



OPEN

# Land subsidence analysis along high-speed railway based on EEMD-Prophet method

Qiu Dongwei<sup>1</sup>, Tong Yuci<sup>1</sup>, Wang Yuzheng<sup>1</sup>, Ding Keliang<sup>1</sup>, Liu Tiancheng<sup>1</sup> & Wan Shanshan<sup>2</sup>✉

Environmental changes and ground subsidence along railway lines are serious concerns during high-speed railway operations. It is worth noting that AutoRegressive Integrated Moving Average (ARMA), Long Short-Term Memory (LSTM), and other prediction methods may present limitations when applied to predict InSAR time series results. To address this issue, this study proposes a prediction method that decomposes the nonlinear settlement time series of feature points obtained through InSAR technology using Ensemble Empirical Mode Decomposition (EEMD). Subsequently, multiple Intrinsic Mode Functions (IMFs) are generated, and each IMF is individually predicted using the Prophet forecasting model. Finally, we employ an equal-weight superimposition method to combine the results, resulting in the prediction of the InSAR settlement time series. The predicted values of each component are subsequently weighted equally and combined to derive the final prediction outcome. This paper selects InSAR monitoring data along a high-speed railway in inland China and uses the proposed method and ARMA and Prophet models to carry out comparative experiments. The experimental results show that compared with the ARMA and Prophet models, the method in this paper improves the root mean square error by 58.01% and 32.3%, and increases the mean absolute error by 62.69% and 33.78%, respectively. The predicted settlement values generated by our method exhibit better agreement with the actual InSAR monitoring values.

High-speed rail plays an important role in connecting cities for socio-economic development. Subsidence in the area along the high-speed railway will have a certain negative impact on the stability and safety of the railway, and then affect the normal operation of the railway<sup>1</sup>. Therefore, to ensure the safety of train operations, it is essential to monitor changes in infrastructure such as high-speed railways and predict their changing trends<sup>2</sup>.

In recent years, many experts and scholars have researched on surface subsidence prediction and achieved meaningful results. Aiming at the problem that improper selection of training samples will directly affect prediction accuracy, Zhou Dingyi<sup>3</sup> constructed a land subsidence prediction method based on a neural network algorithm from the perspective of multiple factors. Zhou Qihang<sup>4</sup> proposed an improved third-order gray prediction model combined with terrain factors and neural network error correction for urban land subsidence prediction. Using a neural network to predict land subsidence will achieve higher precision. However, in the prediction process, there will be problems such as an unstable prediction process, complex parameter setting selection, and only suitable for use in a single environment. Many classic traditional forecasting models, such as ARMA<sup>5</sup> and ARIMA<sup>6</sup>, can also be employed to predict future time series changes. However, the predictive accuracy of these models is often unsatisfactory, and the accuracy needs to be maintained through a rolling forecast. Yang Zhengrong<sup>7</sup> used the glacier deformation field and gray correlation analysis obtained by InSAR technology and revealed the relationship between glacier deformation and various factors such as temperature, precipitation, and slope. Xiong Zhiqiang<sup>8</sup> used the statistical properties of curve fitting to select the best curve model from several candidate curve models to predict the subsidence time series. Shi Xianlin<sup>9</sup> used the regenerative settlement information combined with InSAR technology and Terzaghi theory and the corresponding geological parameters of the artificial airport to successfully deduce the predicted settlement curve of the offshore artificial island airport. Chen Yi<sup>10</sup> used the LSTM neural network to successfully establish a time series InSAR prediction model when analyzing the Beijing Capital Airport using the InSAR monitoring results. Since spatial heterogeneity can interfere with the accuracy of large-scale land subsidence prediction, Liu Qinghao<sup>11</sup> proposed

<sup>1</sup>School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, Beijing 100044, China. <sup>2</sup>School of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China. ✉email: wss@bucea.edu.cn

the HLSTM method that accounts for spatial heterogeneity, thereby mitigating the influence of spatial heterogeneity on the prediction results of large-scale land subsidence. However, using the above method combined with InSAR to predict the subsidence field needs to establish a complex model and consider the influence of various factors. If some factors are not taken into consideration, it may cause significant errors in the prediction results and other shortcomings<sup>12</sup>.

The Prophet model<sup>13</sup> is a time series forecasting model based on an additive model. In 2017, Taylor et al. proposed a new signal-processing method and released an open-source software package. This model employs a generalized additive model to conduct Bayesian curve fitting of nonlinear trends on time series data, thus offering remarkable flexibility and robustness<sup>14</sup>. The EEMD<sup>15</sup> is based on the empirical mode decomposition<sup>16</sup> (EMD) method, which effectively addresses the modal aliasing of the IMF obtained via EMD decomposition by superimposing Gaussian white noise<sup>17</sup>. The core concept is to perform multiple decompositions on different scales of the input data. By averaging the outcomes of each decomposition, the randomness and problems with local extreme values in EMD decomposition are reduced, and the decomposition results are more stable.

Aiming to address the aforementioned issues, this study proposes a prediction method that decomposes InSAR nonlinear time series features before predicting them. Specifically, the settlement time series obtained by InSAR is firstly decomposed into intrinsic mode functions using the EEMD method, which can characterize local features at different times. Then, the Prophet model is used to predict each intrinsic mode function. Finally, the predicted results of each intrinsic mode function are combined with equal weights to enhance the prediction accuracy.

## Method

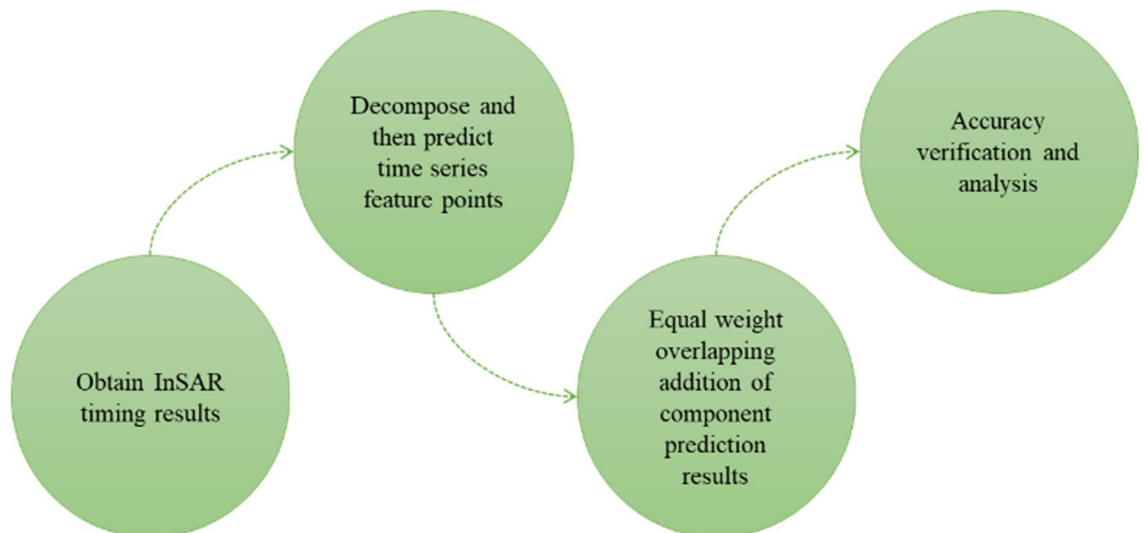
### Main idea

The method proposed in this paper is based on the SBAS-InSAR technology to predict the settlement time series characteristics of InSAR settlement. SBAS-InSAR was originally proposed by Berardino et al.<sup>18</sup>. This technology optimizes the combination of images acquired at different times, and its main process is shown in the literature<sup>19</sup>. In our approach, we first perform EEMD decomposition of the SBAS-InSAR time series results. Subsequently, we make separate predictions for each of the decomposition results, and finally, we apply a novel method involving equal-weight superposition. Our method combines the strong adaptability, fast prediction speed, and high prediction effect of the Prophet model with the EEMD method's characteristics of not requiring prior conditions. Figure 1 illustrates the flow chart of our method for predicting InSAR nonlinear subsidence time-series data, and the main steps are as follows:

- (1) Conduct SBAS -InSAR processing on the study area to obtain the InSAR subsidence time series results.
- (2) Perform EEMD decomposition on the nonlinear time series results to obtain IMF components, and use the Prophet model to predict each IMF component separately.
- (3) Superimpose the prediction results of each IMF with equal weight to obtain the prediction results of the settlement point.
- (4) Evaluate and analyze the prediction accuracy of the settlement point.

### SBAS-InSAR Processing Workflow

SBAS-InSAR is a monitoring technology based on traditional D-InSAR technology, which is further deepened to obtain the time series characteristics of surface deformation. The algorithm selects  $N + 1$  SAR images, sets an appropriate spatiotemporal baseline consistent with PS-InSAR, generates  $N$  small baseline set differential interference pairs, and performs phase unwrapping after generating differential interference patterns.



**Figure 1.** Technical flow chart.

During this process, we set Multilook Number to 4 where Range Looks is 4 and Azimuth Looks is 1. This reduces the impact of coherence spots and achieves a good balance between signal-to-noise ratio and spatial resolution. In addition, in order to eliminate atmospheric noise, Goldstein filtering is used, and its window size is set to 5\*5. The selection of Unwrapping Method Type is Minimum Cost Flow, where Unwrapping Decomposition Level is 1 and Unwrapping Coherence Threshold is 3. These parameter choices were guided by careful consideration of the trade-offs between computational efficiency and preserving deformation signals in the presence of atmospheric effects.

It is assumed that within the ordered time  $(t_1, t_2, \dots, t_n)$  of the interest area,  $N + 1$  SAR images of the same orbit are acquired, and the auxiliary images are registered to the common main image according to the interference conditions, and then the auxiliary images are registered to the common main image according to the interference conditions. The complex phase solution is performed on the image, and  $M$  differential interference patterns are finally generated. Among them,  $M$  needs to meet the following conditions:

$$\frac{(N + 1)}{2} \leq M \leq N \left( \frac{N + 1}{2} \right) \quad (1)$$

During the interference process, in the  $j$ th differential interference pattern generated by the two SAR images acquired at  $t_1$  and  $t_2$  ( $t_1 > t_2$ ) as the main and auxiliary images, the interference phase of any pixel can be expressed as:

$$\begin{aligned} \delta\phi_j(x, r) &= \phi(t_1, x, r) - \phi(t_2, x, r) \\ &\approx \frac{4\pi}{\lambda} [d(t_1, x, r) - d(t_2, x, r)] + \Delta\phi_{jtop} + \Delta\phi_{jatm} + \Delta\phi_{noise} \end{aligned} \quad (2)$$

In the formula,  $\lambda$  is the wavelength,  $d(t_1, x, r)$  and  $d(t_2, x, r)$  are the cumulative deformation quantities,  $\Delta\phi_{jtop}$  is due to the reference DEM Terrain phase error caused by inaccuracy,  $\Delta\phi_{jatm}$  is the error phase caused by atmospheric delay,  $\Delta\phi_{noise}$  is the error phase caused by the influence of noise. After removing the terrain phase error, the error caused by atmospheric delay, and the error caused by the influence of noise, through simplified formula (3), the interference phase can be expressed as:

$$\delta\phi_j(x, r) = \phi(t_1, x, r) - \phi(t_2, x, r) \approx \frac{4\pi}{\lambda} [d(t_1, x, r) - d(t_2, x, r)] \quad (3)$$

To obtain the sedimentation sequence, the average phase velocity can be expressed as:

$$V_j = \frac{\phi_j - \phi_{j-1}}{t_1 - t_{j-1}} \quad (4)$$

In the formula, the phase value of the  $j$ th interference pattern can be expressed as:

$$\sum_{k=t_{B,j}=1}^{t_{A,j}} (t_k - t_{k-1}) v_k = \delta\phi \quad (5)$$

The deformation of the time series calculated by the SBAS-InSAR algorithm can be expressed as:

$$A_v = \delta\phi \quad (6)$$

In the formula,  $A$  is the  $MXN$  matrix, and  $80$  is the vector representing the interference phase value. The SBAS-InSAR technical data processing flow chart is shown in Fig. 2.

### Prophet prediction model based on EEMD

The Prophet prediction method, based on EEMD, combines the advantages of both techniques. It can better capture the nonlinear, non-constant, and periodic changing characteristics of the time series while effectively removing noise and interference through decomposition. This approach enhances prediction accuracy while preserving the important information of the original signal. In line with the proposed EEMD-Prophet prediction method, the original signal of the InSAR monitoring point is firstly decomposed into several IMF and residual components using the EEMD method, specifically:

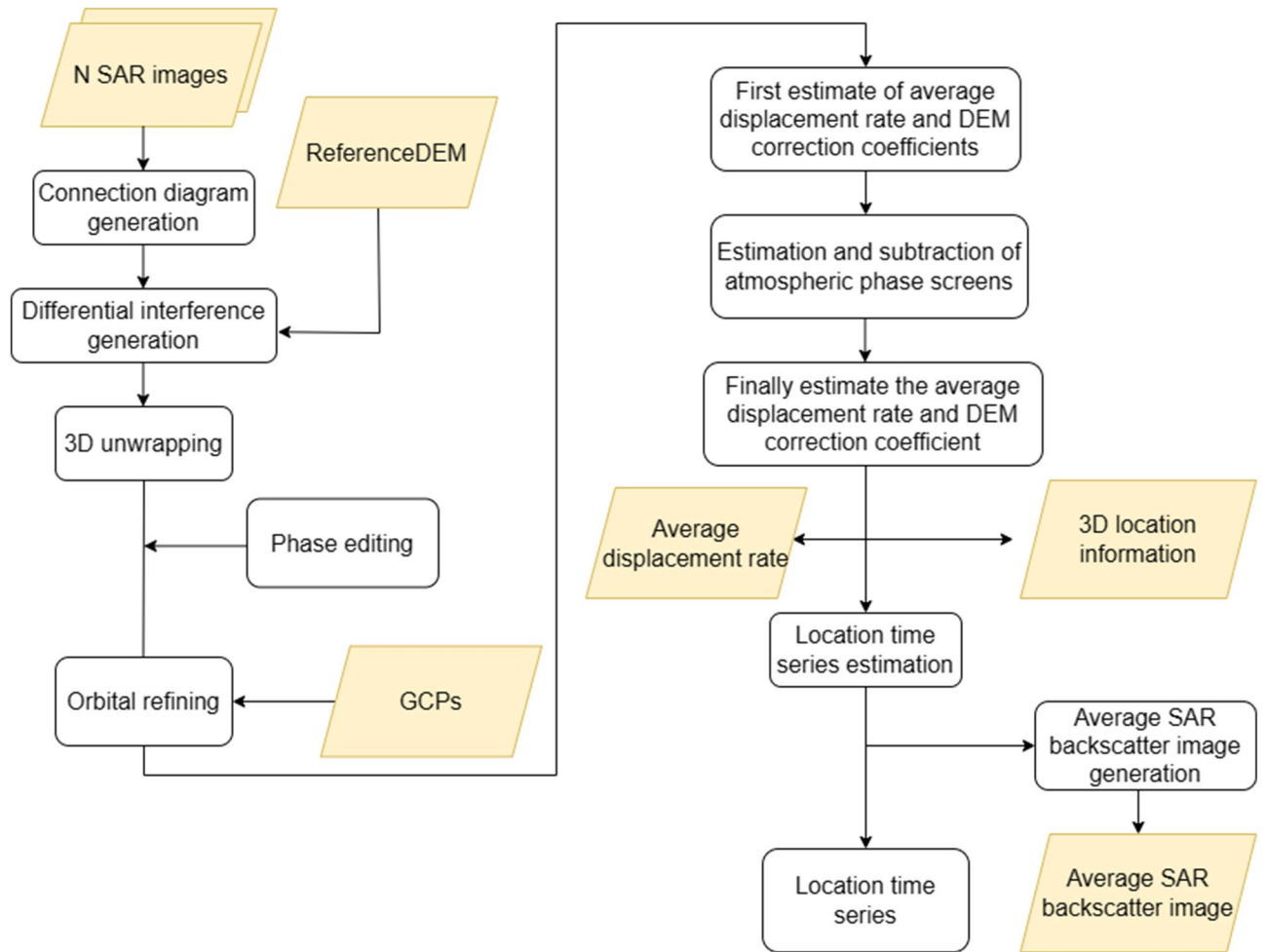
(1) Add a set of Gaussian white noise  $n(t)$  to the InSAR settlement time series signal feature  $Q(t)$  to obtain an overall signal  $X(t)$

$$X(t) = Q(t) + n(t) \quad (7)$$

(2) Decompose  $X(t)$  by EMD first to obtain the corresponding IMF of each order, and its formula is expressed as:

$$X(t) = \sum_{i=1}^n \text{IMF}_i(t) + \text{res}_i(t) \quad (8)$$

(3) Add a set of different white noise signals  $n_j(t)$  to the InSAR settlement time series feature  $Q(t)$  to obtain another overall signal  $X_j(t)$ . After performing EMD decomposition on  $X_j(t)$ , obtain the corresponding IMF components of each order using its formula expression:



**Figure 2.** SBAS technical processing flow.

$$X_j(t) = \sum_{i=1}^n \text{IMF}_{ji}(t) + \text{res}_j(t) \tag{9}$$

(4) Repeat the above steps, and calculate the mean value of each IMF obtained as the IMF of the final signal, and its formula is expressed as:

$$\text{IMF}_j(t) = \frac{1}{N} \sum_{i=1}^n \text{IMF}_{ji}(t) \tag{10}$$

where  $\text{IMF}_i(t)$  denotes the IMF for each specific order, whereas  $\text{res}_i(t)$  corresponds to the associated residual component.  $N$  signifies the total count of iterations involving white noise addition.  $t$  is the time for obtaining InSAR results, and  $i, j$  are the number of IMFs obtained by decomposition.

According to the principle of the Prophet model, each IMF obtained by decomposing the characteristics of the InSAR subsidence timing signal by the EEMD method is adaptively decomposed into:

$$y_{imf}(t) = g_{imf}(t) + s_{imf}(t) + h_{imf}(t) + \varepsilon_t \tag{11}$$

In Eq. (11), the Prophet model decomposes each IMF signal time series feature  $y_{imf}(t)$  into individual elements: the time series data trend item  $g_{imf}(t)$ , the oscillation period item  $s_{imf}(t)$ , the holiday mutation term  $h_{imf}(t)$ , and the residual term  $\varepsilon_t$ . The decomposed trend item  $g_{imf}(t)$  represents the changing function of the nonlinear growth component within the time series. When analyzing land subsidence time series, the trend item is commonly captured using a logistic regression function<sup>20</sup>, which can be represented as:

$$g_{imf}(t) = \frac{m}{1 + e^{[-k(t-l)]}} \tag{12}$$

In Eq. (12),  $k$  represents the settlement growth rate,  $l$  denotes the displacement, and  $m$  represents the upper limit of the settlement trend. As time  $t$  increases,  $g_{imf}(t)$  gradually converges towards  $m$ . However, the time series of surface subsidence is subject to changes, requiring consideration of the dynamic subsidence growth

rate  $k$  and the adaptive offset. Additionally, the changing growth rate needs to be taken into account. Hence, Eq. (12) is revised as:

$$g_{imf}(t) = \frac{m}{1 + e^{-(k+\alpha(t)^T\delta) [t - (l+\alpha(t)^T\gamma)]}} \quad (13)$$

This Eq. (13) describes the variation of the trend component  $g_{imf}(t)$  in the ground subsidence time series with respect to time  $t$ . In this context, the growth rate  $k$  and the offset  $l$  are dynamically changing with time, influenced by the parameters  $\alpha(t)$ ,  $\delta$ , and  $\gamma$ . This dynamic modeling allows the trend to adapt over time to better fit the observed data. The symbol  $t$  stands for time, signifying the time point. The symbol  $\alpha(t)$  is a function related to time  $t$ . This function represents parameters that vary over time. The symbol  $\delta$  is a parameter vector, signifying the variation in the growth rate  $k$  over time. The symbol  $\gamma$  is a parameter vector, representing the variation in the displacement  $l$  over time.

The oscillating period item  $s_{imf}(t)$  includes seasonal periodic oscillations and small trend changes, and the Fourier series construction of the settlement time series data can obtain  $s_{imf}(t)$ , the structural form can be expressed as:

$$s_{imf}(t) = \sum_{n=1}^N \left( a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right) \right) \quad (14)$$

In Eq. (14),  $T$  represents the length of the timing cycle, while  $2n$  denotes the anticipated number of cycles in the model. Typically, the land subsidence time series does not exhibit a holiday mutation term  $h_{imf}(t)$ , and thus, its influence is disregarded. The residual term  $\varepsilon_t$  adheres to a normal distribution and can be interpreted as unexpected random noise and varying trends. The technical approach of the EEMD-based Prophet prediction method is illustrated in Fig. 3.

## Study areas and deformation characteristics

### Study areas and data

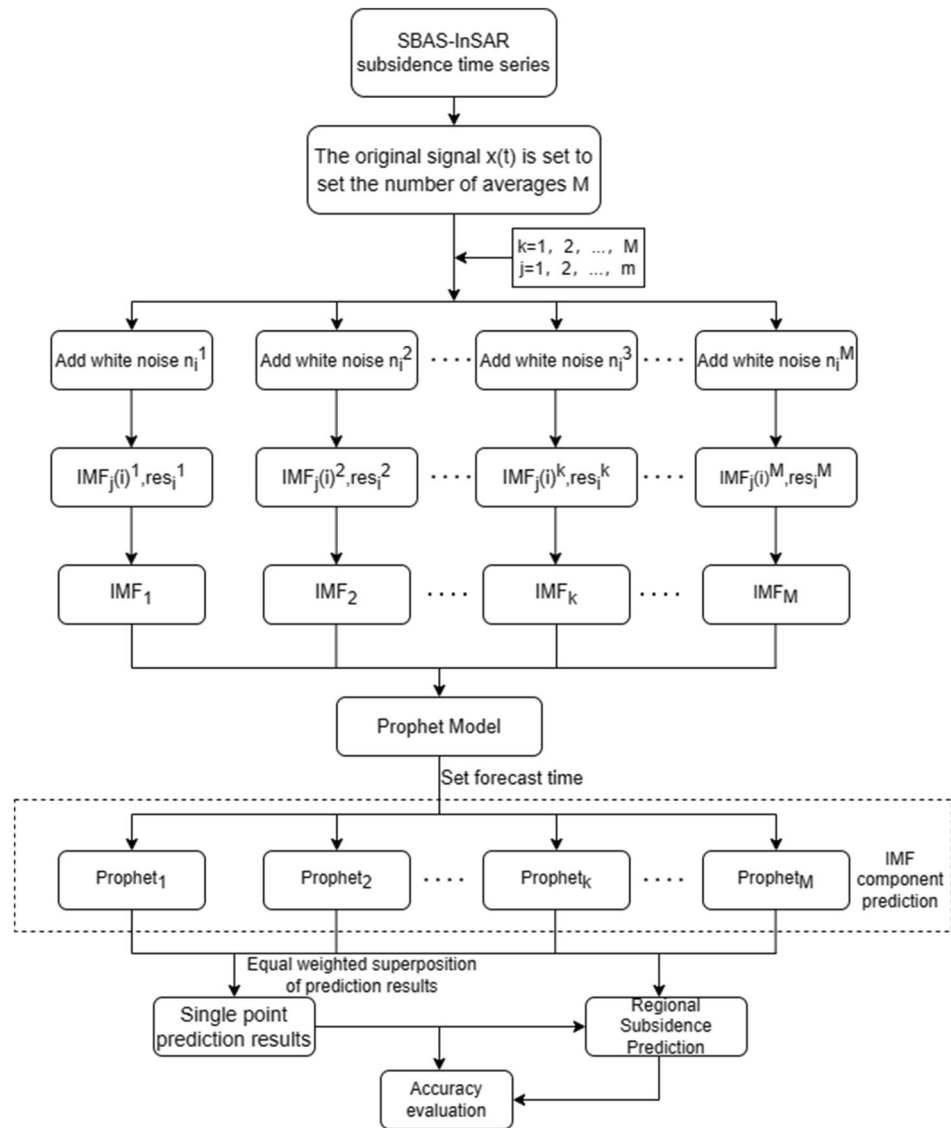
The Taiyuan-Jiaozuo high-speed railway is situated in the inland region of eastern mainland China. It serves as a vital transportation link connecting Jiaozuo City in Henan Province to Taiyuan City in Shanxi Province. This railway line commenced operations on December 12, 2020, spanning a total length of 358.764 km and encompassing 13 intermediate stations. Designed for a maximum speed of 250 km per hour, the route traverses a diverse terrain comprising plains, hills, mountains, rivers, and other intricate landscapes, each presenting unique geological characteristics<sup>21</sup>. Notably, the railway passes through certain coal mine goaf areas, introducing additional uncertainties to train operation safety. Given the stringent requirements for track stability in high-speed rail systems, even minor surface deformations pose substantial risks to the overall operation. Figure 4 illustrates the geographical location of the research area, specifically focusing on the section of the study section spanning from Jinzhong Station to Taigu East Station along the Taijiao high-speed railway.

The data utilized in this study comprises ESA Sentinel-1A data, obtained during the satellite image period spanning from January 09, 2021, to October 07, 2022, encompassing a total of 40 scene data. The corresponding timestamps for each SAR image are presented in Table 1. The dataset is based on C-band observations with an ascending orbit direction and an incident angle of 39.27°. The imaging mode employed is the interferometric wide mode IW, with VV polarization. The pixel resolution is 5m in the range direction and 20m in the azimuth direction. The Global Navigation Satellite System (GNSS) monitoring stations are primarily deployed along the section from Jinzhong Station to Taigu East Station, as illustrated in Fig. 6, which depicts the spatial distribution of 35 observation stations. Monitoring activities commenced in July 2021.

### Regional deformation feature analysis

In the experiment, the SAR image captured on November 29, 2021, was selected as the primary reference image. Ground Control Points (GCP) generated by PS-InSAR were utilized for orbit refinement and re-flattening, thereby improving the accuracy of SBAS-InSAR processing<sup>22</sup>. Figure 5 illustrates the deformation rate results obtained by processing the research area using SBAS-InSAR and overlaying them on Google Earth Pro (V7.3.6.9345; software available at <https://www.google.com/earth/about/versions/#download-pro>). The histogram's peak in the upper right corner indicates that a majority of the area experiences no deformation, which aligns with the actual conditions of the selected area, providing evidence for the reliability of the SBAS-InSAR monitoring results utilizing GCP points generated by PS-InSAR. The most significant subsidence area is located at the intersection of the two railways, exhibiting an average subsidence rate of -38 mm/year. In contrast, the main uplifted areas are primarily concentrated in the southwest region, specifically the Taigucheng District, and the villages in the northwest region, particularly in the south of Liujiabao Township. These uplifted areas have average subsidence rates of 15 mm/year and 20 mm/year, respectively.

There are three distinct deformations observed along the studied railway area, as depicted in Fig. 6. These subsidence areas are marked as A, B, and C, while three InSAR subsidence feature points are selected in each area and labeled as 1, 2, and 3, respectively. The selection of InSAR settlement feature point 1 in each area is determined using the nearest neighbor method with respect to the GNSS monitoring station near the railway. The time series of settlement measurements for each feature point in every region are presented in Fig. 7. The time series analysis reveals that the settlement trends and cumulative subsidence of the three feature points within area A, located at different positions, are primarily consistent. This suggests a large subsidence range within area A, corroborating the findings depicted in Fig. 5. Notably, characteristic point A-1, situated near the center of the

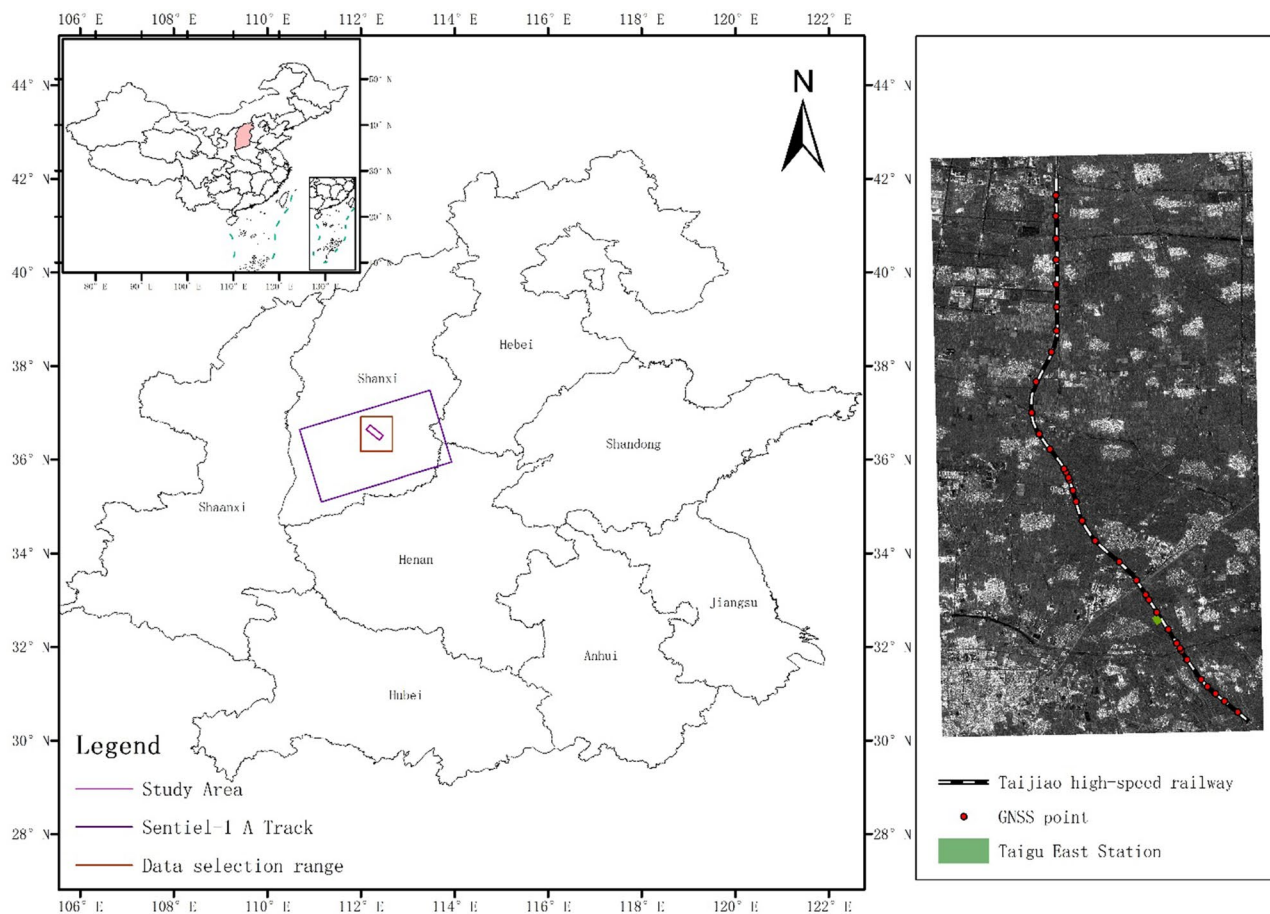


**Figure 3.** Technology Roadmap.

subsidence area and close to the GNSS station, exhibits a maximum cumulative settlement of 110 mm. Feature points A-2 and A-3, situated at the periphery of the subsidence zone, display cumulative subsidence of 90 mm and 80 mm, respectively. During the monitoring period of the three feature points in area A, the cumulative settlement trend demonstrates a sharp increase, followed by a gentler rate, and then another sharp increase. The period of gentler cumulative settlement for the three feature points spans from August 2021 to February 2022, indicating a consistent subsidence trend along the railway lines in area A. The railway section in area B encompasses both uplifted and subsided regions. Among these, feature point B-1 experiences significant cumulative subsidence of 75 mm, while feature point B-2, situated between B-1 and B-3, demonstrates cumulative subsidence of 25 mm. Feature point B-3 represents an uplifted area with a relatively minor cumulative uplift. Area C is predominantly characterized by uplift, albeit at a slower rate compared to the sinking trend observed in area A. The maximum cumulative uplift within this area is observed at feature point C-1, reaching 22 mm, while featuring points C-2 and C-3 exhibit cumulative uplifts of 10 mm each.

Figure 8a represents the time series data after the coordinate conversion of the GNSS data using Eq. (16). It showcases the unified coordinate frame for the results monitored by the GNSS base stations near feature points A-1, B-1, and C-1, as well as the InSAR results. The results show that the Mean Absolute Error (MAE) of InSAR monitoring settlement data of feature point A-1 is 4.61 mm, at feature point B-1 is 3.83 mm, and at feature point C-1 is 2.28 mm. To enhance the persuasiveness of the results, we selected the remaining GNSS monitoring points and employed the nearest neighbor method to choose the corresponding InSAR points to validate the accuracy of the InSAR results. Due to the constraints of our research scope and our specific focus on areas of interest, there were situations where InSAR results were unavailable near some GNSS stations or the distance to the nearest InSAR feature point exceeded a certain threshold, rendering a direct comparison unfeasible. As





**Figure 4.** Taijiao high-speed railway research section.

Year	SAR image acquisition time											
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
2021	9	14	10	15	21	14	8	1	6	12	5	11
	21	26	22	27				13	30	24	29	23
								25				
2022	4	9	5	10	4	9	3	8	1	7		
	28	21	17	22	28	21	27		13			

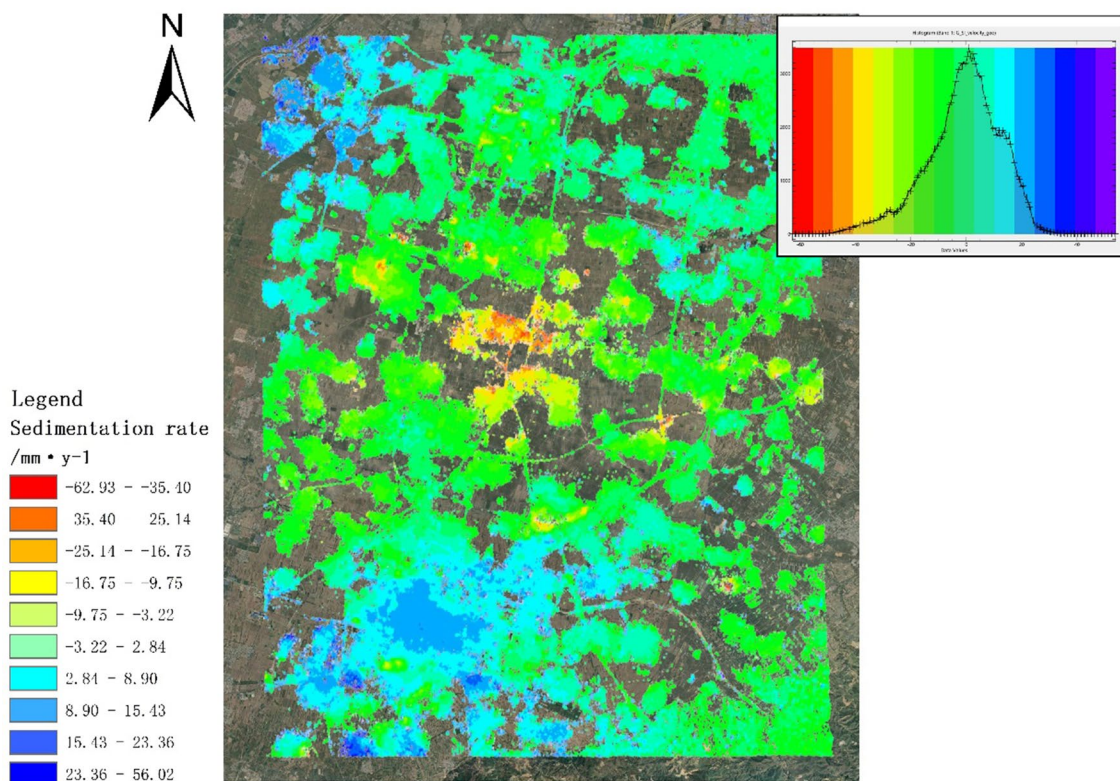
**Table 1.** The acquisition time of SAR image data from 2021 to 2022.

as a result, we selected the locations and monitoring data of 16 eligible GNSS points and compared them with nearby InSAR feature points using the nearest neighbor method. In Fig. 8b, the time series results of the coordinate transformation are presented for 16 randomly selected GNSS monitoring points along the railway and the adjacent InSAR feature points.

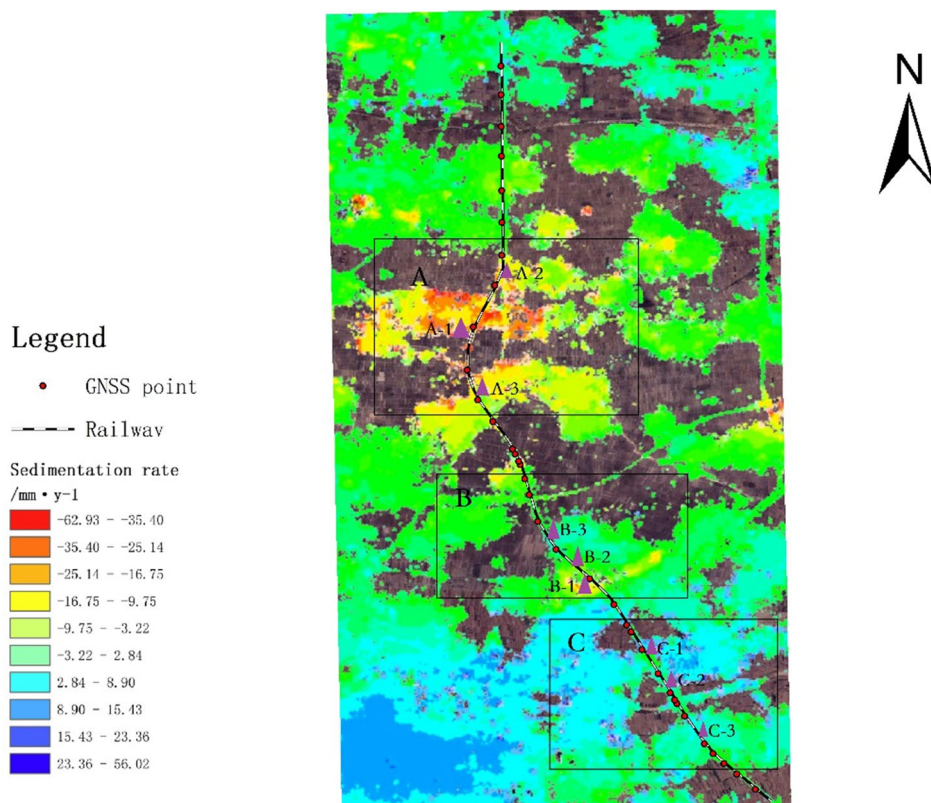
Table 2 shows the MAE and Root Mean Square Error (RMSE) of InSAR points near different GNSS stations. We can find that except for point 3, the MAE and RMSE of the InSAR points are both small. These results affirm the reliability and precision of SBAS-InSAR technology for monitoring settlement data. Therefore, this paper considers the InSAR monitoring data along the railway as the real observation data and conducts retrospective prediction research.

### Experiment Experimental scheme

In this study, we utilized three prediction models: the ARMA model, the Prophet model, and the EEMD-Prophet prediction method, to predict and compare the InSAR monitoring points along the high-speed railway and its surrounding areas.

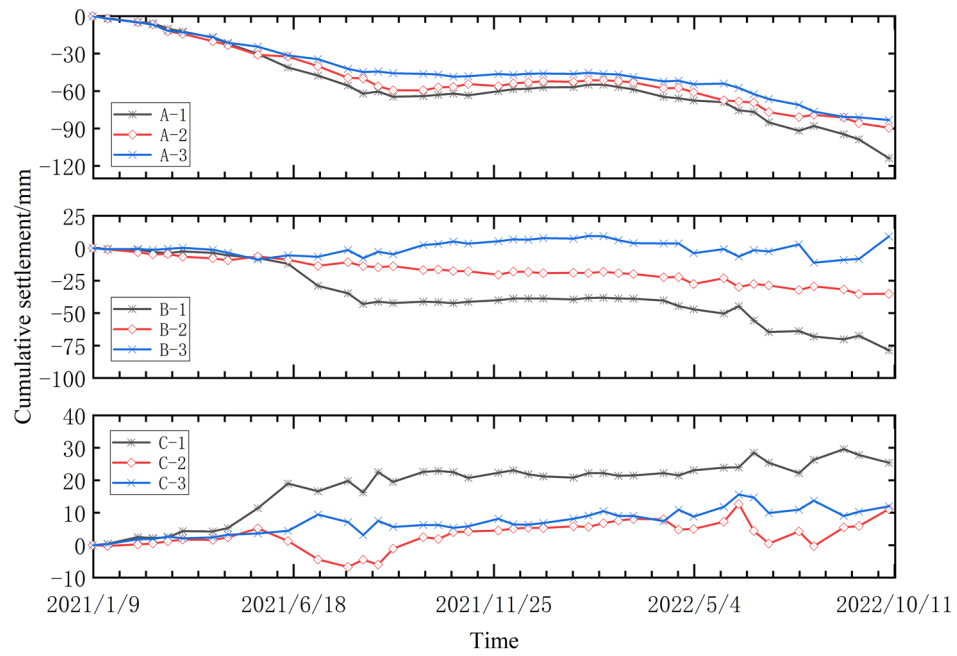


**Figure 5.** Deformation rate and histogram of the study area (the color red indicates subsidence, the color blue indicates an uplift and The color green represents relatively stable regions).



**Figure 6.** Distribution of GNSS stations along the railway and areas A, B and C.





**Figure 7.** Settlement of feature points in ABC region.

Firstly, we utilized GNSS monitoring points along the railway in this area and applied the nearest neighbor method to select the corresponding InSAR settlement point. Subsequently, we conducted a comparative analysis of the settlement trend and the accuracy of InSAR monitoring.

Secondly, we utilized the three prediction models to forecast and compare the feature points along the railway. Initially, we predicted and analyzed the feature points located close to the selected GNSS monitoring points from the previous step and compared the results with the corresponding GNSS monitoring data. Subsequently, for the feature points situated at a greater distance from the GNSS monitoring station, we employed a retrospective prediction method for comparative analysis.

Thirdly, at a regional scale, we selected all InSAR points near Taigu East Railway Station along the railway line. Using the retrospective prediction method, we employed the three prediction methods to generate land subsidence predictions and created a subsidence map. We then compared this map with the original InSAR land subsidence map.

The BeiDou Navigation Satellite System (BDS) measurement and monitoring points deployed along the railway line are utilized to evaluate the accuracy of surface subsidence monitoring using through InSAR technology. BDS monitoring station can measure both horizontal and vertical deformations. The displacement of the BDS base station can be converted into the Line-of-Sight (LOS) direction of InSAR<sup>23</sup>. This conversion is represented by Eq. (15):

$$d_{LOS} = d_N \sin(\theta_{inc}) \sin(\alpha_h) - d_E \sin(\theta_{inc}) \cos(\alpha_h) + d_U \cos(\theta_{inc}) \tag{15}$$

In Eq. (15),  $d_{LOS}$  represents the satellite line of sight direction,  $\alpha_h$  is the satellite radar heading angle,  $\theta_{inc}$  is the satellite radar incident angle,  $d_N$ ,  $d_E$ , and  $d_U$  are the north, east, and displacement in the upward direction. Assuming that the horizontal displacement of the BDS station is negligible<sup>24</sup>, it can be simplified as follows:

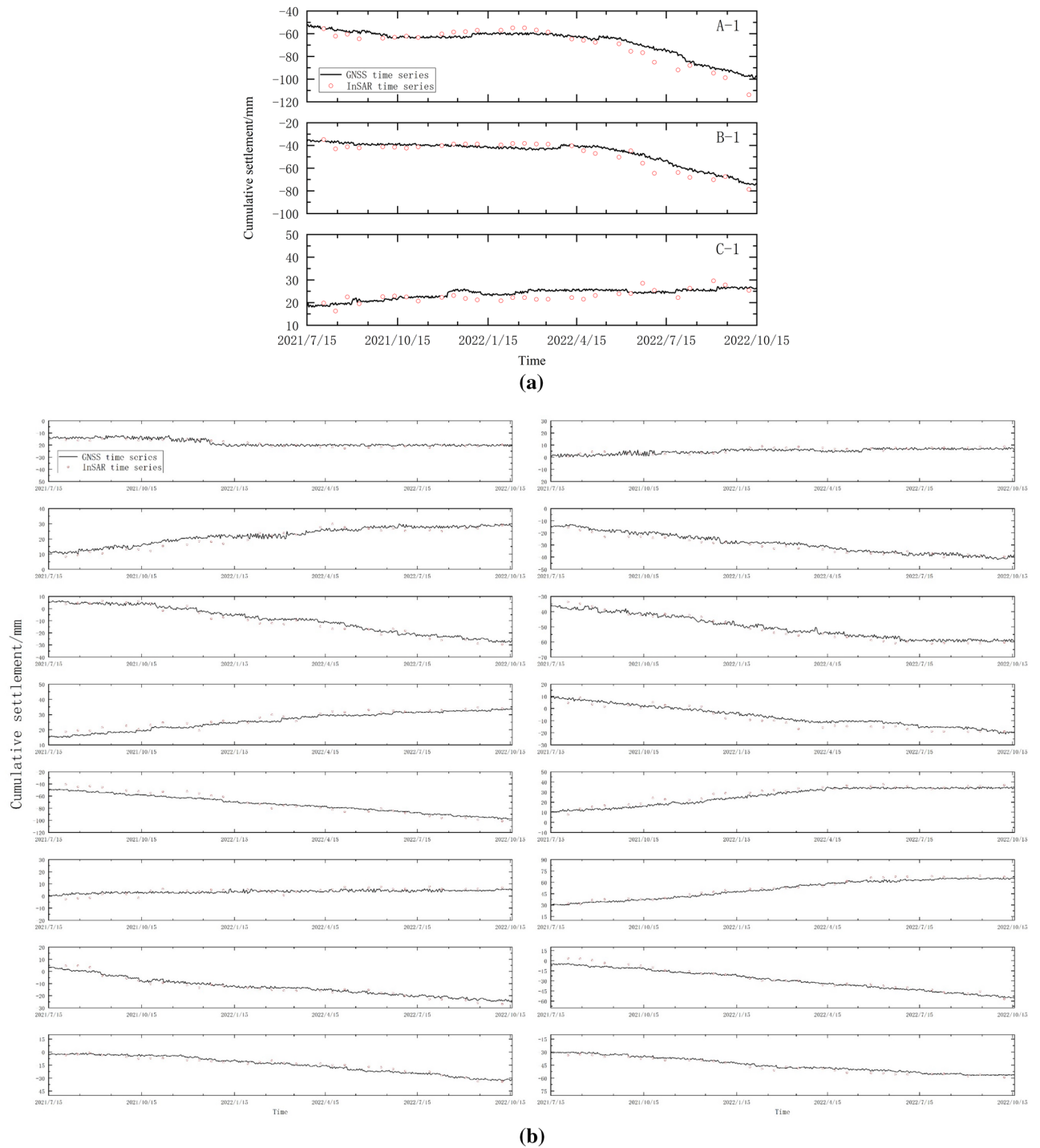
$$d_{LOS} = d_U \cos(\theta_{inc}) \tag{16}$$

To assess the prediction accuracy of feature points, we employ RMSE and MAE as evaluation metrics. The two parameters are defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{Z} - Z)^2}{N}} \tag{17}$$

$$MAE = \frac{\sum_{i=1}^N |\hat{Z} - Z|}{N} \tag{18}$$

where  $\hat{Z}$  is the predicted value,  $Z$  is the measured value.  $RMSE \in (0, +\infty]$ , the closer the matter is to 0, the more consistent the expected value is with the measured value.  $MAE \in (0, +\infty]$ , the closer the value is Closer to 0, then it reflects the combination model is better.



**Figure 8.** (a) Deformation time series of GNSS and InSAR unified coordinate frame of feature points (b) Railway along the GNSS and InSAR feature points unified coordinate system deformation time series.

For intra-regional prediction,  $RMSE_{zone}$  and coefficient of determination  $R^{25}$  are used to evaluate the overall prediction accuracy of the model. The two parameters are defined as follows:

$$RMSE_{zone} = \sqrt{\frac{\sum_{i=1}^N (Z_t(x_i, y_i) - Z_{I,t}(x_i, y_i))^2}{N}} \tag{19}$$

Points	MAE/mm	RMSE/mm	Points	MAE/mm	RMSE/mm
1	2.15	2.54	9	4.85	5.65
2	1.54	1.74	10	2.90	3.27
3	16.36	16.87	11	1.95	2.31
4	2.46	2.78	12	2.53	2.89
5	2.28	2.56	13	2.12	2.39
6	2.23	2.57	14	3.16	4.23
7	2.19	2.46	15	2.33	2.75
8	3.12	3.43	16	2.15	2.54

**Table 2.** InSAR accuracy evaluation near GNSS points at different locations.

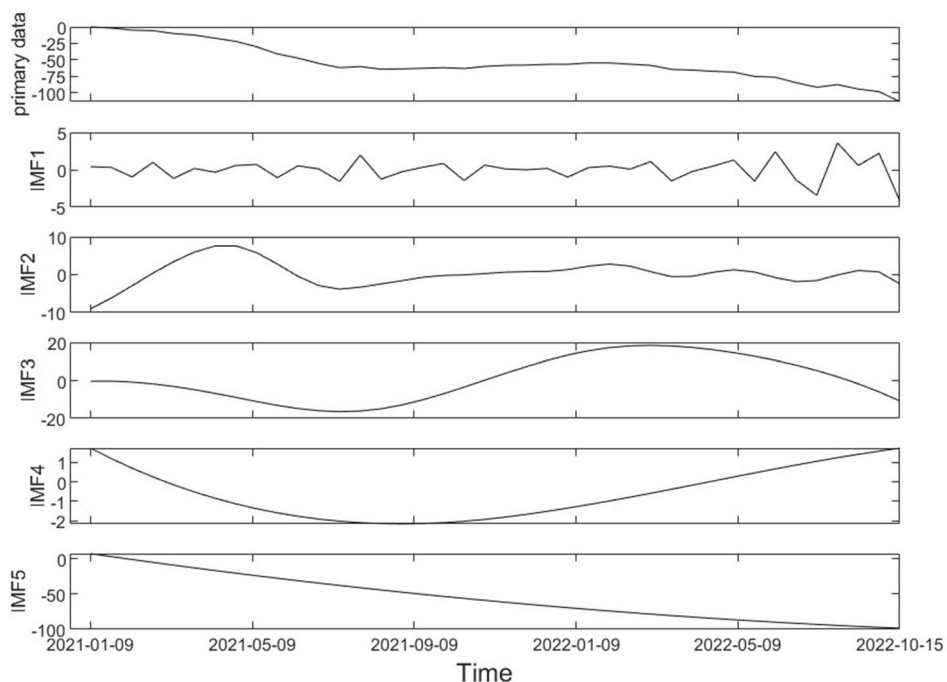
$$R^2(Z_t, \hat{Z}_{1,t}) = 1 - \frac{\sum_{i=0}^N (Z_t(x_i, y_i) - \hat{Z}_{1,t}(x_i, y_i))^2}{\sum_{i=0}^N (Z_t(x_i, y_i) - \bar{Z}_{1,t}(x_i, y_i))^2} \tag{20}$$

Among them,  $Z_t(x_i, y_i)$  represents the predicted settlement value at time  $t$  and position  $(x_i, y_i)$ , while  $Z_{1,t}(x_i, y_i)$  represents the observed settlement value obtained from InSAR monitoring at time  $t$  and position  $(x_i, y_i)$ . The coefficient of determination,  $R^2$ , quantifies the degree to which the independent variable can explain the changes in the dependent variable. A value closer to 1 indicates a stronger fit of the predictive model.  $\hat{Z}_{1,t}(x_i, y_i)$  represents the true value of the observed data, this article refers to the InSAR observed value.  $\bar{Z}_{1,t}(x_i, y_i)$  is the mean of the observed data.  $Z_t$  is the predicted value of the model.

### Single point forecasting and analysis

The subsidence time series data of feature points obtained by SBAS-InSAR technology are used as actual values for retrospective prediction research. Among them, the monitoring data from the initial 35 periods were selected as the sample data to predict the settlement for the next five periods, and the actual monitoring results were compared with the predicted results.

Firstly, EEMD decomposition is performed on the time series data of SBAS-InSAR monitoring points, resulting in the extraction of multiple IMF components and a residual component from the time series signal. This process allows us to capture the local characteristics of the monitoring points at different times, enabling the interpolation and prediction of their time series. By subjecting InSAR time series data to EEMD, the original signal can be approximated, and its components, referred to as IMF, can be extracted. Figure 9 presents a schematic diagram of the EEMD-decomposed IMF for a specific timing monitoring point of SBAS-InSAR. The time



**Figure 9.** EEMD decomposition InSAR timing.

series acquired through SBAS-InSAR technology is decomposed into five components, with each IMF component reflecting distinct information and containing oscillation characteristics of the original sequence. IMF1 represents the component with the highest frequency, while the frequency decreases sequentially for the subsequent IMF components. IMF5, being the component with the lowest frequency, captures the trend information of the original sequence. The downward trend observed in IMF5 aligns with the overall trend of the original sequence.

Table 1 clearly shows the SAR satellite has encountered data loss during a specific time interval, resulting in gaps in the subsidence time series used for SBAS-InSAR analysis of surface subsidence. These missing images present significant challenges to accurately assessing subsidence. In this study, the selected subsidence period spans from August 8, 2022, to October 7, 2022. Leveraging the periodicity of Sentinel-1A remote sensing images acquired every 12 days, we can infer the missing image timestamps by referencing to the information presented in Table 1. Specifically, the temporal gaps correspond to August 20, 2022, and September 25, 2022. Based on this premise, the initial selection for verification of the EEMD-Prophet prediction method includes feature points A-1, B-1, and C-1, which are in proximity to the GNSS base station. These selected feature points serve as the basis for conducting prediction experiments using the proposed EEMD-Prophet method. The predicted values for the missing time intervals are then compared with the corresponding GNSS measurements during those periods.

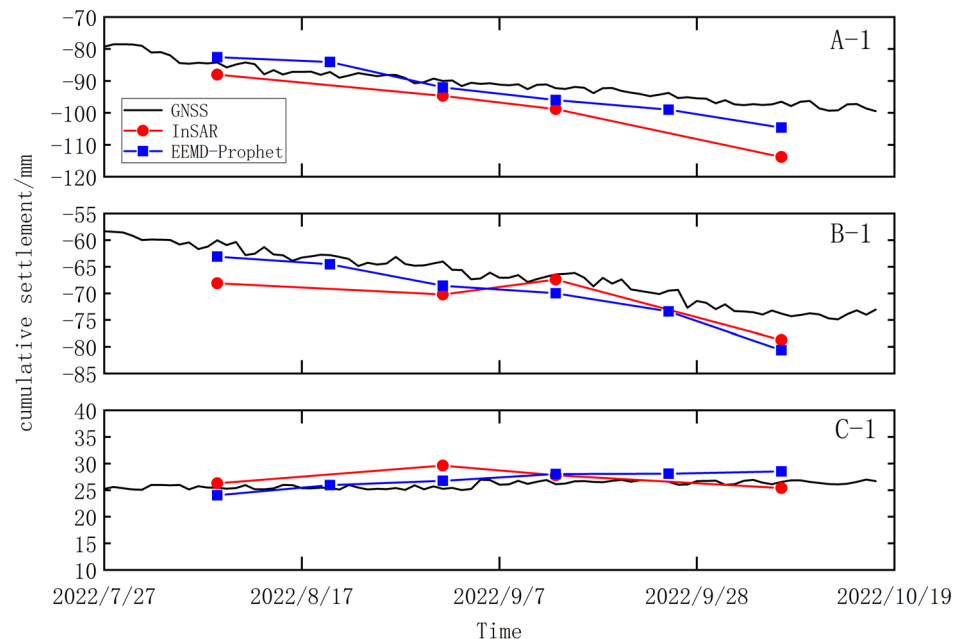
Figure 10 showcases the settlement prediction of feature points A-1, B-1, and C-1 utilizing the EEMD-Prophet method. It is evident that the projected values generated by the EEMD-Prophet method closely align with the GNSS measured values during the initial three prediction periods. However, the accuracy of the predictions gradually diminishes as the forecasting horizon extends. By incorporating span interpolation for the prediction time, the EEMD-Prophet method effectively anticipates the missing image data of SBAS-InSAR. The forecast results and corresponding relative error statistics for feature points A-1, B-1, and C-1 through the implementation of the EEMD-Prophet method are presented in Table 3. Notably, the missing images in InSAR correspond to August 20, 2022, and September 25, 2022.

To validate the effectiveness of the proposed method for settlement prediction along the railway, a single prediction method using EEMD-Prophet is insufficient. In this study, we conducted prediction experiments using the Prophet model, the ARMA model<sup>5,6</sup>, and the EEMD-Prophet model, as proposed in this paper. The experimental results were compared and analyzed to assess the feasibility of the proposed method.

For the purpose of comparative analysis, we selected feature points A-2, A-3, B-2, B-3, C-2, and C-3. These points were chosen because they are situated at a significant distance from the GNSS base station, and the accuracy of SBAS-InSAR technology meets the required standards. Consequently, the InSAR monitoring values of these feature points were employed as the actual observation data, and a retrospective prediction approach was adopted for comparative analysis and verification. The last five periods were designated as the forecast timeframe and compared against the actual observed values.

The prediction results depicted in Fig. 11 illustrate that the EEMD-Prophet prediction values exhibit closer proximity to the actual monitoring values compared to the Prophet model and the ARMA model. Overall, the EEMD-Prophet method demonstrates superior prediction performance. It achieves the highest accuracy during the initial three prediction intervals, gradually decreasing accuracy as the prediction time extends.

To analyze the effectiveness of the EEMD-Prophet model prediction method, various indicators such as RMSE and MAE were employed to compare and analyze the prediction outcomes. As shown in Table 4, the EEMD-Prophet prediction method outperforms the ARMA model and the Prophet model in terms of single-point

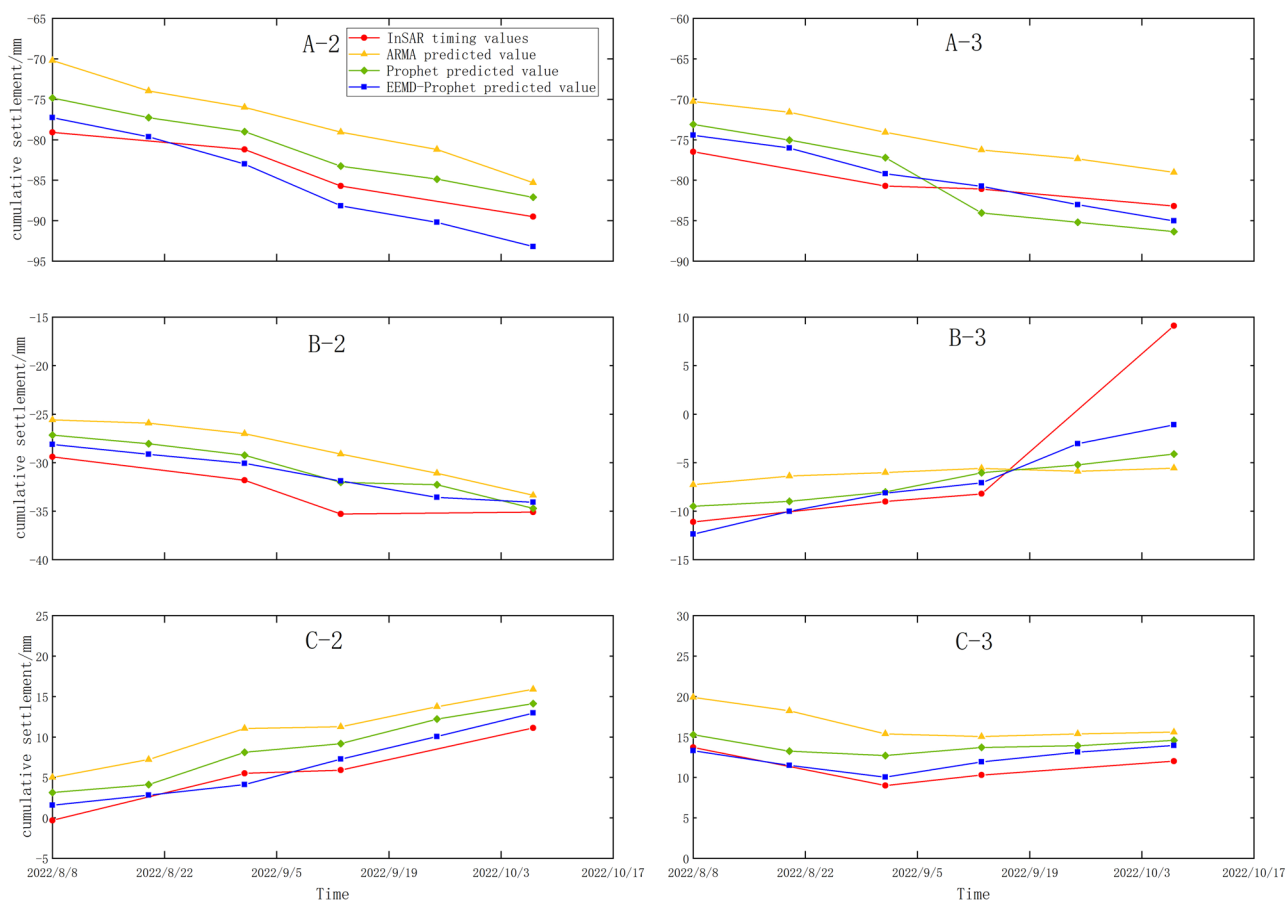


**Figure 10.** Comparison of EEMD-Prophet prediction results with GNSS.



Time	Point A-1		Point B-1		Point C-1	
	Predictive value/mm	Relative error/%	Predictive value/mm	Relative error/%	Predictive value/mm	Relative error/%
2022/8/8	-82.57	1.97	-63.11	5.12	24.02	5.63
2022/8/20	-84.06	3.61	-64.53	2.73	25.95	2.06
2022/9/1	-92.03	2.25	-68.57	7.12	26.76	5.96
2022/9/13	-96.00	4.01	-69.97	5.34	28.01	7.29
2022/9/25	-99.04	5.62	-73.41	5.65	28.09	5.74
2022/10/7	-104.60	8.37	-80.62	9.22	29.50	7.34
Average value		4.31		5.86		5.67

**Table 3.** Predictive value and relative error of EEMD-Prophet combined model.



**Figure 11.** Comparison of prediction models.

Modle point	ARMA		Prophet		EEMD-Prophet	
	RMSE/mm	MAE/mm	RMSE/mm	MAE/mm	RMSE/mm	MAE/mm
A-2	6.48	6.23	2.95	2.82	2.57	2.45
A-3	5.54	3.64	3.25	2.16	1.58	0.96
B-2	4.44	2.76	2.38	1.42	2.09	1.24
B-3	7.84	4.02	6.76	2.99	5.18	2.24
C-2	4.80	3.50	3.11	2.05	1.64	1.08
C-3	5.34	3.48	2.92	1.87	1.38	0.83
average value	5.74	3.94	3.56	2.22	2.41	1.47

**Table 4.** Prediction accuracy of different models.

prediction accuracy. The average RMSE and MAE values of the EEMD-Prophet method are recorded as 2.41 mm and 1.47 mm, respectively. The RMSE improves by 58.01% and 32.3% compared to the ARMA model and the Prophet model, respectively. Similarly, the MAE exhibits improvements of 62.69% and 33.78% when compared to the ARMA model and the Prophet model, respectively.

### Regional forecast and analysis

In Sect. 4.2, we employed the EEMD-Prophet, Prophet, and ARMA models to predict the settlement time series at a single feature point. In this section, we focus on the prediction of settlement at Taigu East Station, a significant area along the Taijiao high-speed railway. The geographical location of Taigu East Station is depicted in Fig. 4. Furthermore, Fig. 12 displays the region encompassing Taigu East Station and its 455 corresponding InSAR monitoring points in Google Earth Pro (V7.3.6.9345; <https://www.google.com/earth/about/versions/#download-pro>). These monitoring points are utilized for prediction and evaluation using the EEMD-Prophet, Prophet, and ARMA models. Figure 13 is a visual representation of ground settlement generated through Kriging interpolation<sup>26</sup> after predicting the monitoring points in the Taigu East Station area.

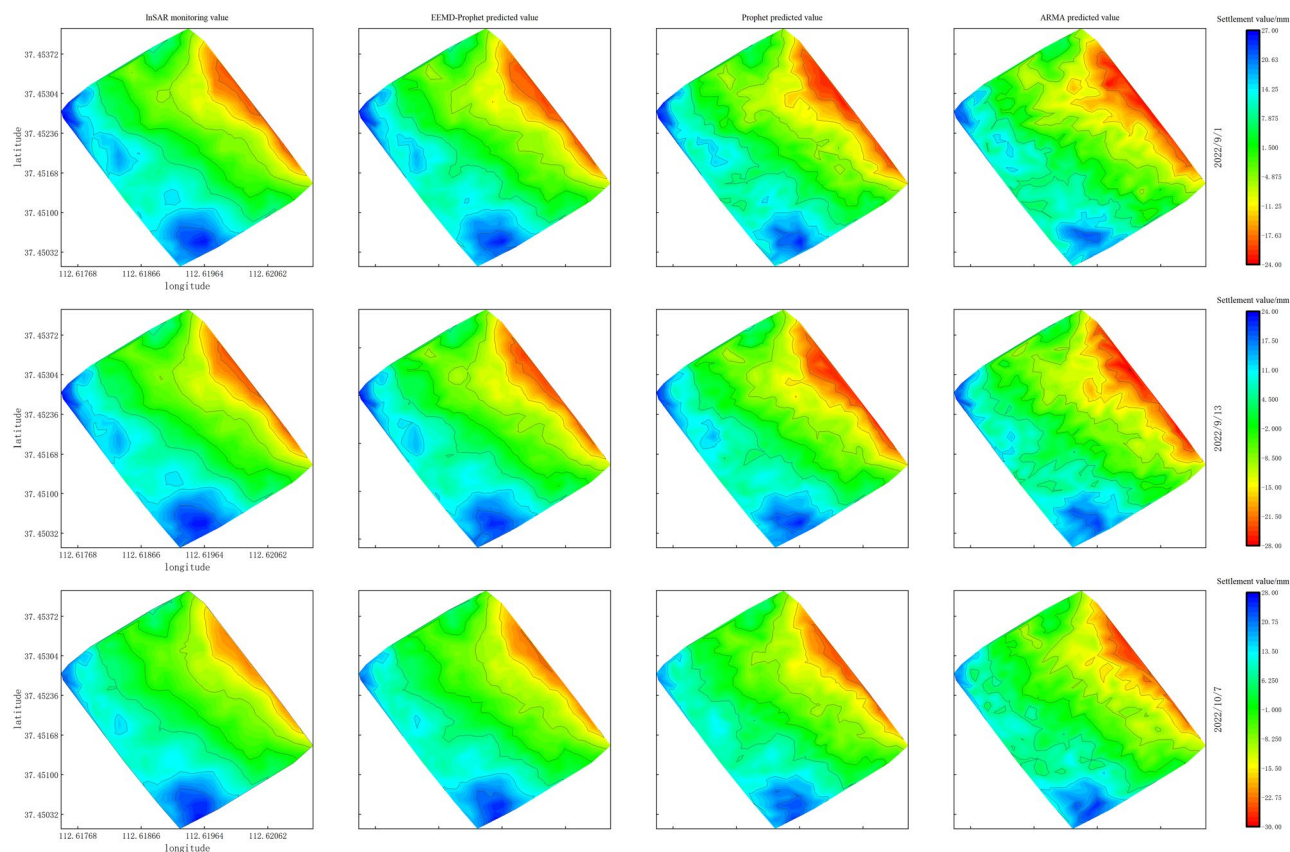
Upon examination, two prominent uplift areas and one notable subsidence area are identified within the vicinity of Taigu East Station. The subsidence area is situated at the intersection of the railway line and Taigu East Station. This observation suggests that, once trains commence operations after railway construction is completed, they exert pressure on the foundation, resulting in ground subsidence. Of particular note, the proximity of the railway station to the railway line indicates that the train load on this site may be substantial, resulting in noticeable settlement.

Figure 13 presents the settlement prediction of the Taigu East Station area using different prediction models, resulting in a ground settlement map. The EEMD-Prophet prediction model exhibits consistency with the actual InSAR land subsidence. However, as the prediction time increases, the established subsidence map gradually deviates from the actual InSAR ground subsidence map. To assess the accuracy and precision of each prediction method for regional prediction, we utilize the  $RMSE_{zone}$ , coefficient of determination  $R^2$ , and the time required for Taigu East Station prediction. Table 5 displays the average indicators of the prediction models at three prediction time points.

When examining the results, it becomes evident that the EEMD-Prophet prediction model outperforms both the ARMA model and the Prophet model in terms of regional prediction accuracy, achieving an  $RMSE_{zone}$  of 3.23 mm. Compared to the ARMA model and the Prophet model, the EEMD-Prophet model yields an improve in  $RMSE_{zone}$  by 43.53% and 12.47%, respectively. Moreover, the coefficient of determination for the EEMD-Prophet model is closer to 1, reaching 0.79, surpassing the other two prediction models. While the EEMD-Prophet model requires an additional 22s for EEMD decomposition on top of the Prophet model, its running speed exhibits a 12.45% improvement compared to the calculation time of the ARMA model.



**Figure 12.** Distribution of InSAR monitoring points at Taigu East Station.



**Figure 13.** Land subsidence of Taigu East Station regional prediction model.

Evaluation index	RMSE <sub>zone</sub>	R <sup>2</sup>	Time
EEMD-Prophet	3.23mm	0.79	7min28s
Prophet	3.69mm	0.74	7min06s
ARMA	5.72mm	0.68	8min32s

**Table 5.** Evaluation indicators of the model in regional prediction.

## Conclusions

This paper utilizes 40 SAR images to monitor surface subsidence in the study area through the SBAS method. By analyzing the 40-period time-series subsidence changes and cumulative subsidence, the SBAS monitoring values are employed as the time-series data sample for retrospective forecasting research. For forecasting research, the EEMD-Prophet forecasting model, the Prophet model, and the ARMA model are employed. Overall, the combined model demonstrates higher prediction accuracy compared to the Prophet model and the ARMA model.

In the study of SBAS-InSAR settlement, this paper introduces the EEMD-Prophet model, which combines EEMD decomposition and Prophet prediction, to interpolate and predict InSAR monitoring time series. The proposed method exhibits strong adaptability and provides favorable prediction results for analyzing similar subsidence time series signal data. However, it should be noted that the decomposition and prediction process of the EEMD-Prophet model introduces an additional workload. Throughout the research process, it is crucial to consider that settlement changes along the high-speed railway are influenced by a variety of factors. Therefore, there is a need for further investigation into the development of an efficient and self-adaptive prediction method capable of capturing sudden changes over time.

In addition to the methods explored in this study, it's worth noting that machine learning techniques, such as LSTM, have shown promise in the prediction of ground subsidence. LSTM, with its ability to capture complex temporal dependencies in data, can be particularly useful in modeling and predicting land subsidence time series, especially in urban areas and regions influenced by anthropic activities. While the EEMD-Prophet model demonstrated enhanced predictive accuracy, combining it with machine learning methods could potentially offer even more robust results, especially in scenarios where sudden changes or non-linear patterns in subsidence need to be captured in a timely manner. This integration of techniques could represent an exciting avenue for further research in the field of land subsidence analysis and prediction.

## Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 24 May 2023; Accepted: 1 January 2024

Published online: 06 January 2024

## References

1. Qiu, D. *et al.* Settlement monitoring data fusion approach for high-speed railways based on GNSS and InSAR. *J. Appl. Remote Sens.* **17**(3), 034507–034507 (2023).
2. Luo, Q., Li, J. & Zhang, Y. Monitoring subsidence over the planned jakarta-bandung (indonesia) high-speed railway using sentinel-1 multi-temporal InSAR data. *Remote Sens.* **14**(17), 4138. <https://doi.org/10.3390/rs14174138> (2022).
3. Zhou, D., Zuo, X. & Zhao, Z. Constructing a large-scale urban land subsidence prediction method based on neural network algorithm from the perspective of multiple factors. *Remote Sens.* **14**(8), 1803. <https://doi.org/10.3390/rs14081803> (2022).
4. Zhou, Q., Hu, Q., Ai, M., Xiong, C. & Jin, H. An improved GM (1, 3) model combining terrain factors and neural network error correction for urban land subsidence prediction. *Geomat. Nat. Hazards Risk.* **11**(1), 212–229. <https://doi.org/10.1080/19475705.2020.1716860> (2020).
5. Cui, Z. D. & Ren, S. X. Prediction of long-term settlements of subway tunnel in the soft soil area. *Nat. Hazards* **74**, 1007–1020 (2014).
6. Liu, W. F. *et al.* Time series forecasting fusion network model based on prophet and improved LSTM. *Comput. Mater. Contin.* **74**(2). <https://doi.org/10.32604/cmc.2023.032595> (2022).
7. Yang, Z., Xi, W., Yang, Z., Shi, Z. & Qian, T. Monitoring and prediction of glacier deformation in the meili snow mountain based on InSAR technology and GA-BP neural network algorithm. *Sensors.* **22**(21), 8350. <https://doi.org/10.3390/s22218350> (2022).
8. Xiong, Z. *et al.* Settlement prediction of reclaimed coastal airports with InSAR observation: A case study of the Xiamen Xiang'an International Airport, China. *Remote Sens.* **14**(13), 3081. <https://doi.org/10.3390/rs14133081> (2022).
9. Shi, X. *et al.* Monitoring and predicting the subsidence of dalian jinzhou bay international airport, china by integrating InSAR observation and Terzaghi consolidation theory. *Remote Sens.* **14**(10), 2332 (2022).
10. Chen, Y. *et al.* Prediction of InSAR deformation time-series using a long short-term memory neural network. *Int. J. Remote Sens.* **42**(18), 6919–6942 (2021).
11. Liu, Q., Zhang, Y., Wei, J., Wu, H. & Deng, M. HLSTM: Heterogeneous long short-term memory network for large-scale InSAR ground subsidence prediction. *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.* **14**, 8679–8688 (2021).
12. Moretto, S., Bozzano, F. & Mazzanti, P. The role of satellite InSAR for landslide forecasting: Limitations and openings. *Remote Sens.* **13**(18), 3735. <https://doi.org/10.3390/rs13183735> (2021).
13. Taylor, S. J. & Letham, B. Forecasting at scale. *Am. Stat.* **72**(1), 37–45 (2018).
14. Tang, J., Li, Y., Yang, D. & Ding, M. An approach for predicting global ionospheric TEC using machine learning. *Remote Sens.* **14**(7), 1585. <https://doi.org/10.3390/rs14071585> (2022).
15. Wu, Z. & Huang, N. E. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **1**(01), 1–41 (2009).
16. Huang, N. E. *et al.* The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceed. R. Soc. Lond. Series Math. Phys. Eng. Sci.* **1998**(454), 903–995. <https://doi.org/10.1098/rspa.1998.0193> (1971).
17. Tang, H. *et al.* A new hybrid forecasting model based on dual series decomposition with long-term short-term memory. *Int. J. Intell. Syst.* (2023).
18. Berardino, P., Fornaro, G., Lanari, R. & Sansosti, E. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Trans. Geosci. Remote Sens.* **40**(11), 2375–2383 (2002).
19. Zhang, P., Guo, Z., Guo, S. & Xia, J. Land subsidence monitoring method in regions of variable radar reflection characteristics by integrating PS-InSAR and SBAS-InSAR techniques. *Remote Sens.* **14**(14), 3265. <https://doi.org/10.3390/rs14143265> (2022).
20. Zhou, Z. Q., Zhou, S. J. & Tao, R. Surface subsidence prediction based on EWT-Prophet method. *J. Geod. Geodyn.* **42**(03), 247–252 (2022).
21. Wang, S., Shi, K., Chen, W., Chen, Y. & Rabe, M. Deformation stability analysis of foundation of Taijiao high-speed railway above the mined-out area affected by rainfall. *Acta Montan. Slovaca* **27**(2), 553–562 (2021).
22. Feng, Y. *et al.* Automatic selection of permanent scatterers-based GCPs for refinement and reflattening in InSAR DEM generation. *Int. J. Digital Earth* **15**(1), 954–974. <https://doi.org/10.1080/17538947.2022.2070678> (2022).
23. Yan, H. *et al.* Fusion of spatially heterogeneous GNSS and InSAR deformation data using a multiresolution segmentation algorithm and its application in the inversion of slip distribution. *Remote Sens.* **14**(14), 3293. <https://doi.org/10.3390/rs14143293> (2022).
24. Lei, K. C. *et al.* Three-dimensional surface deformation characteristics based on time series InSAR and GPS Technologies in Beijing, China. *Remote Sens.* **13**(19), 3964. <https://doi.org/10.3390/rs13193964> (2021).
25. Li, H., Qin, C., He, W., Sun, F. & Du, P. Learning and inferring the diurnal variability of cyanobacterial blooms from high-frequency time-series satellite-based observations. *Harmful Algae* **123**, 102383. <https://doi.org/10.1016/j.hal.2023.102383> (2023).
26. Bajni, G., Camera, C. A., Brenning, A. & Apuani, T. Assessing the utility of regionalized rock-mass geomechanical properties in rockfall susceptibility modelling in an alpine environment. *Geomorphology* **415**, 108401 (2022).

## Author contributions

D.Q. and Y.T. wrote the main manuscript text. Y.W. and T.L. performed the data analysis. K.D. completed the design and implementation of theoretical methods. S.W. reviewed and edited the manuscript.

## Funding

Supported by the National Natural Science Foundation of China under Grant 61902016; the Postgraduate Education and Teaching Quality Improvement Project of the Beijing University of Civil Engineering and Architecture (BUCEA), China (No. J2022005); the BUCEA Postgraduate Innovation Project, China (No. PG2023119).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to W.S.



**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024