to produce a subsurface image without first correlating the traces from a receiver array at the acquisition surface. Instead, the raw data can be extrapolated to all subsurface depth levels before correlation with a shot profile depth migration algorithm. Redatuming the data before correlation does not change the underlying seismic experiment as long as sufficiently accurate propagators and velocity models are used. Direct migration is a new definition of appropriate initial conditions for a specific data acquisition strategy. After defining the correct initial conditions for the wavefields used in migration, the same shot profile migration algorithms developed for active seismic data can be used to process the passive data.

With modeled transmission data, I found very little difference in results obtained by direct migration compared to images produced by conventional migration of synthesized data. Substantial computational cost can often be saved by processing transmission wavefields with direct migration rather than first synthesizing shot gathers. This is because migration is such an expensive processing algorithm, and correlating every trace with every other trace squares the size of the data volume to migrate. Therefore, these cost considerations are more important for passive data collected with large arrays with more than several thousand sensors.

Passive seismic data recorded over 36 hours was processed to produce images in this thesis. I used approximately 2300 hydrophone pressure sensors from the Valhall array in the Norwegian North Sea. Correlation processing to synthesize shot gathers did not reveal subsurface information. Instead, the results were dominated by a direct arrival through the water column from production noise at a facility 40 km away. That facility flares produced gas whereas the Valhall production facility does not. I believe this is the energy source that gives rise to the only interpretable event on the correlated gathers. The event is stable in results produced with passive data volumes collected one year apart.

Direct migration of the Valhall passive data produced noisy, but interpretable subsurface images. Images produced with active seismic data were available for direct comparison. Events as deep as 2 km were present in the passive images that seemed to correspond to events in the active image. However, migrating modeled data based on the far-field noise extracted from the synthesized gathers produced almost identical images. Identifying the dominant seismic energy in the records was accomplished by analysis of deconvolved synthesized shot gathers. Without modeling the effects of this noise in all of the processing schemes used with this data, results would have been over interpreted. The coherent noise events caused artifacts with all the character of real subsurface events seen in the active data. The image volumes produced by synthesizing shot gathers followed by migration were quite different from direct migration images. Because these two processing options are mathematically equivalent for data from a broad distribution of subsurface sources, the difference between the results indicates that the field data do not contain the distribution of subsurface sources required for successful imaging.

The passive imaging experiments described herein used a unique data volume provided by BP. This volume of data recorded more data with more receiver stations than any other passive seismic experiment ever fielded. Because the sensors are permanently installed on the sea floor, data can be continuously collected to continue analyzing the effectiveness of passive imaging. Processing 1.5 days of passive data did not produce images of subsurface structure. It is possible that similar processing with more data could improve on my results. The failure of my experiments, however, provides little encouragement to invest in the substantial processing required to analyze passive seismic records.

I believe that successful subsurface imaging at Valhall with passive seismic data is predicated on removing the contaminating coherent noise from the gas flare at the Ardmore production facility. The noise is so much more energetic than any local seismic energy that my attempts at removing it from the data have been unsuccessful. To improve the chance of successful imaging, the removal should be performed before correlation. However, the noise is continuous and unidentifiable on the raw records, and is not isolated from body waves in the Radon or Fourier domains. Data collected when the flare is extinguished for maintenance would be ideal.

Aside from collecting data when the flare is not active, it may be possible to remove the noise from the raw data with a signal processing strategy. The noise is very broad-band, and continuous in time. This combination will make both frequency-domain and time-domain approaches challenging. A large body of work exists concerning statistic and deterministic signal processing that may address removing the noise from the data. Specifically, echo removal techniques seem likely candidates for initial tests. Many of these techniques operate on the properties of the autocorrelation of the signal. This could be beneficial, or a cause for total failure since the autocorrelation of the noise is so dominant. Multi-dimensional processing strategies are also available. Because the spatial characteristics of the noise is well known, I believe that at least 2D (time and inline direction) algorithms will be required for success. All of these techniques are computationally expensive, often involving chains of cost-function minimizations.

Some form of principal component analysis (PCA) might work well to separate the strong noise arrival from the data. Because the flare noise is so energetic, it should occupy the higher-order principal components from the PCA. The lower-order principal components contain the data components most responsible for the variability within the data. There will be a large amount of background random noise in the data that might be suppressed by also discarding the lowest-order principal components. Geologic signal should lie somewhere in center of the PCA domain. Again, I believe that utilizing the multi-dimensional character of the flare noise, especially in the inline direction, will provide a substantial improve in the quality of result.

An alternative use for the passive recording capabilities of the array is as a global seismology station. Processing data from teleseismic arrivals to image deep structure should be possible and useful. At such low frequencies however, the areal footprint of the array is likely too small to support array-based processing techniques. However, it is possible that similar arrays may be installed at other reservoirs in the region. Recording at multiple arrays simultaneously may provide interesting data sets to investigate lower-frequency phenomenon. Also, a second array may provide more areal information to attack the gas flare, which will undoubtedly be observed elsewhere in the region.

Direct migration is an example of simple seismic processing accomplished with the imaging condition of shot profile migration. In this case, the imaging condition is exactly the desired processing calculation. However, using accurate mathematical propagators to model the physics of wave propagation allows us to perform any processing step after extrapolation in the image space. Statics corrections are a simple data processing example based on this principle.

Direct migration of data due to unknown sources is analogous to imaging reflectors with multiple energy. In terms of a multiple problem, correlation models primaries from first order multiples. The most common operator encountered in the literature concerning multiples is convolution. Convolution models first order multiples from primaries. Correlation and convolution are adjoint processes that either image primaries with multiples or predict multiples from primaries. Image-space multiple prediction is the adjoint of direct migration of passive seismic data. The two algorithms are fundamentally linked by the mathematical similarities of the two operations.

Image-space multiple prediction auto-convolves shot gathers at every depth level in the image domain after extrapolation and sums the result from all shots to produce the final result. After extrapolation, the redundant time axis can be discarded to keep only the space axes. Traditional data-space multiple prediction can be distilled to component processes: Convolve, sum. Convolving traces after extrapolation through the course of the shot profile migration algorithm does not invalidate the accuracy of the multiple prediction. The accuracy of the propagator is not important when producing a multiple prediction in the image space because the conventional image of reflectors is produced with the same propagator.

Events in the multiple prediction correspond exactly to the multiples in the conventional image even if both volumes are inaccurate. The events in the two volumes are always directly comparable since the same kinematic modeling was used to create both image volumes.

Bibliography

Alvarez, G. and A. Guitton, 2006, Simultaneous adaptive matching of primaries and multiples with non-stationary filters: SEP–**125**, 61–77.

Alvarez, G., B. Biondi, and A. Guitton, 2004, Attenuation of diffracted multiples in angle domain common image gathers: 74th Ann. Internat. Mtg., Soc. of Expl. Geophys., Expanded Abstracts, 1301–1304.

Anstey, N. A. and P. Newman, 1966, The sectional auto-correlogram and the sectional retro-correlogram: Geophys. Prosp., **14**, no. 4, 389–426.

Artman, B. and J. C. Shragge, 2003, Passive seismic imaging: AGU Fall Meeting, Eos Transactions of the American Geophysical Union, Abstract S11E–0334.

Artman, B., D. Draganov, C. P. A. Wapenaar, and B. Biondi, 2004, Direct migration of passive seismic data: 66th Conference and Exhibition, EAGE, Extended abstracts, P075.

Artman, B., 2006, Imaging passive seismic data: Geophysics, **71**, no. 4, SI177–SI187.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, no. 11, 1514–1524.

Berkhout, A. J. and D. J. Verschuur, 1997, Estimation of multiple scattering by iterative inversion, part I: Theoretical considerations: Geophysics, **62**, no. 05, 1586–1595.

Berkhout, A. J. G. and D. J. E. Verschuur, 2006, Multiple removal and wavelet deconvolution in the inverse data space: SEG Technical Program Expanded Abstracts, **25**, no. 1, 2684–2688.

Biersteker, J., 2001, MAGIC: Shell's surface multiple attenuation technique: SEG, 71st Annual International Meeting, 1301–1304.

Biondi, B. and W. W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, no. 5, 1283–1298.

Biondi, B., 2006, 3D seismic imaging:, number 14 in Investigations in Geophysics Society of Exploration Geophysicists.

Borcea, L., G. Papanicolaou, and C. Tsogka, 2003, Theory and applications of time reversal and interferometric imaging: Inverse Problems, **19**, S139–S164.

Claerbout, J. and S. Fomel, 2002, Image Estimation by Example: Geophysical soundings image construction: Class notes, http://sepwww.stanford.edu/sep/prof/index.html.

Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: Geophysics, **33**, no. 02, 264–269.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, no. 03, 467–481.

Cole, S. P., 1995, Passive seismic and drill-bit experiments using 2-D arrays: Ph.D. thesis, Stanford University.

Derode, A., E. Larose, M. Campillo, and M. Fink, 2003, How to estimate the Green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves: Applied Physics Letters, **83**, no. 15, 3054–3056.

Draganov, D., K. Wapenaar, and J. Thorbecke, 2003, Passive seismic imaging in the presence of white noise sources: numerical simulations: SEG Technical Program Expanded Abstracts, **22**, no. 1, 90–93.

Draganov, D., K. Wapenaar, W. Mulder, and J. Singer, 2006, Seismic interferometry on background-noise field data: SEG Technical Program Expanded Abstracts, **25**, no. 1, 590–594.

Gebara, J., D. Dolan, et al., 2000, Assessment of offshore platforms under subsidence–Part i: Approach: Journal of Offshore Mechanics and Arctic Engineering, **122**.

Guitton, A., A. Valenciano, D. Bevc, and J. Claerbout, 2006, Robust illumination compensation for shot-profile migration: 68th Annual Conference and Exhibition, EAGE, Expanded Astracts, P265.

Guitton, A., 2002, Shot-profile migration of multiple reflections: SEG Technical Program Expanded Abstracts, **21**, no. 1, 1296–1299.

Guitton, A., 2005, Multiple attenuation in complex geology with a pattern-based approach: Geophysics, **70**, no. 4, V97–V107.

Hampson, D., 1986, Inverse velocity stacking for multiple elimination: Inverse velocity stacking for multiple elimination:, Soc. of Expl. Geophys., 56th Ann. Internat. Mtg., Session:S6.7.

Hohl, D. and A. Mateeva, 2006, Passive seismic reflectivity imaging with ocean-bottom cable data: SEG Technical Program Expanded Abstracts, **25**, no. 1, 1560–1564.

Kessinger, W., 1992, Extended split-step Fourier migration: SEG, 62nd Annual International Meeting, 917–920.

Krishnakumar, K., 1989, Micro-genetic algorithms for stationary and non-stationary function optimization.: SPIE: Intelligent control and adaptive systems, **1196**, 289–296.

Lee, D., I. M. Mason, and G. M. Jackson, 1991, Split-step fourier shot-record migration with deconvolution imaging: Geophysics, **56**, no. 11, 1786–1793.

Liu, F., D. W. Hanson, et al., 2006, Toward a unified analysis for source plane-wave migration: Geophysics, **71**, no. 4, S129–S139.

Louie, J. N., 2001, Faster, better: Shear-wave velocity to 100 meters depth from refraction microtremors arrays: Bulletin of the Seismological Society of America, **91**, no. 2, 347–364.

Malcolm, A. E., M. V. de Hoop, and H. Calandra, 2006, Identification of image artifacts due to internal multiples: Submitted to Geophysics.

Morozov, I. and K. Dueker, 2003, Depth-domain processing of teleseismic receiver functions: Bulletin of the Seismological Society of America, **93**, 1984–1993.

Muijs, R., 2005, Wavefield decomposition and imaging of multicomponent seabed seismic data: Ph.D. thesis, Utrecht University.

Owens, T. J. and G. Zandt, 1985, The response of the continental crust-mantle boundary observed on broadband teleseismic receiver functions: Geophysical Research Letters, **12**, 705–708.

Rickett, J. E. and P. C. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, no. 03, 883–889.

Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, no. 02, 426–436.

Rondenay, S., M. G. Bostock, and J. C. Shragge, 2001, Multiparameter two-dimensional inversion of scattered teleseismic body waves 3. Application to the Cascadia 1993 data set: Journal of Geophysical Research, **106**, no. B12, 30795–30807.

Sava, P. and A. Guitton, 2003, Multiple attenuation in the image space: 73rd Ann. Internat. Mtg., Soc. of Expl. Geophys., Expanded Abstracts, 1933–1936.

Sava, P. and A. Guitton, 2005, Multiple attenuation in the image space: Geophysics, **70**, V10–V20.

Shan, G. and A. Guitton, 2004, Migration of surface-related multiples: tests on the Sigsbee2B dataset: 74th Ann. Internat. Mtg., Soc. of Expl. Geophys., Expanded Abstracts, 1285–1288.

Shan, G., 2003, Source-receiver migration of multiple reflections: 73rd Ann. Internat. Mtg., Soc. of Expl. Geophys., Expanded Abstracts, 1008–1011.

Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller, 2005, High-resolution surface-wave tomography from ambient seismic noise:, **307**, 1615–1618.

Shragge, J. C., B. Artman, and C. Wilson, 2006, Teleseismic shot-profile migration: Geophysics, **71**, no. 4, SI221–SI229.

Tsai, C. J., 1985, Use of autoconvolution to suppress first-order long-period multiples: Geophysics, **50**, no. 09, 1410–1425.

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 09, 1166–1177.

Wapenaar, C. P. A., J. Thorbecke, and D. Draganov, 2004, Relations between reflection and transmission responses of three-dimensional inhomogeneous media: Geophysical Journal International, **156**, 179–194.

Wapenaar, K., D. Draganov, and J. Robertsson, 2006, Introduction to the supplement on seismic interferometry: Geophysics, **71**, no. 4, SI1–SI4.

Weglein, A. B., 1999, Multiple attenuation: An overview of recent advances and the road ahead (1999): The Leading Edge, **18**, no. 1, 40–44.

Zhang, L., 1989, Reflectivity estimation from passive seismic data: Stanford Exploration Project- Annual Report, **60**, 85–96.