FULL PAPER

A method for correcting InSAR interferogram errors using GNSS data and the K-means algorithm

Huineng Yan[1], Wujiao Dai[1*] , Wenbin Xu[1], Qiang Shi[4], Kai Sun[1], Zhigang Lu[2,3] and Rui Wang[2,3]

Abstract

Correcting interferometric synthetic aperture radar (InSAR) interferograms using Global Navigation Satellite System (GNSS) data can effectively improve their accuracy. However, most of the existing correction methods utilize the difference between GNSS and InSAR data for surface fitting; these methods can effectively correct overall long-wavelength errors, but they are insufficient for multiple medium-wavelength errors in localized areas. Based on this, we propose a method for correcting InSAR interferograms using GNSS data and the K-means spatial clustering algorithm, which is capable of obtaining correction information with high accuracy, thus improving the overall and localized area error correction effects and contributing to obtaining high-precision InSAR deformation time series. In an application involving the Central Valley of Southern California (CVSC), the experimental results show that the proposed correction method can effectively compensate for the deficiency of surface fitting in capturing error details and suppress the effect of low-quality interferograms. At the nine GNSS validation sites that are not included in the modeling process, the errors in the ascending track 137A and descending track 144D are mostly less than 15 mm, and the average root mean square error values are 11.8 mm and 8.0 mm, respectively. Overall, the correction method not only realizes effective interferogram error correction, but also has the advantages of high accuracy, high efficiency, ease of promotion, and can effectively address large-scale and high-precision deformation monitoring scenarios.

Keywords InSAR, GNSS, Error correction, K-means algorithm, Interferogram quality, High-precision deformation time series
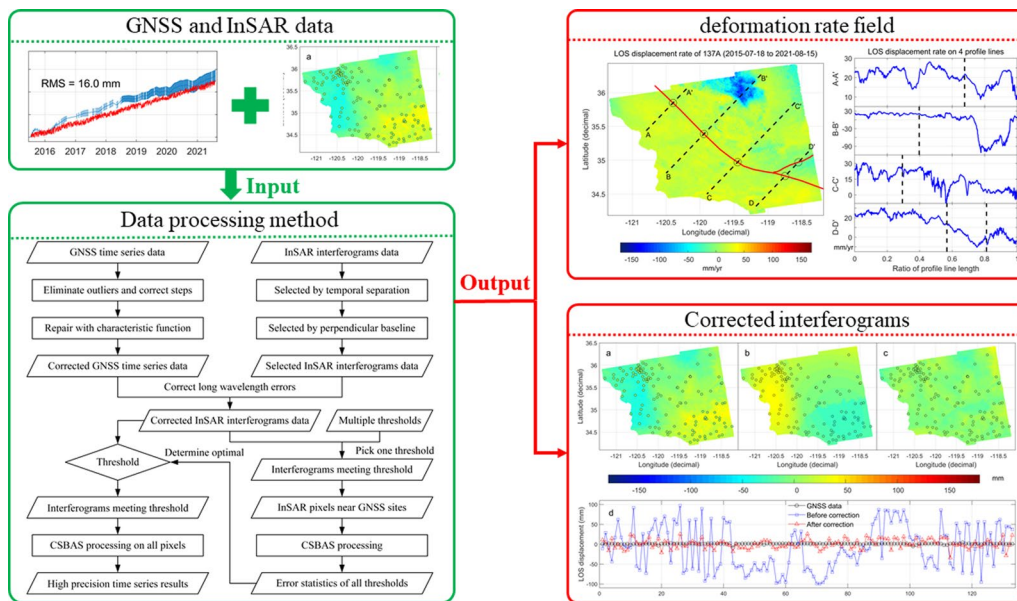
*Correspondence:
Wujiao Dai
wjdai@csu.edu.cn
Full list of author information is available at the end of the article

Graphical Abstract

Introduction

As two current advanced deformation monitoring technologies, interferometric synthetic aperture radar (InSAR) and Global Navigation Satellite System (GNSS) are the main sources of high-precision deformation monitoring data (Takada et al. 2018; Xu et al. 2018; Yan et al. 2022b; Zhang et al. 2023; Zhu et al. 2023). GNSS sites are mostly deployed with deep anchors in bedrock or anchored through reinforced concrete structures and stable buildings (Xi et al. 2021; Xu et al. 2022); thus, these monitoring sites have fixed location properties. Furthermore, their errors can be suppressed by performing multisite combination networking calculations and error modeling (Dai et al. 2019; Guns et al. 2022). However, InSAR technology is limited by its own imaging and positioning principles, resulting in a lack of an absolute reference frame for monitoring data, while the achieved monitoring precision is easily affected by factors such as orbital and atmospheric errors (Yang et al. 2020; Lee and Shirzaei 2023).

To overcome the serious effect of atmospheric errors on InSAR data, correcting long- and medium-wavelength errors in InSAR interferograms via GNSS data is an effective method. Neely et al. (2019) utilized a quadratic surface to fit the differences between GNSS data and interferograms, after recovering the fitted difference into interferograms, applied time series analysis processing to obtain the corresponding InSAR deformation time series.

Weiss et al. (2020) projected the horizontal velocity field obtained by interpolating GNSS displacement data in the line of sight (LOS) direction, then removed the difference between the GNSS and InSAR data by fitting a second-order polynomial to obtain a high-precision deformation rate field. Xu et al. (2021) used Gaussian low-pass filtering to process interpolation results concerning the difference between GNSS data and interferograms, and after removing this difference, the corrected interferograms were processed with the coherence-based small baseline subset (SBAS) method and atmospheric phase correction to obtain a high-precision InSAR deformation time series.

The above studies employed high-precision GNSS data and directly utilized the difference between GNSS and InSAR data to correct interferograms. When the spatial distribution of errors is long-wavelength, these methods can effectively suppress the error effect. However, when multiple error blocks are clustered in an interferogram due to medium-wavelength errors, the corrected results obtained after surface fitting are prone to spatial error aggregation, which leads to low precision in the deformation time series. In addition, with the development of cloud computing and cloud platforms for InSAR data processing (Lazeckỳ et al. 2020; González-Jiménez et al. 2023), additional publicly released interferogram products have become available, greatly reducing the computational volume and difficulty of InSAR data processing.

Yan *et al. Earth, Planets and Space*   (2024) 76:51

Page 3 of 19

However, since the interferograms released on cloud platforms are processed in batches with the same solution programs and scripts, some low-quality interferograms are produced with dynamic atmospheric errors, and they are reflected in the spatial aggregation of non-tectonic deformation signals. For such interferogram products, it is difficult to ensure their correction accuracy with surface fitting.

Therefore, to fully utilize the publicly available interferogram data from cloud platforms and to improve the validity of GNSS-corrected interferograms, it is necessary to consider the spatial distributions of errors (e.g., using K-means spatial clustering algorithms) and evaluate the quality of the produced interferograms through exact indicators. By combining these two considerations, a correction method for InSAR data is studied in this paper, to reduce the long- and medium-wavelength errors in a single interferogram, improve the precision of the InSAR deformation time series results, also to reduce the computational complexity and application difficulty induced by GNSS-corrected interferogram errors.

In this paper, we utilize GNSS data and K-means algorithm to correct the long- and medium-wavelength errors of the interferograms, by controlling the quality of the corrected interferograms and performing the time series analysis process, we obtain high-precision InSAR deformation time series results. Then apply the proposed method to the Central Valley of Southern California (CVSC) and present the specific implementation process. Finally, the experimental results are discussed, where the advantages and future research directions of the developed correction method are summarized.

Data processing method

To obtain high-precision InSAR deformation time series, correcting long- and medium-wavelength errors and performing quality control on the corrected interferograms are necessary. For error correction purposes, we consider the spatial distribution characteristics of the observed errors through the K-means algorithm; for quality control, we use the residuals of the detrended term deformation as the evaluation indicators. In addition, we further improve the computational efficiency of the method by using GNSS data to assist in quickly finding the optimal quality threshold.

During the implementation of the correction method, the steps, outliers and modeling errors in the GNSS data are first processed, and the interferograms are selected according to the spatio-temporal baseline threshold. Then, the K-means algorithm and the difference between the GNSS data and interferograms are combined to correct the interferograms. Next, a set of quality thresholds is selected, the time series processing (only for InSAR

pixels that are adjacent to a GNSS site, we choose 15×15 pixels around each GNSS site) errors of the interferograms selected by each threshold are counted, and the optimal quality threshold is determined based on error minimization. Finally, high-quality interferograms are selected based on the optimal quality threshold, and SBAS processing is performed on all the pixels to obtain high-precision deformation time series results. The detailed data processing procedure is shown in Fig. 1.

GNSS data processing
Step terms, outlier terms, and observation errors inevitably appear in GNSS displacement data due to observation data quality issues, external environmental changes, instrumentation replacement, and error correction modeling (Amiri-Simkooei et al. 2019; Ohno et al. 2021; Agnieszka and Dawid 2022). We adopt the following strategies to address these different error terms:

a) Step terms

Step terms are caused mainly by instrument replacement and external environmental changes (Lahtinen et al. 2022). We identify the times at which steps occur from the step event log file, calculate the difference between the medians of the 30 days of time series data before and after the identified time, and recover the difference in the GNSS time series data if it is greater than the threshold.

b) Outlier terms

Outliers in a GNSS time series significantly deviate from the normal variation range (Wen et al. 2021). We use the mean plus and minus three times the standard deviation (STD) as a criterion for judging outliers and rejecting them directly.

c) Modeling errors

Observation errors are commonly found in GNSS displacement data (Li et al. 2018). By combining previous research results (Dong et al. 2002; Bogusz and Klos 2016) and analyzing deformation time series data acquired from several GNSS sites, we find that the deformation characteristic function can effectively describe the GNSS displacement, and the calculation formula is as follows:
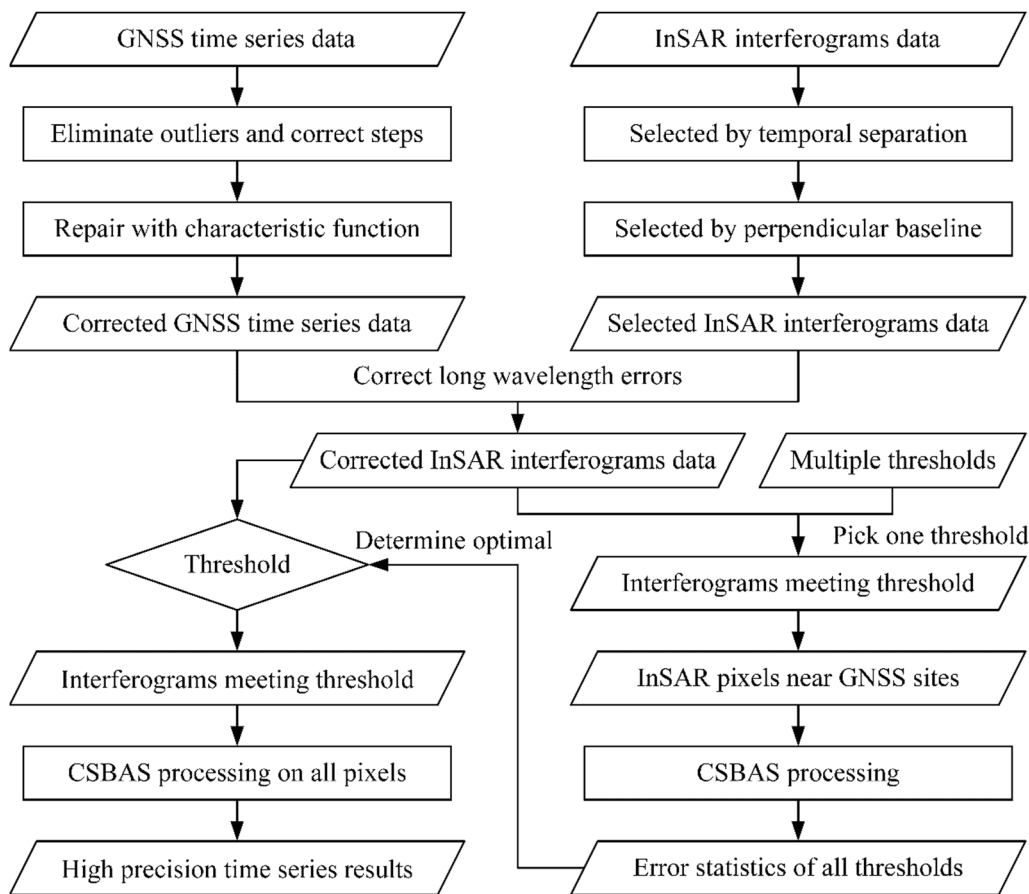
Yan *et al. Earth, Planets and Space* (2024) 76:51

Page 4 of 19

Fig. 1 Flowchart of InSAR and GNSS deformation data processing

$$
\begin{aligned}
y = {} & a_0 + a_1 \times t + a_2 \times \sin(2\pi f_1 t) + b_2 \\
& \times \cos(2\pi f_1 t) + a_3 \times \sin(2\pi f_2 t) \\
& + b_3 \times \cos(2\pi f_2 t) + a_4 \times \sin(2\pi f_3 t) \\
& + b_4 \times \cos(2\pi f_3 t) + \varepsilon,
\end{aligned}
\tag{1}
$$

where y is the GNSS observation, a_0, a_1 are the regression coefficients corresponding to the constant and velocity terms; $a_2, b_2, a_3, b_3, a_4, b_4$ are the regression coefficients corresponding to various periodic terms (annual, semiannual, and seasonal periodical terms); t is the observation time; f_1, f_2, f_3 represent the frequency values corresponding to the annual, semiannual, and seasonal periodical terms ($f_1 = 1/365, f_2 = 1/182.5, f_3 = 1/91.25$), respectively; and ε is the residual.

We first assume that the GNSS displacements include the various periodic terms in Eq. (1) and fit all the coefficients using the least-squares method (Levenberg 1944). Then, the *t*-distribution test statistics corresponding to the coefficients are constructed, and insignificant periodic terms are removed, as shown in Additional file 1: Fig. S1. The GNSS displacements are refitted by

the least-squares method using the updated deformation characteristic function, and the GNSS data are subsequently processed according to the fitted weights as follows: if a weight is less than the threshold value, the weighted average of the observation and the fitted result is used (Grigg and Spiegelhalter 2007); otherwise, the observation is left unmodified. The weighted averaging formula is as follows:

$$
y\prime = y p^2 + \widehat{y}\left(1 - p^2\right) p \le p_t,
\tag{2}
$$

where $y\prime$ is the weighted average, y is the GNSS observation, $\widehat{y}$ is the model prediction, p is the fitting weight, and p_t is the weight threshold.

According to the above method, the step terms, outlier terms and modeling errors in the GNSS displacement data are processed separately, to obtain "clean" GNSS displacement data for the subsequent data processing step.

InSAR interferogram data processing

Due to the influence of atmospheric errors and the limitations of error correction models, several long- and

medium-wavelength errors and topography-related errors are often exist in interferograms and cannot be eliminated (Li et al. 2019; Morishita and Kobayashi 2022; Yan et al. 2022a). In addition, the deformation information in interferograms is relative to the reference point. To correct the errors in interferograms and ensure that they are tightly aligned with the GNSS reference frame, we follow the basic principles of the methods adopted by Neely et al. (2019) and Xu et al. (2021), and carry out error correction based on the difference between the GNSS data and interferogram data. The specific process is shown in Fig. 2.

First, an interferogram is selected based on the time span and vertical baseline length. The GNSS displacement is projected in the InSAR LOS direction as follows (Kobayashi et al. 2018):

$$d_{\text{GNSS}} = \mp d_n \sin(\alpha)\sin(\theta) \pm d_e \cos(\alpha)\sin(\theta) \mp d_u \cos(\theta), \quad (3)$$

$$\Delta d_{\text{GNSS}}(t_{ij}) = d_{\text{GNSS}}(t_{ij}) - d_{\text{InSAR}_{\text{GNSS}}}(t_{ij}), \quad (4)$$

where d_{GNSS} is the LOS displacement at the GNSS site, positive values indicate surface displacements toward the satellite, and negative values indicate motion away from the satellite. d_n, d_e and d_u represent the displacement components of the GNSS in the north, east and vertical directions, respectively. θ is the radar incidence angle, and α is the satellite heading angle (clockwise from the north is positive). $\Delta d_{\text{GNSS}}(t_{ij})$ is the difference between the GNSS data and interferogram data from time t_i to time t_j, $d_{\text{GNSS}}(t_{ij})$ and $d_{\text{InSAR\_GNSS}}(t_{ij})$ are the LOS displacements at the GNSS site and the InSAR pixels near the GNSS site from time t_i to time t_j, respectively. For the plus-minus signs in Eq. (3), the upper sign represents the right-looking direction, and the lower sign represents the left-looking direction.

Then, the differences between the GNSS data and interferogram data are spatially clustered using the K-means algorithm (MacQueen 1967; Arthur and Vassilvitskii 2007). The K-means algorithm divides the data into K groups according to the setting, through randomly selects K clustering centers, calculates the distances from each object to the clustering centers, assigns each object to the clustering center that is closest to it, and repeats the process until the clustering centers no longer change. Considering the number of medium-wavelength
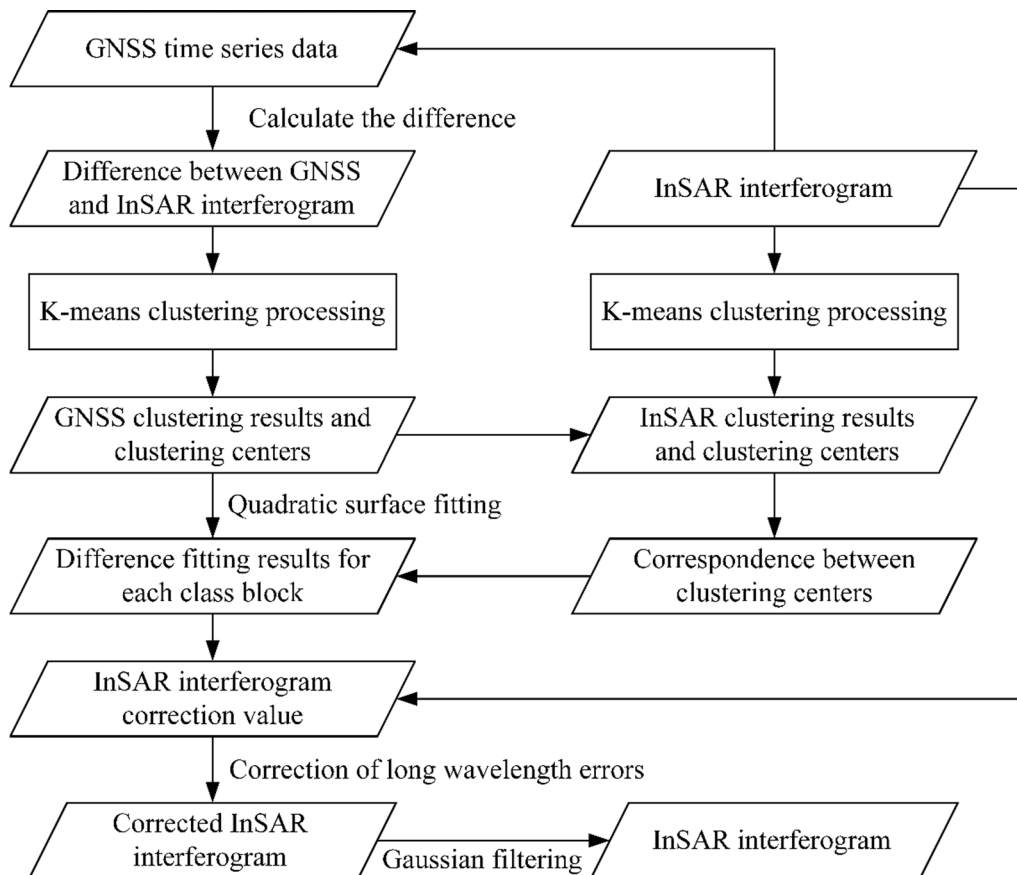

Fig. 2 Flowchart for correcting long- and medium-wavelength errors in InSAR interferograms

atmospheric errors in the interferograms and ensuring that the number of GNSS sites in each cluster is sufficient for quadratic surface fitting (greater than 7), we loop through the number of clusters to be selected, compute the difference between the corrected interferograms and the GNSS data, and determine the optimal clusters based on difference minimization.

A quadratic surface is fitted to the difference between the GNSS and InSAR data for the interferogram spanning from t_i to time t_j in each clustered block. The coefficient terms are estimated by the least-squares method. The formula is as follows:

$$
\begin{bmatrix}
\Delta d_{\mathrm{GNSS}}(t_{ij}, L_1, B_1) \\
\Delta d_{\mathrm{GNSS}}(t_{ij}, L_2, B_2) \\
. \\
. \\
. \\
\Delta d_{\mathrm{GNSS}}(t_{ij}, L_m, B_m)
\end{bmatrix}
=
\begin{bmatrix}
1\,L_1\,B_1\,L_1B_1\,L_1{}^2B_1\,L_1B_1{}^2\,L_1{}^2B_1{}^2 \\
1\,L_2\,B_2\,L_2B_2\,L_2{}^2B_2\,L_2B_2{}^2\,L_2{}^2B_2{}^2 \\
. \\
. \\
. \\
1\,L_m\,B_m\,L_mB_m\,L_m{}^2B_m\,L_mB_m{}^2\,L_m{}^2B_m{}^2
\end{bmatrix}
\begin{bmatrix}
a_{00} \\
a_{10} \\
a_{01} \\
a_{11} \\
a_{21} \\
a_{12} \\
a_{22}
\end{bmatrix}
+
\begin{bmatrix}
\varepsilon_1 \\
\varepsilon_2 \\
. \\
. \\
. \\
\varepsilon_m
\end{bmatrix},
\tag{5}
$$

where $L_1, B_1, \ldots, L_m, B_m$ denote the longitudes and latitudes of the GNSS sites in the current cluster block, $a_{00}, a_{10}, \ldots, a_{22}$ denote the coefficient terms of the quadratic surface, and $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$ denote the fitting residuals.

Substituting the corresponding longitudes and latitudes of the InSAR pixels into Eq. (5), the differences between the GNSS and InSAR data at arbitrary pixels in a specific block can be obtained.

Then, the deformation values of the interferograms are spatially clustered using the K-means algorithm (MacQueen 1967; Arthur and Vassilvitskii 2007), and the number of blocks is consistent with the number of spatial clusters formed for the GNSS data. Based on the correspondence between the spatial clustering centroids of the interferogram and the GNSS data, the difference fitting results corresponding to each clustering block are combined to obtain the correction for the current interferogram.

To weaken the block boundary effect exhibited by the corrections, we apply Gaussian low-pass filtering to the corrected interferograms. After comparing the filtering effects produced under different wavelengths, we adopt 80 km as the filtering wavelength (which is consistent with that adopted by (Xu et al. 2021)) and then apply corrections to the interferograms.

By processing each interferogram via the above method, corrected LOS deformation values can be obtained, and the reference frame of each interferogram is consistent with the GNSS data.

Determination of the optimal quality threshold

High-quality interferograms are essential for obtaining high-precision displacement time series via SBAS processing. To minimize the impacts of interferograms with low data quality on the corresponding SBAS processing results, we develop criteria and methods for assessing the quality of interferograms.

Studies conducted by Solari et al. (2020), Li and Bürgmann (2021), and Liu et al. (2023) have shown that the STD of the deformation values can be used as an indicator of interferogram quality. Considering that different time spans result in interferograms containing varying degrees of deformation information (Suo et al. 2015; Wang et al. 2022), and that the influence of errors in the interferograms is significantly weakened after performing correction using GNSS data, we use the average of the absolute values obtained after removing the first-order deformation trend term as the quality indicator (Q_i). Due to the short time span of the selected interferograms, the theoretical values of the deformation residuals after removing the trend term are small, while large Q_i values generally correspond to low-quality interferograms with significant atmospheric error residuals or discontinuity error introduced by the K-means clustering algorithm.

The first-order trend term is derived from the sum of the deformation values of all interferograms divided by the total time span. The quality indicator is defined as follows:

$$
v_j = \sum_{i=1}^{N} d_{i,j} \Big/ \sum_{i=1}^{N} t_{i,j}, \tag{6}
$$

$$
\widetilde{d}_{i,j} = d_{i,j} - v_j \times t_{i,j}, \tag{7}
$$

$$
Q_i = \frac{1}{n} \sum_{j=1}^{n} \mathrm{abs}\left(\widetilde{d}_{i,j}\right), \tag{8}
$$

where v_j denotes the first-order trend term of the j th pixel, N denotes the number of interferograms, $d_{i,j}$ and $t_{i,j}$ denote the deformation value and time span corresponding to the j th pixel in the i th interferogram,

respectively, and $\tilde{d}_{i,j}$ denotes the deformation value of the j th pixel in the i th interferogram after removing the first-order trend term. Q_i denotes the quality indicator of the i th interferogram, which is effective if the temporal phase change is stationary throughout the image and is useful for identifying significant uncorrected residuals in interferograms. n denotes the number of pixels, and *abs* denotes the operation used to calculate the absolute value of a term.

To determine the optimal quality threshold, a simple method is to determine a set of quality thresholds within a certain range, select a threshold in turn to pick the interferograms for SBAS processing (for all pixels) and statistical error calculation, finally, determine the optimal value based on the errors corresponding to different thresholds. However, if the eligible interferograms need to be selected for SBAS processing to every selected threshold, then the number of required computations is certainly unacceptable. Considering the SBAS processing strategy utilized for each selected quality threshold, the final accuracy evaluation is performed on the InSAR pixels in the vicinity of the GNSS site. Among the interferograms that satisfy the quality threshold, to prevent anomalous variations in individual pixels from affecting the solution, a certain range of InSAR pixels near the GNSS site is selected for SBAS processing; this range can include 10×10 or 15×15 pixels around the GNSS site. The processed results are subsequently compared with the GNSS time series at each site to evaluate the achieved SBAS accuracy. This processing strategy not only greatly reduces the computational burden, but also does not affect the SBAS accuracy evaluation corresponding to each quality threshold.

Based on the methodology described above, we divide the process of determining the quality threshold into four steps:

1) Calculate the first-order trend term using all the corrected interferograms (Eq. 6).
2) Calculate the mean of the absolute values of all the corrected interferograms after removing the first-order trend term (Eq. 8), and determine a series of alternative thresholds.
3) Process each alternative quality threshold in turn as follows: first, select the interferograms according to the current quality threshold; then, select the InSAR pixels near the GNSS sites for SBAS processing; finally, record the root mean square error (RMSE) between the obtained SBAS time series results and the GNSS displacement.

4) Based on the RMSE of the SBAS processing results obtained with different thresholds, select the threshold with the smallest RMSE as the optimal threshold.

To reduce unnecessary computational complexity, the third and fourth steps are accomplished using a two-step approach: the first step involves selecting a series of quality thresholds with large gradients and determining the approximate range of the optimal threshold; the second step involves selecting a series of quality thresholds with smaller gradients within the approximate range and determining the optimal threshold from this series.

SBAS processing

SBAS processing can be used to obtain displacement time series, and it has been applied to monitor the deformation trends of geologic hazards such as earthquakes (Liu et al. 2021; Jafari et al. 2023), volcanoes (Doke et al. 2021; Festa et al. 2022), and landslides (Luo et al. 2020; Khan et al. 2023). The basic principle is to consider the deformation value as the product of the deformation rate and the time span (Berardino et al. 2002; Lanari et al. 2004). Then, the observational equation is constructed by combining the data from several interferograms and back-calculating the deformation rate, and finally, the results for the deformation time series are obtained by integrating the deformation rate. To increase the resolution without significant spatial jumps and to avoid the problem of rank deficiency, smoothness constraints are added to make the equation have full rank (López-Quiroz et al. 2009; Osmanoğlu et al. 2016). The specific formula is as follows:

$$\begin{bmatrix} T \\ \lambda L \end{bmatrix} V = \begin{bmatrix} D \\ 0 \end{bmatrix}, \qquad (9)$$

where T denotes the length of the time span corresponding to the interferogram, λ denotes the smoothing factor (determined by the grid search method), L denotes the smoothing operator, V and D denote the displacement rate and deformation value within the time span corresponding to the interferogram, respectively, and 0 denotes the zero vector.

To increase the number of solvable InSAR pixels, we adopt the temporally connected SBAS (CSBAS) method proposed by Neely et al. (2019) for determining the deformation rate at all independently observed moments in the observation equation, which is calculated using the following formula:

$$
\begin{bmatrix}
t_2 - t_1 & 0 & & 0 & 0 & \cdots & 0 \\
t_2 - t_1 & t_3 - t_2 & & 0 & 0 & \cdots & 0 \\
& & \vdots & & & & \\
0 & 0 & 0 \; 0 \cdots & & & & t_p - t_{p-1} \\
-\lambda & \lambda & 0 \; 0 \cdots & 0 & & & 0 \\
0 & -\lambda & \lambda \; 0 \cdots & 0 & & & 0 \\
& & \vdots & & & & \\
0 & 0 & 0 \; 0 \cdots & -\lambda & & & \lambda
\end{bmatrix}
\begin{bmatrix}
v_{12} \\
v_{23} \\
\cdot \\
\cdot \\
\cdot \\
v_{n-1,n}
\end{bmatrix}
=
\begin{bmatrix}
D_1 \\
D_2 \\
\cdot \\
\cdot \\
\cdot \\
D_p \\
0 \\
0 \\
\cdot \\
\cdot \\
\cdot \\
0
\end{bmatrix},
\tag{10}
$$

where t_1, t_2, ..., t_p are the timing of SAR image acquisitions; $v_{n-1,n}$ is the deformation rate between t_{n-1} and t_n; and D_1, D_2, ..., D_p are the deformation values of the pixels in the interferogram.

After constructing the observation equation according to Eq. (10), least-squares inversion is performed using the Moore–Penrose pseudo-inversion method to obtain the deformation rate of the InSAR pixels for each observation period; then, the deformation time series results are obtained by executing time-domain integration on the deformation rate.

Accuracy evaluation

To evaluate the accuracy of the processing method, we selected several GNSS sites that were not included in the modeling and compared the InSAR deformation time series results obtained by the processing method. In addition, the InSAR deformation rate results are compared
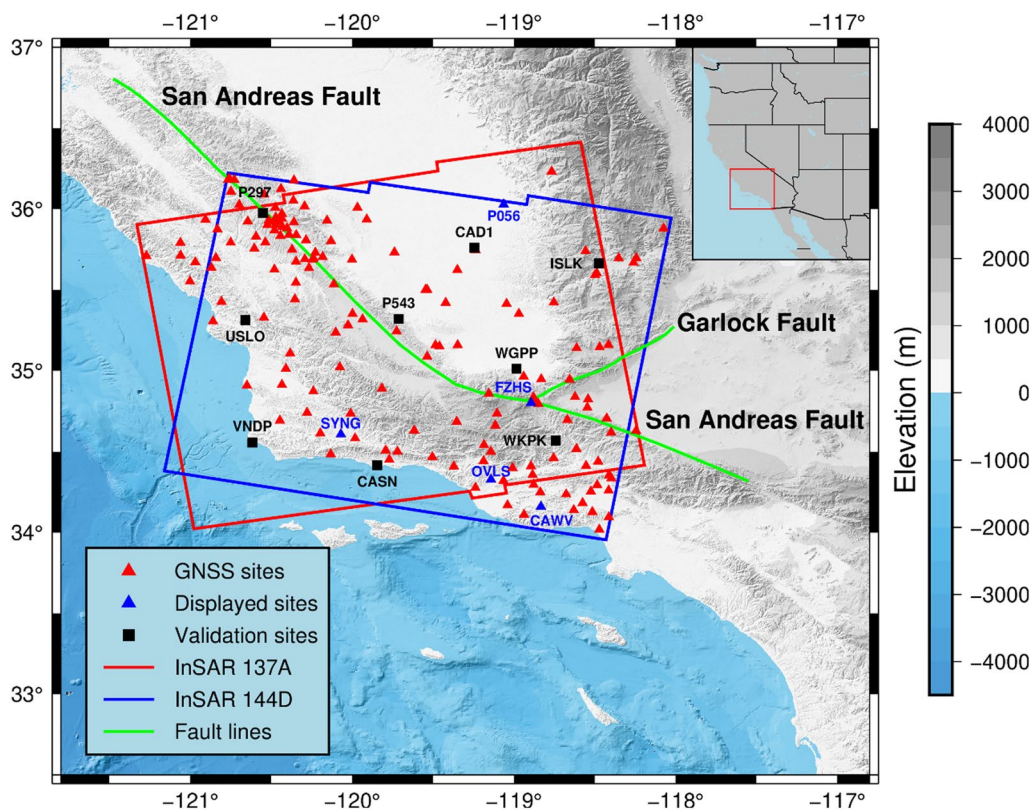

Fig. 3 Topographic and spatial distribution of GNSS sites (red triangles and blue triangles indicate modeling sites, and blue triangles indicate displayed sites, black squares indicate validation sites), InSAR data (red lines indicate ascending track 137A, blue lines indicate descending track 144D) and major faults (green lines) covering the Central Valley of Southern California. The red box in the upper right inset depicts the study area

Yan *et al. Earth, Planets and Space* (2024) 76:51

Page 9 of 19

with those of previous studies, focusing on the magnitude and spatial distribution of the deformation.

Data processing experiment

Using our proposed method, we processed InSAR and GNSS data from the CVSC.

Data sources

The interferograms and GNSS data used here were released by scientific institutions. The employed interferograms were published by LiCSAR (Yu et al. 2018; Morishita et al. 2020; Lazeckỳ et al. 2020), and they include an ascending track frame ID 137A_05534_131822 (137A) and a descending track frame ID 144D_05501_131413 (144D), where 137A spans from 2015-07-18 to 2021-08-15 and includes 234 scenes, 144D spans from 2015-11-03 to 2021-11-07 and includes 242 scenes. The utilized GNSS data were published by MAGNET (Blewitt et al. 2016), and they

include single-day deformation time series and step event log files (http://geodesy.unr.edu/) for 171 GNSS sites; 9 of these data are used for leave-one-out validation and are not included in the modeling process. The spatial distributions of the InSAR and GNSS data are shown in Fig. 3.

Data processing experiment for 137A

The GNSS displacement data are first processed according to the method introduced in Sect. "GNSS data processing". Based on the difference between the treatment results and the output of the MIDAS algorithm (Blewitt et al. 2016) and referring to other scholars' studies (Santamaría-Gómez et al. 2014), the weighting threshold is set to 0.8. The results obtained before and after conducting processing at the CAWV site are shown in Fig. 4, and the other four selected GNSS sites are shown in Additional file 1: Fig. S2. To verify the reliability of the GNSS data processing results, the deformation rate estimation results are compared with those calculated by
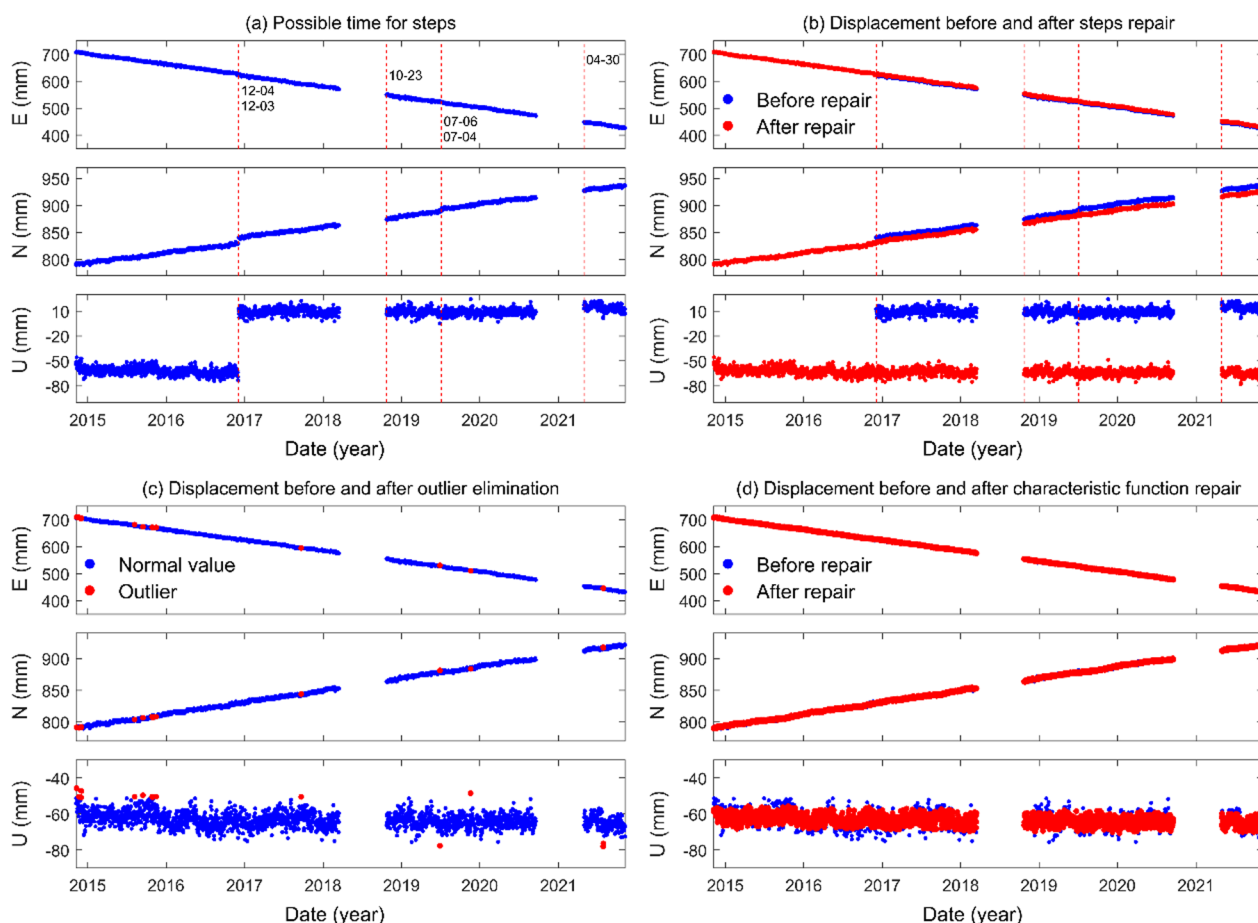

Fig. 4 Results of processing deformation time series data at CAWV GNSS site. Subfigure **a** represents the identified time points at which the step may occur, subfigure **b** represents the deformation time series before and after the step term is repaired, subfigure **c** represents the deformation time series before and after the outliers rejection, and subfigure **d** represents the deformation time series before and after the anomalies are repaired
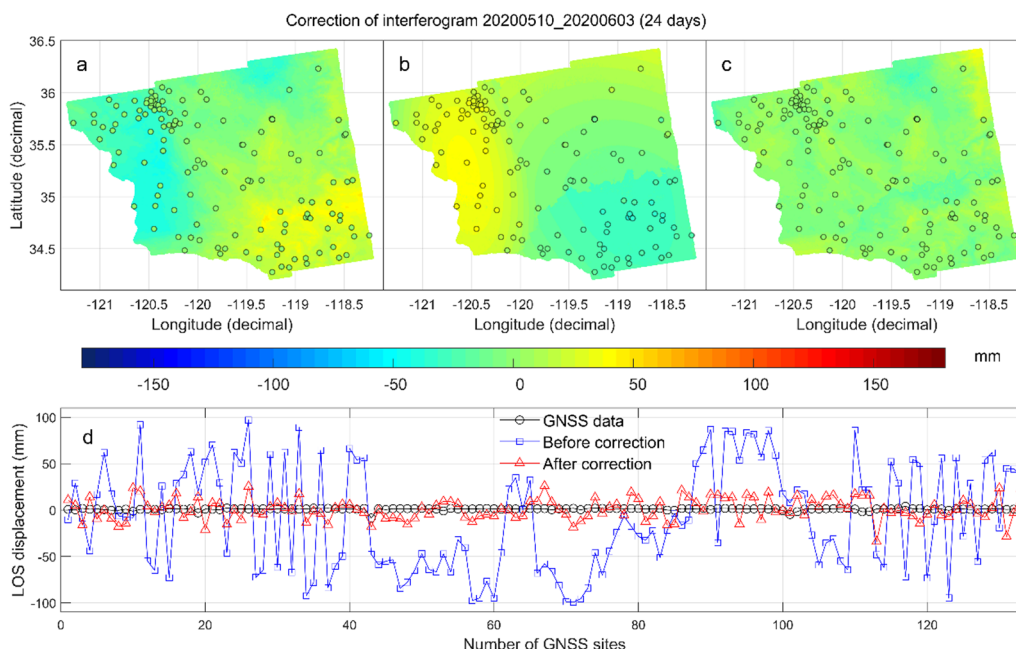
Fig. 5 Error correction results of the 2020-05-10 to 2020-06-03 interferogram for ascending track 137A. Subfigures **a**–**c** represent the before-correction, correction amount, and after-correction InSAR interferogram, respectively. Warm colors (i.e., positive values) indicate surface displacement toward the satellite, and cold colors (i.e., negative values) indicate motion away from the satellite. In subfigures **a**–**c**, circles represent locations of GNSS sites, and filled colors represent the deformation values. Subfigure **d** represents the LOS deformation values of the GNSS sites (black open circles connecting line), and the deformation values of the 15×15 InSAR pixels before correction (blue squares connecting line) and after correction (red triangles connecting line), the numbers, names and latitude/longitude coordinates of the GNSS sites are shown in Table S2 in the Additional file 1
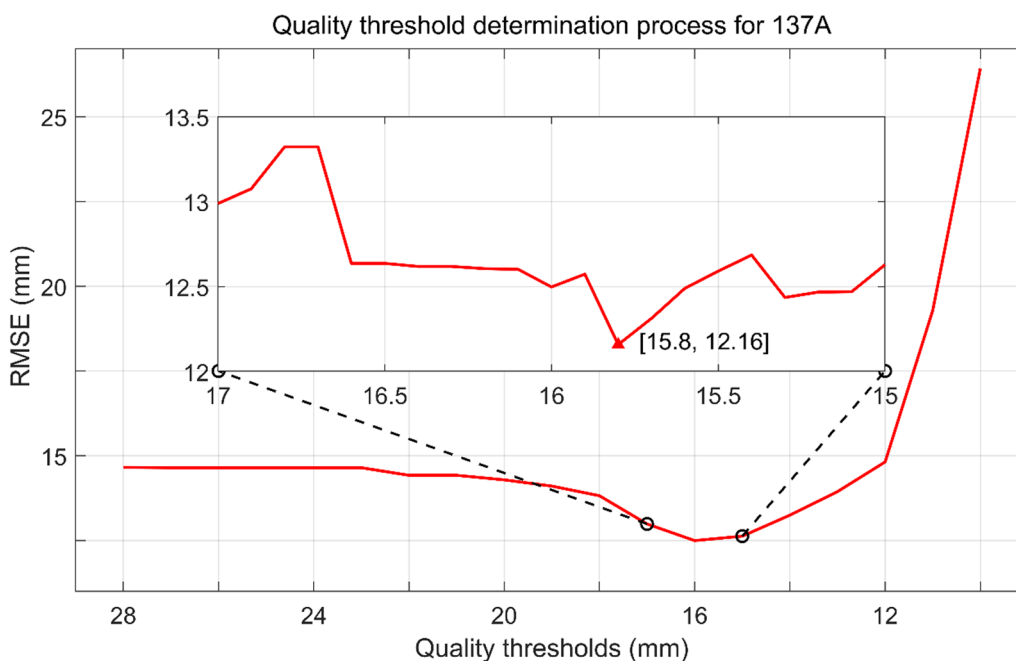

Fig. 6 Process of determining the quality threshold for ascending track 137A. The horizontal axis represents the quality threshold, the vertical axis represents the RMSE between the InSAR time series results and GNSS corresponding to different quality thresholds

the MIDAS algorithm (Blewitt et al. 2016) and plotted in Additional file 1: Fig. S3.

Then, all interferograms with time spans that are less than 100 days and vertical baselines with lengths that are less than 150 m are processed. After browsing through all the interferograms, we find that the maximum number of significant atmospheric error clusters in the study area is 4, so we vary the number of K-means clustering blocks from 1 to 4. According to the method introduced in Sect. "InSAR interferogram data processing", the exact number of clustering blocks is determined by minimizing the difference between the correction results and the GNSS data. If the optimal number is determined to be 1, the method is equivalent to the traditional integral surface fitting method, which means that the interferogram correction effect is better than that of the traditional method. The interferogram for the period from 2020-05-10 to 2020-06-03 is selected for visualization purposes, and the results are plotted in Fig. 5. The corresponding correction generation process is shown in Additional file 1: Fig. S4. In addition, the interferograms for 1 cluster, 2 clusters, 3 clusters and 4 clusters are also selected, and the associated results are shown in Additional file 1: Fig. S5.

Next, the quality threshold is determined according to the method introduced in Sect. "Determination of the optimal quality threshold", and the quality threshold determination process is plotted in Fig. 6. The

first-order deformation trend term is plotted in Additional file 1: Fig. S6, and the numerical distribution of the STD values and the average of the absolute values of the signals remaining after removing the first-order deformation trend term are plotted in Additional file 1: Fig. S7.

After determining the quality threshold, all interferograms satisfying the threshold condition are processed with CSBAS to obtain their deformation time series and deformation rate results. A comparison of the InSAR deformation time series with the GNSS deformation time series at the nine leave-one-out validated GNSS sites is shown in Fig. 7. To further compare the accuracies of the processing results, the distributions of RMSE between all the GNSS sites and the InSAR deformation time series are plotted in Additional file 1: Fig. S8. The processed deformation rate field and the rate variations for the four equally spaced selected profiles are plotted in Fig. 8.

Data processing experiment for 144D

The same processing procedure is used for the 144D data. The interferogram from 2020-05-10 to 2020-06-03 is selected for processing, and the interferogram results obtained before and after processing are shown in Fig. 9. The correction generation process is shown in Additional file 1: Fig. S9. In addition, one interferogram each for 1 cluster, 2 clusters, 3 clusters and 4 clusters are selected,
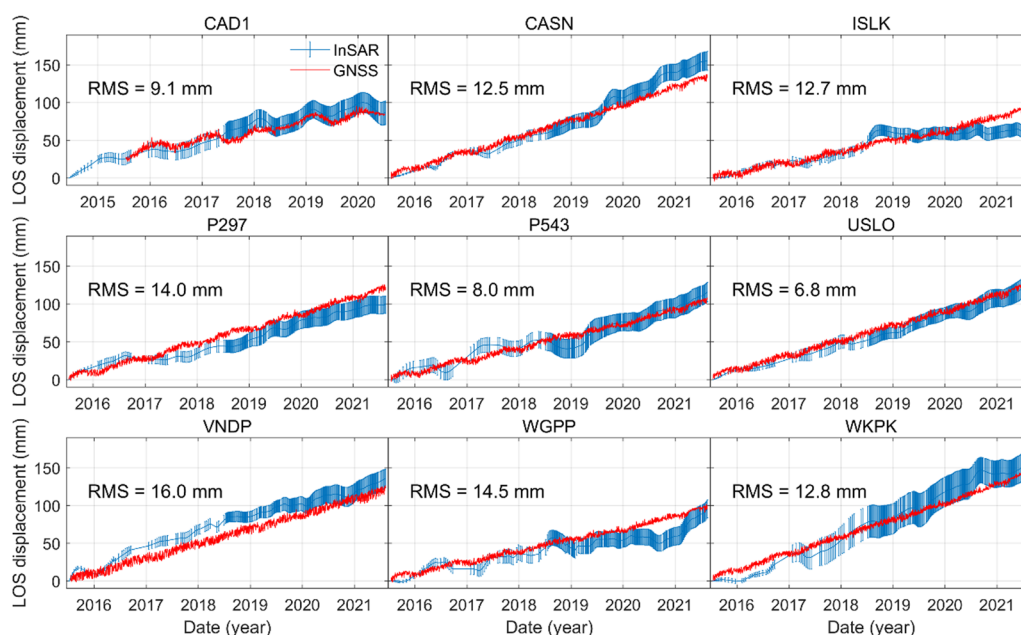

Fig. 7 Comparison of the LOS deformation time series of InSAR pixels in 137A (blue vertical lines connecting line) and GNSS site (red line) on nine leave-one-out validation sites. The midpoint and length of the vertical axis of the InSAR data represent the mean and STD value of the deformation values, respectively, corresponding to the InSAR pixels within a 15 by 15 box around the GNSS site
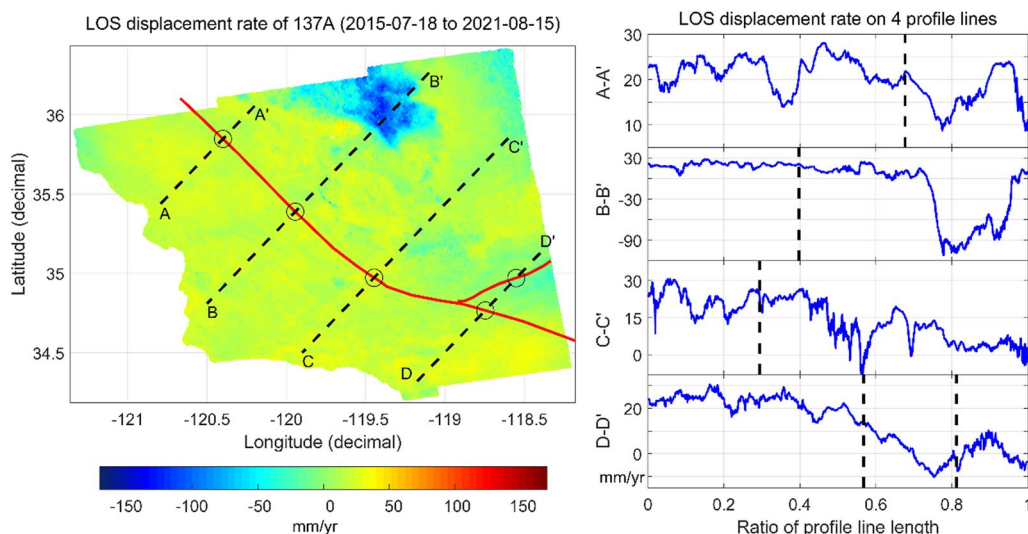
Fig. 8 LOS deformation rate of 137A (left) and rate variation on four cross-fault profile lines (right). For the left subfigure, the black dashed lines indicate the location of the fault profile line, the solid red lines indicate the locations of the major faults, and the circles represent the intersection of the fault and the profile line (corresponding to the black dashed line in the right subfigure). For the right subfigure, the blue lines represent the deformation rate value on the profile lines, the number marked on the abscissa axis represents the length ratio on the profile line
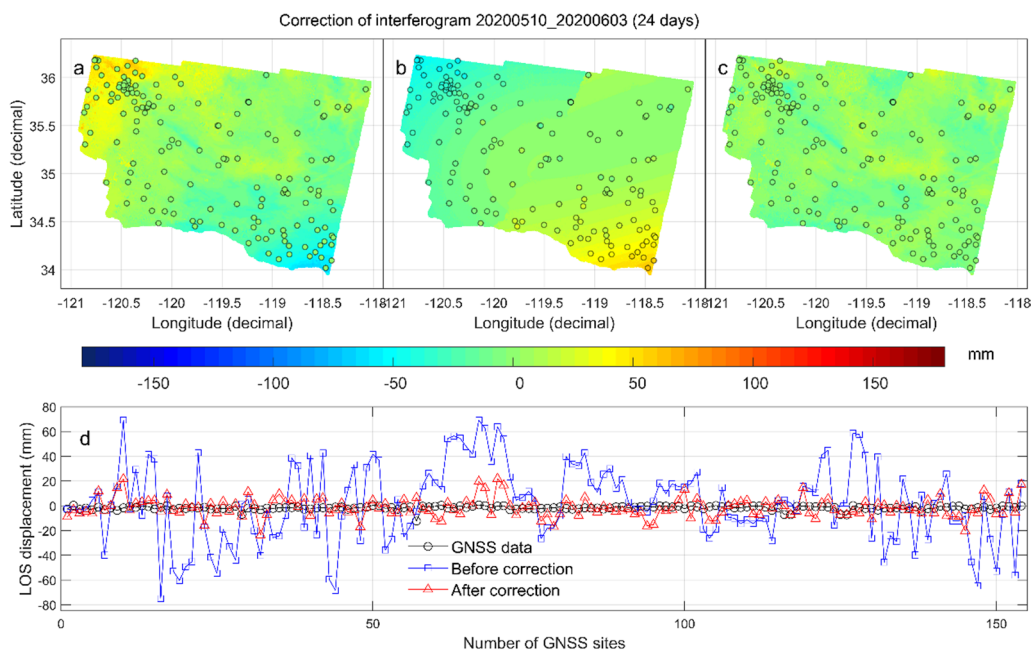

Fig. 9 Error correction results of the 2020-05-10 to 2020-06-03 interferogram for descending track 144D. Subfigures **a**–**c** represent the before-correction, correction amount, and after-correction InSAR interferogram, respectively. Warm colors (i.e., positive values) indicate surface displacement toward the satellite, and cold colors (i.e., negative values) indicate motion away from the satellite. In subfigures **a**–**c**, circles represent locations of GNSS sites, and filled colors represent the deformation values. Subfigure **d** represents the LOS deformation values of the GNSS sites (black open circles connecting line), and the deformation values of the adjacent InSAR pixel points before correction (blue squares connecting line) and after correction (red triangles connecting line), the numbers, names and latitude/longitude coordinates of the GNSS sites are shown in Tab. S3 in the Additional file 1
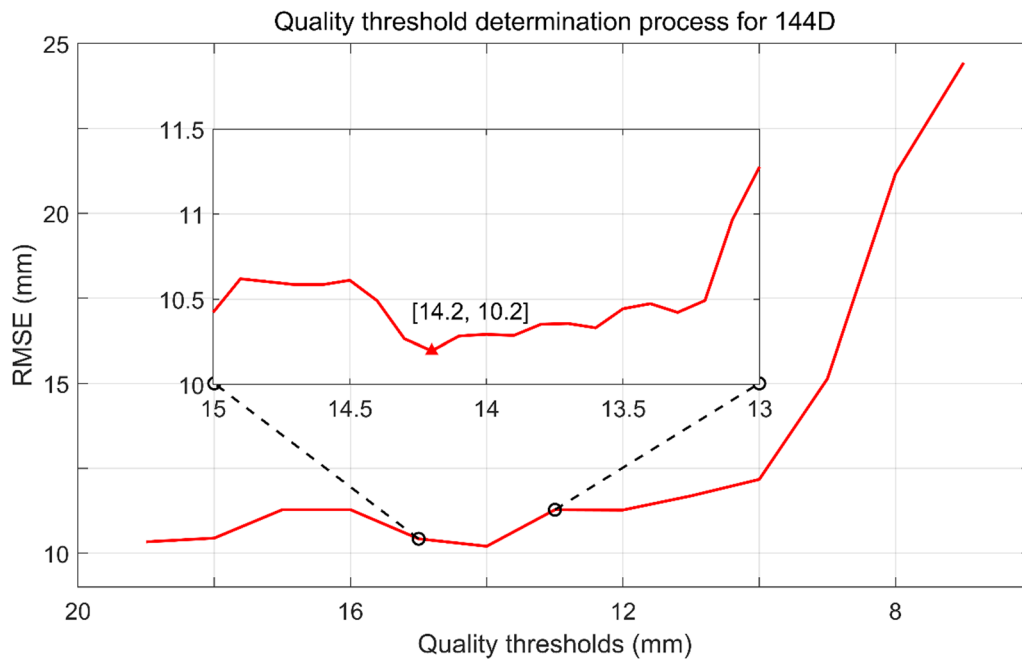
Fig. 10 Process of determining the quality threshold for descending track 144D. The horizontal axis represents the quality threshold, the vertical axis represents the RMSE between the InSAR time series results and GNSS corresponding to different quality thresholds
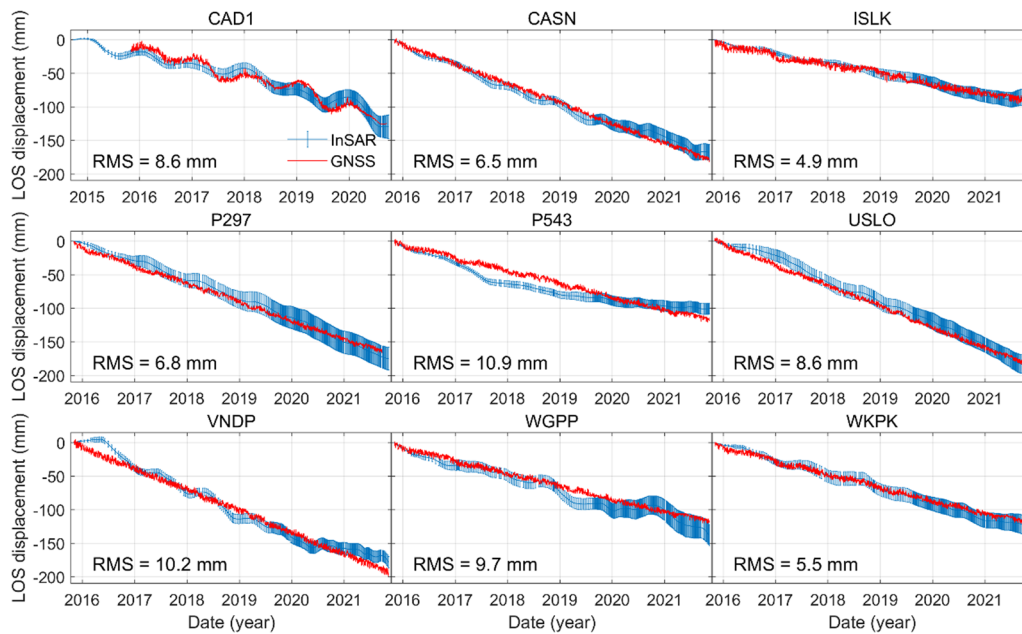

Fig. 11 Comparison of the LOS deformation time series of InSAR pixels in 144D (blue vertical lines connecting line) and GNSS site (red line) on nine leave-one-out validation sites. The midpoint and length of the vertical axis of the InSAR data represent the mean and STD value of the deformation values, respectively, corresponding to the InSAR pixels within a 15 by 15 box around the GNSS site

and the results produced before and after processing are shown in Additional file 1: Fig. S10.

The process of determining the quality threshold for 144D is plotted in Fig. 10. The first-order deformation trend term obtained by stacking the corrected
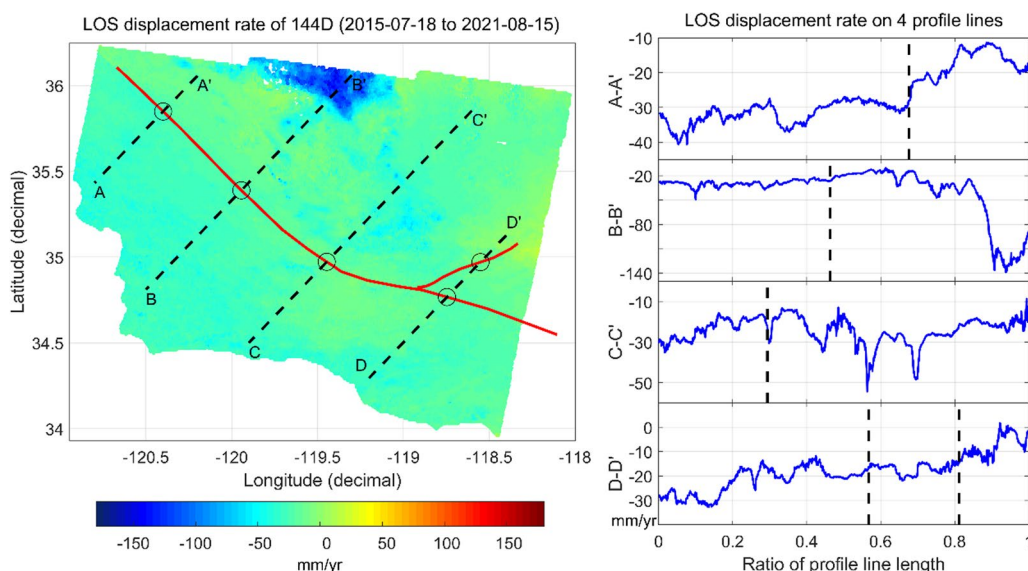
Fig. 12 LOS deformation rate of 144D (left) and rate variation on four cross-fault profile lines (right). For the left subfigure, the black dashed lines indicate the location of the fault profile line, the solid red lines indicate the locations of the major faults, and the circles represent the intersection of the fault and the profile line (corresponding to the black dashed line in the right subfigure). For the right subfigure, the blue lines represent the deformation rate value on the profile lines, the number marked on the abscissa axis represents the length ratio on the profile line

interferograms is plotted in Additional file 1: Fig. S11, and the numerical distributions of the STD values and the average absolute values of the signals remaining after removing the first-order deformation trend term are plotted in Additional file 1: Fig. S12.

The InSAR deformation time series results and the GNSS data for the nine leave-one-out validation GNSS sites are plotted in Fig. 11. The RMSE between all the GNSS sites and the InSAR deformation time series results are plotted in Additional file 1: Fig. S13. The processed deformation rate field and rate variations for the four equally spaced profiles are plotted in Fig. 12.

Discussion
As shown in Fig. 4, Additional file 1: Figs. S1 and S2, after treating the step terms, outlier terms and modeling errors, the deformation time series of the GNSS sites are repaired according to the influences of external events, individual outliers and observation noise. As shown in Additional file 1: Fig. S3, the estimated deformation rates are highly consistent with those calculated by the MIDAS algorithm; most differences are less than 1 mm/year, and some minor numerical differences may be caused by the different time periods of the deformation data.

As shown in Figs. 5, 9, Additional file 1: Figs. S5, S10, the theoretical deformation values are small due to the short time intervals corresponding to the interferograms (generally less than 10 mm); this is also illustrated by the results derived from the GNSS data. However,

due to factors such as atmospheric errors, the interferograms actually exhibit larger deformation values in some local areas (even larger than 50 mm), and our proposed method can effectively account for the spatial distribution of the errors (Additional file 1: Fig. S4 and Fig. S9). For correction results obtained with different numbers of clusters (Additional file 1: Figs. S5 and S10), the result produced with one cluster is equivalent to that of traditional surface fitting correction. Since the K-means correction method selects the optimal number of clusters based on the difference between the corrected interferogram and the GNSS data, it is able to obtain interferograms with the best results compared to those yield the traditional correction method, and it also exhibits improved feasibility.

The corrected interferograms are obtained by using a combination of the corrections from different clustering results (Additional file 1: Figs. S4 and S9). The long- and medium-wavelength errors are significantly weakened, which is reflected by the more continuous spatial distribution of the deformation values; moreover, the difference between the deformation values of the GNSS sites and interferogram data is also significantly reduced. From Additional file 1: Figs. S4 and S9, it can be seen that the Gaussian filtering has some suppression of the larger value variation region and discontinuity errors (especially in the red ellipse in Additional file 1: Figs. S4g and S9g), with values around 15 mm. As shown in Figs. 5b and 9b, the corrections obtained by the K-means clustering

method have some discontinuities at the boundaries, but since the atmospheric errors exhibit similar spatial distributions, these two values can offset each other to some extent. The complementarity of the errors and corrections is reflected in the overall consistency of the deformation being significantly enhanced, and the apparent numerical anomalies in localized areas (in the southwest, northwest, and southeast corners) are effectively suppressed.

After correcting for long- and medium-wavelength errors in the interferograms, we remove the first-order deformation trend term from the interferograms. The deformation trend terms for 137A and 144D (Additional file 1: Figs. S6 and S11) indicate that deformations in these areas occur mainly in the central region due to agricultural irrigation activities (Carlson et al. 2020), whereas the deformations in other regions are less pronounced. A comparison between the distributions of the STD values and the average of the absolute values of the remaining signals (Additional file 1: Figs. S7 and S12) reveals that the overall distribution characteristics of the two indicators are similar, and considering the quantitative characterization of the deformation signal intensity, we adopt the average of the absolute values as the quality threshold. A comparison of the average absolute values obtained for 137A and 144D (Additional file 1: Figs. S7 and S12) reveals that the spatial distributions of the quality thresholds for 137A and 144D are similar, which are mainly concentrated in the middle (137A: approximately 10 mm; 144D: approximately 7 mm), while the numbers of interferograms corresponding to larger and smaller quality thresholds on the two sides are smaller.

As shown in Figs. 6 and 10, the gradient of the quality thresholds defined in the first step is relatively sparse (with a value of 1), with initial values ranging from 10 to 28 mm for 137A and from 7 to 19 mm for 144D. The quality threshold determination processes for 137A and 144D exhibit the same variation pattern: at larger quality thresholds, the RMSE between the InSAR and GNSS data corresponding to different quality thresholds are almost unchanged, because no interferograms are rejected or because the number of rejected interferograms is small; as the quality threshold continues to decrease, some low-quality interferograms are rejected, and the RMSE between the InSAR and GNSS data decreases significantly. This situation continues until the RMSE reaches the minimum point; when the quality threshold is below the optimal value, the RMSE starts to increase, because too many interferograms and even some high-quality interferograms are rejected, resulting in a gradual increase in the difference between the InSAR and GNSS data. As shown in the vignettes of Figs. 6 and 10, after determining the quality threshold corresponding to the

minimum RMSE, we use one gradient above and below this threshold as the upper and lower boundaries, respectively, and set the threshold gradient to 0.1 to obtain a new series of thresholds. Finally, we determine the optimal thresholds for 137A and 144D as 15.8 and 14.2, respectively.

Based on the optimal quality threshold determined in the previous step, we select the interferograms that satisfy these threshold conditions for CSBAS processing. As shown in Figs. 7 and 11, after excluding the low-quality interferograms, the deformation time series trends corresponding to the InSAR pixels are consistent with those of the validated GNSS sites, and the evolution trend of the deformation details with time is also consistent with that of the GNSS data. The RMSE between the InSAR time series and GNSS data for 137A are concentrated within 15 mm (except for the VNDP site), while the RMSE for 144D are mostly concentrated within 10 mm. By calculating the RMSE between the InSAR deformation time series results and all GNSS site data (Additional file 1: Figs. S8 and S13), we find that most of the GNSS sites have RMSE below 15 mm, among which the average RMSE for all GNSS sites are 12.9 mm for 137A and 10.6 mm for 144D.

The deformation rate fields in Figs. 8 and 12 show that the spatial distribution of the deformation rate obtained from 137A is similar to that obtained from 144D, and both the large-scale deformation trends and the local deformation details largely correspond. Overall, the deformation information is mainly concentrated in the central part of the study area, where the impact of agricultural irrigation activities leads to greater deformation (Carlson et al. 2020), while the deformation rates in other regions are lower. No obvious boundary discontinuity appears in the deformation rate field because the clustering results of the interferograms are correlated with the atmospheric errors, which randomly vary in terms of their spatial distribution; thus, the boundary discontinuities in the corrected interferograms also present random distributions. When a large number of corrected interferograms (137A: over 800; 144D: over 700) are selected for SBAS processing, these discontinuities are effectively suppressed and are ultimately not significantly manifested in either the displacement time series (Figs. 7 and 11) or the deformation rate field (Figs. 8 and 12).

A comparison of the deformation rate field results (shown in the left panels of Figs. 8 and 12) with those of Xu et al. (2021) reveals that their spatial distributions accurately match, including spatially heterogeneous deformation signals with larger amplitudes in the northern region, as well as the details concerning the differences between the deformation of the two sides of the fault. This further validates the effectiveness of the

proposed method in generating high-precision InSAR deformation rate results. The numerical magnitude difference is mainly due to the different time ranges and GNSS reference frames used in this paper and by Xu et al. (2021).

Regarding the corresponding LOS deformation rate variations across the fault profiles (shown in the right panels of Figs. 8 and 12), the deformation rates of 137A and 144D have significant steps on both sides of the faults, especially at the 1st and 3rd intersection points of the San Andreas fault and the intersection point of the Garlock fault (approximately 5 mm), which is in agreement with the findings of previous studies conducted by other scholars (Dolan et al. 2016; Scharer and Yule 2020; Sui et al. 2023). In addition, since the experiments in this paper use publicly released interferogram data from scientific research institutions, the cumbersome process of InSAR data processing is avoided, the computational efficiency of acquiring high-precision InSAR deformation time series is improved, and the application difficulty for the GNSS correction of interferogram errors is decreased.

Summary

To solve the current problems including the large amount of InSAR data processing required and the low correction accuracy achieved when using the GNSS to correct InSAR errors, this paper proposes an interferogram data processing method based on the K-means algorithm and GNSS data. The GNSS data published by MAGNET and the interferogram data published by LiCSAR for the CVSC are processed by this method, and the experimental results are discussed and verified with the research results of other scholars. The experimental results show that our proposed method has excellent data processing capabilities and can effectively correct long- and medium-wavelength errors in interferograms, although the discontinuity error will be additionally introduced by the K-means clustering algorithm, but by avoiding the low-quality interferograms and carrying out the times series analysis processing, we can obtain high-precision InSAR deformation time series results; moreover, the obtained InSAR time series are highly consistent with the GNSS displacement data, so they can ultimately provide reliable and high-precision deformation information for geological hazard deformation monitoring and other projects.

The main advantages of our proposed method are as follows. (1) The K-means algorithm is utilized to consider the spatial distribution of interferogram errors, effectively correcting the influences of long- and medium-wavelength errors. (2) Quality evaluation criteria for interferograms and a corresponding effective filtering method are proposed, instead of performing time series processing on all corrected interferograms (Neely et al. 2019; Xu et al. 2021); this strategy suppresses the influence of low-quality interferograms on InSAR deformation time series, and contributes to obtaining high-precision and refined deformation monitoring results under situations with high amounts of interferogram data. (3) Utilizing the SBAS processing results obtained for InSAR pixels near GNSS sites to measure the accuracy corresponding to different quality thresholds, the computational workload required during the process of determining the optimal quality thresholds is greatly reduced. (4) Making full use of publicly released interferogram data and avoiding the cumbersome InSAR data processing procedure, not only improves the efficiency of data processing, but also lowers the application difficulty of correcting interferogram errors by using GNSS data. (5) The proposed method has a high degree of automation, and it can effectively utilize the existing publicly released GNSS and interferogram data to provide reliable data support for large-scale, small-gradient, high-precision deformation monitoring and geological hazard census scenarios.

Our current research mainly focuses on the application in the InSAR deformation time series. In the future, we plan to establish specific identification and correction methods for the discontinuity errors introduced by the K-means clustering algorithm, so as to obtain better correction results for individual InSAR interferograms, and further improving the applicability and effectiveness of the correction method.

Abbreviations

InSAR	Interferometric synthetic aperture radar
GNSS	Global Navigation Satellite System
CVSC	The Central Valley of Southern California
LOS	Line of sight
SBAS	Small baseline subset
STD	Standard deviation
Q_i	Quality indicator
RMSE	Root mean square error
CSBAS	Temporally connected SBAS
137A	Ascending track frame ID 137A_05534_131822
144D	Descending track frame ID 144D_05501_131413

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40623-024-01999-5.

Additional file1: Figure S1. Example of *t*-distribution test for processing GNSS data. The first subfigure represents the fitting of GNSS data using the initial deformation eigenfunction, the second subfigure represents the fitting of GNSS data using the *t*-distribution test for the updated deformation eigenfunction, and the third subfigure represents the difference between the two. **Figure S2.** Processing results of deformation time series data of four GNSS sites. The blue dots indicate the deformation time series before error repair, and the red dots indicate the deformation time series after error repair. **Figure S3.** Comparison of the deformation rate values calculated in this paper (blue line) for the GNSS sites with the results calculated by the MIDAS algorithm (red line). **Figure S4.** Interferogram

Yan *et al. Earth, Planets and Space*   (2024) 76:51

Page 17 of 19

of 2020-05-10 to 2020-06-03 correction amount generation process (3 clusters). The three subfigures (**a**–**c**) in the first row represent the plane fitting results corresponding to the three clustering results of the GNSS data, where the red triangles represent the locations of the cluster centers, the black stars represent the locations of the GNSS stations, and the colors correspond to the values of the deformation corrections of the plane fitting. The second row of subfigures from left to right represent (**d**–**g**): the spatial clustering results of the InSAR interferograms (red triangles represent the locations of the cluster centers, and different colors represent different clusters), the deformation correction values generated according to the correspondence between the GNSS and InSAR cluster centers, the deformation correction values after Gaussian filtering, and the differences due to Gaussian filtering (the red ellipses mark the regions with significant differences). **Figure S5** Error correction results for four interferograms with different number of clusters (one to four clusters) for 137A. The subfigures in each row from left to right represent the original interferogram, the amount of interferogram correction, and the corrected interferogram, respectively. Warm colors (i.e., positive values) indicate surface displacement toward the satellite, and cold colors (i.e., negative values) indicate motion away from the satellite. Circles represent locations of GNSS sites, and filled colors represent the deformation values. **Figure S6.** First order deformation trend term of 137A. Warm colors (i.e., positive values) indicate surface displacement toward the satellite, and cold colors (i.e., negative values) indicate motion away from the satellite. **Figure S7.** Numerical distribution of STD values and average absolute values, corresponding to the remaining signals after removing the first order deformation trend term from the corrected interferogram of 137A. **Figure S8.** Histogram of deformation RMSE between GNSS sites and 137A InSAR time series. **Figure S9.** Interferogram of 2019-03-23 to 2019-04-28 correction amount generation process (2 clusters). The three subfigures (**a** and **b**) in the first row represent the plane fitting results corresponding to the three clustering results of the GNSS data, where the red triangles represent the locations of the cluster centers, the black stars represent the locations of the GNSS stations, and the colors correspond to the values of the deformation corrections of the plane fitting. The second row of subfigures from left to right represent (**c**–**f**): the spatial clustering results of the InSAR interferograms (red triangles represent the locations of the cluster centers, and different colors represent different clusters), the deformation correction values generated according to the correspondence between the GNSS and InSAR cluster centers, the deformation correction values after Gaussian filtering, and the differences due to Gaussian filtering (the red ellipses mark the regions with significant differences). **Figure S10.** Error correction results for four interferograms with different number of clusters (one to four clusters) for 144D. The subfigures in each row from left to right represent the original interferogram, the amount of interferogram correction, and the corrected interferogram, respectively. Warm colors (i.e., positive values) indicate surface displacement toward the satellite, and cold colors (i.e., negative values) indicate motion away from the satellite. Circles represent locations of GNSS sites, and filled colors represent the deformation values. **Figure S11.** First order deformation trend term of 144D. Warm colors (i.e., positive values) indicate surface displacement toward the satellite, and cold colors (i.e., negative values) indicate motion away from the satellite. **Figure S12.** Numerical distribution of STD values and average absolute values, corresponding to the remaining signals after removing the first order deformation trend term from the corrected interferogram of 144D. **Figure S13.** Histogram of deformation RMSE between GNSS sites and 144D InSAR time series. **Table S1.** Statistics of *t*-distribution test for each parameter. **Table S2.** Number, name and latitude/longitude coordinates of the GNSS modeling sites used in 137A. **Table S3.** Number, name and latitude/longitude coordinates of the GNSS modeling sites used in 144D.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]School of Geosciences and Info-Physics, Central South University, Changsha 410083, Hunan, China. [2]School of Resources and Civil Engineering, Gannan University of Science and Technology, Ganzhou 341000, Jiangxi, China. [3]Ganzhou Key Laboratory of Remote Sensing for Resource and Environment, Ganzhou 341000, Jiangxi, China. [4]School of Marine Technology and Geomatics, Jiangsu Ocean University, Lianyungang 222005, Jiangsu, China.

References

Agnieszka W, Dawid K (2022) Modeling seasonal oscillations in GNSS time series with complementary ensemble empirical mode decomposition. GPS Solut 26(4):1–13. https://doi.org/10.1007/s10291-022-01288-2

Amiri-Simkooei A, Hosseini-Asl M, Asgari J, Zangeneh-Nejad F (2019) Offset detection in GPS position time series using multivariate analysis. GPS Solut 23(1):1–12. https://doi.org/10.1007/s10291-018-0805-z

Arthur D, Vassilvitskii S (2007) K-means++ the advantages of careful seeding. In: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. pp 1027–1035

Berardino P, Fornaro G, Lanari R, Sansosti E (2002) A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. IEEE Trans Geosci Remote Sens 40(11):2375–2383. https://doi.org/10.1109/TGRS.2002.803792

Blewitt G, Kreemer C, Hammond WC, Gazeaux J (2016) MIDAS robust trend estimator for accurate GPS station velocities without step detection. J Geophys Res Solid Earth 121(3):2054–2068. https://doi.org/10.1002/2015JB012552

Bogusz J, Klos A (2016) On the significance of periodic signals in noise analysis of GPS station coordinates time series. GPS Solut 20(4):655–664. https://doi.org/10.1007/s10291-015-0478-9

Yan *et al. Earth, Planets and Space*   (2024) 76:51

Page 18 of 19

Carlson G, Shirzaei M, Werth S, Zhai G, Ojha C (2020) Seasonal and long-term groundwater unloading in the Central Valley modifies crustal stress. J Geophys Res Solid Earth 125(1):e2019JB018490. https://doi.org/10.1029/2019JB018490

Dai W, Yan H, Chen B, Liu B (2019) Regional spatio-temporal zenith tropospheric delay modelling using independent component analysis. Surv Rev 51(369):544–552. https://doi.org/10.1080/00396265.2018.1515812

Doke R, Mannen K, Itadera K (2021) Observing posteruptive deflation of hydrothermal system using InSAR time series analysis: an application of ALOS-2/PALSAR-2 data on the 2015 phreatic eruption of Hakone volcano, Japan. Geophys Res Lett 48(19):e2021GL094880. https://doi.org/10.1029/2021GL094880

Dolan JF, McAuliffe LJ, Rhodes EJ, McGill SF, Zinke R (2016) Extreme multi-millennial slip rate variations on the Garlock fault, California: strain super-cycles, potentially time-variable fault strength, and implications for system-level earthquake occurrence. Earth Planet Sci Lett 446:123–136. https://doi.org/10.1016/j.epsl.2016.04.011

Dong D, Fang P, Bock Y, Cheng M, Miyazaki S (2002) Anatomy of apparent seasonal variations from GPS-derived site position time series. J Geophys Res Solid Earth 107(B4):ETG9. https://doi.org/10.1029/2001JB000573

Festa D, Bonano M, Casagli N, Confuorto P, De Luca C, Del Soldato M, Lanari R, Lu P, Manunta M, Manzo M et al (2022) Nation-wide mapping and classification of ground deformation phenomena through the spatial clustering of P-SBAS InSAR measurements: Italy case study. ISPRS J Photogramm Remote Sens 189:1–22. https://doi.org/10.1016/j.isprsjprs.2022.04.022

González-Jiménez M, Guardiola-Albert C, Ezquerro P, Aguilera H, Béjar-Pizarro M, Naranjo-Fernández N, Bru G, Herrera G (2023) Analysis of aquifer-system deformation in the Doñana Natural Space (Spain) using unsupervised cloud-computed InSAR data and wavelet analysis. Water Resour Res. https://doi.org/10.1029/2022WR033858

Grigg O, Spiegelhalter D (2007) A simple risk-adjusted exponentially weighted moving average. J Am Stat Assoc 102(477):140–152. https://doi.org/10.1198/016214506000001121

Guns K, Xu X, Bock Y, Sandwell D (2022) GNSS-corrected InSAR displacement time-series spanning the 2019 Ridgecrest, CA Earthquakes. Geophys J Int 230(2):1358–1373. https://doi.org/10.1093/gji/ggac121

Jafari M, Aflaki M, Mousavi Z, Walpersdorf A, Motaghi K (2023) Coseismic and post-seismic characteristics of the 2021 Ganaveh earthquake along the Zagros foredeep fault based on InSAR data. Geophys J Int 234(2):1125–1142. https://doi.org/10.1093/gji/ggad127

Khan SD, Faiz MI, Gadea OC, Ahmad L (2023) Study of land subsidence by radar interferometry and hot spot analysis techniques in the Peshawar Basin, Pakistan. Egypt J Remote Sens Space Sci 26(1):173–184. https://doi.org/10.1016/j.ejrs.2023.02.001

Kobayashi T, Morishita Y, Yarai H (2018) SAR-revealed slip partitioning on a bending fault plane for the 2014 Northern Nagano earthquake at the northern Itoigawa-Shizuoka tectonic line. Tectonophysics 733:85–99. https://doi.org/10.1016/j.tecto.2017.12.001

Lahtinen S, Jivall L, Häkli P, Nordman M (2022) Updated GNSS velocity solution in the Nordic and Baltic countries with a semi-automatic offset detection method. GPS Solut 26(1):1–12. https://doi.org/10.1007/s10291-021-01194-z

Lanari R, Mora O, Manunta M, Mallorquí JJ, Berardino P, Sansosti E (2004) A small-baseline approach for investigating deformations on full-resolution differential SAR interferograms. IEEE Trans Geosci Remote Sens 42(7):1377–1386. https://doi.org/10.1109/TGRS.2004.828196

Lazeckỳ M, Spaans K, González PJ, Maghsoudi Y, Morishita Y, Albino F, Elliott J, Greenall N, Hatton E, Hooper A et al (2020) LiCSAR: an automatic InSAR tool for measuring and monitoring tectonic and volcanic activity. Remote Sens 12(15):2430. https://doi.org/10.3390/rs12152430

Lee J-C, Shirzaei M (2023) Novel algorithms for pair and pixel selection and atmospheric error correction in multitemporal InSAR. Remote Sens Environ 286:113447. https://doi.org/10.1016/j.rse.2022.113447

Levenberg K (1944) A method for the solution of certain non-linear problems in least squares. Q Appl Math 2(2):164–168

Li Y, Bürgmann R (2021) Partial coupling and earthquake potential along the Xianshuihe Fault, China. J Geophys Res Solid Earth 126(7):e2020JB021406. https://doi.org/10.1029/2020JB021406

Li Y, Xu C, Yi L, Fang R (2018) A data-driven approach for denoising GNSS position time series. J Geodesy 92(8):905–922. https://doi.org/10.1007/s00190-017-1102-2

Li Z, Cao Y, Wei J, Duan M, Wu L, Hou J, Zhu J (2019) Time-series InSAR ground deformation monitoring: atmospheric delay modeling and estimating. Earth Sci Rev 192:258–284. https://doi.org/10.1016/j.earscirev.2019.03.008

Liu F, Elliott J, Craig T, Hooper A, Wright T (2021) Improving the resolving power of InSAR for earthquakes using time series: a case study in Iran. Geophys Res Lett 48(14):e2021GL093043. https://doi.org/10.1029/2021GL093043

Liu H, Xie L, Zhao G, Ali E, Xu W (2023) A joint InSAR-GNSS workflow for correction and selection of interferograms to estimate high-resolution interseismic deformations. Satell Navig 4(1):14. https://doi.org/10.1186/s43020-023-00105-6

López-Quiroz P, Doin M-P, Tupin F, Briole P, Nicolas J-M (2009) Time series analysis of Mexico City subsidence constrained by radar interferometry. J Appl Geophys 69(1):1–15. https://doi.org/10.1016/j.jappgeo.2009.02.006

Luo X, Wang C, Long Y, Yi Z (2020) Analysis of the decadal kinematic characteristics of the Daguangbao landslide using multiplatform time series InSAR observations after the Wenchuan earthquake. J Geophys Res Solid Earth 125(12):e2019JB019325. https://doi.org/10.1029/2019JB019325

MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Oakland, CA, USA, pp 281–297

Morishita Y, Kobayashi T (2022) Three-dimensional deformation and its uncertainty derived by integrating multiple SAR data analysis methods. Earth Planet Space 74(1):16. https://doi.org/10.1186/s40623-022-01571-z

Morishita Y, Lazecky M, Wright TJ, Weiss JR, Elliott JR, Hooper A (2020) LiCSBAS: an open-source InSAR time series analysis package integrated with the LiCSAR automated Sentinel-1 InSAR processor. Remote Sens 12(3):424–452. https://doi.org/10.3390/rs12030424

Neely WR, Borsa AA, Silverii F (2019) GInSAR: a cGPS correction for enhanced InSAR time series. IEEE Trans Geosci Remote Sens 58(1):136–146. https://doi.org/10.1109/TGRS.2019.2934118

Ohno K, Ohta Y, Kawamoto S, Abe S, Hino R, Koshimura S, Musa A, Kobayashi H (2021) Real-time automatic uncertainty estimation of coseismic single rectangular fault model using GNSS data. Earth Planet Space 73:1–18. https://doi.org/10.1186/s40623-021-01425-0

Osmanoğlu B, Sunar F, Wdowinski S, Cabral-Cano E (2016) Time series analysis of InSAR data: methods and trends. ISPRS J Photogramm Remote Sens 115:90–102. https://doi.org/10.1016/j.isprsjprs.2015.10.003

Santamaría-Gómez A, Gravelle M, Wöppelmann G (2014) Long-term vertical land motion from double-differenced tide gauge and satellite altimetry data. J Geodesy 88:207–222. https://doi.org/10.1007/s00190-013-0677-5

Scharer KM, Yule D (2020) A maximum rupture model for the southern San Andreas and San Jacinto faults, California, derived from paleoseismic earthquake ages: observations and limitations. Geophys Res Lett 47(15):e2020GL088532. https://doi.org/10.1029/2020GL088532

Solari L, Bianchini S, Franceschini R, Barra A, Monserrat O, Thuegaz P, Bertolo D, Crosetto M, Catani F (2020) Satellite interferometric data for landslide intensity evaluation in mountainous regions. Int J Appl Earth Obs Geoinf 87:102028–102043. https://doi.org/10.1016/j.jag.2019.102028

Sui S, Shen W, Holt W, Kim J (2023) Crustal architecture across Southern California and its implications on San Andreas Fault development. Geophys Res Lett 50(8):e2022GL101976. https://doi.org/10.1029/2022GL101976

Suo Z, Zhang J, Li M, Zhang Q, Fang C (2015) Improved InSAR phase noise filter in frequency domain. IEEE Trans Geosci Remote Sens 54(2):1185–1195. https://doi.org/10.1109/TGRS.2015.2476355

Takada Y, Sagiya T, Nishimura T (2018) Interseismic crustal deformation in and around the Atotsugawa fault system, central Japan, detected by InSAR and GNSS. Earth Planet Space 70(1):1–12. https://doi.org/10.1186/s40623-018-0801-0

Wang Y, Dong J, Zhang L, Zhang L, Deng S, Zhang G, Liao M, Gong J (2022) Refined InSAR tropospheric delay correction for wide-area landslide identification and monitoring. Remote Sens Environ 275:113013–113030. https://doi.org/10.1016/j.rse.2022.113013

Weiss JR, Walters RJ, Morishita Y, Wright TJ, Lazecky M, Wang H, Hussain E, Hooper AJ, Elliott JR, Rollins C et al (2020) High-resolution surface velocities and strain for Anatolia from Sentinel-1 InSAR and GNSS data. Geophys Res Lett 47(17):e2020GL087376. https://doi.org/10.1029/2020GL087376

Wen W, Zhang G, Hsu L-T (2021) GNSS outlier mitigation via graduated non-convexity factor graph optimization. IEEE Trans Veh Technol 71(1):297–310. https://doi.org/10.1109/TVT.2021.3130909

Yan *et al. Earth, Planets and Space*   (2024) 76:51

Page 19 of 19

Xi R, He Q, Meng X (2021) Bridge monitoring using multi-GNSS observations with high cutoff elevations: a case study. Measurement 168:108303–108315. https://doi.org/10.1016/j.measurement.2020.108303

Xu W, Wu S, Materna K, Nadeau R, Floyd M, Funning G, Chaussard E, Johnson CW, Murray JR, Ding X et al (2018) Interseismic ground deformation and fault slip rates in the greater San Francisco Bay Area from two decades of space geodetic data. J Geophys Res Solid Earth 123(9):8095–8109. https://doi.org/10.1029/2018JB016004

Xu X, Sandwell DT, Klein E, Bock Y (2021) Integrated Sentinel-1 InSAR and GNSS time-series along the San Andreas fault system. J Geophys Res Solid Earth. https://doi.org/10.1029/2021JB022579

Xu K, Gan W, Wu J, Hou Z (2022) A robust method for 3-D surface displacement fields combining GNSS and single-orbit InSAR measurements with directional constraint from elasticity model. GPS Solut 26(2):1–16. https://doi.org/10.1007/s10291-022-01226-2

Yan H, Dai W, Liu H, Gao H, Neely WR, Xu W (2022a) 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):3293–3310. https://doi.org/10.3390/rs14143293

Yan H, Dai W, Xie L, Xu W (2022b) Fusion of GNSS and InSAR time series using the improved STRE model: applications to the San Francisco Bay Area and Southern California. J Geodesy 96(7):1–20. https://doi.org/10.1007/s00190-022-01636-7

Yang Z, Li Z, Zhu J, Wang Y, Wu L (2020) Use of SAR/InSAR in mining deformation monitoring, parameter inversion, and forward predictions: a review. IEEE Geosci Remote Sens Mag 8(1):71–90. https://doi.org/10.1109/MGRS.2019.2954824

Yu C, Li Z, Penna NT, Crippa P (2018) Generic atmospheric correction model for interferometric synthetic aperture radar observations. J Geophys Res Solid Earth 123(10):9202–9222. https://doi.org/10.1029/2017JB015305

Zhang S, Fan Q, Niu Y, Qiu S, Si J, Feng Y, Zhang S, Song Z, Li Z (2023) Two-dimensional deformation monitoring for spatiotemporal evolution and failure mode of Lashagou landslide group, Northwest China. Landslides 20(2):447–459. https://doi.org/10.1007/s10346-022-01979-4

Zhu J, Liu Z, Fu H, Zhou C, Zhou Y, Wang H, Xie Y (2023) High-resolution sub-canopy topography mapping via TanDEM-X DEM combined with future P-band BIOMASS PolInSAR data. J Geodesy 97(12):114. https://doi.org/10.1007/s00190-023-01807-0

Publisher's Note