Chapter 1

Introduction

Acquisition of reflection seismic data is a resource-intensive process: both source and receiver positions vary along the surface of the earth, sampling data along five different axes, including the finite recording time. Fully sampling all of these axes is logistically unrealistic for many reasons: a finite number of active recording channels, surface obstacles, experimental design, and limited capital. Meanwhile, many algorithms used in the processing of reflection seismic data are designed for regularly-sampled, unaliased data with a wide aperture. As such, interpolating reflection seismic data is a common practice. The interpolation method used will often vary depending on the type of acquisition. Let us now discuss common types of acquisitions.

ACQUISITION DESIGNS

Methods of acquiring surface reflection seismic data vary considerably. Marine acquisition has typically been constrained by the use of ships such that data are typically collected along parallel lines, as both the sources and receivers are towed along the same axis, usually by the same ship. In contrast, land acquisition is typically done with intersecting source and receiver lines, with considerable deviation from this regular grid caused by surface obstacles and logistical issues. Other methods, such as ocean-bottom cable or ocean-bottom sensor surveys have still different issues. I now

discuss the two most common types of geometry that are in part addressed in this thesis. For a more thorough description of acquisition geometries, multiple textbooks are available (Vermeer, 2002; Biondi, 2006).

Land acquisition geometry

Multiple geometries of land data acquisition have been proposed, with the cross-swath method as one of the most popular. In this geometry, both the sources and receivers are placed along intersecting lines. The receiver locations are typically less influenced by surface conditions, since it is possible to place receivers where access for sources is restricted, it is easier to plant a geophone than it is to position either a vibroseis truck or a drill for shot holes. An example of a 3D survey in South America with this geometry is shown in Figure 1.1, here source locations, denoted by an 'x', are more sporadically placed than are the receiver locations, denoted by points. A large gap exists in the bottom-left of the survey, most likely caused by a surface obstacle that prevents sampling with both source and receiver. This geometry has great variability in both source-receiver space as well as midpoint-offset space. The near offsets, for example, typically contain a wider azimuth range than do the far ones.

While the sampling issues in 3D are more clearly evident, the differences between receiver and source sampling in Figure 1.1 are also applicable for a 2D survey. An example of a 2D field acquisition geometry is shown in Figure 1.2. The aspect ratio in Figure 1.2 is severely distorted, as the axes are stretched by a factor of 500 on the 'y' axis relative to the 'x' axis. This example again illustrates that in land geometries sources are more unevenly distributed than are receivers, and that surface obstacles also cause problems (albeit less dramatic) in 2D acquisition. While land acquisition contains great variability, the marine case is less irregular because the ship tows the array in a continuous fashion, as discussed next.

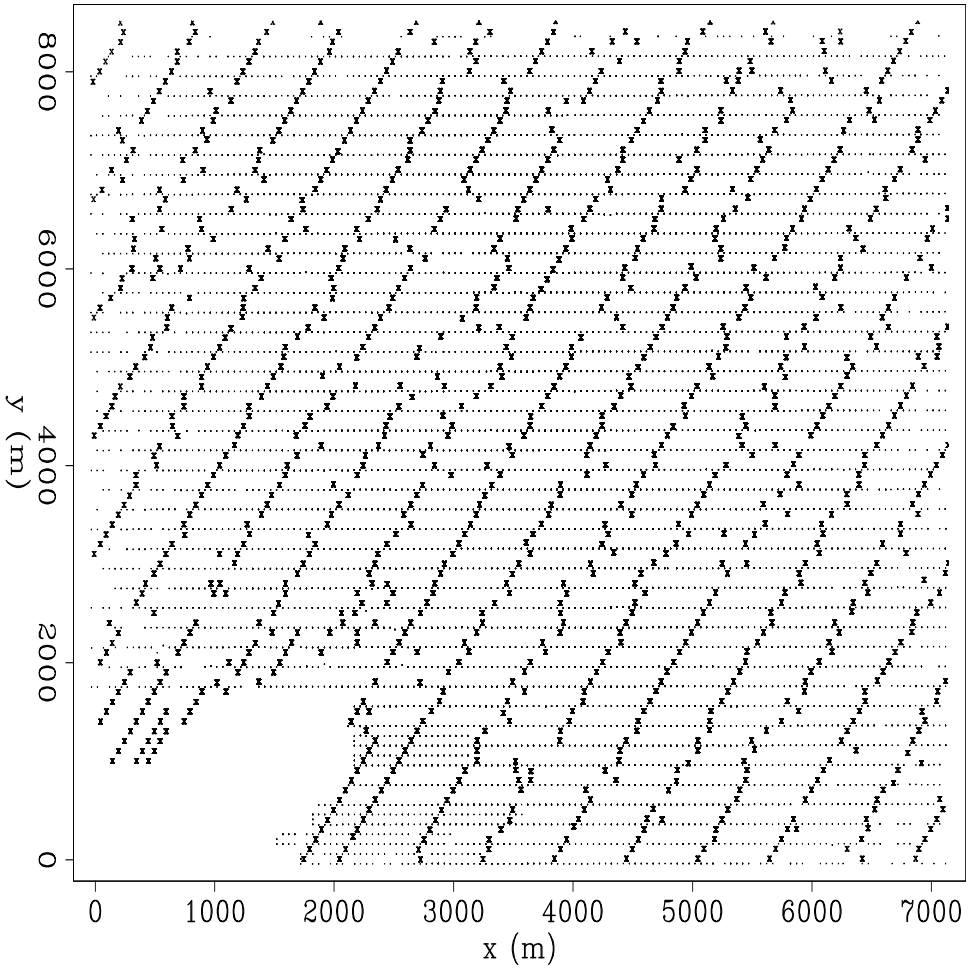

Figure 1.1: An example of the cross-swath geometry from a 3D land survey. The points are receiver locations while each 'x' denotes a source position. The receiver sampling is more regular than the source sampling. **ER** Intro/. landgeom3

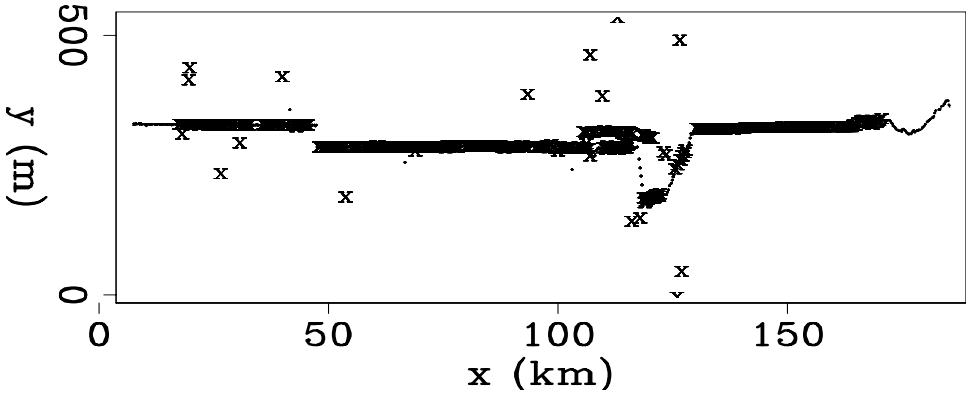

Figure 1.2: An example of receiver and source sampling for a 2D land survey. There are slight deviations along the 'y' direction for the receivers and much greater variability in source sampling. **ER** Intro/. landgeom2

Marine data geometry

Sampling in marine geometry is more regular than that in land geometry. This is largely because both the receivers and sources are towed by the same ship, so the offsets between the source and the receivers are roughly equal for each shot. One systemic gap in the data is created by the fixed distance between the source and the first receiver on the cable, leaving a near-offset gap that exists in every shot in both 2D and 3D marine surveys. Also in both 2D and 3D surveys the source interval is typically a multiple of the receiver interval since the air-gun source takes a while to recharge and the ship has traveled more than one receiver interval in the interim. To increase source sampling (and reduce the cost of acquisition) a second source is often added near the first one, creating what is referred to as *flip-flop shooting*, wherein the two sources are fired alternately. A map containing source locations for a 3D marine survey is shown in Figure 1.3. In this map, the two closely-spaced parallel lines are the alternating flip-flop sources for a single sail line. We can also see the comparatively sparse source distribution in the crossline (y) direction compared to the inline (x) direction. This usually results in severe aliasing in the crossline source direction.

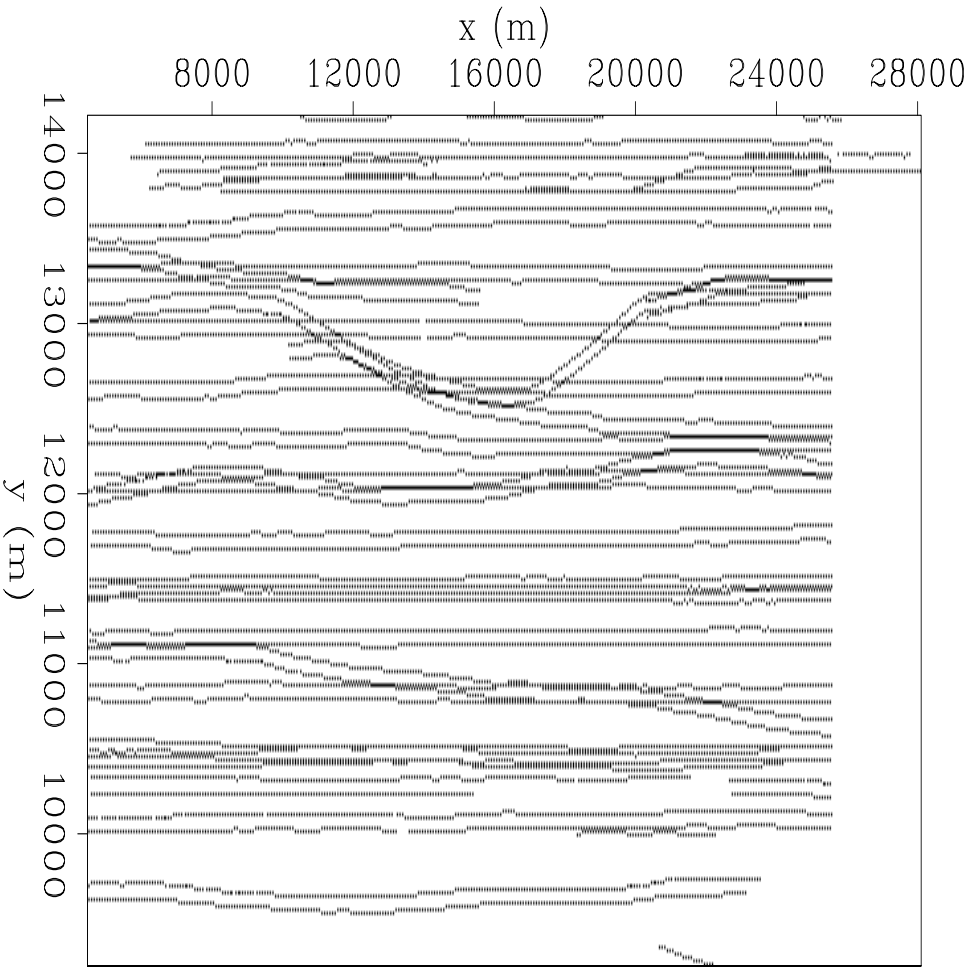

Figure 1.3: Source locations for a 3D marine survey. The ship sailed in approximately straight lines. The parallel lines are due to flip-flop shooting. Sources are much better sampled in the inline (horizontal) than the crossline (vertical) direction. **NR** Intro/. marinegeom1

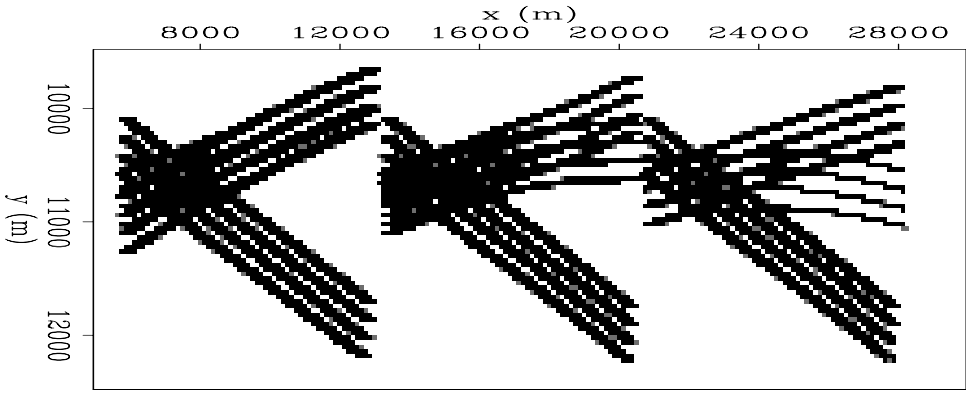

Figure 1.4: Different sets of receiver locations put on the same map for various source positions. The crossline offsets vary considerably from sail line to sail line, even if the sources are nearby. **NR** Intro/. marinegeom2

In addition to the sparse crossline source sampling, the crossline offset axis is also undersampled. Figure 1.4 shows a map view containing the positions of receiver cables for three points along three different sail lines. The three sets of cables on the bottom of the figure are relatively consistent, but the feathering is on the order of several hundred meters. The three overlapping sets of receiver cables at the top of the figure show that even when the source locations are near one another, in this case all near $y = 11500$ m, the cable feathering can be considerably different, as the ocean currents and previous ship movement that determine the cable feathering are functions of space *and* time.

Since the number of channels (both governed by length of cable and number of cables) is limited, a trade-off exists between the acquisition aperture, the total area covered by the recording array, and the sampling within that recording array. Interpolating data allows wider apertures to be used, provided that the interpolation method is able to capture information from the potentially aliased data.

THE NEED FOR INTERPOLATION

Many data-processing algorithms, including wave-equation migration and many multiple-removal methods, require regular, complete, and densely-sampled data as input, but as the previous section illustrates, actual sampling in the field is far from ideal. Surface obstacles and logistical requirements cause irregular sampling of sources and receivers, in both land and marine data. Acquisition design also results in systematic gaps in the data, such as the near-offset gap in marine data. Finite capital results in a trade-off of sampling density and aperture size, for example when adequately dense sampling along all axes is traded for larger sampling areas, resulting in spatially aliased data, particularly in the crossline direction in marine data.

Traditionally, many 2D algorithms had been applied on 3D data, one example of which is surface-related multiple elimination (SRME) (Verschuur et al., 1992). 2D SRME is effective on data from areas where the subsurface has mostly cylindrical symmetry with little variation in the crossline direction. When applied to data with multiples generated from out of the inline plane, the multiple prediction degrades. Two examples of this are shown in Figures 1.5 and 1.7. Figure 1.5a is a close-up of water-bottom multiples in field marine data (courtesy of CGGVeritas) generated from a sea floor with a varying dip in the cross-line direction, while Figure 1.5b is an example of the multiple model generated from a single receiver cable using 2D SRME. The right side of the prediction matches the multiples present in the data reasonably well, excluding the expected amplitude and phase differences. The left side of the prediction, however, is kinematically quite different from the observed multiples in the data. The predicted multiples arrive later than the actual multiples, since the actual multiple arrivals reflect from out-of-plane in the shallower water bottom.

Another example, this time synthetic, of shortcomings in 2D SRME is in Figure 1.7. Figure 1.7a shows synthetic marine data containing a horizontal water bottom and many diffractions and diffracted multiples from out-of-plane. Figure 1.7b shows the 2D multiple prediction from auto-convolution of these data. While the cloud of diffracted multiples is in roughly the correct place, the predicted diffracted multiples

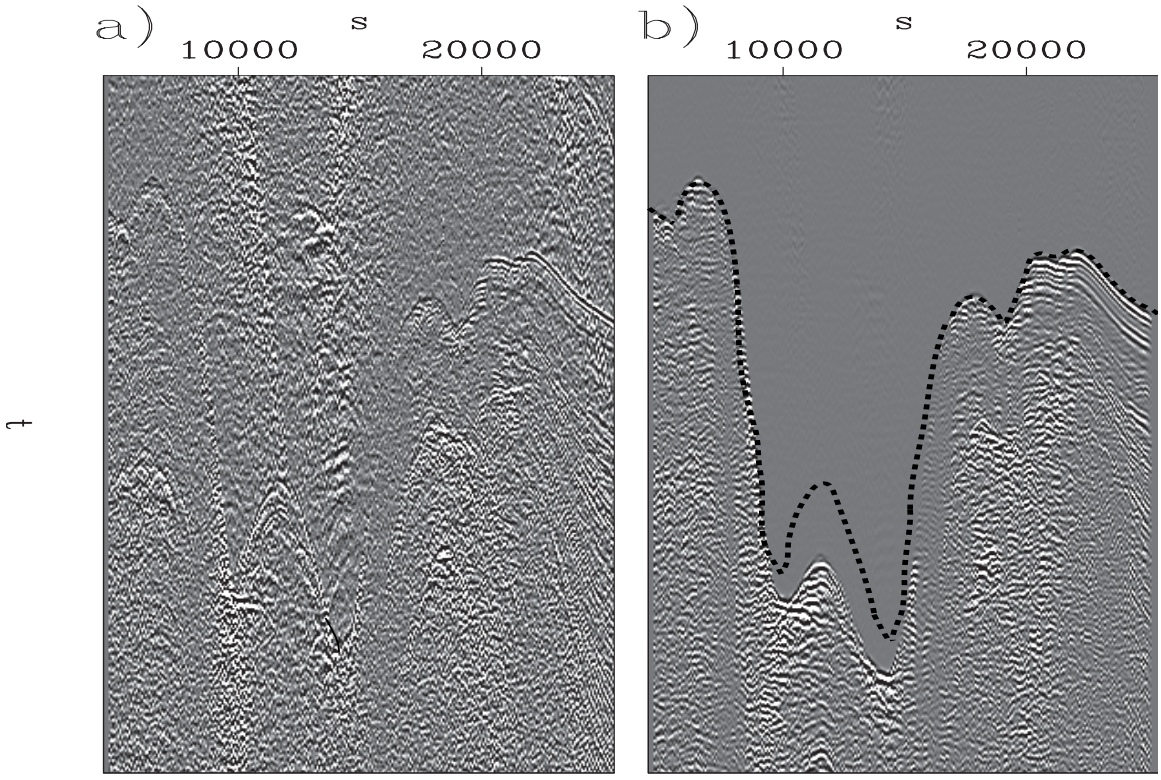

Figure 1.5: The failure of 2D multiple prediction in a 3D world: a) inline constant-offset section of original data; b) 2D multiple prediction for the same region, with the arrival times of the multiples from the actual data superimposed as a dashed line. **NR** Intro/. srmprann

are aliased and the positions of the many diffractions that make up the cloud of multiples are incorrect. The aliasing results from insufficient inline source sampling that is one-third that of the receiver sampling. The mispositioning of the events within the cloud of diffracted multiples is attributable to the 2D nature of the algorithm, in which all events are assumed to originate from within the vertical 2D plane at zero crossline-offset. Implementing 3D SRME on these data requires that a source is present at every receiver location, as described for the 2D case in Figure 1.6would require considerable crossline receiver and source interpolation as well as extrapolation of crossline offsets.

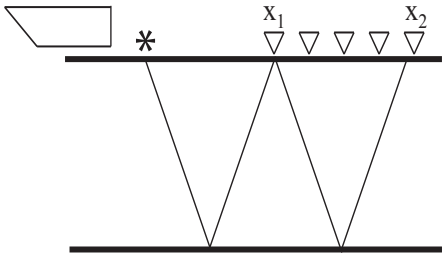

Figure 1.6: Schematic for surface-related multiple prediction. A free-surface multiple recorded at x_2 with a bounce point at x_1. This raypath can be expressed as two primaries, one with the original source and a receiver at x_1, the second with a source at x_1 and a receiver at x_2. The source and receiver sampling should be equal. **NR** Intro/. srmecartoon

EXISTING INTERPOLATION METHODOLOGIES

Interpolation has a long history, with many classic methods such as Gaussian polynomial and sinc-based (Shannon, 1949) interpolation. I place methods developed for reflection seismic data into two main categories: those that use mainly a-priori information about the physics of the recorded data, and those that use mainly statistical information gathered from the data.

The first category contains many methods, some based upon normal move-out, others on more sophisticated imaging operators such as prestack partial migration (Chemingui, 1999; Fomel, 2001; Biondi and Vlad, 2002; Clapp, 2003; Baumstein,

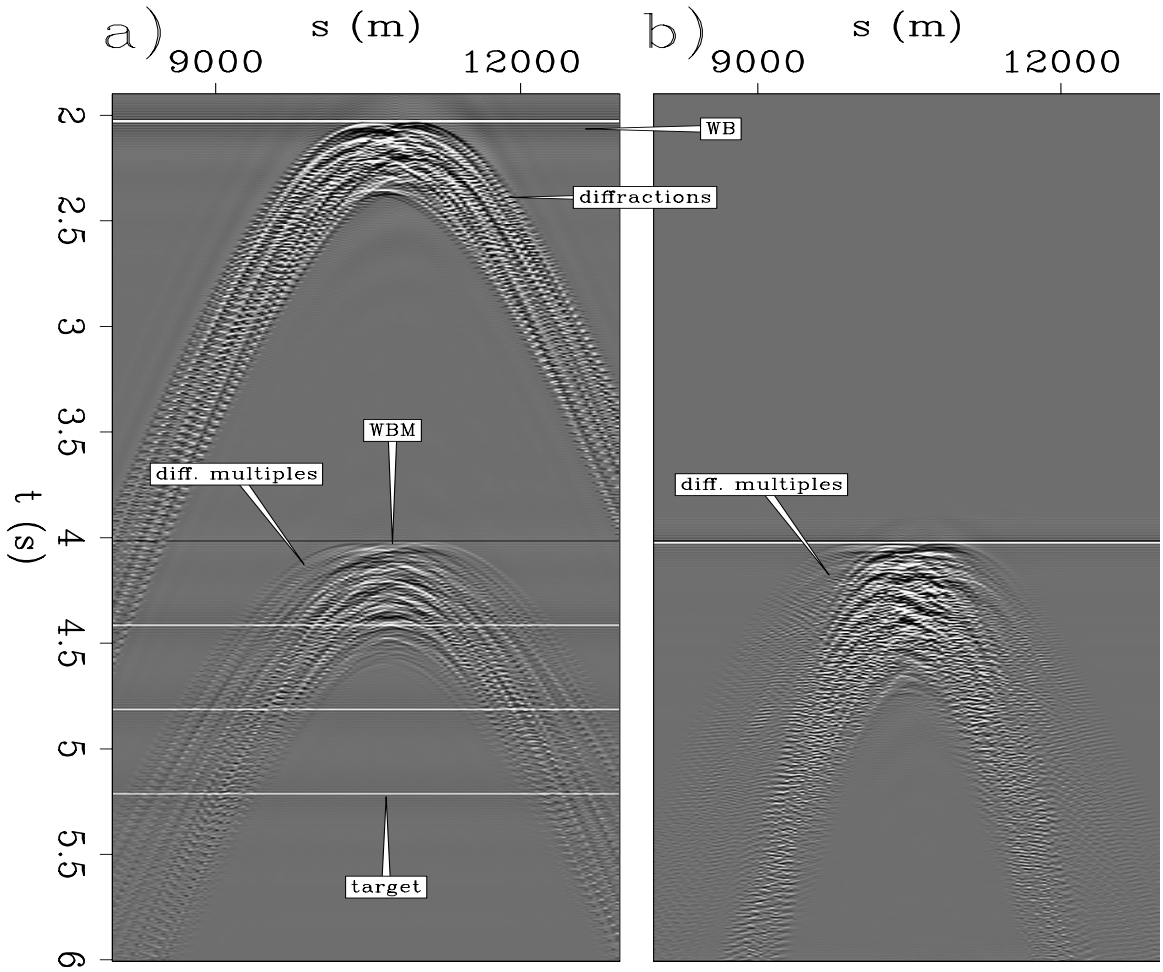

Figure 1.7: Synthetic marine data: a) original data; b) multiple prediction. The diffracted multiples created from out-of-plane in (a) are not accurately predicted by a 2D SRME prediction shown in (b), although the water-bottom multiple is accurately predicted. **CR** Intro/. srmpsann

2004), and others of full migration and demigration of data (Trad, 2002) or use of a surrogate for migration operators (Verschuur and Berkhout, 2005). These methods have several benefits, the first of which is the ability to extrapolate as well as interpolate data. Another benefit is that many of these methods can deal with irregular sampling of data, but the main drawback with these methods is that the events in the recorded data do not necessarily conform to the physics of the imaging operator. This can be because of an incorrect assumed velocity, the presence of energy (such as ground roll) that is not predictable by the imaging method, multiple reflections, or coherent noise.

Statistical methods use only information contained in the data to interpolate missing data. Statistical methods include a myriad of transform-based methods, such as Fourier, Radon, wavelet (Wang and Li, 1994), or curvelet (Thomson and Herrmann, 2006) transforms. Fourier transform-based methods exist in many forms (Gulunay and Chambers, 1997). Examples include non-uniform (Schonewille and Duijndam, 1998), minimum reweighted norm (Liu and Sacchi, 2004), Projection onto Convex Set (POCS) (Abma and Kabir, 2005), sparse (or high-resolution) methods (Zwartjes and Gisolf, 2007), and antileakage methods (Xu et al., 2005). Radon-based methods are also commonly used, often in the form of a high-resolution radon-based approach (Sacchi and Ulrych, 1995), and more recently a shifted-apex radon transform (Hargreaves and Trad, 2005). Most of these methods can interpolate irregularly-sampled data, but are limited in that they typically do not faithfully interpolate aliased data.

Another group of non-physically-based statistical methods are filter-based methods. These filters include dip filters and prediction or prediction-error filters. With dip filters, a single dip is estimated within a window of data, from which a filter is created that is then used to interpolate the unknown data in that dip direction (Fomel, 2002). While this method can cope with irregularly-sampled and spatially-aliased data, extension to treating multiple interfering dips is not straightforward and is considerably more complicated. Unlike a single local-dip filter, prediction (Spitz, 1991) or prediction-error (Claerbout, 1992) filters can estimate data in the presence

of multiple simultaneous slopes. Since these filters are larger (i.e., they have many more coefficients) than do two-column dip filters, their estimation requires regularly-sampled training data, typically estimated from a lower-frequency sampling of existing data. More recently, nonstationary prediction-error filters have been used to interpolate data in the time-space (t-x) domain (Crawley, 2000), instead of the previous patch-based approach (Claerbout, 1992).

MOTIVATION AND CONTRIBUTIONS

In this thesis, I create new nonstationary prediction-error filter (PEF)-based interpolation methods applicable to multiple applications, including problems of irregular data, large systemic gaps in acquisition, and higher-dimensional interpolation to the extent necessary for 3D SRME. I do this by adapting the choice of data used to estimate the PEF to each particular problem.

Methods to interpolate irregular data are typically limited to unaliased data or to a single dip, while prediction-error filters can capture multiple aliased slopes, but require regularly-sampled data. My contribution is a method to estimate a nonstationary PEF from irregularly-sampled data, using multiple regridded copies of the data as training data (Curry, 2003). While the more coarsely-gridded copies of the data are not exact, they produce a superior result than when using no information from the data or from using a single regridded copy of the data. This method succeeds in more accurately interpolating multiple simultaneous changing slopes that are irregularly sampled than do previous approaches.

Marine data contain a near-offset gap between the air gun source and the nearest towed receiver, but this missing near-offset information is essential for use in SRME methods. Most methods to recover these near offsets rely on the curvature of the primaries in nearby data, either in a Radon-transform-based approach or in a more straightforward NMO-based prediction wherein the stacking velocity is derived from the nearby recorded primaries. Similarly, a conventional approach using prediction-error filters would rely on using the nearby primaries as training data. In my approach,

I generate training data for a nonstationary PEF by cross-correlating traces within each shot to generate a *pseudoprimary* dataset (Curry and Shan, 2006). Instead of simply substituting these data into the near-offsets or manually matching them to the recorded data using deconvolution or matching filters, I use these as training data for a nonstationary PEF. I estimate this PEF both in time, offset, and source position, as well as in frequency, offset, and source position, and compare the results. Estimating nonstationary PEFs in the frequency domain provides a much faster, more parallelizable result.

3D SRME requires dense sampling of both receivers and sources in the inline and crossline directions. Currently, most algorithms for generating these data are based upon normal moveout (Levin, 2002), dip moveout (Baumstein and Hadidi, 2006), azimuth moveout (Matson and Abma, 2005), or a migration/demigration approach (Weisser and Taylor, 2006). While SRME is a purely data-driven approach with no required velocity model, these methods still require some sort of velocity estimate. I generate data densities suitable for 3D SRME by interpolation using nonstationary PEFs. Instead of estimating the filter in time and space, I estimate it in frequency and space by applying the Spitz (1991) approximation that links data at coarse sampling and low frequencies to data at finer sampling and higher frequencies, but instead of the patch-based approach of Spitz I estimate a single nonstationary PEF for each frequency. This new approach of nonstationary interpolation of frequency slices is much faster than both the nonstationary t-x based approach and the previous patch-based approach of Spitz in f-x. I apply this approach to interpolate both inline sources and crossline receivers in two, three, and four dimensions, and compare the results for the different choices of dimensionality and domain of the PEFs. I also iteratively interpolate the data to generate data with increases in density by factors of four and six in multiple dimensions, and show how the data degrades as the iterative interpolation proceeds.

In this thesis, I focus mainly on examining and comparing the interpolated data. In practice, these data are used in subsequent processing stages to produce a final migrated image and ultimately, a geologic interpretation. For the land data example, I

show a subsequent processing step used in velocity estimation, while in the subsequent marine data examples, the subsequent step is multiple prediction.

THESIS OVERVIEW

Chapter 2: PEFs and Interpolation

This chapter is a review of the background theory for this thesis. I review estimation of a prediction-error filter on fully-sampled training data, and interpolation of data with an estimated prediction-error filter. I show the extent to which use of inexact training data can degrade the interpolation result. I then review the estimation of nonstationary prediction-error filters, how to interpolate data using them, and demonstrate the interpolation with perfectly-known training data on a 3D synthetic model, comparing it to a patch-based approach.

Chapter 3: Interpolation of irregularly-sampled data

The previous chapter assumes that the training data for the PEF are completely known. I review how to exclude unknown data from the PEF estimation by weighting equations containing unknown data to zero, and then propose a new method of estimating a PEF when the data are irregularly-sampled and PEF estimation was previously impossible, with all equations weighted to zero. I use multiple copies of the data placed on grids that vary in both size and origin as training data, estimate both stationary and non-stationary PEFs on synthetic examples, and then interpolate to create a regular source-offset cube from an irregularly-sampled prestack 2D land data set from Colombia.

Chapter 4: Interpolation of near offsets with multiples

Here, I fill in the systemic near-offset gap in marine data by using information contained in free-surface multiple reflections. I cross-correlate traces within each shot

using the known principle that primaries correlated with free-surface multiples create pseudo-primaries, with one receiver acting as a virtual source. With the near-offsets of these pseudo-primaries as training data for non-stationary PEFs, I use the estimated PEFs to interpolate traces in the large near-offset gap. I test this method on the Sigsbee2B dataset and a large near-offset gap, using nonstationary filters in both time and, for the first time, in frequency and compare the results. I then apply this same approach to field data with a smaller gap. Finally, using a 3D synthetic model, I investigate the issues with using this approach in 3D, both in terms of crossline variability of 3D field data and the severely limited source distribution in the crossline direction.

Chapter 5: Nonstationary f-x interpolation of prestack 3D data

One common application of interpolation and regularization methods is for 3D surface-related multiple elimination, in which the data need to be regularly and fully sampled so that there is a source at every receiver location in both the inline and crossline directions. Data need to be created in the inline source direction as well as both the crossline receiver and crossline source directions.

After first reviewing Spitz's observation that the wavenumber spectra of a plane wave in the f-x domain at a single frequency and spatial sampling with that of a higher frequency and finer spatial sampling are linked, I then use coarser-sampled lower frequencies as training data for a complex-valued nonstationary PEF that I use to interpolate a higher frequency to a finer sampling. I demonstrate this method on multidimensional plane waves as well as a 3D synthetic model.

I apply the method to interpolate 3D prestack synthetic data containing many diffracted multiples. I investigate the different domains on which these filters can be estimated, as well as the dimensionality of the filter for interpolating both inline sources and crossline receivers. For iteratively interpolating data to the extent required for 3D SRME, I continue this investigation for a 3D field data example with

only four cables, pushing the nonlinear approach of interpolation previously interpolated data to its breaking point.

Chapter 6: Conclusions

I summarize the different approaches to generating training data for a PEF used to interpolate data, and highlight some future avenues of research. This non-stationary PEF-based approach is very generalizable, and can be tailored to other applications by varying the training data and domain used, and the tools present in inverse theory.