doi:10.1007/s11589-012-0859-x

A new approach for high fidelity seismic data recovery by fractal interpolation*

Hongyan Liu[1,♣] Tongjiang He[1] Yukun Chen[1] and Xinfu Li[2]

[1] *Earthquake Administration of Tianjin Municipality, Tianjin* 300201, *China*
[2] *School of Geophysics and Information Technology, China University of Geosciences, Beijing* 100083, *China*

Abstract Recovering accurate data is important for both earthquake and exploration seismology studies, when data are sparsely sampled or partially missing. We present a method that allows for precise and accurate recovery of seismic data using a localized fractal recovery method. This method requires that the data are self-similar on local and global spatial scales. We present examples that show that the intrinsic structure associated with seismic data can be easily and accurately recovered by using this approach. This result, in turn, indicates that seismic data are indeed self-similar on local and global scales. This method is applicable not only for seismic studies, but also for any field studies that require accurate recovery of data from sparsely sampled datasets with partially missing data. Our ability to recover the missing data with high fidelity and accuracy will qualitatively improve the images of seismic tomography.

Key words: **fractal interpolation; seismic data recovery; high-fidelity; seismic tomography**

CLC number: P315.3 **Document code**: A

1 Introduction

In seismic studies, it is crucial to have a sufficient amount of data that represent the study area, and allow for meaningful conclusions to be drawn. However, in practice, seismic data are usually collected from sparsely sampled stations. In many cases, data may be collected from seismic arrays with missing traces or stations because of reasons such as: malfunctioning instrumentation, presence of obstacles, presence of no-access areas, feathering, dead traces, and/or economic considerations. Most multi-trace data processing algorithms cannot adequately handle irregular sampling problems. These difficulties make it challenging to obtain detailed information on subterranean structures, and to avoid spatial aliasing. Numerical recovery of seismic data has therefore become a valuable tool in amending unevenly distributed datasets. In particular, interpolating seismic traces or recovering seismic data from one specific seismometer is common when reconstructing seismic datasets, but unfortunately, conventional recovery methods typically fail to provide details that are useful in revealing the fine structures presented in seismic wavefields. Thus, it is needed to develope a new recovery approach that can provide detailed, accurate recovery of seismic traces.

Conventional interpolation approaches (Larner et al., 1981; Kabir and Verschuur, 1995; Hindriks and Duijndam, 2000; Trad et al., 2003; Wang, 2002; Liu and Sacchi, 2004; Zwartjes and Sacchi, 2007) are valid for the de-aliasing of seismic events, seismic data are usually self-similar on both local and global scales. Standard interpolation recovery approaches are not optimal methods for dealing with data that exhibit local-global self-similarity. Using fractal methods to recover seismic data may yield ideal results, since fractals are themselves self-similar on local and global spatial scales. In contrast, non-fractal interpolation recovery methods are easily adapted to smooth data sets, but usually ignore detailed structures and local properties of more complicated data sets, including data sets with irregularities, or sampling gaps.

Conventional fractal recovery methods based on the Random Fractional Brown Motion model (Barns-

* **Received** 11 May 2012; **accepted in revised form** 20 July 2012; **published** 10 August 2012.
♣ **Corresponding author. e-mail**: liuhongyan_012@163.com

ley and Demko, 1985; Fan et al., 2005) assign the same value for the vertical scaling factor, which plays the most important role in the whole recovery process throughout the entire interpolated interval. This generalization makes it difficult to obtain the important information that carries details about the local properties of the data. In addition, conventional fractal interpolation methods do not have any explicit expression, so the approach is often a complicated, iterative process.

2 Basic formulas

In this paper, we employ explicitly localized, fractal interpolation functions within interpolating intervals that use explicit-variable vertical scaling factors throughout the whole data recovery process. Consequently, we are able to successfully remove the abovementioned limitations of conventional fractal recovery methods by using affine transform and iterative function systems. The basic idea is that the localized fractal interpolation function used to recover the missing seismic data is derived from the concepts of affine transform and the iterated function systems. We assume here that the data are self-similar on local and global spatial scales, and that the vertical scaling factors, which play the most important role during the recovery process, are explicitly represented by the variable endpoints of the local interpolating intervals (Barnsley and Demko, 1985; Singer and Zajdler, 1999; Chu and Chen, 2003; Navascués and Sebastián, 2004; Fan et al., 2005). Generally, the two-dimensional affine transform (map) is defined by:

$$\omega_n \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_n & 0 \\ c_n & d_n \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_n \\ f_n \end{pmatrix} = \begin{pmatrix} L_n(x) \\ F_n(x,y) \end{pmatrix}, \quad n = 1, 2, \cdots N, \tag{1}$$

where ω_n is a contracting mapping. The zero in the coefficient matrix guarantees that the resulting linear fractal interpolation function will be a single value (Barnsley 1986). There are N maps for the N+1 interpolation points, and the interpolation points are required to be single-valued with respect to their first index.

Usually, $L_n(x)$ and $F_n(x, y)$ are defined as localized linear functions that have the expression

$$L_n(x) = a_n x + e_n$$

and

$$F_n(x, y) = c_n x + d_n y + f_n, \tag{2}$$

respectively. Each affine map is constrained to map the endpoints of the set of interpolation points to two consecutive interpolation points, which are

$$\omega_n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix}, \omega_n \begin{pmatrix} x_N \\ y_N \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix}, \quad n = 1, 2, \cdots N, \tag{3}$$

where (x_{n-1}, y_{n-1}) and (x_n, y_n) are two consecutive interpolation points with $x_n > x_{n-1}$. (x_0, y_0) and (x_N, y_N) are the endpoints associated with the interpolation points. Once the interpolation points are chosen, the free parameters of each affine map are then determined.

The adjustable parameter $d_n(|d_n| < 1)$, which is called the vertical scaling factor, plays a key role in the data recovery.

3 Time-domain implementation

If we assume that $y_{\max}$ and $y_{\min}$ are the maximum and minimum over the interval $[x_{n-l}, \quad x_{n+l}]$, and ζ is determined by $\zeta = 1.0 + \text{random}(n)$, where random$(n)$ is a random deviate generated from a uniform (0, 1) distribution, then the vertical scaling factor d_n can be represented by

$$d_n = \frac{y_n - y_{n-1}}{\zeta \sqrt{(y_{\max} - y_{\min})^2 + (y_n - y_{n-1})^2}}. \tag{4}$$

Because random(n) is unique at the nth point, the vertical scaling factor d_n is unique. Apparently, the vertical scaling factor d_n calculated in this way ranges from $-\sqrt{2}/2$ to $\sqrt{2}/2$. From equations (1), (2), (3) and (4), the high fidelity, localized, fractal recovery approach can be constructed.

Feng and Zhou (1998) discussed the stability of one approach that is used in this paper. Wang and Fan (2011) discussed the analytical characteristics of fractal interpolation functions with function-vertical scaling factors. Their results guaranteed that our method is feasible in recovering the seismic data.

To increase the method's efficiency, we introduce the inorder traversing binary tree algorithm into the recovery process (see appendix).

4 Influence of the vertical scaling factors to the accuracy of fractal recovery

In order to investigate how the vertical scaling factors influence the accuracy of fractal recovery proposed in this paper, we use the Ricker wavelet as the test func-

tion to analyze the variation tendency of fractal recovery errors with the variation of vertical scaling factors.

We carry out the experiments with fixed and varying vertical scaling factors. The results demonstrate that the explicit expression of the vertical scaling factors significantly improves the accuracy of fractal recovery. The expression of the wavelet is

$$f(t) = [1 - 2(\pi f_0 t)^2]\mathrm{e}^{-(\pi f_0 t)^2}, \qquad (5)$$

where the constant parameter f_0 is the central frequency of the wavelet, and t is the time.

Table 1 gives the values of the fixed vertical scaling factors and the corresponding recovery errors. Figure 1a shows the reference curve and the recovered curves computed by the fractal recovering approach with different fixed vertical scaling factors. The recovery error increases proportionally with an increase in the vertical scaling factors. Figure 1b illustrates the relationship between the recovery errors and the values of the vertical scaling factors.

Figure 1a shows that the recovered curves deviate from the theoretical curve with an increase in the vertical scaling factors, which indicates that the recovery errors increase proportionally with an increase in vertical scaling factors. Figure 1b demonstrates this observation more clearly.

Table 1 Vertical scaling factors and the corresponding recovery errors for the Ricker wavelet

d_n	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.90
Error (%)	0.016	0.06	0.12	0.20	0.30	0.46	0.70	1.13	2.14	3.41

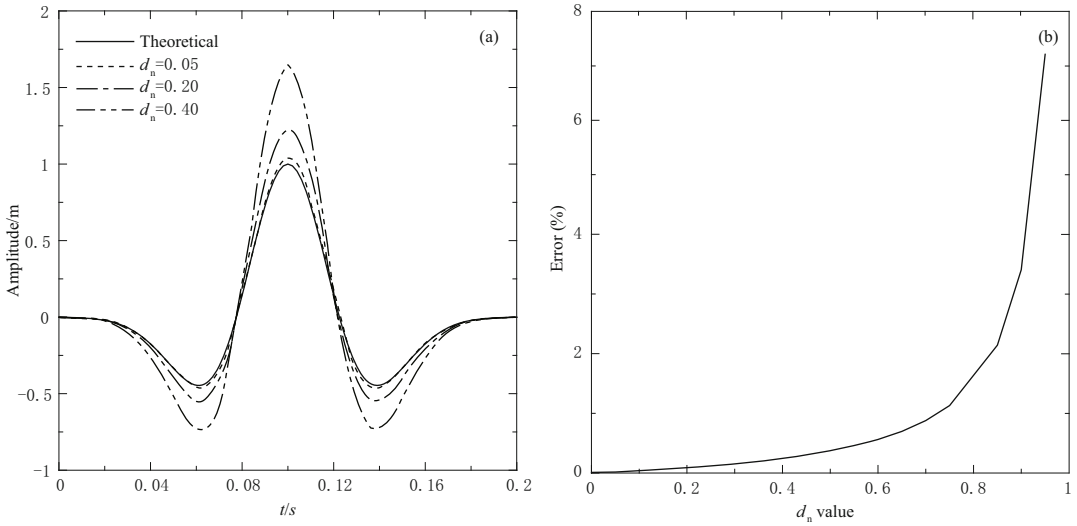

Figure 1 Recovery errors of the Ricker wavelet. (a) Reference curve and the recovered curves computed by fractal recovering approach with different fixed vertical scaling factors. The solid line is the reference curve and the dashed lines are the recovered curves with different vertical scaling factors. (b) Variation of the recovery errors to the vertical scaling factors.

In order to compare the accuracy of the data recovery using the explicit vertical scaling factors shown in equation (4) with that of the fixed vertical scaling factors, we still use the Ricker wavelet as an example. Figure 2a shows the recovery curve (dashed line) with explicit vertical scaling factors, and Figure 2b shows how the vertical scaling factors vary with different interpolating points.

Figure 2 shows that the recovery accuracy is very good, even if the vertical scaling factors change dramatically with the change of the interpolating points.

5 Numerical tests

To demonstrate the validity and the reliability of the fractal recovery method, we show a typical recovery example using the recovery approach for the step function and the derivative of a Gauss wavelet. Figures 3a and 3b compare the original curves and the recovered curves for the step curve and the derivative curve of

the Gauss function, respectively. The dashed lines were obtained using the localized fractal recovery approach presented here, and the solid lines are the exact curves. Although the new recovery method appears simple, it is both accurate (Figure 1) and efficient. This method is not only suitable for unsmoothed (Figure 3a) and smoothed (Figure 3b) data structures, but accurately recovers the local properties of a data set (Figure 3a).

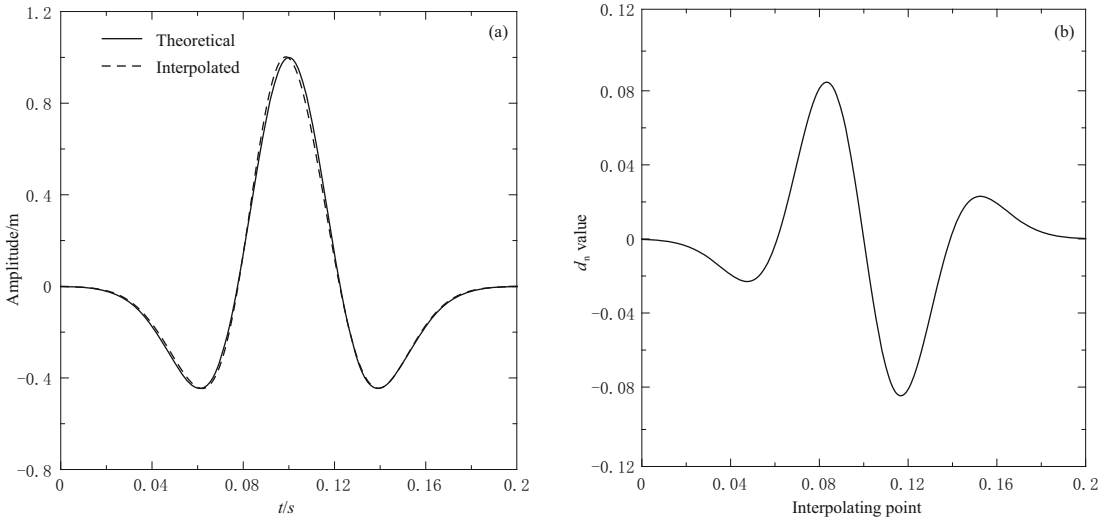

Figure 2 Recovery accuracy of the explicit vertical scaling factors. (a) Fractal recovery curve of the Ricker function with explicit vertical scaling factors. (b) Variation of the vertical scaling factor with the change of interpolating points.

As is known, the heterogeneous model is more realistic than a plane-layered model, and is also more complex in application. To evaluate the capability of our fractal method in the presence of heterogeneity, we designed a heterogeneous model (Figure 4a) and synthesized common-shot seismograms. The physical parameters of the heterogeneous model are given in Table 2. The source is located at the surface in the center of the model, and the receivers are symmetrically located on both sides of the source. The original (solid line) and the recovered (dashed line) seismograms are shown in Figure 4b. As can be seen from the figure, the recovered seismogram accurately resembles the original seismogram. To confirm the similarity between the two

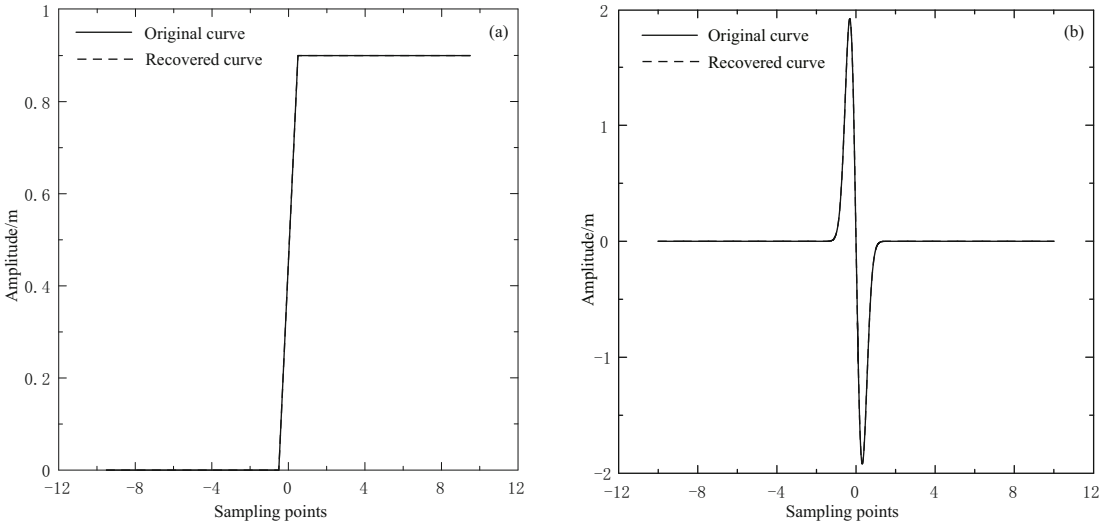

Figure 3 Testing the accuracy of the localized fractal recovery method presented in this paper. (a) Original data curve (solid line) and recovered data curve (dashed line). (b) Original Gauss wavelet (solid line) and recovered wavelet (dashed line).

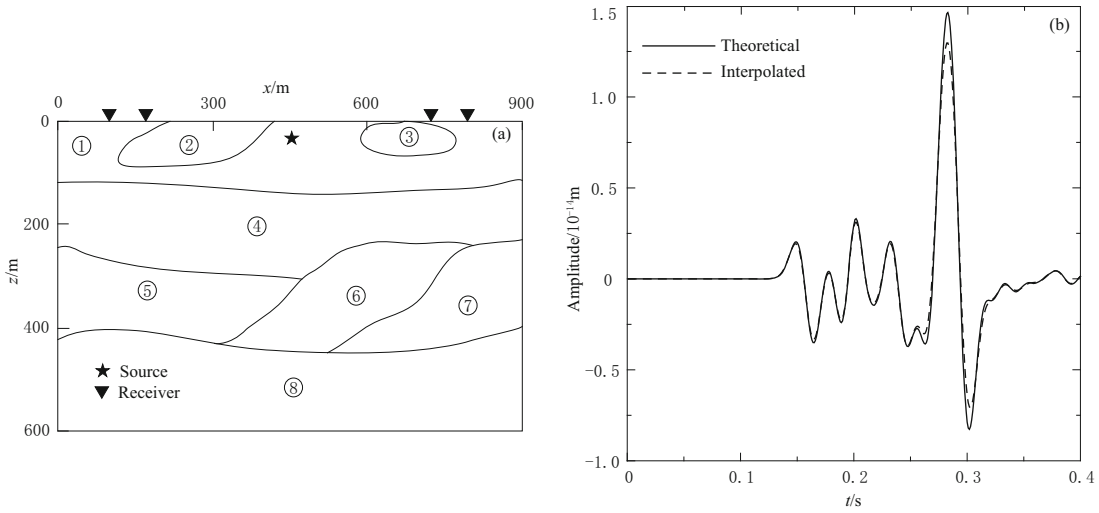

Figure 4 (a) Heterogeneous model (Zhang et al., 2006). (b) Originally synthesized (solid line) and recovered (dotted line) common shot seismograms.

Table 2 The physical parameters of the heterogeneous model (Zhang et al., 2006)

No. of media	$C_{11}/10^{10}$Pa	$C_{13}/10^{10}$Pa	$C_{33}/10^{10}$Pa	$C_{55}/10^{10}$Pa	α_0/m·s^{-1}	β_0/m·s^{-1}	ε	δ	ρ/kg·m^{-3}
①	1.881 6	0.652 8	1.881 6	0.614 4	2 800	1 600	0.00	0.00	2 400
②	2.348 9	0.786 9	2.060 5	0.667 0	2 900	1 650	0.07	0.03	2 450
③	2.388 0	0.828 2	2.132 1	0.683 3	2 950	1 670	0.06	0.03	2 450
④	3.123 3	0.964 3	2.402 5	0.810 0	3 100	1 800	0.15	0.08	2 500
⑤	4.006 4	1.229 5	2.861 7	0.949 9	3 350	1 930	0.20	0.10	2 550
⑥	3.832 2	1.320 2	2.777 0	0.940 0	3 300	1 920	0.19	0.17	2 550
⑦	3.562 5	1.169 1	2.873 0	0.973 4	3 350	1 950	0.12	0.09	2 560
⑧	3.369 6	1.076 4	3.369 6	1.146 6	3 600	2 100	0.00	0.00	2 560

Note: α_0, β_0, ε, δ are Thomsen's parameters obtained from Cij, ρ is density.

seismograms, we computed the cross-correlation coefficient between the recovered seismogram, which was purposefully erased from the common-shot record, and the original one. The resulting coefficient of 0.9912 demonstrates the similarity between the two. In addition, the coefficient of determination is 0.880 2, further indicating a good correlation between the recovered seismogram and the original one. These two tests confirm that our method of recovering the lost data is robust.

6 Real data tests

We now demonstrate an application of the explicit fractal recovery approach in recovering high-resolution seismic data from sparsely and non-uniformly distributed seismic stations. The example illustrated here uses seismic data from the Jiyang Depression and the Southern Bohai Bay Basin in Northern China (Zhao et al., 2004). We selected records from a seismic event recorded at ten seismic stations. The seismogram is shown in Figure 5a. The intervals between the adjacent two stations range from 5.28 km to 15.79 km, and the average spacing between stations is 10 km. Location information from the seismic event is given in Table 3.

The seismic data are band-pass-filtered between 0.05 Hz and 4 Hz. Instrumental responses as well as mean and linear trends are removed by data filtering. Because the seismogram is complicated and highly irregular, conventional recovery approaches cannot reproduce the local properties and detailed structure of the dataset. We eliminated the even traces in the seismogram section to construct a sparsely sampled seismogram (Figure 5b) for recovery purposes. Figure 5c shows the seismogram section after recovery with the newly developed fractal recovery approach. Comparing the original and recovered seismogram sections, we find that the two are essentially identical in amplitude, phase and waveform. To demonstrate the

Table 3 Parameters of the seismic event used in this paper

Event No.	Origin time (UTC)		Lat./°N	Long./°E	Depth/km	M_W
	Date a-mo-d	Time h:min:s				
010102	2001-01-02	07:30:04	6.75	126.81	33	6.4

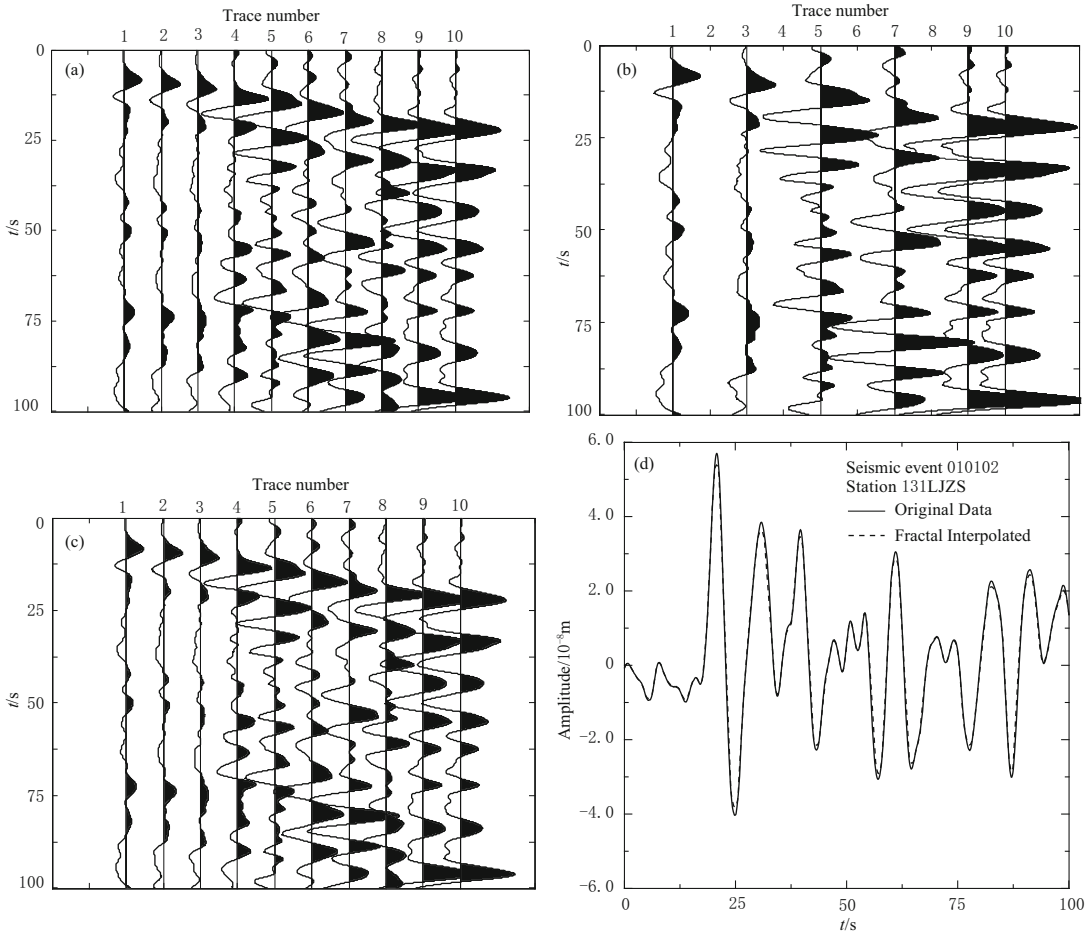

Figure 5 Recovery of nonuniform seismic data. (a) Original seismogram section from seismic traces of event 010102 at seismic stations 143SLX, 137HJZ, 131LJZS, 125XFX, 119YFZ, 113YZC, 107XZX, 101QDZ, 95QDZ and 87SZZ. (b) Sparse seismogram section (the even traces are erased from panel a). (c) Recovered seismogram section, including the even traces from Figure 5a and the interpolated traces. (d) Comparison of original seismogram (solid line) at Station 131LJZS and recovered seismogram (dashed line).

precision and accuracy of the recovery method, we compared a recovered trace with its original trace (which was erased in the initial section). This comparison yields a near-perfect match (Figure 5d). Correlation analysis for this recovered trace and its original trace show that the determination coefficient R^2 is 0.985. These results indicate that the fractal recovery method can accurately deal with the problems associated with complicated data structures, provided that the dataset is self-similar on local and global spatial scales. Such a result would normally be impossible by using conventional data recovery methods. Also, since the recovery is essentially identical to the original data, the conclusion can be made that the seismic data have the local and global self-similarity property associated with fractal geometry. Actually, a detailed recovery method of complex data sets is realized by using the high-fidelity fractal approach proposed in this paper.

7 Discussion and conclusions

In summary, we have proposed a new method for recovering seismic data based on a localized fractal that is particularly adept at amending unevenly distributed

datasets. Although this method is a data-driven algorithm that does not require any geological or geophysical assumptions, the seismic data naturally contain geological or geophysical information of the studied area, and are self-similar on local and global spatial scales. The new approach is numerically stable and easier to implement. The method is robust and able to generate near-perfect results, as we have demonstrated with real-world examples. These numerical results illustrate that the use of our high-fidelity fractal recovery scheme is a valid and efficient approach, not only for seismic wave field de-aliasing, but also for detailed, amplitude-preserving recovery of seismic data from sparsely and unevenly distributed seismic traces.

Results shown in Figure 5 demonstrate the usefulness of the localized fractal recovery method, and in particular, its ability to obtain detailed information for high-resolution seismic imaging of the Earth's interior. Indeed, this method is valid for any research field that requires detailed, accurate recovery of sparsely sampled and irregular data sets, provided that the data exhibits local and global self-similarity. Also it is noted that our approach only consider the first-order interpolation recovery. Futher studies are needed for the effects on data that may be found with higher orders. Extending the approach to higher dimensions may yield a more precise, reliable, and economical method for high-resolution imaging and high-precision wavefield reconstructions.

The examples illustrate that the fractal recovery algorithm provides significant improvements over the conventional methods. We expect that this algorithm can be extended to reconstruct data from missing seismic datasets with even larger gaps, and this is a question that will be investigated in the near future.

Acknowledgements All of the real seismic data used in this work were collected by the Broadband Seismic Array Laboratory, Institute of Geology and Geophysics, Chinese Academy of Sciences. This work was supported by the Spark Program of Earthquake Sciences (Grant No. XH13002).

References

Barnsley M F and Demko S (1985). Iterated function system and the global construction of fractals. *Proc R Soc Lond A* **399**(1817): 243–275.

Barnsley M F (1986). Fractal functions and interpolation. *Constr Approx* **2**(1): 303–329.

Chu H T and Chen C C (2003). On bounding boxes of iterated function system attractors. *Comput Graph* **27**(2): 407–414.

Fan Y H, Luan Y C, Wang Y, Liang Q K and Wang G X (2005). The method study of combination of fractal interpolation and linear interpolation. *Science of Surveying and Mapping* **30**(2): 76–77, 80 (in Chinese with English abstract).

Feng Z G and Zhou Q S (1998). About the stability of a kind of fractal interpolation. *J of Anqing Teachers College (Natural Science)* **4**(3): 18–21 (in Chinese with English abstract).

Hindriks K and Duijndam A J W (2000). Reconstruction of 3D seismic signals irregularly sampled along two spatial coordinates. *Geophysics* **65**(1): 253–263.

Kabir M M N and Verschuur D J (1995). Restoration of missing offsets by parabolic Radon Transformation. *Geophys Prosp* **43**(2): 347–368.

Larner K, Gibson B and Rothman D (1981). Trace interpolation and the design of seismic surveys. *Geophysics* **46**(3): 407–415.

Liu B and Sacchi M D (2004). Minimum weighted norm interpolation of seismic records. *Geophysics* **69**(6): 1 560–1 568.

Navascués M A and Sebastián M V (2004). Generalization of Hermite functions by fractal interpolation. *J Approx Theory* **131**(1): 19–29.

Singer P and Zajdler P (1999). Self-affine fractal functions and wavelet series. *J Mat Anal Appl* **240**(2): 518–551.

Trad D, Ulrych T and Sacchi M (2003). Latest views of the sparse Radon transform. *Geophysics* **68**(1): 386–399.

Wang Y H (2002). Seismic trace interpolation in the *f-x-y* domain. *Geophysics* **67**(3): 1232–1239.

Wang H Y and Fan Z L (2011). Analytical characteristics of fractal interpolation functions with function vertical scaling factors. *Acta Math Sin* **54**(1): 147–158.

Zwartjes P M and Sacchi M D (2007). Fourier reconstruction of nonuniformly sampled, aliased seismic data. *Geophysics* **72** (1): V21–V32.

Zhao L, Zheng T Y and Xu W W (2004). Modeling the Jiyang Depression, North China, using a wave-field extrapolation finite-difference method and waveform inversion. *Bull Seism Soc Am* **94**(2): 988–1 001.

Zhang M G, Huang Z Y, Li X F, Wang M Y, Xu G Y(2006). Pre-stack full wavefield inversion for elastic parameters of TI media. *J Geophys Eng* **3**(1): 90–99.

Appendix

The following codes illustrate how we realize the inorder traversing binary tree algorithm.

```
{ InitStack(stack) // Construct a null stack
Push(stack, Y_BT) // Push the root point into the stack, top=top+1
While(not empty(stack))
    {while(GetTop(stack)!=null)
    push(stack, GetTop(stack)→lchild) // Push the left sub-tree into the stack
    stack(top)=Pop(stack) // move upward the top of the stack
    If(not empty(stack))
        {visit(GetTop(stack))
    stack(top)=Pop(stack)
    push(stack, GetTop(stack)→rchild) // Push the right sub-tree into the stack
}
    }
}
```