Robust reconstruction of aliased data using autoregressive spectral estimates

Mostafa Naghizadeh and Mauricio D. Sacchi

ABSTRACT
Autoregressive modeling is used to estimate the spectrum of aliased data. A region of spectral support is determined by identifying the location of peaks in the estimated spatial spectrum of the data. This information is used to pose a Fourier reconstruction problem that inverts for a few dominant wavenumbers that are required to model the data. Synthetic and real data examples are used to illustrate the method. In particular, we show that the proposed method can accurately reconstruct aliased data and data with gaps.

INTRODUCTION
During recent years, interpolation and reconstruction of seismic data has become an important topic for the seismic processing community. In general, logistic and economic constraints dictate the spatial sampling of seismic surveys. Wave fields are continuous, in other words, seismic energy reaches the surface of the earth everywhere in our area of study. The process of acquisition records a finite number of spatial samples of the continuous wave field generated by a finite number of sources. The latter leads to a regular or irregular distribution of sources and receivers. Many important techniques for removing coherent noise and imaging the earth interior have stringent sampling requirements which are often not met in real surveys. In order to avoid information losses, the data should be sampled according to the Nyquist criterion (Vermeer, 1990). When this criterion is not honored, reconstruction can be used to recover the data to a denser distribution of sources and receivers and mimic a properly sampled survey (Liu, 2004).

Seismic data reconstruction via signal processing approaches is a topic of current research interest in exploration seismology. During the last decade, important advances have been made in this area. Nowadays, signal processing reconstruction algorithms based on Fourier synthesis operators can cope with multidimensional sampling as demonstrated by several authors (Duijndam et al., 1999; Liu et al., 2004; Zwartjes and Gisolf, 2006; Schonewille et al., 2009). These methods are based on signal processing principles, they do not require information about the subsurface and, in addition, are quite robust in situations were the optimality condition under which they were designed are not completely satisfied (Trad, 2009).

Signal processing methods for seismic data reconstruction often rely on transforming the data to other domains. The latter can be achieved via the Fourier transform (Sacchi and Ulrych, 1996; Sacchi et al., 1998; Duijndam et al., 1999; Liu et al., 2004; Xu et al., 2005; Zwartjes and Gisolf, 2006), the Radon transform (Darche, 1990; Verschuur and Kabir, 1995; Trad et al., 2002), the local Radon transform (Sacchi et al., 2004; Wang
et al., 2009) and the curvelet transform (Hennenfent and Herrmann, 2008; Herrmann and Hennenfent, 2008). Another group of signal processing interpolation methods rely on prediction error filtering techniques (Spitz, 1991) and (Porsani, 1999). Interpolation with prediction filters operates in the frequency-space ($f$-$x$) domain. Low frequency data components in a regular spatial grid are used to estimate the prediction filters needed to interpolate high frequency data components. An equivalent interpolation method in the frequency-wavenumber ($f$-$k$) domain was introduced by Gulunay (2003) and often referred as $f$-$k$ interpolation. Prediction filters are also the cornerstone of $f$-$x$ noise attenuation method (Canales, 1984).

Spitz (1991) showed how one could extract prediction filters from spatial data at low frequencies to reconstruct aliased spatial data. This idea was expanded by Naghizadeh and Sacchi (2007) and used it to reconstruct data with irregular distribution of traces on a grid. The latter is named multi-step autoregressive reconstruction. The multi-step autoregressive reconstruction method is a combination of a Fourier reconstruction method (Duijndam et al., 1999; Sacchi et al., 1998; Liu and Sacchi, 2004; Zwartjes and Gisolf, 2006) and $f$-$x$ interpolation (Spitz, 1991) and it can be summarized as follows:

1. The low frequency (spatially unaliased) portion of data is restored using Fourier reconstruction. In general, the low-frequency portion of the $f$-$k$ spectrum of the seismic data is not contaminated with aliased energy (commonly known as wrap-around energy). Therefore, band-limited Fourier operators can robustly reconstruct the low frequency portion of the data (Duijndam et al., 1999). In this paper, the low-frequency portion of the data is reconstructed via a Fourier reconstruction algorithm called minimum weighted norm interpolation (Liu and Sacchi, 2004; Sacchi and Liu, 2005; Liu et al., 2004; Trad, 2009).

2. Prediction filters of all frequencies are extracted from already regularized low frequency spatial data. This is an interesting property of the multi-step autoregressive method which allows to estimate prediction filter for high frequencies by just using data from unaliased low frequencies. This guarantees that the estimated prediction filters for the high frequencies will only contain information from the unaliased data.

3. The estimated prediction filters are used to reconstruct missing spatial samples in the aliased portion of the spectrum.

Stages 1) and 2) are estimation stages and 3) is the reconstruction stage. In this paper we propose a new and robust method to solve the reconstruction stage. In the original formulation of the multi-step autoregressive reconstruction method (Naghizadeh and Sacchi, 2007) the reconstruction stage uses prediction filters extracted from low frequencies to reconstruct spatial data in the aliased band. In this article, on the other hand, we use the spectrum derived from the predictor filters to estimate the dominant wavenumbers needed to reconstruct the data.

The multi-step autoregressive method is a hybrid reconstruction technique conceived to overcome the shortcomings of Fourier interpolation techniques (Duijndam et al., 1999; Liu and Sacchi, 2004) and $f$-$x$ prediction filter interpolation (Spitz, 1991). Fourier interpolation methods can fail in the presence of aliased data. On the other hand, they can handle irregularly sampled data quite well. Methods based on $f$-$x$ prediction filters can
cope with alias but, they were designed to work on regularly sampled data. The multistep autoregressive reconstruction method successfully combines the advantages of both of the aforementioned reconstruction methods (Naghizadeh and Sacchi, 2007). In the original formulation of multi-step autoregressive reconstruction the prediction filters derived from low frequencies were used to reconstruct spatial data in the aliases portion of the spectrum. This methodology works quite well when missing traces do not form large gaps. In the presence of large gaps and/or data extrapolation problems we have found that a more robust way of reconstructing seismic data is by using the spectrum of the data, computed from the prediction filters, rather than using the prediction filters themselves. In essence, we use low-frequency estimators of prediction filters to compute the amplitude spectrum of the data at potentially aliased high-frequency spatial data. The methodology entails estimating the dominant wave-numbers that are needed to model the data at a given monochromatic temporal frequency. This also implies that the data are composed of a finite number of dominant wavenumbers. To validate this assumption, the interpolation must be carried out in small spatial windows. In addition, it is important to point out that our method interpolates missing data on a regular grid rather than reconstructing data with observations on an irregular grid (Xu et al., 2005).

Our article shows with synthetic examples that the proposed method can handle gaps and extrapolation problems much better than our original formulation of multi-step autoregressive reconstruction (Naghizadeh and Sacchi, 2007, 2010). For completeness, we also applied the proposed reconstruction algorithm to a real data set.

THEORY

Let \(d(x_h, f)\) represent the seismic data in the \(f-x\) domain where \(x_h\) indicates given spatial positions of the available seismic traces. The subindex \(h\) indicates the indices of the available spatial samples. Consider spatial data at low frequencies \(f \in [f_{\text{min}}, f_{\text{max}}]\), where \(f_{\text{min}}\) and \(f_{\text{max}}\) denote the minimum and maximum (unaliased) frequencies in the data. In general, due to the band-limited nature of the seismic wavelet, we consider \(f_{\text{min}} > 0\). In addition, we consider that we have reconstructed equally spaced data \((x_n = (n-1)\Delta x, n = 1, 2, \ldots, N)\) in the aforementioned frequency band. In other words, \(d(x_h, f)\) was used to estimate \(d(x_n, f)\) for the frequencies \(f \in [f_{\text{min}}, f_{\text{max}}]\). This is not a difficult task and can be achieved via various methods including minimum weighted norm interpolation (Liu and Sacchi, 2004). A short outline of minimum weighted norm interpolation is provided in Appendix A. Now, the prediction filters of all frequencies can be estimated using the multi-step autoregressive method (Naghizadeh and Sacchi, 2007, 2010) via forward and backward prediction equations

\[
d(x_n, f) = \sum_{m=1}^{M} P_m(\alpha_f) d(x_{n-am}, f) + \epsilon_f(x_n, f), \quad \begin{cases} n = \alpha M + 1, \ldots, N, \\ f \in [f_{\text{min}}, f_{\text{max}}], \end{cases}
\]

\[
d^*(x_n, f) = \sum_{m=1}^{M} P_m(\alpha_f) d^*(x_{n+am}, f) + \epsilon_b(x_n, f), \quad \begin{cases} n = 1, \ldots, N - \alpha M, \\ f \in [f_{\text{min}}, f_{\text{max}}], \end{cases}
\]

where the symbol * indicates conjugate, \(P_m\) is the \(m\)-th component of the prediction filter with length equal to \(M\) and \(\alpha = 1, 2, \ldots, \alpha_{\text{max}}\) is the step factor used to extract the prediction filter for frequency \(\alpha f\) from frequency \(f\). The signals \(\epsilon_f(x_n, f)\) and \(\epsilon_b(x_n, f)\) denote forward and backward modeling errors, respectively. Notice that the final output of
the multi-step autoregressive algorithm is the estimated prediction filters for the frequencies \( f \in [f_{\text{min}}, \alpha f_{\text{max}}] \). The parametric autoregressive spectrum of the data can be computed from the prediction filter via the following expression (Ulrych and Bishop, 1975; Marple, 1987)

\[
S_{\text{AR}}(\kappa, f) = \frac{\sigma^2}{|1 - \sum_{m=1}^{M} P_m(f)e^{-i2\pi m\kappa}|^2}, \quad -\frac{1}{2} \leq \kappa \leq \frac{1}{2},
\]

(3)

where, \( \kappa \) is the normalized wavenumber and \( \sigma^2 \) is the noise variance. The parametric autoregressive spectrum (equation 3) of the data is a continuous function with peaks at the dominant wavenumbers present in the data. The location and number of peaks are accurately represented. At the same time, the spectral amplitudes are not a direct estimate of the amplitude of the signal in the \( f-x \) domain (Marple, 1987). The key information needed for data reconstruction, however, is the location and number of dominant wavenumbers (Xu et al., 2005; Schonewille et al., 2009; Ozdemir et al., 2008). A basic sample-by-sample comparison algorithm can identify the location of the peaks in the parametric autoregressive spectrum. Considering a discrete axis of wavenumbers \( \kappa = (k_1, k_2, \ldots, k_q) \), the location of the spectral peaks are found by the following algorithm

\[
\Lambda(k_j, f) = \begin{cases} 
1 & \text{if } S_{\text{AR}}(k_{j-1}, f) < S_{\text{AR}}(k_j, f) > S_{\text{AR}}(k_{j+1}, f) \\
0 & \text{otherwise} 
\end{cases}.
\]

(4)

The function \( \Lambda \) will be equal to one at the location of the spectral peaks and zero everywhere else. Uncertainties in the estimation of the prediction filter are handled by widening the position of the spectral peaks via convolution of a boxcar function \( B(k, f) \) with \( \Lambda(k, f) \). Then the region of spectral support is defined by the set of wavenumbers \( k_j(f), j = 1, \ldots, NK(f) \), where \( B(k, f) \ast \Lambda(k, f) = 1 \). We can now consider the problem of finding the Fourier coefficients that represent the available data

\[
d(x_h, f) = \sum_{j=1}^{NK(f)} D(k_j(f), f)e^{i2\pi k_j x_h}.
\]

(5)

The latter is solved by minimizing the following cost function

\[
J(f) = \sum_h \|d(x_h, f) - \sum_{j=1}^{NK(f)} D(k_j(f), f)e^{i2\pi k_j x_h}\|^2.
\]

(6)

The minimization of equation 6 leads to the standard least-squares Fourier reconstruction solution (Duijndam et al., 1999). It is important to stress that we are inverting for a few wavenumbers and therefore, in general, the problem is over-determined and stable. The proposed algorithm can be summarized as follows:

1. Transform the data from \( t-x \) to \( f-x \) domain and reconstruct the low frequency portion of the data using a Fourier reconstruction method. We have adopted for this purpose minimum weighted norm interpolation (Liu and Sacchi, 2004).

2. Use the multi-step autoregressive reconstruction algorithm (Naghizadeh and Sacchi, 2007) to extract the prediction filters of all frequencies from the low frequency reconstructed data.
3. Compute the parametric autoregressive spectrum of the data using expression 3.

4. Identify the location of spectral peaks and define regions of spectral support.

5. Solve Equation 6 to estimate \( D(k, f) \) and to predict data at new spatial locations.

6. Use the inverse Fourier to transform \( D(k, f) \) to \( d(x, f) \).

7. Transform the \( f\)-\( x \) data to the \( t\)-\( x \) domain.

Steps 3-6 are the new contributions proposed in this article. For data that depend on more than one spatial dimension one can utilize multidimensional autoregressive operators and their associated spectra (Kumaresan and Tufts, 1981). Naghizadeh and Sacchi (2009a) provide a detailed formulation of multidimensional autoregressive spectrum estimation and its application to seismic data reconstruction.

SYNTHETIC TESTS

We first analyze the performance of the proposed algorithm using synthetic data examples. Amplitude spectra are portrayed in terms of normalized frequency and normalized wavenumber. Normalized frequency and wavenumber axes are obtained by considering \( \Delta t = 1 \) and \( \Delta x = 1 \), respectively. This means that in order to obtain frequency axes in Hz and Cycles/m one must divide the normalized ones by \( \Delta t \) and \( \Delta x \), respectively. We will make a clear distinction of methodologies in order to make clear our comparisons. We will utilize the following methods:

**Method A.** The minimum weighted norm interpolation method (Liu and Sacchi, 2004) is used to reconstruct the data for all the temporal frequencies. The latter will be shown to only work when the problem entails reconstructing data that have undergone random sampling. Minimum weighted norm interpolation is described in Appendix A.

**Method B.** The minimum weighted norm interpolation is used to reconstruct the unaliased part of the spectrum, then multi-step autoregressive reconstruction is utilized to estimate prediction filters. The reconstruction of aliased data components uses prediction filters following the algorithm described in Naghizadeh and Sacchi (2007) and Naghizadeh and Sacchi (2010). In other words, the prediction filters are used to define a system of equations where missing samples become the unknown variables.

**Method C.** The minimum weighed norm interpolation is used to reconstruct the unaliased part of the spectrum, then multi-step autoregressive reconstruction is utilized to estimate prediction filters. The reconstruction of the aliased data components is implemented with the algorithm proposed in this paper. In other words, prediction filters from the unaliased portion of the data are used to identify the alias-free wavenumbers needed to model the data. In essence, this is the new methodology introduced in this paper.

We will also compare Methods A, B and C for different sampling scenarios. In particular, we will analyze the effect of random, and regular sampling with and without the presence of large gaps.
Randomly sampled data

The first synthetic data example consists of three severely aliased linear events (Figure 1a). We have eliminated 70% of the traces from the original data to obtain a section with randomly missing traces (Figure 1b). Figures 1c and 1d depict the $f$-$k$ spectra of Figures 1a and 1b, respectively. First, the missing traces are reconstructed using minimum weighted norm interpolation for all frequencies (Method A). The results are provided in Figure 2a.

Figure 2b shows the low frequency unaliased part of the data reconstructed via minimum weighted norm interpolation ($f \in [0.03, 0.07]$). These data were used to reconstruct the aliased portion of the data using Method B (Figure 2c) and Method C (Figure 2d). The amplitude spectrum associated to Figures 2a-d are provided by Figures 2e-h.

Figure 3 shows the parametric spectrum of the data (solid line with solid circles) for the normalized frequency 0.3 in conjunction with the region of spectral support (Dashed line). It is important to stress that the spectrum portrayed by this figure was computed using prediction filters estimated from unaliased low frequencies. This explains the presence of only 3 dominant wavenumbers.

Regular data decimation

The previous example illustrates a case with randomly missing samples. In this case, interpolation/reconstruction becomes an easy task because aliased events are distorted by random sampling (Hennenfent and Herrmann, 2008; Naghizadeh and Sacchi, 2009b). We now consider the more difficult case of regularly decimated data. In this case, we have eliminated every other traces of the original data (Figure 1a) to create a section with regularly missing traces (Figure 4a). The missing data in Figure 4a is reconstructed using minimum weighted norm interpolation (Method A). The results are provided by Figure 4b. Methods B and C were also used to reconstruct the data with results portrayed in Figures 4c and d. Again, the spectra of Figures 4a-d is provided by Figures 4e-h. The minimum weighted norm interpolation (Method A) fails to reconstruct the missing data. The artificial events created by regular sampling (alias) are of the same amplitude of the original spectrum of the data. Therefore, the minimum weighted norm interpolation method fails to eliminate the unwanted events (Naghizadeh and Sacchi, 2009b). On the other hand, Methods B and C were able to regularize the data. Figure 5 shows the autoregressive spectrum of the data (solid line with solid circles) as well as the derived region of spectral support (dashed line) for the normalized frequency 0.3. The identified dominant wavenumbers are the same as the ones in the original spectrum of data.

Data with gap

Next we examine the reconstruction of a data set with a gap (Figure 6a and 6e). The reconstruction with minimum weighted norm interpolation (Method A) is presented in Figures 6b and 6f. Reconstruction with the original formulation of multi-step autoregressive reconstruction (Method B) is provided in Figures 6c and 6g. The reconstruction using Method C are provided by Figures 6d and 6h. Figure 7 shows the autoregressive spectrum of the data (solid line with solid circles) as well as the region of spectral support (dashed line).
Figure 1: 2D synthetic example. a) Original data. b) Data with 70% of traces randomly removed. c) and d) are the f-k spectra of (a) and (b), respectively.
Figure 2: Reconstruction of the data portrayed in Figure 1b. a) Minimum weighted norm interpolation is applied to all the temporal frequencies to reconstruct the data (Method A). b) Minimum weighted norm interpolation results for the normalized frequency band \( f \in [0.03, 0.07] \). c) Multi-step autoregressive Reconstruction where prediction filters extracted from \( f \in [0.03, 0.07] \) were used to reconstruct the aliased portion of the data (Method B). d) Reconstruction using Method C; prediction filters extracted from \( f \in [0.03, 0.07] \) are used to determine the dominant wavenumbers needed to reconstruct the aliased portion of the data. e), f), g), and h) are the \( f-k \) spectra of (a), (b), (c), and (d), respectively.

Figure 3: The autoregressive spectrum of the data at normalized frequency \( f = 0.3 \) for the data in Figure 2 (Solid line). The prediction filter was estimated from unaliased low-frequency data (Figure 2b). The region of spectral support (dashed line) highlights the wavenumber that were used to reconstruct the data at \( f = 0.3 \) using the least-squares method.
Figure 4: a) Data with regularly missing traces. b) Minimum weighted norm interpolation is applied to all the temporal frequencies to reconstruct the data (Method A). c) Multi-step autoregressive reconstruction where prediction filters extracted from $f \in [0.03, 0.07]$ are used to reconstruct the aliased portion of the data (Method B). d) Reconstruction using Method C; prediction filters extracted from $f \in [0.03, 0.07]$ are used to determine the dominant wavenumbers needed to reconstruct the aliased portion of the data. e), f), g), and h) are the $f$-$k$ spectra of (a), (b), (c), and (d), respectively.

Figure 5: The autoregressive spectrum of the data at normalized frequency $f = 0.3$ for the data in Figure 4a (Solid line). The prediction filter was estimated from unaliased low frequencies. The region of spectral support (dashed line) highlights the wavenumber that were used to reconstruct the data at $f = 0.3$ using the least-squares method.
Figure 6: a) Data with a gap in the center of section. b) Minimum weighted norm interpolation is applied to all the temporal frequencies to reconstruct the data (Method A). b) Multi-step autoregressive reconstruction where prediction filters extracted from $f \in [0.03, 0.07]$ are used to reconstruct the aliased portion of the data (Method B). d) Reconstruction using Method C; prediction filters extracted from $f \in [0.03, 0.07]$ are used to determine the dominant wavenumbers needed to reconstruct the aliased portion of the data. e), f), g), and h) are the $f$-$k$ spectra of (a), (b), (c), and (d), respectively.

Extrapolation exercise

It is also interesting to investigate the case of data extrapolation; another important facet of the reconstruction problem. Figure 8a was created by eliminating 7 traces from each side of the original data (Figure 1a). Eliminating traces from the sides of the data widens the original spectrum of data in the $f$-$k$ domain (Figure 8e). Figures 8b and 8f shows the reconstruction using minimum weighted norm interpolation (Method A). Reconstructed data using methods B and C are displayed by Figures 8c and 8d, respectively. The spectra of Figures 8c and 8d are shown in Figures 8c and 8d, respectively. The autoregressive spectrum of the data and the derived region of spectral support for the normalized frequency 0.3 is shown in Figure 9.
Figure 7: The autoregressive spectrum of the data at normalized frequency $f = 0.3$ for the data in Figure 6 (Solid line). The prediction filter was estimated from unaliased low frequencies. The region of spectral support (dashed line) highlights the wavenumber that were used to reconstruct the data at $f = 0.3$ using the least-squares method.

Figure 8: a) Data that require trace extrapolation. b) Minimum weighted norm interpolation is applied to all the temporal frequencies to reconstruct the data (Method A). c) Multi-step autoregressive reconstruction where prediction filters extracted from $f \in [0.03, 0.07]$ are used to reconstruct the aliased portion of the data (Method B). d) Reconstruction using Method C; prediction filters extracted from $f \in [0.03, 0.07]$ are used to determine the dominant wavenumbers needed to reconstruct the aliased portion of the data. e), f), g), and h) are the $f$-$k$ spectra of (a), (b), (c), and (d), respectively.
Figure 9: The autoregressive spectrum of the data (Solid line with solid circles) and region of spectral support (dashed line) for the normalized frequency 0.3 for the data in Figure 8.

**Mixture of regular decimation and gap**

Figure 10a shows an example of mixed sampling. First, the data was regularly decimated and, then, central traces were removed to simulate a gap. Figures 10b, 10c, and 10d show the reconstructed data using minimum weighted norm interpolation (Method A), Method B and Method C, respectively. The f-k panel of Figures 10a-d are depicted in Figures 10e-f, respectively. It is interesting to notice that the minimum weighted norm interpolation method was only able to recover a decimated version of the data in the gap. Method B, on the other hand, successfully interpolates the data outside the gap. Method C overcomes the shortcomings of both minimum weighted norm interpolation and multi-step autoregressive reconstruction methods and recovers all the missing traces. The autoregressive spectrum of the data and region of spectral support or the normalized frequency 0.3 is shown in Figure 11.

**Extrapolation and interpolation of regularly decimated data**

The next example represents a section with decimated traces as well as missing traces on the sides (Figure 12a). In this case the reconstructed data using minimum weighted norm interpolation (Method A) (Figure 12b) and multi-step autoregressive Reconstruction (Figure 12c) are not fully recovered. Method C, on the other hand, recovers all the missing data. The f-k panels of Figures 12a-d are displayed by Figures 12e-h, respectively. Figure 13 shows the autoregressive spectrum of data (solid line) and the region of spectral support (dashed line) for the normalized frequency 0.3.

**REAL DATA EXAMPLES**

We apply the proposed technique to the reconstruction of a shot record from a marine data set from the Gulf of Mexico. About 50% of the traces were randomly removed from the original section (Figure 14a) to simulate a section with missing traces (Figure 14b). We used minimum weighted norm interpolation to reconstruct only the low frequency portion of the data for normalized frequencies in the range 0.02 to 0.06 (Figure 14c). After extracting
Figure 10: a) Data that require gap filling and interpolation. b) Minimum weighted norm interpolation is applied to all the temporal frequencies to reconstruct the data (Method A). b) Multi-step autoregressive reconstruction where prediction error filters extracted from $f \in [0.03, 0.07]$ are used to reconstruct the aliased portion of the data (Method B). d) Reconstruction using Method C; prediction filters extracted from $f \in [0.03, 0.07]$ are used to determine the dominant wavenumbers needed to reconstruct the aliased portion of the data. e), f), g), and h) are the f-k spectra of (a), (b), (c), and (d), respectively.

Figure 11: Autoregressive spectrum of the data (Solid line with solid circles) and region of spectral support (dashed line) for the normalized frequency 0.3 for the data in Figure 10d.
Figure 12: a) Data that require extrapolation and interpolation. b) Minimum weighted norm interpolation is applied to all the temporal frequencies to reconstruct the data (Method A). b) Multistep autoregressive reconstruction where prediction error filters extracted from $f \in [0.03, 0.07]$ are used to reconstruct the aliased portion of the data (Method B). d) Reconstruction using Method C; prediction filters extracted from $f \in [0.03, 0.07]$ are used to determine the dominant wavenumbers needed to reconstruct the aliased portion of the data. e), f), g), and h) are the f-k spectra of (a), (b), (c), and (d), respectively.

Figure 13: The autoregressive spectrum of the data (Solid line with solid circles) and region of spectral support (dashed line) for the normalized frequency 0.3 for the data shown in Figure 12.
Figure 14: 2D real data example from the Gulf of Mexico. a) A window of the original shot gather. b) Data with randomly missing traces. c) Minimum weighted norm interpolation of the low frequency portion of the data. d) Reconstructed data using the method outlined in this paper (Method C). e), f), g), and h) are the f-k panel of (a), (b), (c), and (d), respectively.

Prediction filters from low frequency data components using the multi-step autoregressive reconstruction algorithm, the complete data spectrum was reconstructed using Method C proposed in this paper (Figure 14d). For this particular test, the length of the prediction filters is set to \( L = 3 \). Figures 14e-h show the f-k panel of Figures 14a-d, respectively.

Figure 15a shows the original data in Figure 14a after decimation of every other traces. The data in Figure 15a is reconstructed via the proposed method C. Figures 15e and 15f shows the f-k panel of Figures 15a and 15b, respectively. Figure 15c shows the original data in Figure 14a after placing a gap in the middle of the section. The method described in this paper (Method C) yields the reconstruction provided in Figure 15d. Figures 15h and 15g shows the f-k panel of Figures 15c and 15d, respectively. Figures 16a, 16b, and 16c show the error panels between the interpolated data in Figures 14d, 15b, and 15d and the original data in Figure 14a, respectively. It is clear that it is more difficult, in the real data case to interpolate large gaps. Notice that we have reconstructed the complete window rather than only the missing traces. It is clear, however, that it is possible to reinsert the original traces in the reconstructed volume.
Figure 15: 2D real data example from Gulf of Mexico. a) Regular decimation of the original data in Figure 14a. b) Reconstruction of (a) using the method proposed in this paper (Method C). c) Original data of Figure 14a with a gap in the center of the gather. d) Reconstruction of (c) using the method proposed in this paper (Method C). e), f), g), and h) are the f-k panel of (a), (b), (c), and (d), respectively.

Figure 16: a), b), and c) are the difference sections between the interpolated data in Figures 14d, 15b, and 15d and the original data in Figure 14a, respectively.
CONCLUSIONS

Parametric spectral analysis can accurately determine the location of the dominant wavenumbers at each single frequency of the seismic data. The latter permits us to define regions of spectral support that can be used to define a beyond alias, stable and robust Fourier reconstruction. The proposed reconstruction method overcomes the shortcomings of classical Fourier reconstruction methods in the case of aliased and regularly decimated data. In addition, our examples show that it improves the original formulation of the multi-step autoregressive reconstruction algorithm at the time of reconstructing large gaps of missing data. One of the objectives of our tests was to show that the proposed method (method C) works well under diverse choices of sampling function. In our synthetic examples method C was successful for all of the sampling functions while methods A and B failed in some of the examples.

It is important to stress that prediction filters from the unaliased low frequency portion of the data are used to estimate the spatial spectra at high (and potentially aliased) frequencies. Therefore, the region of spectral support estimated by our procedure corresponds to true seismic events. In essence, we rely on low frequency information to regularize aliased data (Spitz, 1991). The procedure outlined in this paper is an improvement to the multi-step autoregressive Reconstruction method (Naghizadeh and Sacchi, 2007, 2010) previously proposed to reconstruct irregularly sampled data using prediction filters. The new algorithm provides the following improvements: stability at the time of reconstructing severely decimated data and accurate reconstruction in the presence of large gaps of missing data.

ACKNOWLEDGEMENTS

We would like to acknowledge financial support of the sponsors of the Signal Analysis and Imaging Group at the University of Alberta.

APPENDIX A

MINIMUM WEIGHTED NORM INTERPOLATION

The minimum weighted norm interpolation method is described in detail in Liu (2004) Liu and Sacchi (2004) and Trad (2009). For completeness, we have added a short outline of the method.

Interpolation of band-limited data with missing samples can be summarized in the following inversion scheme:

\[
\text{Minimize } \|Gd - d_{\text{obs}}\|_W^2 + \mu\|d\|_W^2
\]

where \(\|.\|_W^2\) indicates a weighted norm, \(\mu\) is the trade-off parameter, and \(G\) is the sampling matrix which maps desired data samples \(d \equiv d(x_n, f)\) to available samples \(d_{\text{obs}} \equiv d(x_h, f)\) at a given frequency \(f\). The transpose of sampling matrix, \(G^T\), fills the position of missing samples with zeros. A regularization norm can be selected in the wavenumber domain as follows:

\[
\|d\|_W^2 = \sum_{k \in K} \frac{D_k^*D_k}{W_k^2},
\]

(A-2)
where \( D_k \) indicates the coefficients of the Fourier transform of the vector of spatial data \( d \). The values of \( W_k \) determine the type of interpolation. For band-limited minimum weighted norm interpolation a diagonal matrix is defined as:

\[
\Upsilon_k = \begin{cases} 
W_k^2 & k \in \mathcal{K} \\
0 & k \notin \mathcal{K}
\end{cases}
\]  

where \( \mathcal{K} \) indicates the region of support of the Fourier transform. The pseudoinverse of \( \Upsilon \) is defined as:

\[
\Upsilon_k^\dagger = \begin{cases} 
W_k^{-2} & k \in \mathcal{K} \\
0 & k \notin \mathcal{K}
\end{cases}
\]  

For band-limited minimum norm interpolation, the values of \( W_k \) are equal to one, while for minimum weighted norm interpolation, their values must be iteratively updated to find an optimal reconstruction. The minimizer of the cost function (A-1) is given by

\[
\hat{d} = F^H \Upsilon FG^T (GFH \Upsilon FG^T + \alpha I)^{-1} d^{obs},
\]  

where \( F \) is the Fourier matrix, \( \alpha \) is trade-off parameter, \( I \) is the identity matrix, \( T \) and \( H \) stand for transpose and Hermitian transpose operators, respectively. For further details see Liu (2004) and Liu and Sacchi (2004).

**REFERENCES**


