The Fractal Characteristics of Ground Subsidence Caused by Subway Excavation

Yongjun Qin¹,*, Peng He¹, Jiaqi Zhang² and Liangfu Xie¹

¹ School of Civil Engineering and Architecture, Xinjiang University, Urumqi 830047, China; p_he@stu.xju.edu.cn (P.H.); xieliangfu@xju.edu.cn (L.X.)
² Xinjiang Transportation Investment (Group) Co., Urumqi 830016, China; jq15026035043@163.com
* Correspondence: qyjjg@xju.edu.cn

Abstract: The issue of uneven ground settlement caused by the excavation of subway tunnels represents a significant challenge in the design and construction of subway projects. This paper examined the fractal characteristics of surface settlement caused by subway excavation, employing wavelet transform and fractal theory. Firstly, the noise reduction effects of different wavelet functions, decomposition levels, threshold functions, and threshold selection rules were evaluated using the SNR and RMSE. Subsequently, 291 data points were derived from 18 interpolation points through fractal interpolation, representing a utilization of only 18% of the original data, to enhance the original monitoring data information by a factor of 2.94. The average relative error between the fractal interpolation results and the monitoring data was approximately 13%, which was indicative of a high degree of accuracy. Finally, the fractal dimension of the monitoring curves under different monitoring frequencies was calculated using the box-counting method. The denoising effect of the dbN wavelet function was found to be superior to that of the symN wavelet function in the denoising process of subway construction surface deformation monitoring data. Furthermore, the hard threshold method was observed to be more effective than the soft threshold method. As the monitoring frequency decreased, the fractal dimension of the deformation curves showed an overall decreasing trend. This indicated that high-frequency monitoring could capture more details and complexity of the surface settlement, while low-frequency monitoring led to a gradual flattening of the curves and a decrease in details.

Keywords: ground subsidence; wavelet de-noising; fractal interpolation; fractal dimension; deformation monitoring

1. Introduction

The increasing urbanization and expansion of urban rail transport networks have led to a rise in the number of people opting to travel by metro [1]. However, the construction of metro systems inevitably causes surface settlement, which has significant impacts on building safety, underground facilities, and transport. This phenomenon has become one of the key challenges to sustainable urban development. Surface settlement is a complex geographic process, and its time series data contain key information about dynamic change characteristics and internal mechanisms. Therefore, accurately analyzing and understanding the fractal characteristics of the time series of surface subsidence is of great theoretical and practical value for predicting surface subsidence.

Constructing mathematical models is an important method for studying surface settlement. However, it mainly relies on linear models and assumptions, which limits the understanding of the complexity of surface settlement to some extent [2–4]. In recent years, more and more scholars have carried out related research using neural network algorithms, etc. Zhang et al. [5] proposed an intelligent model mixing particle swarm optimization and random forests for predicting the creep index $C_\alpha$, which is better than the
current empirical model of $\alpha$. Savvides, A.-A. et al. [6] proposed a set of neural networks for estimating the settlement damage of the foundations under cohesive soil as well as the soil damage with volumetric and bias stresses and strains. In addition, researchers have started to use fractal theory to analyze surface settlement more deeply. The term ‘fractal’ was first proposed by the American mathematician B.B. Mandelbrot [7,8] and it has since been widely used in the study of linear geographic information. S.S. Quadri et al. [9,10] analyzed the relationship between the mass of the moving landslide and the fractal dimensions from a mechanical point of view. He et al. [11] derived the landslide displacement and Hurst exponent using fractal theory, which provided insights into the landslide deformation law and predicted future deformation evolution trends; Li et al. [12] constructed a prediction model of the landslide deformation period using the incremental displacement Hurst index; Liu et al. [13] investigated the fractal characteristics of slope crack development in different stable states; Gui et al. [14] investigated the multiple fractal characteristics of creeping landslide displacement curves by using the multiple fractal theory. Many scholars have studied the relationship between fractal dimension number and tunnel deformation: Wang et al. [15] investigated the settlement and deformation of hard rock tunnels using Koch fractal theory. Zhang et al. [16] analyzed the impact of nodal parameters on the fractal dimension of a tunnel’s surrounding rock. Xu et al. [17] examined the fractal characteristics of the surrounding rock deformation during tunnel construction. Hu et al. [18] calculated the fractal dimension of the motion trajectory curve of the landslide monitoring point to assess the development stage of the landslide. The focus of most scholars is on the qualitative relationship between the fractal dimension and the deformation process. As research continues to deepen, researchers increasingly adopt a quantitative approach to studying fractal features. Gao et al. [19] calculated the fractal dimension of the pore structure according to the solid mass fractal model and statistically analyzed the change in pore volume. They then analyzed the fractal characteristics of the pore structure of the coal body. Zhang et al. [20] calculated the correlation dimension of the acoustic emission time series of gas-bearing coals using the G-P algorithm and phase space reconstruction theory, subsequently employing this to analyze the fractal features of the aforementioned coals. Li et al. [21] employed the mercury intrusion pore method and low-field nuclear magnetic resonance method to quantitatively analyze the pore fracture structure and its fractal features. Song et al. [22] employed the box-counting method and Sierpinski carpet analytical solution to analyze the fractal features of shale pore structure at multiple scales.

Although previous studies have made progress in understanding the relationship between fractal dimensions and geological deformation processes, the accuracy and reliability of the analysis are often affected by noise, missing data, and outliers in monitoring data. Fractal interpolation, as an interpolation method based on fractal theory, addresses these challenges by filling in missing data points while correcting and optimizing abnormal data. This method effectively deals with irregularities and missing data. Zuo et al. [23] used R/S and V/S analysis to derive the Hurst index of tunnel surface deformation time series based on fractal theory. They then applied fractal interpolation to predict the trend of tunnel deformation evolution. Lai et al. [24] proposed a fractal interpolation method to fill in missing data using a sequence of self-similar features. Jiang [25] introduced the Cuckoo Search-Designated Fractal Interpolation Functions (CS-DFIFs) method and the CS-DFIFs-winners combining (CS-DFIFs-WC) method to estimate missing values. Liu et al. [26] proposed a method for accurately recovering seismic data using the local fractal recovery method, which enables easy and accurate recovery of the intrinsic structure associated with the seismic data.

In addition, the time series data for surface settlement are often affected by various types of noise, posing challenges for accurately extracting fractal features. Wavelet noise reduction emerges as an effective signal processing method capable of removing random noise from the data while preserving the main features of the signal, making it a more effective approach [27,28]. Wang et al. [29] analyzed the dynamic deformation data of
high-rise buildings monitored in the field using wavelet noise reduction, concluding that wavelet analysis has an advantage over traditional signal processing techniques in non-stationary monitoring signal processing. Xie et al. [30] observed a significant enhancement in the first-to-first characteristics of signals after noise reduction and compared the spectra of microseismic signals before and after wavelet noise reduction. Li et al. [31] employed wavelet noise reduction to preprocess hydropower station landslide monitoring data, revealing that analyzing the deformation time series after noise reduction better describes the deformation trend of landslides. Regarding foundation tunnels, Qin et al. [32] utilized wavelet noise reduction on data related to underground station settlement and combined it with a BP neural network to predict underground settlement, resulting in improved accuracy. Yang et al. [33] employed a combination of wavelet analysis and neural networks to predict the settlement of deep foundation pits, achieving an error between the predicted and actual values of from 1% to 5%, demonstrating the effectiveness of wavelet noise reduction in signal processing. Li et al. [34] applied wavelet noise reduction to a section of the Wuhan Yangtze River Tunnel online monitoring data, effectively reducing noise while retaining the original data mutation and aligning the monitoring data more closely with the real trend of change. Kong X [35] also utilized wavelet transform to reduce noise in shield tunnel monitoring data, providing reliable results for diagnosing the tunnel’s health.

The existing research faces two main issues. Firstly, surface settlement data are often affected by noise interference, missing data, and errors resulting from monitoring instruments, manual input, and environmental factors. Secondly, there is a lack of in-depth studies that utilize surface settlement deformation information to guide monitoring work. This paper addresses the aforementioned issues by examining the fractal properties of surface settlement using monitoring data from the North Gate Station of Urumqi Railway Line 1. Firstly, the noise reduction efficacy of the model is compared under different wavelet functions, two thresholds, and four threshold selection rules. Subsequently, the optimal noise reduction model for metro deformation monitoring data is derived, effectively reducing data noise. Secondly, the issue of a substantial number of continuously missing original monitoring data is addressed using the fractal interpolation method. Thirdly, the fractal dimensions of the surface deformation monitoring curves under different monitoring frequencies are calculated using the box-counting method. The results demonstrate that all of them are greater than 1, indicating that the surface deformation curves caused by underground construction exhibit fractal characteristics, i.e., self-similarity.

2. Theoretical Foundation

2.1. Wavelet Transform

Wavelet analysis represents a relatively novel mathematical approach to time series analysis that builds upon the principles of Fourier analysis. Its ability to effectively localize both time and frequency enables it to effectively reconcile the challenge of preserving signal locality while suppressing noise [36,37].

It is widely acknowledged that monitoring data are subject to noise pollution, with the signal being represented by \( s(i) \) and the noise by \( q(i) \). The former is the real signal with low frequency, while the latter is the noise signal with high frequency [38]. The wavelet transform was employed to separate the high-frequency noise signal from the relatively stable low-frequency signal. The following steps are required for the decomposition of a monitoring signal into \( N \) layers using a wavelet function [39–41]: (1) select the appropriate wavelet function to decompose the monitoring signal in \( N \)-layer; (2) choose the appropriate threshold to process; (3) reconstruct the low-frequency signal and the high-frequency signal after threshold processing.

Let \( \Psi(t) \in L^2(R) \) and its Fourier transform be \( \tilde{\Psi}(\omega) \). When the Fourier transform of the wavelet function, \( \tilde{\Psi}(\omega) \), satisfies the tolerance condition, \( \Psi(t) \) is said to be a fundamen-
Fractal Interpolation is an interpolation method based on the theory of the Iterated Function System (IFS), which is frequently employed to fit non-smooth curves and interpolate non-smooth data. It provides a novel tool for data fitting. In contrast to the traditional interpolation function, which generates smooth curves, fractal interpolation can construct non-smooth curves with fractal dimensions between 1 and 2 by iterating the changes in the parameters of the function system. Consequently, the fractal interpolation method is well-suited to the study of frequently fluctuating time series, with the iterated function system (IFS) representing the most critical component.

Fractal interpolation [45–48] produces an iterated function system (IFS) from a given set of interpolated points. The mapping from the set of interpolated points to two consecutive interpolated points is constrained by several conditions. The following endpoint conditions are satisfied:

\[
\begin{align*}
\varphi_i(x_0) &= \left( \frac{a x_0 + b}{d} \right) + \left( \frac{c e_i + d f_i}{f} \right) \\
\varphi_i(x_n) &= \left( \frac{a x_n + b}{d} \right) + \left( \frac{c e_i + d f_i}{f} \right)
\end{align*}
\]

The parameter \( b_i \) is zero in IFS theory, thus, the remaining map parameters may be determined using the constraints (3) and (4):

\[
\begin{cases}
    a_i = \frac{x_n - x_{i-1}}{x_n - x_0} \\
    c_i = \frac{y_n - y_{i-1}}{x_n - x_0} - \frac{d_i y_0 - y_0}{x_n - x_0} \\
    e_i = \frac{y_0 y_{i-1} - x_n y_0}{x_n - x_0} - \frac{d_i y_0 - y_0}{x_n - x_0} + \frac{d_i y_0 - y_0}{x_n - x_0}
\end{cases}
\]

where \( i = 0, 1, 2, \ldots, n, x_0 < x_1 < \ldots < x_n \).

After all parameters have been found, the fractal interpolation function may be constructed.

2.3. Fractal Dimension

Fractal dimensions are metrics that characterize a fractal pattern or set by quantifying its complexity as the ratio of changes in detail to changes in scale. They are employed to characterize phenomena from the abstract to the real. Of the various fractal dimensions that have been proposed, the most commonly used are the Hausdorff dimension and the box dimension. The latter is based on the notion of measure and is more intuitively defined.

\[
\varphi_i(t) = \frac{1}{\sqrt{|a|}} \Psi \left( \frac{t - b}{a} \right), \quad a, b \in \mathbb{R}
\]
in mathematical theory. Furthermore, the box dimension is more readily computable than other fractal dimensions due to its countable stability. In this paper, we primarily employ the Box-Counting method to calculate the fractal dimension.

The Box-Counting method, as described in [49–53], is a widely utilized approach for determining fractal dimensions in the field of linear geographic information. This approach determines the fractal dimensions through the relationship between the scale \( r \) and the corresponding number of boxes \( N(r) \) covering the target object. The fractal dimension is defined as follows:

\[
D = \lim_{n \to \infty} \frac{\ln[N(1/r)]}{\ln(r)}
\]  

where \( D \) is the fractal dimension, \( r \) is box size, and \( N(1/r) \) is the total number of boxes where the attractor intersects the box with side length \( r \).

3. Processes

3.1. Project Overview

This study focuses on Urumqi Urban Rail Transit Line 1, currently the sole operational metro line in the city. The project represents a significant investment in infrastructure and forms the backbone line running through Urumqi in a north–south direction. It connects all of the city’s key sectors and facilitates economic and social progress. Located at the six-way intersection, Beimen Station is an important hub in Urumqi’s bustling urban matrix. The area is densely populated with commercial, cultural, and administrative activities, highlighting the station’s role in alleviating surface traffic congestion and improving the efficiency of public transport.

Geographically, Beimen Station is located in a highly urbanized and densely populated area, where the construction and operation of the metro are critical for relieving traffic congestion and improving public transport efficiency. The complex geological conditions, together with the dense urban fabric, require a focused investigation of the impact of metro construction on ground stability. The cut-and-cover construction method used for the station is common in urban environments but poses significant ground stability challenges due to the extensive excavation required.

To monitor the impact of the subway excavation on the stability of the surrounding ground, 30 monitoring points were strategically placed around Beimen Station, covering adjacent areas and areas close to the line, as shown in Figure 1. These points generate comprehensive data on surface changes over time, providing an important resource for analyzing subsidence patterns. In particular, the DB-02-01 monitoring point is located northwest of the North Gate Station, at the intersection of Guangming Road and Jiefang North Road. The surrounding area is complex. The main building on the north side is the Morning Post Building, with a total of 33 floors, its raft foundation depth is about 6.5 m, and the structural skin is 10 m away from the foundation pit, while on the west side is the brick house of the Bayi Theater, with a total of 1 floor, and it is only 2 m away from the enclosing skin of the auxiliary foundation pit, which is very likely to cause ground settlement or cracks in the upper house during the construction process. In addition, this point is subject to the double instability of tunnel excavation and foundation pit construction, so its data are crucial to understanding the regional settlement pattern.
3.2. Evaluating Indicator

In general, the key to denoising quality lies in the selection of wavelet function, decomposition level, and threshold. Therefore, Signal-to-Noise Ratio (SNR) and Root Mean Square Error (RMSE) are selected as indicators to evaluate the denoising effect, and then the best wavelet function, decomposition level, and threshold are obtained. SNR is used to measure the degree of noise reduction of the model, and RMSE can be used to measure the deviation between the original data and the noise reduction data. The larger the SNR, the smaller the RMSE, and the better the denoising effect.

\[
\text{SNR}(dB) = 10 \log_{10} \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right) = 10 \log_{10} \left( \frac{A_{\text{signal}}}{A_{\text{noise}}} \right)
\]  

(7)

\[
\text{RMSE} = \frac{\sum_{i=1}^{N} (A_{\text{signal}} - A_{\text{denoised}})^2}{N}
\]

(8)

where \((P_{\text{signal}}, A_{\text{signal}})\) denotes pure signal, \((P_{\text{noise}}, A_{\text{noise}})\) denotes pure noise signal, and \(A_{\text{denoised}}\) denotes noise-reduced signal. \(N\) denotes the sample size.
3.3. Wavelet Transform

We collected the surface deformation monitoring data of the DB-02-01 monitoring point within 290 days, as illustrated in Figure 2. The monitoring frequency ranges from one to fifteen days, resulting in a total of 99 data points.

![Figure 2. The monitoring data of DB-02-01.](image)

3.3.1. The De-Noising Model

To identify the optimal denoising model, a series of comparisons are conducted between different wavelet functions, decomposition levels, threshold functions, and threshold selection rules. The signal-to-noise ratio and root-mean-square error are employed as evaluation metrics.

A. Wavelet function

The classical wavelet functions include the Haar wavelet, Mexican hat wavelet, Daubechies (db) wavelet, Morlet wavelet, and Symlet (sym) wavelet. In this paper, only dbN and symN wavelets (N: 1–10), which are widely used in engineering, are considered. To identify a suitable wavelet function, the DB-02-01 monitoring data were denoised using dbN and symN wavelets while ensuring that all other conditions remained consistent. The corresponding SNR and RMSE are shown in Figure 3 and Table 1.

![Figure 3. The de-noising effects of different wavelet functions.](image)
Table 1. The average de-noising effects of different wavelet functions.

<table>
<thead>
<tr>
<th>Wavelet Function</th>
<th>SNR</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>db N</td>
<td>25.26</td>
<td>0.11</td>
</tr>
<tr>
<td>sym N</td>
<td>23.35</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Figure 3 illustrates that, for N values between 1 and 5, the noise reduction effect of the dbN wavelet function and the symN wavelet function is essentially comparable, with more pronounced differences becoming apparent. For N values between 6 and 10, the symN wavelet function exhibits greater stability in its noise reduction effect. It can be observed that the noise reduction effect of the symN wavelet function is superior to that of the dbN wavelet function when N = 6. However, when N = 7–10, the noise reduction effect of the dbN wavelet function is more pronounced than that of the symN wavelet function, although it exhibits greater fluctuations.

However, concerning the average noise reduction effect, as illustrated in Table 1, the dbN wavelet function exhibits a greater SNR and RMSE, accompanied by superior noise reduction compared to the symN wavelet function. Consequently, in this paper, the dbN wavelet function is selected.

B. Decomposition level

In the process of signal denoising, there will be a decomposition level with the best denoising effect. The appropriate decomposition level can be chosen to better separate the signal and noise. When the decomposition level is too large, the thresholding process will cause the effective signal information to be filtered out too much, resulting in a worse denoising effect and a slower calculation rate. Conversely, when the decomposition level is too small, the result of denoising the signal is not ideal. Consequently, in this paper, under the condition that all other variables remain unchanged, the original monitoring data are decomposed into 1–12 levels using the dbN wavelet function to identify the optimal decomposition level. The subsequent figure illustrates the corresponding signal-to-noise ratio and effective signal-to-noise ratio for each level of decomposition.

As illustrated in Figure 4, it can be observed that, as the decomposition level increases, the SNR declines, the RMSE rises, and the noise reduction effect diminishes gradually. At Lev = 1, the SNR is at its maximum and the RMSE is at its minimum. Therefore, the decomposition level in this paper is 1.

![Figure 4](image-url)

**Figure 4.** The de-noising effects of different decomposition levels. (a) SNR for different decomposition levels; (b) RMSE for different decomposition levels.
C. Threshold

The selection of the threshold value and threshold function is of particular importance in the process of thresholding noise-containing signals. The wavelet coefficient of the noise signal is smaller than that of the effective signal. The selection of the threshold rule is primarily concerned with setting the threshold and filtering out signals with wavelet coefficients below the threshold, thereby inhibiting the noise to the greatest extent possible. The wavelet threshold function is selected to filter out noise signals with a wavelet coefficient below the threshold. This is achieved by filtering out the noise coefficient contained in the wavelet coefficient through the appropriate threshold function, thus obtaining the required effective signal. The required effective signal is obtained through the application of a suitable threshold. This paper compares the noise reduction effect under different threshold functions and different threshold rules to identify the most suitable threshold for the given data. The results are shown in Figure 5.

![Figure 5](image)

**Figure 5.** The de-noising effects of different thresholds. (a) SNR and RMSE for different threshold functions; (b) SNR and RMSE for different threshold selection rules.

From Figure 5, it can be seen that, in terms of threshold function, the hard threshold has a higher SNR and lower RMSE, and the noise reduction effect is more pronounced than that of the soft threshold. In terms of the threshold selection rule, the results indicate that rigrsure has a higher SNR and lower RMSE than the other three rules, and thus a superior noise reduction effect. Therefore, hard threshold and rigrsure denoising rules are selected in this paper.

In summary, a model has been developed to denoise the raw monitoring data. The wavelet function is dbN, the threshold function is hard threshold, the threshold selection rule is rigrsure, and the decomposition level is 1.

3.3.2. The De-Noising Effect

Once the denoising model had been determined, a Matlab program was developed using MATLAB R2018a to denoise the original monitoring data of DB-02-01.

Upon analysis, it was observed that there was significant contamination in the monitoring data between days 90 and 290 (as depicted in Figure 6). In particular, during the initial 90 days, the rock and soil disturbance caused by foundation pit excavation was substantial, resulting in less obvious contamination of deformation information. However, following the completion of the excavation of the foundation pit and the stabilization of the rock and soil, the deformation information becomes more susceptible to contamination. Notably, noise pollution is present during periods of relatively stable settlement (days 90–290). The denoised data aligns more accurately with the actual deformation patterns.
4. Fractal Interpolation Simulation

In this section, we considered the exploration of deformation characteristics. There is a significant and continuous absence in the original monitoring data [25], necessitating the application of interpolation to obtain daily surface deformation. The interpolation methods include traditional interpolation and fractal interpolation. The traditional interpolation method connects two data points with a smooth curve, which ignores the fluctuation characteristics of the data. Fractal interpolation enables the mapping of the entire fluctuation characteristics of the data to two interpolation points, thereby ensuring self-similarity between the local fluctuation and the overall deformation [54,55]. After a comprehensive comparison of the advantages and limitations of the two methods, the latter was selected to obtain daily surface deformation.

4.1. The Algorithm

We derived the fractal interpolation function of the de-noised data using the Matlab program. The algorithm is as follows:

1. Choose the interpolation point \((x_i, y_i)\), where \(i = 0, 1, \ldots, n\);
2. Compute the contraction factor \(d_i\) by analytic [56]. In general, the factor \(d_i\) lies in the interval \((-1, 1)\);
3. Compute the map parameters with (4);
4. Store the data set generated by the map, which is empty at the beginning;
5. The fractal interpolation function of de-noised data has been searched.

4.2. The Effect

Figure 2 illustrates the surface deformation to have undergone two distinct stages, namely the “obvious deformation period” and the “unclear deformation period”. Over the course of 100 days, the surface deformed significantly, with a maximum deformation of 6.44 mm. Over the subsequent 100–290 days, the surface deformation tended to stabilize.

Following the principle of random midpoint [57,58], we separately selected nine interpolation points in the “obvious deformation period” and “unclear deformation period”, and used the Matlab program to calculate the fractal interpolation of the curves in the above two periods. The results are shown in Figures 7 and 8.
Figure 7. The fractal interpolation effect in the “obvious deformation period”. (a) Interpolation effect of the first iteration; (b) interpolation effect of the second iteration; (c) interpolation effect of the third iteration; (d) interpolation effect of the fourth iteration.

As illustrated in Table 2, 291 data points were derived from 18 interpolation points through fractal interpolation in this study, representing a 2.94-fold increase in the original monitoring data information, with approximately 18% of the original monitoring data utilized. This equates to the daily monitoring of surface deformation. Furthermore, as illustrated in Figure 9, the mean relative error between the fractal interpolation results and the monitoring data is approximately 13%, thereby demonstrating that the fractal interpolation methodology employed in this study is highly accurate. It is important to note that the relative errors between the fractal interpolation results and the monitoring data are greater when the monitoring periods are 13–26 days, 194–222 days, and 244–264 days, with values of 58.31%, 33.87%, and 29.99%, respectively. This is due to the fact that the raw monitoring curves in these three time periods exhibit pronounced spikes and significant fluctuations in the data (as illustrated in Figure 6), which increases the error.
Figure 8. The fractal interpolation effect in the “unclear deformation period”. (a) Interpolation effect of the first iteration; (b) interpolation effect of the second iteration; (c) interpolation effect of the third iteration; (d) interpolation effect of the fourth iteration.

Table 2. Comparison between the fractal interpolation results and the monitoring data.

<table>
<thead>
<tr>
<th>Time Period/d</th>
<th>Monitoring Data (Quantity)</th>
<th>Interpolation Results (Quantity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–13</td>
<td>0.0324 (2)</td>
<td>0.0384 (13)</td>
</tr>
<tr>
<td>13–26</td>
<td>0.0818 (5)</td>
<td>0.1295 (13)</td>
</tr>
<tr>
<td>26–40</td>
<td>1.8905 (6)</td>
<td>2.0088 (14)</td>
</tr>
<tr>
<td>40–50</td>
<td>3.8863 (6)</td>
<td>4.3266 (10)</td>
</tr>
<tr>
<td>50–63</td>
<td>4.6987 (6)</td>
<td>4.8323 (13)</td>
</tr>
<tr>
<td>63–75</td>
<td>5.2813 (4)</td>
<td>5.6643 (12)</td>
</tr>
<tr>
<td>75–89</td>
<td>2.5576 (8)</td>
<td>2.5132 (14)</td>
</tr>
<tr>
<td>89–100</td>
<td>−0.4147 (9)</td>
<td>−0.3984 (11)</td>
</tr>
<tr>
<td>100–124</td>
<td>0.2445 (16)</td>
<td>0.2314 (25)</td>
</tr>
<tr>
<td>124–148</td>
<td>0.8864 (9)</td>
<td>0.9232 (24)</td>
</tr>
<tr>
<td>148–170</td>
<td>0.8730 (6)</td>
<td>0.9480 (22)</td>
</tr>
<tr>
<td>170–194</td>
<td>0.3345 (6)</td>
<td>0.3196 (24)</td>
</tr>
<tr>
<td>194–222</td>
<td>0.4895 (3)</td>
<td>0.6553 (28)</td>
</tr>
<tr>
<td>222–244</td>
<td>0.3798 (3)</td>
<td>0.3528 (22)</td>
</tr>
<tr>
<td>244–264</td>
<td>0.3385 (4)</td>
<td>0.2370 (20)</td>
</tr>
<tr>
<td>264–290</td>
<td>0.1087 (7)</td>
<td>0.1184 (26)</td>
</tr>
</tbody>
</table>
1. Plot Von Koch curve, as shown in Figure 10;

2. Gray processing of the target object, the result is as shown in Figure 11;

3. By changing the scale \( r \), we could obtain the total box counts needed to cover the target object;

4. Linearly fit \( \ln(r) - \ln(N(1/r)) \), as shown in Figure 12. The slope of \( \ln(r) \) and \( \ln(N(1/r)) \) is regarded as the fractal dimension.

5. Fractal Dimension Feature

5.1. The Algorithm

We adopted the Box-Counting method to calculate the fractal dimension of the Von Koch curve using the Matlab program. The algorithm is as follows:

1. Plot Von Koch curve, as shown in Figure 10;
2. Gray processing of the target object, the result is as shown in Figure 11;
3. By changing the scale \( r \), we could obtain the total box counts needed to cover the target object;
4. Linearly fit \( \ln(r) - \ln(N(1/r)) \), as shown in Figure 12. The slope of \( \ln(r) \) and \( \ln(N(1/r)) \) is regarded as the fractal dimension.

Figure 9. The relative error between monitoring data and interpolation results under different periods.

Figure 10. Von Koch curve.

Figure 11. Von Koch curve grayscale.
It can be found that the slope of \( \ln(r) \) and \( \ln(N(1/r)) \) are 1.3406. That is, the fractal dimension of the Von Koch curve is 1.3406, which differs from the theoretical value of 1.26 \([59]\) by 6.4%. It is considered that the algorithm to calculate fractal dimension is accurate.

5.2. The Fractal Characteristics

To investigate the fractal characteristics of ground subsidence caused by subway excavation, we obtained the deformation curve with different monitoring frequencies: once a day, once every two days, once every four days, once every six days, once every eight days, once every ten days, once every twelve days, once every fourteen days, once every sixteen days, once every eighteen days, once every twenty days, and once a month (that is, \( f = 2n, n = 1–10 \)), as shown in Figure 13. As can be seen from the graph, the curve becomes flatter as the monitoring interval increases. Accordingly, less and less detail and complexity can be captured in the curve.

We used the algorithm mentioned in Section 5.2 to calculate the fractal dimension of these deformation curves, the result is listed in Table 3.

Table 3. Fractal dimension for the deformation curve with different monitoring frequencies.

<table>
<thead>
<tr>
<th>Monitoring Frequency</th>
<th>Fractal Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 d/time</td>
<td>1.2418</td>
</tr>
<tr>
<td>2 d/time</td>
<td>1.2231</td>
</tr>
<tr>
<td>4 d/time</td>
<td>1.2261</td>
</tr>
<tr>
<td>6 d/time</td>
<td>1.2098</td>
</tr>
<tr>
<td>8 d/time</td>
<td>1.2109</td>
</tr>
<tr>
<td>10 d/time</td>
<td>1.2055</td>
</tr>
<tr>
<td>12 d/time</td>
<td>1.1998</td>
</tr>
<tr>
<td>14 d/time</td>
<td>1.2087</td>
</tr>
<tr>
<td>16 d/time</td>
<td>1.2139</td>
</tr>
<tr>
<td>18 d/time</td>
<td>1.2026</td>
</tr>
<tr>
<td>20 d/time</td>
<td>1.2283</td>
</tr>
<tr>
<td>30 d/time</td>
<td>1.1921</td>
</tr>
</tbody>
</table>
We used the algorithm mentioned in Section 5.2 to calculate the fractal dimension of these deformation curves, the result is listed in Table 3.

<table>
<thead>
<tr>
<th>Monitoring Frequency</th>
<th>Fractal Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 time/day</td>
<td>1.2418</td>
</tr>
<tr>
<td>2 time/day</td>
<td>1.2231</td>
</tr>
<tr>
<td>4 time/day</td>
<td>1.2261</td>
</tr>
<tr>
<td>6 time/day</td>
<td>1.2098</td>
</tr>
<tr>
<td>8 time/day</td>
<td>1.2109</td>
</tr>
<tr>
<td>10 time/day</td>
<td>1.2055</td>
</tr>
<tr>
<td>12 time/day</td>
<td>1.1998</td>
</tr>
<tr>
<td>14 time/day</td>
<td>1.2087</td>
</tr>
</tbody>
</table>

We found that the deformation and the fractal dimension have a similar fluctuation pattern. That is, when the monitoring frequency is once every 2 days, 12 days, and 18 days, the deformation trend turns, and the corresponding fractal dimension curve also appears as an “inflection point” (Figures 14 and 15). It can be found from Figure 15 that when the monitoring frequency is once every 2 days, the fractal dimension rises slightly; when the monitoring frequency is once every 12 days and 18 days, the fractal dimension rises significantly. Therefore, to monitor the key information of surface deformation, you should ensure that the monitoring period is less than 12 d in the field monitoring. If you want to capture subtle changes in deformation, the monitoring frequency should be once every 2 days.
The deformation varies with monitoring frequency. 

The fractal dimension varies with monitoring frequency. 

In general, the larger the monitoring frequency is, the less the fractal dimension. Hence, we derive the expression of the fitting function: \( y = 1.2438 \times x^{-0.0117} \), as shown in Figure 15.

6. Conclusions

This paper presented a denoising model for reducing the noise of monitoring data based on wavelet transform theory. The fractal characteristics of the surface settlement caused by subway excavation were analyzed by applying fractal theory. The noise reduction effect of the model was evaluated by comparing it under different wavelet functions, decomposition levels, threshold functions, and threshold rules. The optimal noise reduction model for surface settlement data was identified. In the fractal interpolation phase, the original monitoring curves were divided into two stages (i.e., the period of significant deformation and the period of insignificant deformation), and nine interpolation points were selected in each stage (i.e., 18 interpolation points were selected from the 99 in the original monitoring data). The effects of missing data and outliers were overcome by using the fractal interpolation method. The fractal dimension of the monitoring curves under different monitoring frequencies was calculated in the fractal characteristics section using the box-counting method. This was followed by an analysis of the relationship between the fractal dimension and surface settlement. The conclusions were as follows:
(1) When utilizing the dbN wavelet function, a decomposition level of 1, the hard threshold method, and the rigrsure threshold rule, the noise reduction model exhibited a high SNR and a low RMSE and exhibited the most effective noise reduction in this instance.

(2) The daily surface deformation data were derived by fractal interpolation, and 291 data points were successfully extracted from 18 interpolated points, which compensated for the effects of missing data and outliers, to a certain extent. This method effectively captured the actual deformation information with an average relative error of only 13%.

(3) The fractal dimension of the monitoring curves was consistently greater than 1, indicating that the surface deformation caused by the construction of the Urumqi subway was self-similar. Additionally, as the monitoring frequency decreased, the fractal dimension exhibited a declining trend, and the surface settlement and deformation curves tended to smooth, accompanied by a concurrent decrease in the actual information that could be captured.

It is important to note that, while the research method employed in this paper is effective in the case of the Urumqi Railway Transit Line 1, further studies are necessary to assess its efficacy in other projects. Additionally, this paper only examines the fractal characteristics of surface settlement under varying monitoring frequencies. However, surface settlement is a complex process, and future research will consider a wider range of factors for improvement.

Author Contributions: Conceptualization, P.H. and Y.Q.; methodology, P.H.; software, P.H. and L.X.; validation, P.H. and J.Z.; formal analysis, P.H. and J.Z.; investigation, P.H. and L.X.; resources, Y.Q.; data curation, P.H. and J.Z.; writing—original draft preparation, P.H.; writing—review and editing, P.H. and J.Z.; visualization, Y.Q. and L.X.; supervision, Y.Q.; project administration, Y.Q.; funding acquisition, Y.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This study has been funded by the Natural Science Foundation of Xinjiang Uygur Autonomous Region, Research on data mining and surface subsidence prediction of subway construction deformation monitoring based on big data (Project Number 2021D01C073).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data generated and analyzed during this study are included in this article.

Conflicts of Interest: Author Jiaqi Zhang is employed by Xinjiang Transportation Investment (Group) Co. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.