

Source and pattern identification of ground deformation based on the framework of Blind Source Separation–Nonnegative Matrix Factorization: a case study in a long-term GPS monitoring mine

hongyu Gu (✉ xy0909040129@126.com)

chengdu center, china geology survey <https://orcid.org/0000-0001-6837-2685>

fengshan Ma

Chinese Academy of Sciences

Liangjun Lin

China Geological Survey

donghui wang

China Geological Survey

yongbo Tie

China Geological Survey

weichang Chen

Chinese Academy of Cultural Heritage

Research Article

Keywords: source and pattern identification, Nonnegative Matrix Factorization, ground deformation, GPS

Posted Date: March 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-232646/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 **Source and pattern identification of ground deformation based on the framework of Blind**
2 **Source Separation--Nonnegative Matrix Factorization: a case study in a long-term GPS**
3 **monitoring mine**

4
5 Hongyu Gu^{*1}, Fengshan Ma², Liangjun Lin^{*4,5}, Donghui Wang¹, Yongbo Tie¹, Weichang Chen³

6
7 1. Chengdu Center, China Geological Survey, Sichuan, Chengdu 610081

8 2. Key Laboratory of Shale Gas and Geoengineering, Institute of Geology and Geophysics Chinese
9 Academy of Sciences, Beijing 100029

10 3. Chinese Academy of Cultural Heritage, Beijing 100029

11 4. Tianjin Center, China Geological Survey, Tianjin 300170

12 5. Xiong'an Urban Geological Research Center of China Geological Survey, Beijing 100037

13
14 Corresponding author: Hongyu Gu & Liangjun Lin

15 E-mail: xy0909040129@126.com; 9781866765@qq.com

16 Address: Yihuanlu Beisanduan No.2, Chengdu, China, 610081

17
18 Fengshan Ma

19 E-mail: 1083471408@qq.com

20 Liangjun Lin

21 E-mail: 9781866765@qq.com

22 Donghui Wang

23 E-mail: 742824902@qq.com

24 Yongbo Tie

25 E-mail: 414429287@qq.com

26 Weichang Chen:

27 E-mail: 289835829@qq.com

28
29 **Abstract**

30 Ground deformation caused by mining affects the safety of underground mining and buildings.
31 High-resolution characterization of ground deformation respond to latent sources is a first step
32 toward improved hazard forecasting. We rely on long-term GPS displacement data and an improved
33 NMF algorithm of BSS to identify the sources driving ground deformation. The NMF identifies
34 three sources (S1, S2, and S3) with distinct spatial and temporal surface deformation patterns and
35 quantitatively reveals the contribution of each source to ground deformation. S1 captures horizontal
36 Y-displacement related to horizontal tectonic stress σ_1 . S2 dominates vertical Z-displacement with
37 a widespread isotropous deformation driven by self-weight body force. S3 controls horizontal X-
38 displacement related to horizontal tectonic stress σ_2 . Besides, we find that independent
39 contributions of these three sources can be resolved from the GPS data. The results show that the
40 sharp change of source contribution and even transformation of dominate source are highly related
41 to the severe deformation belt. This phenomenon is time independent, and helps us to select sites
42 and find the potential risk area at its early stage.

43 **Keywords:** source and pattern identification; Nonnegative Matrix Factorization; ground
44 deformation; GPS

45 **1. Introduction**

46 Ground deformation is an common phenomenon in terms of civil engineering, groundwater
47 pumping, earthquake, landslide and underground mining operations (Cao et al. 2019; Gourmelen et
48 al. 2007; Schultz-Fellenz et al. 2020; Ye et al. 2016; Zhao et al. 2012) and always leads to severe
49 damage of buildings and even collapse of roof caving (Ding et al. 2017; Xia et al. 2019). Typically,
50 there are multiple sources (factors) that can cause rock movement, such as topography, type of rock,
51 structure of rock mass, tectonic stress, mining technology, and co-seismic vibration effects, etc. The
52 combination of these sources and the dominant source vary significantly depending on the local
53 geological conditions. Without greater confidence in how these proposed factors affect ground
54 deformation near surface, we cannot reliably use the results of geodetic analyses to infer shallow
55 deformation, nor can we formulate models to predict deformation in future events.

56 In recent years, these sources have been widely studied by various methods to have a deep
57 understanding of the mechanism of ground deformation. There are four broadly categories.

58 First, numerical model based on finite element or discrete element method. It is a forward
59 method that takes some sources (e.g., stress condition, joint distribution, and mining operations) and
60 their generalized parameters into consideration (Fathi Salmi et al. 2017; Kabwe 2017; Khanal et al.
61 2015; Salmi et al. 2019). The number and kinds of input sources are always decided by prior
62 knowledge and generalized parameters from laboratory. Although numerical model is parameters
63 dependent and sensitive, researchers can, to some extent, identify the dominant source depend on
64 the fluctuation of displacement and stress by inputting different parameters of one source while
65 fixing other parameters at the same time. The second category is classical statistical methods (e.g.,
66 correlation analysis, factor analysis, regression analysis, etc.). These methods try to find the law of
67 ground deformation from observed data(Hui et al. 2018). They also can help to find out parts of the
68 sources leading to ground deformation. However, the hidden sources of massive data with high
69 dimensions cannot be easily interpreted by using these methods (Sahu et al. 2017; Song et al. 2011).
70 The third category is analytical solution based on empirical methods, profile function, and influence
71 function(Díaz-Fernández et al. 2010). These methods take the geometrical and mechanical
72 parameters as input to predict the ground deformation. The type and value of parameters are highly
73 dependent on the prior knowledge of researchers. However, it is unreasonable to predict before we
74 know the mechanism of each potential source causing ground deformation. The forth category is
75 machine learning consists of a series of supervised learning algorithms(Chen et al. 2017; Perrin et
76 al. 2015; Zhou et al. 2019). These methods use a set of data to train the learning model with factors
77 identified by prior knowledge and then to fit the observed data. Therefore, these methods are not
78 self-constraint by data and much dependent on the knowledge of researchers.

79 However, none of these studies try to objectively and quantitatively find out that 1) what kind
80 of and how many sources leading to the ground deformation? 2) what is the independent pattern of
81 each source? 3) how much is the proportion of each source contributing to the deformation at a
82 specific site? and 4) how the dominant source transfers with time among these sources?

83 Mechanism and pattern vary significantly among different sources when causing ground
84 deformation. As a result, the characteristics of ground deformation induced by different sources can
85 be also different in terms of horizontal and vertical displacement. In plain words, *even though we*
86 *did not know who they (sources) are, how many of them there and how they did (causing ground*
87 *deformation), but what they have done was recorded in the data.* Based on this logic, it is possible
88 to find a backward way to separate the sources from the long-term monitored displacement data. If

89 sources can be successfully decomposed from observed data, hidden pattern of each source can be
 90 revealed independently with knowledge of petrology, tectonic geology, fracture mechanics, and rock
 91 mechanics, etc. Most importantly, it is the transformation of dominant sources with time in a
 92 concerned area that is essential to the prediction of ground deformation and safety management.

93 Indeed, sources identification causing temporal and spatial development of ground
 94 deformation is still challenging. In recent years, source identification has received extensive
 95 attention in areas such as signal processing, biomedical engineering, data analysis and data mining.
 96 Nonetheless, the theory and applications are still being developed. Source identification can be
 97 complex because 1) some of source signals may mutually interfere, 2) there may be attenuation
 98 (linearly or nonlinearly) of signals when transmitting through the medium or discontinuity, 3) the
 99 ratios of signal to noise vary in real world, 4) some of sources may be partly collinear or have similar
 100 patterns.

101 The observed displacement data can be seen as the mixture signal of different displacement
 102 patterns with some unknown mixing processes. From this point of view, we can introduce Blind
 103 Source Separation (BSS) in the field of signal processing to find the hidden sources and their
 104 patterns. The aim of BSS is to process observations acquired by sensor arrays in such a way that the
 105 original unknown source signals are extracted by various algorithms without knowing or with
 106 limited information about the characteristics of the transmission channels through which the sources
 107 propagate to the sensors. Therefore, BSS can be used to reveal hidden sources and patterns in a
 108 mixed data set.

109 The basic framework of BSS can be described as a linear mixture model by equation 1.

$$110 \quad \mathbf{X} = \mathbf{AS} + \mathbf{E} \quad (1)$$

111 Where $\mathbf{X} \in R^{m \times n}$ is observed data (mixtures). $\mathbf{A} \in R^{m \times p}$ is an unknown mixing matrix
 112 representing the linear combination of the sources and each row of $\mathbf{S} \in R^{p \times n}$ is a source. \mathbf{E}
 113 represents additive noise or interference introduced during mixing and transmission.

114 Typically, the only known information in BSS is the mixtures \mathbf{X} . We need to determine the
 115 number of sources \mathbf{S} and the solution of mixing matrix \mathbf{A} and \mathbf{S} , i.e., mathematically, we need to
 116 solve equation 2

$$117 \quad \vartheta = \min \|\mathbf{X} - \mathbf{AS}\|_F^2 \quad (2)$$

118 It is noteworthy that such a problem is ill-posed and has an infinite number of solutions. Thus,
 119 extra imposed constraints (e.g., statistical independence, non-negativity, sparsity, maximum mutual
 120 information, etc.) and regularized optimization procedures are often required to minimize the
 121 objective function ϑ and to find an optimal and robust solution. Different constraints are usually
 122 imposed on \mathbf{A} and \mathbf{S} based on particular cases and may lead to different results. Many methods
 123 are included in BSS (e.g., independent component analysis (ICA), sparse component analysis (SCA),
 124 non-negative matrix/tensor factorization (NMF/NTF), etc.). They all can successfully identify the
 125 hidden sources and patterns from observed data set. However, it is the physical meaning and
 126 interpretation of particular case that crucial to determine which method to be used.

127 ICA is one of the early and widely used techniques for BSS. It is an extension of linear principal
 128 component (PCA), where mutually orthogonal assumption is replaced by a stronger statistically
 129 independent assumption of latent sources (Ciaramella et al. 2006). Cohen-Waeber et al. (2018)
 130 applied ICA to InSAR time series in San Francisco East Bay Hills and revealed four distinct spatial
 131 and temporal surface deformation components (sources). However, the performance of ICA is not
 132 good in some areas (e.g., in bio-signal and genomic signal) because the latent sources are partially

133 or entirely correlated. In other words, it is hard and often impossible to verify the statistical
134 assumptions on the sources (Tangirala et al. 2007). Under this situation, much effort has been
135 devoted to find an alternative approach with weaker assumptions. SCA started to emerge at the end
136 of the 1990s and became prominent during the 2000s (Chang et al. 2006; Lee et al. 1999). Unlike
137 ICA, SCA is suited to correlated sources and tries to explain data as a mixture of sparse
138 components (Karoui et al. 2012). In practice, both non-negativity and sparsity are often desirable or
139 necessary when the hidden sources have physical meaning. However, results of ICA and SCA often
140 lead to the subtraction in order to reconstruction of observed data. As a result, it is difficult to
141 interpret the physical meaning of sources.

142 In contrast to ICA, NMF enforces a non-negative constraint exhibiting some degree of natural
143 sparsity on the mixing matrix \mathbf{A} and source matrix \mathbf{S} (Lee and Seung 1999). In other words, NMF
144 does not allow subtraction and combinations, and, therefore, it is often used to quantitatively
145 describe the parts that comprise the entire entity. Furthermore, matrix factorization methods that
146 exploit non-negativity and sparsity constraints usually lead to estimation of the hidden sources with
147 specific structures and physical interpretations, in contrast to other BSS methods. For example, it is
148 the sources that contribute to the development (rather than reduce) of ground deformation that gains
149 our most concerns. Sparsity constraint significantly limits the number of sources to represent the
150 observed ground deformation data. Non-negativity constraint gives the physical interpretation of
151 each source on the deformation at monitoring sites.

152 In this study, the highlight is that without knowing other subsurface conditions and other prior
153 knowledge, we introduce and improve Blind Source Separation (BSS) from the field of signal
154 processing to identify the latent sources and their corresponding patterns based only on a long-term
155 transient displacement data collected by GPS between 2001 and 2017 (measuring interval: 0.5 year).
156 Nonnegative Matrix Factorization (NMF) algorithm (one of BSS algorithms) by Lee and Seung
157 (1999) is applied to decompose the displacement data. Finally, we obtain the following items: 1) the
158 number of identified sources, source patterns and physical meanings, 2) the dynamic of dominant
159 source by comparing the contribution of each source.

160

161 **2. Methods**

162 All the main steps of the method are shown in Fig. 1. We develop an extended NMF framework
163 to decouple the ground deformation data. This method consists of (1) definition of physical problem
164 in a mathematical way, (2) extended NMF with stronger sparsity constraints, (3) initialization of
165 NMF, (4) determination of true number of sources, and (5) Evaluation and selection of robust
166 solution.

167 **2.1 Definitions**

168 Let us consider ground deformation is recorded by m GPS monitoring sites at a fixed
169 measurement frequency. Ground deformation is subjected to the effects of p unknown physical
170 sources that cause rock movement. The locations of the sources are unknown. Physical process of
171 different sources may interfere through the medium and their influences may be distributed or point.
172 In our analysis, there are no assumptions neither on the initial or boundary conditions like in
173 numerical model nor stochastic or probability on physical process. The only assumption in this
174 particular problem is that the value of ground deformation at a motoring site is the proportionally
175 added by unknown sources.

176 Here, we give the mathematical form of this particular problem.

$$177 \quad \mathbf{X}_{m \times n} = \mathbf{A}_{m \times p} \mathbf{S}_{p \times n} + \mathbf{E}_{m \times n} \quad (3)$$

178 Where $\mathbf{X}_{m \times n}$ is observed data by m GPS sites and n monitoring campaigns. $\mathbf{A}_{m \times p}$ is an
 179 unknown mixing matrix representing the linear combination of p sources. $\mathbf{S}_{p \times n}$ is a source
 180 contribution at each discrete time moment. $\mathbf{E}_{m \times n}$ represents the measurement errors or noises. The
 181 number of hidden sources is definitely less than the number of GPS monitoring sites (i.e., $p \leq m$).

182 2.2 NMF with sparsity constraints

183 In order to estimate matrix \mathbf{A} and \mathbf{S} , a cost function ϑ is chosen to quantify the similarity
 184 measure (also referred to as divergence or dissimilarity, distance depending on the probability
 185 distribution of estimated sources and the structure of data) between reconstructed data and observed
 186 data. $\|\cdot\|_F^2$ represents the Frobenius norm which is the squared Euclidean distance. α_A and α_S
 187 are nonnegative regularization parameters (typically, $\alpha_A = \alpha_S \in [0.01, 0.5]$) and the sparseness
 188 increases with an increase in the values of regularization coefficients. Penalty terms $J_A(\mathbf{A})$ and
 189 $J_S(\mathbf{S})$ enforce a certain application-dependent features to get a unique and desired solution.

$$190 \quad \vartheta = D_F(\mathbf{X}|\mathbf{A}\mathbf{S}) = \frac{1}{2} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 + \alpha_A J_A(\mathbf{A}) + \alpha_S J_S(\mathbf{S}) \quad (4)$$

$$191 \quad s. t. \ a_{ij} \geq 0, s_{jt} \geq 0, \forall i, j, t$$

192 In order to impose sparsity, we choose to use ℓ_1 -norm to minimize the cost function

$$193 \quad \vartheta = D_F(\mathbf{X}|\mathbf{A}\mathbf{S}) = \frac{1}{2} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 + \alpha_A \|\mathbf{A}\|_1 + \alpha_S \|\mathbf{S}\|_1 \quad (5)$$

$$194 \quad s. t. \ a_{ij} \geq 0, s_{jt} \geq 0, \forall i, j, t$$

195 To minimize the cost function ϑ , we choose the multiplicative update algorithm (equation 6,
 196 7) developed by [Lee and Seung \(1999\)](#), which updates the a_{ij} while keeping s_{jt} fixed, and then
 197 updates the s_{jt} while keeping a_{ij} fixed. Obviously, all the successive estimates remain positive if
 198 the initial estimate is positive. However, if an element of matrix becomes zero, it remains at zero
 199 for all the successive iterations. To circumvent this problem, we force the values of a_{ij} and s_{jt}
 200 not to be less than a small positive value ε (typically, $\varepsilon = 10^{-16}$).

$$201 \quad a_{ij} \leftarrow a_{ij} \frac{[\mathbf{X}\mathbf{S}^T]_{ij} - \alpha_A}{[\mathbf{A}\mathbf{S}\mathbf{S}^T]_{ij} + \varepsilon} \quad (6)$$

$$202 \quad s_{jt} \leftarrow s_{jt} \frac{[\mathbf{A}^T \mathbf{X}]_{jt} - \alpha_S}{[\mathbf{A}\mathbf{S}\mathbf{S}^T]_{jt} + \varepsilon} \quad (7)$$

203 The iterative algorithm is stopped when there is little or no improvement of cost function
 204 between successive iterations, $\omega = 10^{-10}$.

$$205 \quad \Delta\vartheta = \frac{|D_F^{(k)} - D_F^{(k-1)}|}{D_F^{(k)}} \leq \omega \quad (8)$$

206 2.3 Optimization and Solution

207 Obviously, the final solution is almost a local minima because the global minima is seldom
 208 achievable. In order to achieve a reasonable and optimal solution, another three issues (the first issue
 209 is to develop a problem dependent cost function) in NMF should be stressed: 1) alleviation of the
 210 rotation indeterminacy, 2) determination of the number of sources, 3) initialization of \mathbf{A} and \mathbf{S} .

211 *Alleviation of the rotation indeterminacy:*

212 The eq. 3 can also be presented as $\mathbf{X}_{m \times n} = \mathbf{A}_{m \times p} \mathbf{R} \mathbf{R}^{-1} \mathbf{S}_{p \times n} + \mathbf{E}_{m \times n}$. It is easy to find many
 213 rotational matrixes \mathbf{R} such that $\mathbf{A}_{m \times p} \mathbf{R} = \hat{\mathbf{A}}$ and $\mathbf{R}^{-1} \mathbf{S}_{p \times n} = \hat{\mathbf{S}}$ are also the solutions of eq. 3.
 214 Therefore, the solution of NMF is non-unique. To overcome this, we should keep the input data \mathbf{X} ,
 215 decomposed matrixes \mathbf{A} and \mathbf{S} sufficiently sparse ([Rickard and Cichocki 2008](#)). One common

216 way is to normalize the input data \mathbf{X} to make \mathbf{A} and \mathbf{S} zero-grounded. Besides, it is a necessary
 217 step to find a meaningful approximation of the hidden pattern for ground deformation neglecting
 218 the scaling ambiguity.

219 *Number of sources:*

220 Remember that the number of sources (rank of \mathbf{S} , p) and mixing process is unknown in NMF.
 221 Therefore, the number of sources can be extracted from observed data only. However, this remains
 222 an open issue. There are several methods by trial and error and heuristic techniques for determining
 223 the number of sources. 1) p is determined by the dimension of data, i.e., $p < mn/n + n$, where m
 224 represents the number of observations and n represents the number of monitoring campaigns. 2) p
 225 can be chosen in an ideal noiseless case by increasing the number of sources until the error of
 226 observed and reconstructed data ($\mathbf{E} = \mathbf{X} - \mathbf{X}'$) is zero. 3) [Tangirala et al. \(2007\)](#) proposed a PSV
 227 method based on the incremental variance captured by each source relative to the total power to
 228 determine the value of p . 4) others such as residual analysis can also help to choose the rank([Cuss](#)
 229 [et al. 2016](#)).

230 In this work, a heuristic method proposed by [Minka \(2001\)](#) was used to calculate the
 231 smoothness index defined as:

$$232 \quad SI(k) = \frac{var[\{\tilde{\lambda}_i\}_{i=k+1}^{I-1}]}{var[\{\tilde{\lambda}_i\}_{i=k}^{I-1}]}, \quad k = 1, 2, \dots, I - 2 \quad (9)$$

233 where $\tilde{\lambda}_i = \lambda_i - \lambda_{i+1}$, λ is the eigenvalue of the covariance matrix for observed data in a
 234 descending order ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l > 0$). The sample variance is defined as:

$$235 \quad var \left[\{\tilde{\lambda}_i\}_{i=k}^{I-1} \right] = \frac{1}{I-k} \sum_{i=k}^{I-1} (\tilde{\lambda}_i - \frac{1}{I-k} \sum_{i=k}^{I-1} \tilde{\lambda}_i)^2 \quad (10)$$

236 Finally, the number of sources p can be selected by the following criterion:

$$237 \quad p = \arg \min_{k=1, 2, \dots, I-2} SI(k) \quad (11)$$

238 *Initialization:*

239 The intrinsic of multiplicative update algorithms is nonconvex, even though the cost function
 240 is strictly convex with respect to one set of variables (i.e., either \mathbf{A} or \mathbf{S}). Therefore, it is impossible
 241 to find an analytical-based optimal solution([Rezaei and Boostani 2013](#)). The solution and
 242 convergence are highly depending on the initialization of \mathbf{A} and \mathbf{S} . Initializations nearing the
 243 optimal point in search space can enhance the performance of NMF, while poor initializations often
 244 lead to low convergence, and, in some certain instances, incorrect results. Many initialization
 245 methods have been proposed, such as random, PCA, SVD, divergence-based k-means, and
 246 IGA([Boutsidis and Gallopoulos 2007](#); [Langville et al. 2014](#); [Xue et al. 2008](#)). To obtain a robust
 247 initialization of \mathbf{A} and \mathbf{S} , we use the following steps:

248 1) Generate k initial \mathbf{A} and \mathbf{S} with Monte Carlo stimulation and run the ALS algorithm to
 249 obtain initialized \mathbf{A} and \mathbf{S} ([Bertin et al. 2007](#)) for predetermined rank p in eq. 11.

250 2) Run the proposed algorithm in this paper for each set of initial matrices.

251 **2.4 Evaluation criteria of solutions**

252 All the possible solutions are evaluated on the basis of criteria in terms of relative error (e),
 253 orthogonality (o) and sparsity (u) of identified sources([Hoyer 2004](#)). Each of these criteria reveals
 254 the capability of the demixing method from a specific aspect. The final optimal solution is a tradeoff
 255 among these three criteria, and this can be defined as a Euclidean distance in a 3D rectangular
 256 coordinate system.

$$\zeta = \min(\|\mathbf{v}_i\|^2), i = 1, 2, \dots, p * (k + 1) \quad (12)$$

Where \mathbf{v}_i is a vector composed of the three parameters(e_i, o_i, u_i). e_i, o_i and u_i are defined as the following equation 13 to 15 with a monotonic property. The value of e_i, o_i and u_i ranges $[0, +\infty)$, $[0, +\infty)$, and $[1, +\infty)$, respectively. A lower value of each parameter represents a better performance in terms of a specific aspect.

$$e_i = \|\mathbf{X} - \mathbf{A}\mathbf{S}\|^2 / \|\mathbf{A}\|^2 \quad (13)$$

$$o_i = \|\mathbf{O}_i\|^2, \mathbf{O}_i = (O_1, O_2, \dots, O_l), O_l = \mathbf{S}_q^T \mathbf{S}_h, l = C_p^2, 1 \leq q \neq h \leq p \quad (14)$$

$$\left\{ \begin{array}{l} u_i = \|\mathbf{U}_i\|^2, \mathbf{U}_i = (U_1, U_2, \dots, U_p) \\ \text{where } U_p = \frac{\sqrt{n}-1}{\sqrt{n-(\sum |s_{ph}|) / \sqrt{\sum s_{ph}^2}}}, h = 1, 2, \dots, n \end{array} \right. \quad (15)$$

2.5 Permutation of identified sources

Methods based on NMF cannot give the order of identified sources measured by explained variance like PCA. However, one of our goals is to identify the dominant source that is crucial for the prediction and controlling of ground deformation. The dominant source should explain the most ground deformation, and the secondary source explains the second-most remaining ground deformation. In order to reorder the identified sources, we defined the contribution ξ_l by l th source as the following equation 16. The source with the highest value of ξ_l is the most robust one that can be seen as the dominant source. By using this measurement, sources are reordered in a decreasing order of contribution.

$$\xi_l = \frac{\|A_{ip} S_{pj}\|^2}{\|\mathbf{X}\|^2}, i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (16)$$

3. Application

3.1 Geological setting

Jinchuan No.2 mine (JinChang City of GanSu Province, China) is one of the largest underground nickel mine in the world (Fig. 2a). Mining area lies in a monoclonal structure and more than 180 oblique faults distribute in this area after experiencing each geological tectonic movement and magma intrusion since Luliang Movement of Protozoic in China (Hui et al. 2018). High-level tectonic stress is a prominent characteristic in the mine area. In shallow strata, the axis of the maximum principal stress is horizontal (NE), which is basically perpendicular to the strike of the ore body (strike :N50°W, and dip angle : 40° to 70°). The maximum principal stress axis rotates to SSW-SW direction in deep strata. The dip of maximum principal stress also shows difference from shallow to deep strata. The dip increases up to 40° with the depth increasing (Cai 1999; Zhao et al. 2012). Ultrabasic rocks are the ore-bearing native rock. Surrounding rock comprises of granite, marble, and migmatite. These rock masses have poor mechanical stability because of the high density of discontinuity. For more detail in geological conditions, see Li et al. (2004) and Li et al. (2020a). The mechanized backfill mining technique is adopted in this mine field because of the well-developed geological foliation in the metamorphic strata and poor massif stabilities. After finishing of excavation in an access drift, the void is backfilled by cement paste subsequently and then the next access drift is excavated. Thus, no interval pillar is left and the filling body in the upper layer is served as the artificial roof for mining of the next layer. Underground mining operations have induced large scaled ground deformation

295 and surface fissures after 1999. Therefore, GPS has been used to monitor the ground deformation
296 since 2001. Previous studies have qualitatively assessed that high tectonic stress, mining rate,
297 discontinuity in rock mass, ratio of backfill and mining were the factors significantly affecting the
298 ground deformation (Ma et al. 2015; Zhao et al. 2013).

299 **3.2 Data processing**

300 GPS monitoring network contains a reference net and a deformation monitoring net on the
301 ground surface. Reference net containing 7 benchmarks is set far away from the deformed area.
302 During the long-term period, some of points are destroyed. Therefore, a total of 78 monitoring points
303 with successive time series out of 101 monitoring points are selected in this study. These points
304 distribute along the exploratory lines with a spacing of 50 m between two points on the same line.
305 Besides, an additional monitoring line (H line) is laid out on the upper surface of the ore body, which
306 is perpendicular to the exploratory lines (Fig. 2b). Original data is received by Z-12-type GPS
307 receiver and antenna (Ashtech, USA) (Fig. 2c). The nominal accuracies of the horizontal and
308 vertical displacements are 3 mm+0.5 ppm and 5 mm + 1 ppm, respectively. Square error of
309 processed displacement data is about ± 1.96 mm. For a single monitoring point, displacement data
310 comprises of the increment of X, Y, Z (local coordinate system) and each of them has 30 records.
311 Besides, displacement data is offset without changing the structure of data to ensure there is no
312 negative value in the matrix.

313 **4. Results**

314 The GPS monitoring data is a mixture of a series of unknown sources with dynamic mixing
315 ratios. Our goal is to separate the spatial and temporal patterns of dominate long-term and common-
316 mode. The NMF does this by minimizing cost function with non-negative and sparsity constraints.
317 To seek a trade-off between interpretability and statistical fidelity, the formulated evaluation criteria
318 helps to select the optimal solution from all possible solutions that has the highest resolution of each
319 source pattern.

320 After 100 tries of NMF and a series of optimization criteria by using the parameters of Table
321 1, we obtain the 100 solutions with different accuracies (Fig. 3). The optimal solution is determined
322 by error (e), orthogonality (o) and sparsity (u) according to eq. 12 with the number of iteration steps
323 of 1245. The value of e , o , and u is 0.02, 11.14, and 1.06, respectively. Although the NMF cannot
324 give the accurate solution, the 100 tries show a validation in statistics. The reconstructed data with
325 less noises (Fig. 4c) fits well with original data (Fig. 4a).

326 The final solution consists of three sources (Fig. 4b), namely source 1 (S1), source 2 (S2), and
327 source 3 (S3), which optimally isolate the major spatial and temporal trends underlying the ground
328 deformation time series. For comparison, the amplitude is normalized by the relative maximum. It
329 is worth noting that these three sources are not orthogonal because NMF does not assume the
330 orthogonality of sources. Therefore, the sum of source contribution is not equal to one according to
331 the equation 16. The source contribution to original data (Table 2) shows that S1 has the highest
332 contribution of 0.53, and S3 has the secondary highest contribution of 0.51, and S2 has the lowest
333 contribution of 0.28.

334 S1 is a source that has a high impact on Y displacement and a low impact on X and Z
335 displacement. In contrast, S3 shows a high contribution to X displacement and a low contribution
336 to Y and Z displacement. S2 primarily contributes to the Z displacement and has a low contribution
337 to X and Y displacement (Fig. 4b).

338 **4.1 Spatial pattern of source**

339 The mixing matrix \mathbf{A} represents contribution of the three sources at different monitoring point.
340 Therefore, each column of the mixing matrix \mathbf{A} can be regarded as the spatial pattern of one source.
341 As the element of \mathbf{A} is fixed, the spatial pattern of each source is stable among different monitoring
342 campaigns. This means that these three spatial patterns are time-independent and not to be affected
343 by mining intensity and mining methods. The value of displacement in different monitoring
344 campaign is controlled by the magnitude of source transient. Fig. 5 shows the contour of each
345 column of matrix \mathbf{A} that represent the spatial pattern of S1 (Fig. 5a), S2 (Fig. 5b), and S3 (Fig. 5c),
346 respectively.

347 S1 captures the feature of Y displacement. In general, the value of Y displacement decreases
348 from northwest to the southeast. There are two displacement centers that is nearly divided by line
349 14. The contribution of S1 near line 14 decreases sharply because the contour is dense. This indicates
350 there exists a deformation belt that has different mechanical properties near line 14 causing the sharp
351 change of Y displacement.

352 S2 controlling the Z displacement describes a homogeneously and broadly distributed pattern.
353 A spatial shape like concentric circle shows the effect of gravity and it seems that the displacement
354 does not affected by many kinds of discontinuities. This indicates that NMF has successfully
355 extracted a gravity-induced pattern. Another important conclusion is that the highest contribution of
356 S2 locates at the edge and gradually decrease at the center. This suggests that the vertical
357 displacement at the center is not gravity-induced.

358 S3 controlling the X displacement exhibits two concentric circle shapes at southeast and
359 northwest direction. The contribution of S3 increases from northwest to the southeast. Like the
360 pattern of S1, the two concentric circles are combined by a strong deformation belt distributing near
361 the H line and the center of the two circle are on the opposite side.

362 **4.2 Physical meaning of identified source**

363 In study area, high tectonic stress has been empirically considered as an important factor to the
364 ground deformation(Ma et al. 2015). However, the mechanism of ground deformation induced by
365 tectonic stress has never been revealed and quantitatively supported. Herein, according to the
366 identified patterns by NMF, we can make a forward step to deeply understand the process about
367 how tectonic stress affects ground deformation.

368 In natural conditions, the maximum principal stress σ_1 is perpendicular to the strike of ore
369 body(Ma et al. 2015). The long axis of strain ellipse is NW-SE and the relationship of stress is σ_1
370 $> \sigma_2$ (Fig. 6b). In post-mining condition, we find that the Z displacement at the center of ground
371 deformation in S2 pattern can be simplified as two sources controlled by S1 and S3 because the
372 contribution of S2 is nearly zero. Therefore, the ellipse shape at the center region (Fig. 5b) can be
373 thought as strain ellipse and the long axis is NE-SW. The rotation of ellipse indicates that mining
374 operations has caused the stress rotation in a larger area. At the same time, the relationship of stress
375 changes to $\sigma_2 > \sigma_1$. This can be supported by deformation traces. In detail, by overlying S1 and S3
376 patterns, we find a set of "x" shape strong deformation belts on the surface (Fig. 6a). Through the
377 deformation traces, it is evident that the deformation belt by S3 (left wing) shows a shear property,
378 and the slide direction can be obtained to reveal $\sigma_2 > \sigma_1$. The two quasi-circular closed region at
379 each end of shear belt may be caused by stress concentration (Fig. 5c)(Guo et al. 2020). Deformation
380 belt by S1 (right wing) shows an extension property because the displacement contour line in cross
381 section of "x" is straight and parallel to each other. In fact, as the mining tunnel is perpendicular to
382 σ_1 , broader free faces cause a larger stress release of σ_1 than σ_2 . This results in a series of extension

383 fractures perpendicular to σ_1 on surface (field investigation can be seen in [Ma et al. \(2015\)](#)) and
384 constraint of Y displacement by increasing the normal stress on joint planes. As a result, σ_1 causes
385 a large scale of X displacement at the beginning, and, then, σ_2 becomes the dominate stress in the
386 disturbed stress field which dominates Y displacement. This results in that S1 and S3 explain
387 different volume information of original data (Table 2) with the explanation ratios of $\xi_1/\xi_3=1.05$. It
388 is worth noting that the sum of explanation ξ_1 , ξ_2 , and ξ_3 is larger than 1 because these three sources
389 are not orthogonal causing the existence of partial correlation. Therefore, we conclude that S1 is
390 driven horizontal principal stress σ_2 , and S3 is driven horizontal principal stress σ_1 .

391 S2 describes a homogeneously and broadly distributed pattern that represents the self-weight
392 body force caused by gravity. In the scale of study area, density of rock mass is the only factor affect
393 the value and homogeneity of self-weight body force. We find that the mechanism of self-weight
394 body force on ground deformation is significantly different from empirical knowledge. Specifically,
395 on one hand, surface deformation induced by self-weight body force explains the least information
396 of original data compared to S1 and S3. On the other hand, self-weight body force has a high
397 contribution to Z displacement far away from center while low contribution near the subsidence
398 center (Fig. 5b). Compression state of rock mass is found at the center of ground deformation in
399 many studies ([Ma et al. 2018](#); [Xia et al. 2016](#)). In this study, high confining pressure caused by σ_1
400 and σ_2 produces enough force of friction on joint planes with high dip angle that alleviating the Z
401 displacement. At the same time, Z displacement is dominated by σ_1 and σ_2 because of the
402 existence of the joint planes with lower dip angle. This can be supported by ratios of source
403 contribution on single point such as 14~7 and H~11 (Fig. 9c). Besides, we propose that the high
404 contribution of S2 away from the center may be caused by the extension-subsidence after the stress
405 release of σ_1 and σ_2 .

406 **4.3 Temporal pattern of source**

407 Temporal evolution of displacement is critical for disaster early warning. However, previous
408 studies focus on the displacement evolution responding to the combined actions of ambient
409 sources([Li et al. 2020b](#)). In contrast, in this study, temporal evolution of each displacement
410 component is related to separate source. Fig. 7 shows the temporal pattern of three sources. The
411 temporal pattern does not show periodicity or seasonality related to precipitation indicating that the
412 precipitation induced consolidation deformation can be ignored in this arid region. Another
413 important conclusion from the analyses is that the time lag is not obvious by comparing the curve
414 of three sources (Fig. 7). This indicates these three sources are independent in physical mechanism.

415 The deformation rate of three sources slowly increases with time. In detail, S1 induces the
416 largest increment of deformation rate of 0.0094, and S3 induces a medium value of 0.0086, and S2
417 causes a lowest increment of deformation rate of 0.0056 (Fig. 7). The different increment of
418 deformation rate may be driven by the different physical process. The stress concentration is more
419 remarkable with mining depth increase, and, thus, the displacement controlled by σ_1 and σ_2
420 increases slowly. In contrast, the self-weight body force has no obvious change because it is
421 controlled by gravity and the volume and density of overlying rock mass. Therefore, the increment
422 of deformation rate caused by S2 would be a stable value ideally. However, the discontinuities,
423 especially the joint and other structure planes, contribute to a component of Z displacement when
424 suffering horizontal stress. As the dynamic of stress redistribution, the normal stress may decrease
425 on structure planes at the tensile zone mainly distributed at the edge of deformation area (Fig. 5b)
426 can favor the Z displacement driven by self-weight body force.

427 By comparing the source contribution to X, Y, and Z displacement, we can divide the temporal
428 pattern into three stages (Fig. 8). These three stages reflect the evaluation of ground deformation. A
429 “vibration-acceleration” phenomenon is revealed. During the vibration period, the stress
430 redistribution and adjustment of discontinuities occur in rock mass. The acceleration is a relatively
431 faster process, which is driven by the stress concentration and the degradation of rock mass. Another
432 phenomenon is that each source has a more distinct contribution to its controlling displacement
433 component while keeps a continuous decrease contribution to another two components. A
434 reasonable explanation is that the overlying rock mass is becoming loose, and the confining pressure
435 is decreasing on the plane of discontinuities. Therefore, S3 contributes more X displacement and
436 less Y, and Z displacement along the plane (S1 and S2 have the same mechanism with S3). This
437 conclusion can be used to explain many collapse events in other mining areas (Feng et al. 2019;
438 Waltham et al. 2011).

439 **4.4 Relationship between potential strong deformation belt and the change of “controlling 440 pattern”**

441 Most of sudden disasters are evolved from a long-time slow and micro deformation and the
442 location of final failure always controlled by the previously deformed belts. GPS has successfully
443 applied to detect the slow slip and micro deformation of landslides and mining induced ground
444 deformation (Gourmelen et al. 2007; Klein et al. 2018; LI Xiao 2020). Prediction of deformation is
445 the main issue in previous studies(Chen et al. 2020) . The main process is, firstly, fitting measured
446 data by various methods and, then, predicting the deformation by learned models. However, few of
447 these studies make a forward to reveal what drives the deformation. As we all know, ground
448 deformation is always drove by multi sources in reality. Prediction may be more valuable and
449 reliable on the basis of physical mechanism of each source. In this study, we assume that the GPS
450 data is a mixture of a series of unknown sources with dynamic mixing ratios. As a result, we separate
451 two horizontal tectonic stresses and self-weight body force controlling the characteristics and
452 tendency of ground deformation.

453 On the basis of separated sources, we, firstly, propose a new conception “controlling pattern”
454 that is defined as a combination of sources ordered by their contribution at a certain time. It is
455 dynamic with time and any source can become the dominate source. It is known that different value
456 of confining pressure can cause different shape and location of rupture in the triaxial compression
457 test. Like that, different controlling patterns will lead to different features of ground deformation. In
458 addition, the location of micro deformation by first controlling pattern will play an important role
459 in subsequent deformation process like “buckets effect”. Fig. 9 shows the contribution ratio of S1,
460 S2, and S3 at each point. The relationship of three ratios is stable generally except for several points
461 located on strong deformation belts. For X displacement, the dominate source is S3, and the
462 secondary and the third is S1 and S2, respectively. In the strong deformation belts, the source
463 contribution of S3 decreases sharply while the contribution of S1 increases at points, such as 8~3,
464 10~9, 14~7, 18~6, H-11, and Q~1 (Fig. 9a, 9b). The trend of Y displacement is similar to X
465 displacement. Transformation of dominate source is found in Z displacement (Fig. 9c). The
466 dominate sources S2 transfers to S1 or S3 in strong deformation belts. In addition, these changes
467 are the same trend in each monitoring campaigns, which means we can identify it at an early phase
468 of ground deformation helping to prevent risk like the severe damage of ventilation shaft on line 14
469 (Fig. 2b) in 2005(Ma et al. 2015).

470 **5. Discussion**

5.1 Effect of discontinuity

Displacement is the comprehensive result that can be treated as a transient signal on the surface without caring about the transmit process in rock mass. Here, it is important to emphasize that in the presented analysis, we do not estimate the source transients changed during their propagation, and the effect of many influence factors before the displacement reaching the monitoring point. The NMF only decomposes the manifestation of the source transients in the observed data. Many driving sources and influence factors affect the development of ground deformation (Peng et al. 2018). The driving sources are related to external force loadings that are the source of displacement. Influence factors include lithology and structure of rock mass, and some physical processes that affect the mechanical properties and geometry of rock mass and discontinuity. These influence factors are highly related to the characteristics of displacement components, which covers up the characteristics of driving sources. In general, it is difficult to characterize and separate the effect of each influence factor in a regional scale because we even not understand the mechanic of a single factor, much less the coupling mechanics of all factors. However, it is more meaningful for safety management to know the comprehensive effect of these influence factors. As a result, NMF successfully identify three driving sources and the corresponding spatial and temporal patterns that is a reflection of the comprehensive effect of these influence factors. It is worth noting that driving sources identification is a core step to analyze the mechanism and evolution of ground deformation because they are carriers of these influence factors.

From the analysis in section 4, we find that discontinuities such as joints and faults, may dominant effects on the deformation behaviors of the rock mass in study area. Although NMF cannot give the precise effect of discontinuities, a conclusion can be proposed that discontinuities significantly contribute to rock movement in study area. In fact, field investigations find that there are four groups of dominate joints (Fig. 10a, c) with an average of dip direction is 7.2° (percent: 16.31%), 96.27° (26%), 186.2° (29.37%), and 289.5° (15.61%), respectively. This indicates that joints with E and S dip direction can significantly affect the ground deformation. The consistency between direction of principal stress (Fig. 6a) and E and S dip direction favors the shear movement of rock, therefore, causing the shear belt and extension belt on the ground (Fig. 6a). In addition, S1 and S3 have a high contribution of Z displacement (Fig. 4) while the pattern of S2 shows that S2 has a small contribution to X and Y displacement (Fig. 4), which suggests that the discontinuities with high dip angle are highly developed and, therefore, the rock has a small slide component in X and Y direction. Field investigations also reveal the highly developed of joints with high dip angle. As shown in Fig. 10c, the average dip of four groups of joints is 59.6° , 54.2° , 64.2° , 63° , respectively.

The temporal pattern of three sources are also affected by discontinuity. The characteristic of each stage in Fig. 8 is different. The amplitude of source continuously increases from the first stage to the third stage. For example, S3 has a normalized amplitude of 0.63 in the first stage and increases to 0.72 in the second stage, and 0.82 in the third stage (Fig. 11a). S1 and S2 have the similar changes in amplitude (Fig. 11b, c). In addition, the repeated mining operations can disturb tectonic stress causing the concentration or relaxation of stress but the self-weight body force remains stable. This leads to the oscillation amplitude of S2 is smaller than that of S1 and S3. These temporal patterns reveal that the overlaying rock mass is becoming increasingly sensitive to stress perturbations, which is consistent with the process closing to the volumetric expansion point in uniaxial compression test (Fig.12). During this process, the discontinuity becomes unstable and deforms easily under small stress perturbation, especially shear stress perturbation. However, NMF cannot

515 predict when the ground deformation will evolve to failure (collapse of overlaying strata).

516 **6. Conclusions**

517 Our analyses demonstrate the applicability of a new inverse method for analysis of ground
518 deformation based on NMF algorithm of BSS. The solution is optimized by newly formulated
519 evaluation criteria to reduce the interference among sources and improve the resolution of each
520 pattern. The unknown sources are identified from a set of GPS time series data without any
521 information about the sources, underground mining conditions, and the physical processes
522 impacting the displacement propagation through rock mass.

523 The NMF based method identifies three unique sources causing the observed data. These are
524 listed in a descending order in terms of source contribution to original data: S1 (driven by horizontal
525 principal stress σ_2), S3 (driven by horizontal principal stress σ_1), and S2 (driven by self-weight
526 body force caused by gravity). They appear to be proportionally manifested at the monitoring point.
527 Relative contribution of each source remains the same order as $S1 > S3 > S2$ except for the monitoring
528 points on strong deformation belts where the source contribution changes distinctly and even
529 dominate source transformation occurs. Besides, each source has stable spatial pattern which is
530 time-independent. This allows us to identify the potential risk area at an earlier stage. Finally, the
531 NMF based method can be applied to any real problem where temporal system behavior is
532 monitored at multiple locations.

533 Although NMF is useful to long-term and cyclical driving source, NMF cannot identify some
534 instant driving sources like blasting vibrating load and seismic load. The data sampling frequency
535 (interval: 0.5 year) affects the result of NMF and high frequency sampling may overcome this
536 problem and show more details of sources.

537 **Acknowledgements**

538 We gratefully acknowledge our financial support from the National Science Foundation of
539 China (Grant Nos. 41907174, 41831293, 41772324), and National Key Research Projects of China
540 (2016YFC0402802). The data used in our analysis is included with and described
541 in <https://github.com/guhongyu18/grounddeformation>.

542

543 **Credit author statement**

544 We outline all authors' individual contributions as follows:

545 Hongyu Gu: Conceptualization, Data curation, Formal analysis, Investigation, Visualization,
546 Writing – original draft, Writing – review & editing, and Funding.

547 Fengshan Ma: Funding, Investigation, Project administration, Resources, Supervision,
548 Validation, and Writing – review & editing.

549 Liangjun Lin: Validation, Formal analysis, Revise, and Visualization

550 Donghui Wang: Investigation, Validation.

551 Yongbo Tie: Investigation, Funding.

552 Weichang Chen: Writing – review & editing

553

554 **Declaration of Competing Interest**

555 None.

556 **References**

557 Bertin N, Badeau R, Richard G Blind Signal Decompositions for Automatic Transcription of
558 Polyphonic Music: NMF and K-SVD on the Benchmark. In: Acoustics, Speech and Signal
559 Processing, 2007. ICASSP 2007. IEEE International Conference on, 2007.

560 Bibo D, Yongbin J (2016) Research on Distribution of Geometric Parameters of Structural Plane
561 Metal Mine 4: 58-61.

562 Boutsidis C, Gallopoulos E (2007) SVD based initialization: A head start for nonnegative matrix
563 factorization Pattern Recognition 41:1350-1362

564 Cai MQ (1999) Results and analysis of insitu stress measuremet at deep position of No.2
565 miningarea of Jinchuan Nickkel Mine Chinese Journal of Rock Mechanics & Engineering
566 5:414-418

567 Cao J, Ma F, Guo J, Lu R, Liu G (2019) Assessment of mining-related seabed subsidence using
568 GIS spatial regression methods: a case study of the Sanshandao gold mine (Laizhou,
569 Shandong Province, China) Environmental Earth Sciences 78:26 doi:10.1007/s12665-018-
570 8022-1

571 Chang C, Fung PCW, Hung YS On a Sparse Component Analysis Approach to Blind Source
572 Separation. In, Berlin, Heidelberg, 2006. Independent Component Analysis and Blind Signal
573 Separation. Springer Berlin Heidelberg, pp 765-772

574 Chen B, Gong H, Li X, Lei K, Lin Z, Gao M, Zhou C (2017) Characterization and causes of land
575 subsidence in Beijing, China International Journal of Remote Sensing 38:808-826

576 Chen B, Li Z, Yu C, Fairbairn D, Kang J, Hu J, Liang L (2020) Three-dimensional time-varying large
577 surface displacements in coal exploiting areas revealed through integration of SAR pixel
578 offset measurements and mining subsidence model Remote Sensing of Environment
579 240:111663 doi:https://doi.org/10.1016/j.rse.2020.111663

580 Ciaramella A, Lauro ED, Martino SD, Falanga M, Tagliaferri R (2006) ICA based identification of
581 dynamical systems generating synthetic and real world time series Soft Computing 10:587-
582 606

583 Cohen-Waeber J, Bürgmann R, Chaussard E, Giannico C, Ferretti A (2018) Spatiotemporal
584 Patterns of Precipitation-Modulated Landslide Deformation From Independent Component
585 Analysis of InSAR Time Series Geophysical Research Letters 45:1878-1887
586 doi:10.1002/2017gl075950

587 Cuss CW, Guéguen C, Andersson P, Porcelli D, Kutscher L (2016) Advanced Residuals Analysis for
588 Determining the Number of PARAFAC Components in Dissolved Organic Matter Applied
589 Spectroscopy 70:1-13

590 Díaz-Fernández ME, Álvarez-Fernández MI, Álvarez-Vigil AE (2010) Computation of influence
591 functions for automatic mining subsidence prediction Computational Geosciences 14:83-
592 103 doi:10.1007/s10596-009-9134-1

593 Ding K, Ma F, Guo J, Zhao H, Lu R, Liu F (2017) Investigation of the Mechanism of Roof Caving in
594 the Jinchuan Nickel Mine, China Rock Mechanics and Rock Engineering 51:1215-1226
595 doi:10.1007/s00603-017-1374-0

596 Fathi Salmi E, Nazem M, Karakus M (2017) Numerical analysis of a large landslide induced by
597 coal mining subsidence Engineering Geology 217:141-152
598 doi:https://doi.org/10.1016/j.enggeo.2016.12.021

599 Feng X, Zhang N, Xue F, Xie Z (2019) Practices, experience, and lessons learned based on field
600 observations of support failures in some Chinese coal mines International Journal of Rock
601 Mechanics and Mining Sciences 123:104097
602 doi:https://doi.org/10.1016/j.ijrmms.2019.104097

603 Gourmelen, Noel, Amelung, Falk, Riccardo (2007) Mining-related ground deformation in
604 Crescent Valley, Nevada: Implications for sparse GPS networks Geophysical Research Letters
605 34:252-254

606 Guo S, Qi S, Zhan Z, Ma L, Gure EG, Zhang S (2020) Numerical study on the progressive failure
607 of heterogeneous geomaterials under varied confining stresses Engineering Geology
608 269:105556 doi:https://doi.org/10.1016/j.enggeo.2020.105556

609 Hoyer PO (2004) Non-negative matrix factorization with sparseness constraints Journal of
610 Machine Learning Research 5:1457-1469

611 Hui X, Ma F, Zhao H, Xu J (2018) Monitoring and statistical analysis of mine subsidence at three
612 metal mines in China *Bulletin of Engineering Geology and the Environment*:3983–4001
613 Kabwe E (2017) Mining Sequence Deformation and Failure Behaviour Analysis in the hangingwall
614 and Orebody Rock Formations; A Continuum Approach *Geotechnical and Geological*
615 *Engineering* 35:1453-1473 doi:10.1007/s10706-017-0187-y
616 Karoui MS, Deville Y, Hosseini S, Ouamri A A new spatial sparsity-based method for extracting
617 endmember spectra from hyperspectral data with some pure pixels. In: *Geoscience &*
618 *Remote Sensing Symposium, 2012 IEEE International IEEE, 2012.*
619 Khanal M, Adhikary D, Jayasundara C, Rao B (2015) Numerical Study of Mine Site Specific
620 Multiseam Mining and Its Impact on Surface Subsidence and Chain Pillar Stress
621 *Geotechnical & Geological Engineering* 34:1-19
622 Klein E, Duputel Z, Zigone D, Vigny C, Boy J-P, Doubre C, Meneses G (2018) Deep Transient Slow
623 Slip Detected by Survey GPS in the Region of Atacama, Chile *Geophysical Research Letters*
624 45:12,263-212,273 doi:10.1029/2018gl080613
625 Langville AN, Meyer CD, Albright R, Cox J, Duling D (2014) Algorithms, Initializations, and
626 Convergence for the Nonnegative Matrix Factorization Eprint Arxiv
627 Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization
628 *Nature* 401:788-791
629 Lee TW, Lewicki MS, Girolami M, Sejnowski TJ (1999) "Bind source separation of more sources
630 than mixtures using overcomplete representations," *Signal Processing Letters IEEE* 6:87-90
631 Li G, Ma F, Guo J, Zhao H, Liu G (2020a) Study on deformation failure mechanism and support
632 technology of deep soft rock roadway *Engineering Geology* 264:105262
633 doi:<https://doi.org/10.1016/j.enggeo.2019.105262>
634 Li M, Zhang L, Ding C, Li W, Luo H, Liao M, Xu Q (2020b) Retrieval of historical surface
635 displacements of the Baige landslide from time-series SAR observations for retrospective
636 analysis of the collapse event *Remote Sensing of Environment* 240:111695
637 doi:<https://doi.org/10.1016/j.rse.2020.111695>
638 Li X, Wang SJ, Liu TY, Ma FS (2004) Engineering geology, ground surface movement and fissures
639 induced by underground mining in the Jinchuan Nickel Mine *Engineering geology* 76:93-
640 107
641 Li Xiao ZN, SHENG Zhuping, LI Shouding, HAO Jianming (2020) Sliding mechanisms and fracture
642 genesis of Jiweishan landslide in Wulong#br# chinese *Journal of Rock Mechanics &*
643 *Engineering* 39:1-12 doi:10.13722/j.cnki.jrme.2019.0453
644 Ma F, Gu H, Guo J, Lu R (2018) Analysis of ground deformation based on GPS in Sanshandao
645 gold mine, *China Journal of Nepal Geological Society* 55:7-14
646 Ma F, Zhao H, Yuan R, Guo J (2015) Ground movement resulting from underground backfill
647 mining in a nickel mine (Gansu Province, China) *Natural Hazards* 77:1475-1490
648 Minka TP (2001) Automatic choice of dimensionality for PCA *Media Laboratory Perceptual*
649 *Computing Section Technical Report* 13:556-562
650 Peng X, Yu P, Zhang Y, Chen G (2018) Applying modified discontinuous deformation analysis to
651 assess the dynamic response of sites containing discontinuities *Engineering Geology*
652 246:349-360 doi:<https://doi.org/10.1016/j.enggeo.2018.10.011>
653 Perrin J, Cartannaz C, Noury G, Vanoudheusden E (2015) A multicriteria approach to karst
654 subsidence hazard mapping supported by weights-of-evidence analysis *Engineering*
655 *Geology* 197:296-305 doi:<https://doi.org/10.1016/j.enggeo.2015.09.001>
656 Rezaei M, Boostani R (2013) Using the genetic algorithm to enhance nonnegative matrix
657 factorization initialization *Expert Systems* 31:213-219
658 Rickard S, Cichocki A When is non-negative matrix decomposition unique? In: 42nd Annual
659 Conference on Information Sciences and Systems, CISS 2008, Princeton, NJ, USA, 19-21
660 March 2008, 2008.
661 Sahu SP, Yadav M, Das AJ, Prakash A, Kumar A (2017) Multivariate statistical approach for

662 assessment of subsidence in Jharia coalfields, India Arabian Journal of Geosciences 10:191
 663 Salmi EF, Karakus M, Nazem M (2019) Assessing the effects of rock mass gradual deterioration
 664 on the long-term stability of abandoned mine workings and the mechanisms of post-
 665 mining subsidence – A case study of Castle Fields mine Tunnelling and Underground Space
 666 Technology 88:169-185 doi:10.1016/j.tust.2019.03.007
 667 Schultz-Fellenz ES et al. (2020) High-resolution surface topographic change analyses to
 668 characterize a series of underground explosions Remote Sensing of Environment
 669 246:111871 doi:https://doi.org/10.1016/j.rse.2020.111871
 670 Song S, Zhao X, Xie J, Guan Y (2011) Grey Correlation Analysis and Regression Estimation of
 671 Mining Subsidence in Yu-Shen-Fu Mining Area Procedia Environmental Sciences 10:1747-
 672 1752 doi:https://doi.org/10.1016/j.proenv.2011.09.274
 673 Tangirala AK, Kanodia J, Shah SL (2007) Non-Negative Matrix Factorization for Detection and
 674 Diagnosis of Plantwide Oscillations Industrial & Engineering Chemistry Research 46:801-817
 675 Waltham T, Park HD, Suh J, Yu MH, Kwon HH, Bang KM (2011) Collapses of old mines in Korea
 676 Engineering Geology 118:29-36 doi:https://doi.org/10.1016/j.enggeo.2010.11.007
 677 Xia K, Chen C, Fu H, Pan Y, Deng Y (2016) Mining-induced ground deformation in tectonic stress
 678 metal mines: A case study Bulletin of Engineering Geology & the Environment 210:212-230
 679 Xia K, Chen C, Zheng Y, Zhang H, Liu X, Deng Y, Yang K (2019) Engineering geology and ground
 680 collapse mechanism in the Chengchao Iron-ore Mine in China Engineering Geology
 681 249:129-147 doi:https://doi.org/10.1016/j.enggeo.2018.12.028
 682 Xue Y, Tong CS, Chen Y, Chen W-S (2008) Clustering-based initialization for non-negative
 683 matrix factorization Applied Mathematics & Computation 205:525-536
 684 Ye S, Yue L, Wu J, Yan X, Wang H, Xun J, Teatini P (2016) Three-dimensional numerical modeling
 685 of land subsidence in Shanghai, China Hydrogeology Journal 24:695-709
 686 Zhao H, Ma F, Xu J, Guo J (2012) In situ stress field inversion and its application in mining-
 687 induced rock mass movement International Journal of Rock Mechanics and Mining Sciences
 688 53:120-128 doi:10.1016/j.ijrmms.2012.05.005
 689 Zhao H, Ma F, Zhang Y, Jie G (2013) Monitoring and mechanisms of ground deformation and
 690 ground fissures induced by cut-and-fill mining in the Jinchuan Mine 2, China Environmental
 691 Earth Sciences 68:1903-1911
 692 Zhou C et al. (2019) Quantifying the contribution of multiple factors to land subsidence in the
 693 Beijing Plain, China with machine learning technology Geomorphology 335:48-61
 694 doi:https://doi.org/10.1016/j.geomorph.2019.03.017
 695

Figures

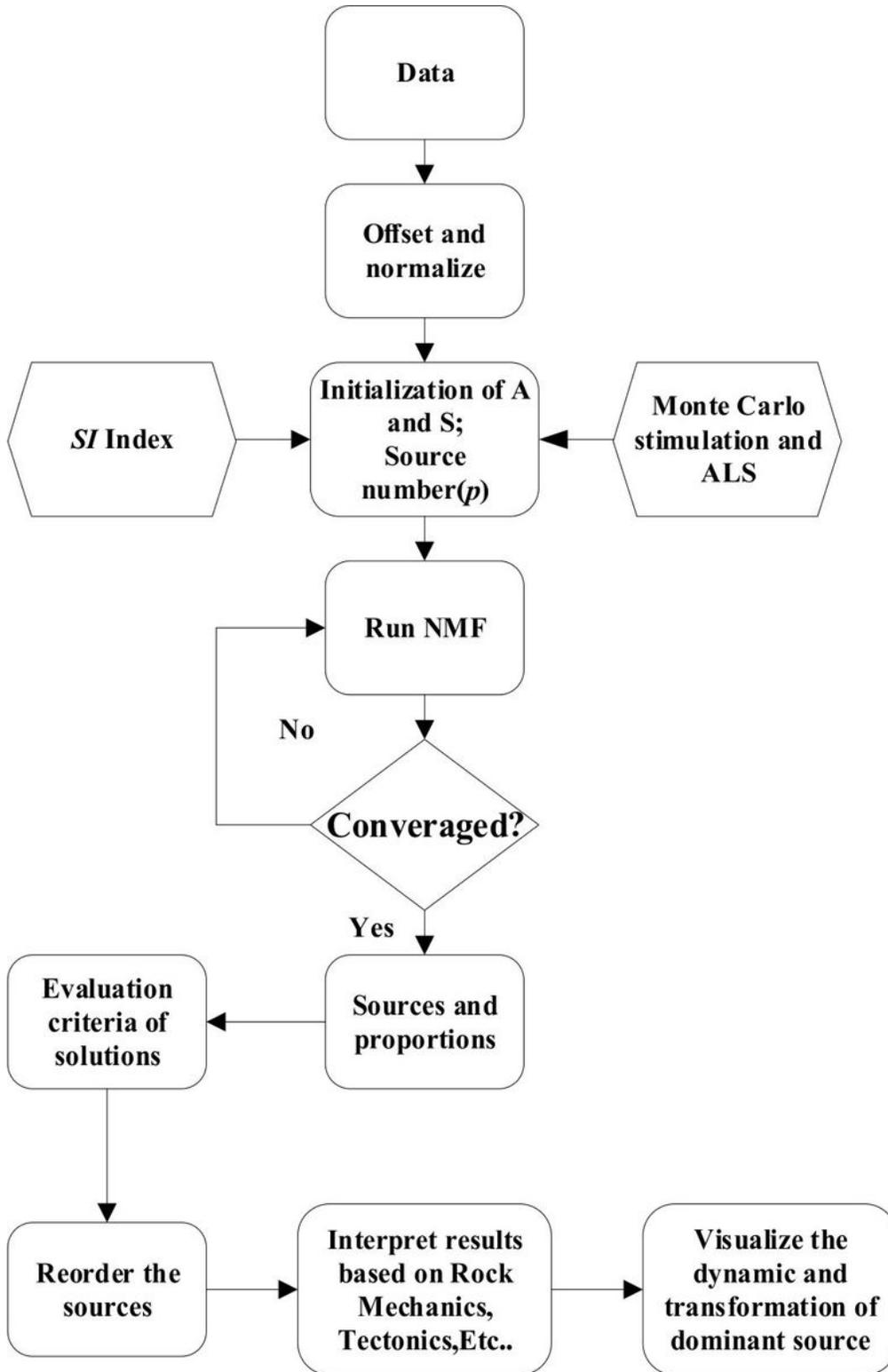


Figure 1

Flow chart of the BSS based on NMF algorithm.

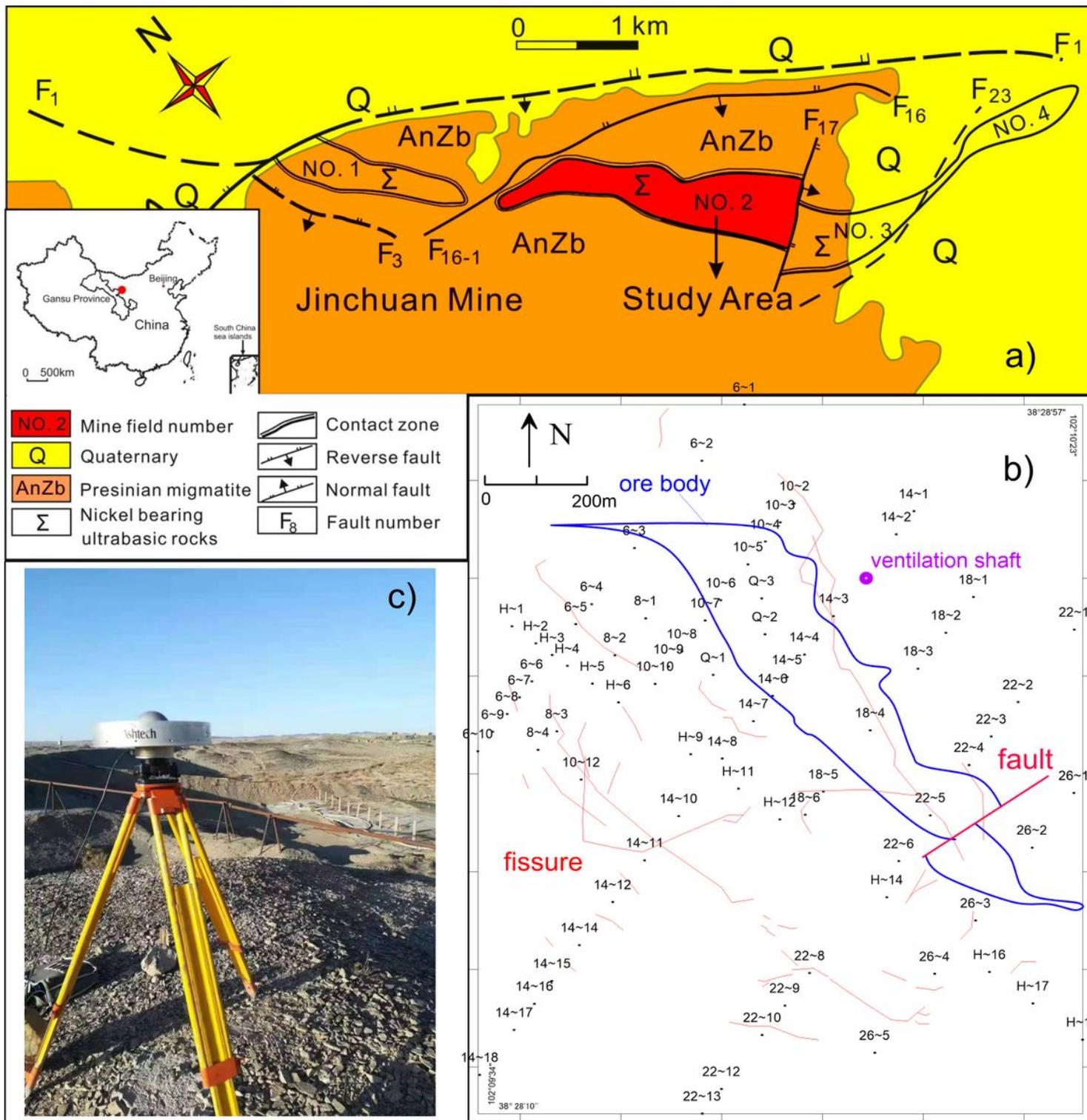


Figure 2

a) location of study area and geological setting, b) distribution of monitoring points, and c) the Z-12-type GPS receiver. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

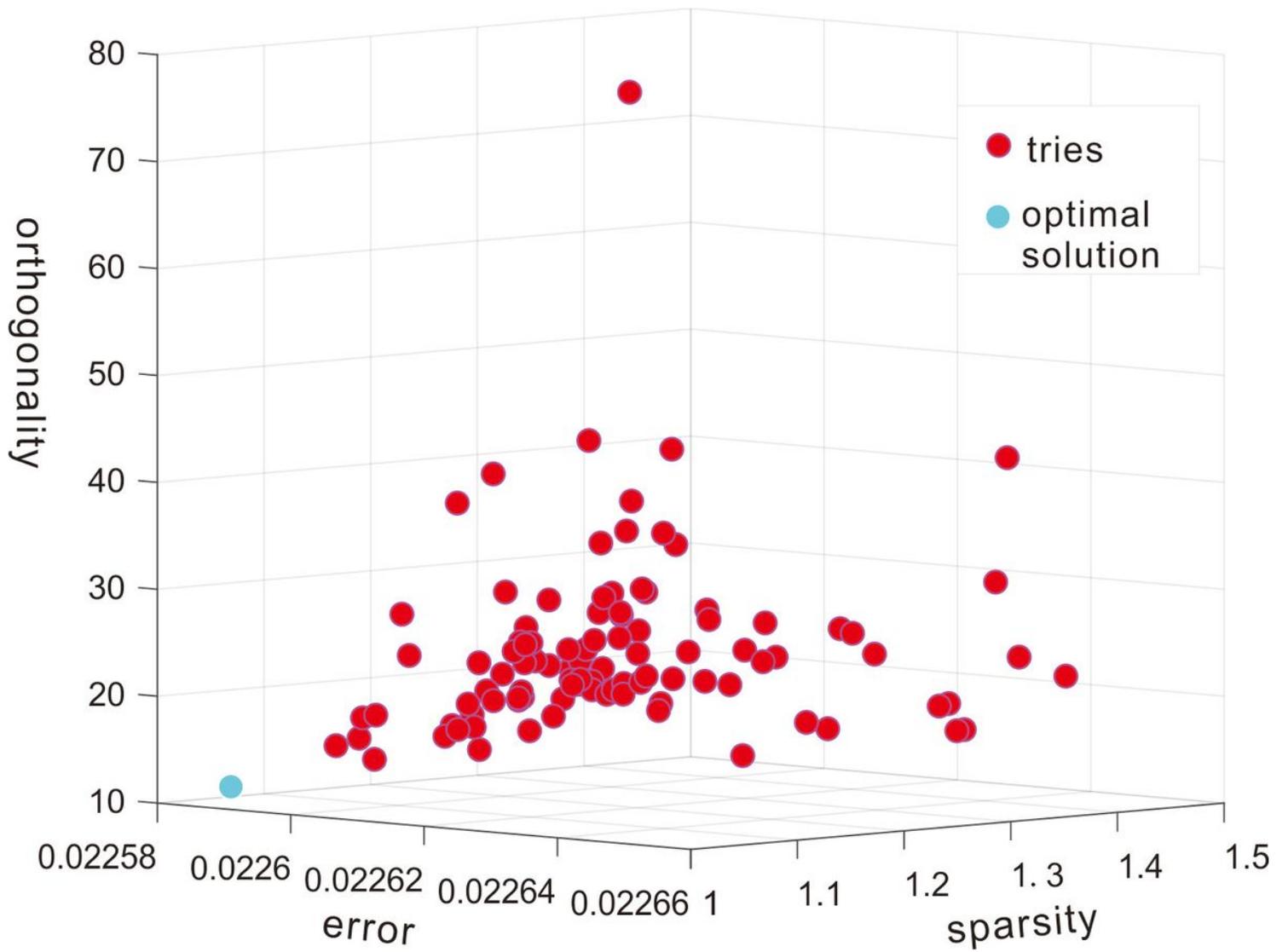


Figure 3

Results of 100 tries of NMF and the optimal solution is in blue.

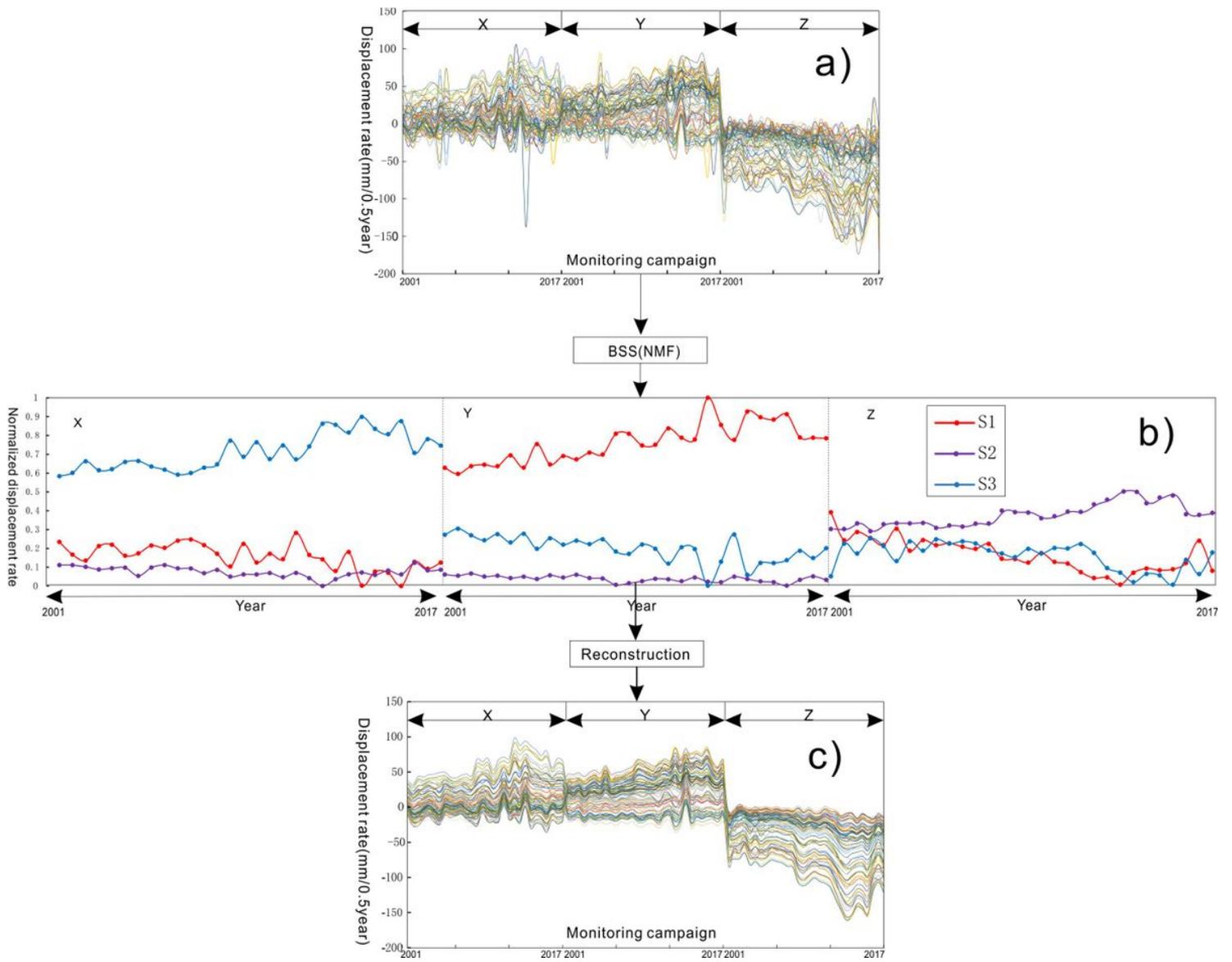


Figure 4

a) original data, b) the identified three sources by NMF, and c) the reconstructed data.

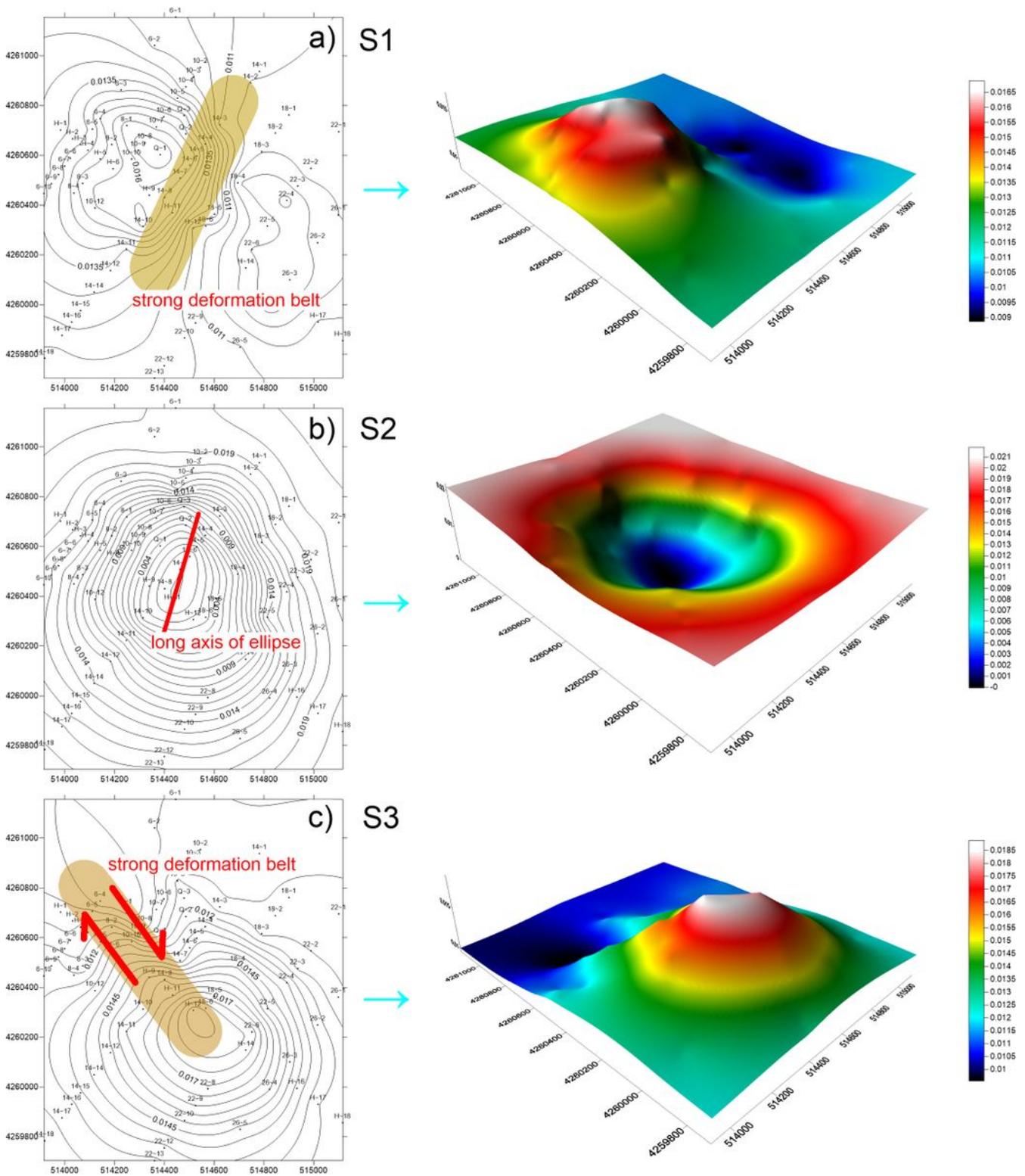


Figure 5

Spatial pattern of three sources including contour map and 3 dimensional graphics. Strong deformation belts are identified by deformation traces.

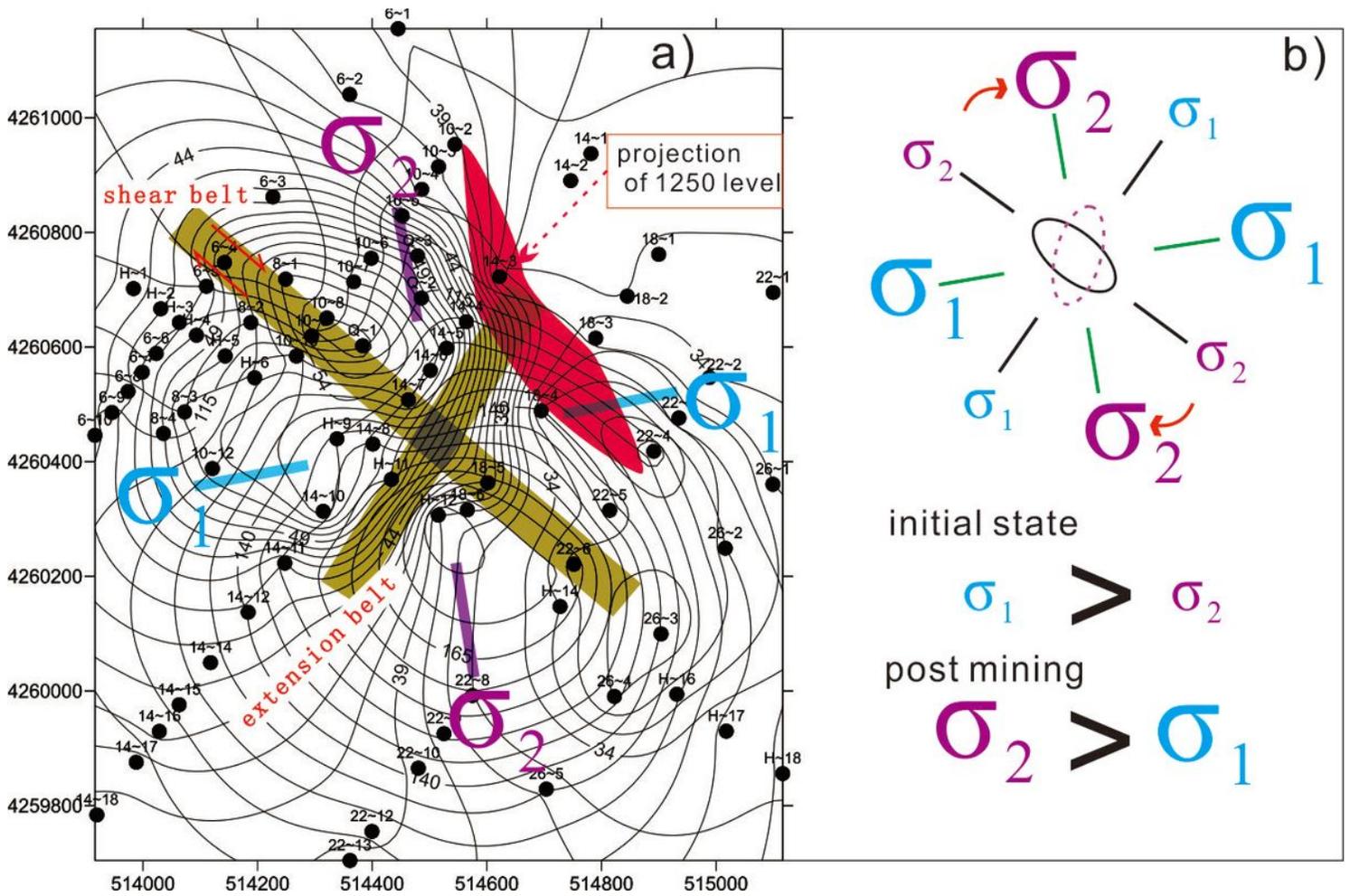


Figure 6

a) stress state is analyzed based on deformation traces by overlying the contour map of S1 and S3. b) stress state before mining and post mining, and the rotation of stress occurs after mining.

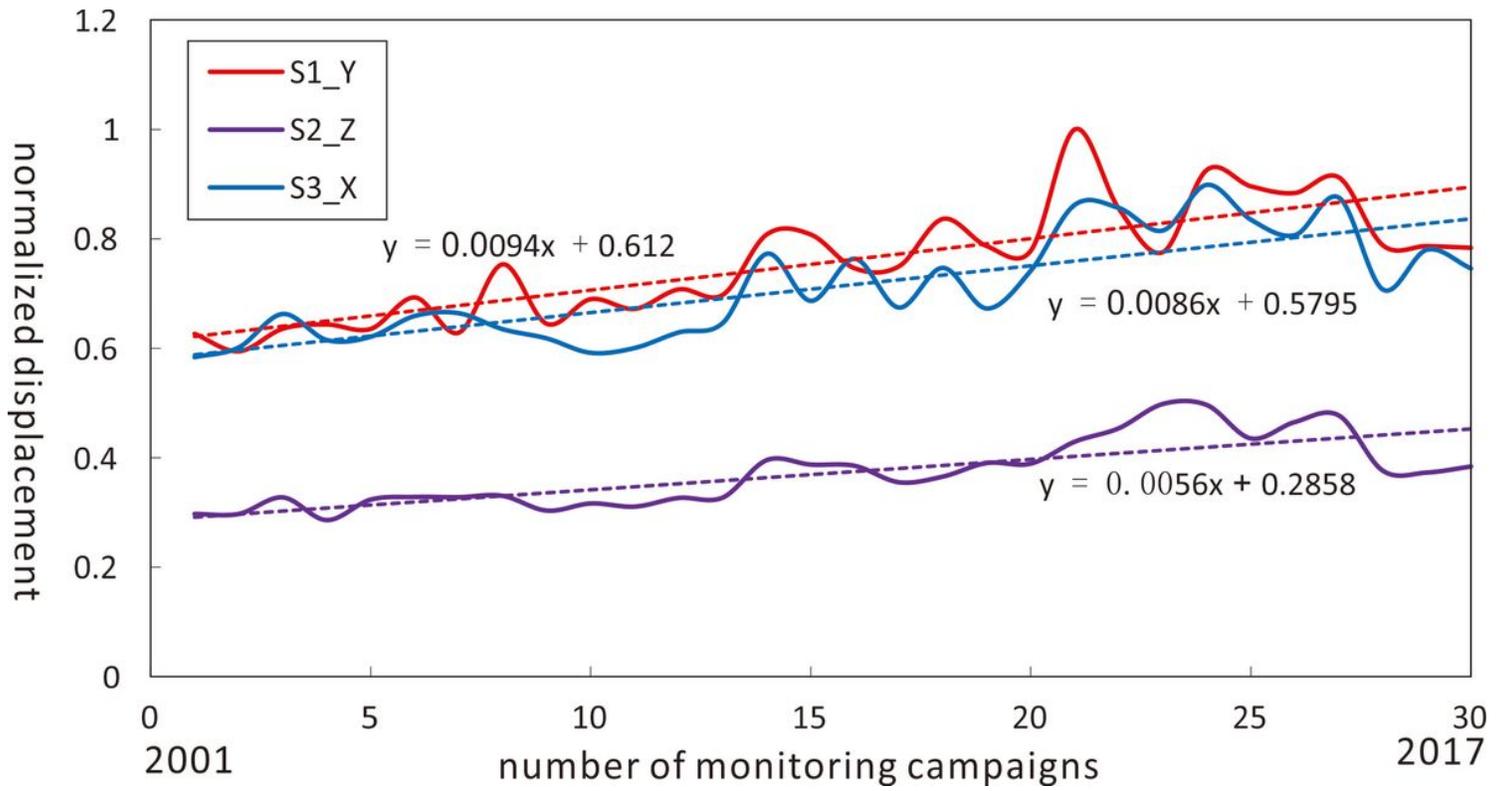


Figure 7

The temporal pattern of three sources.

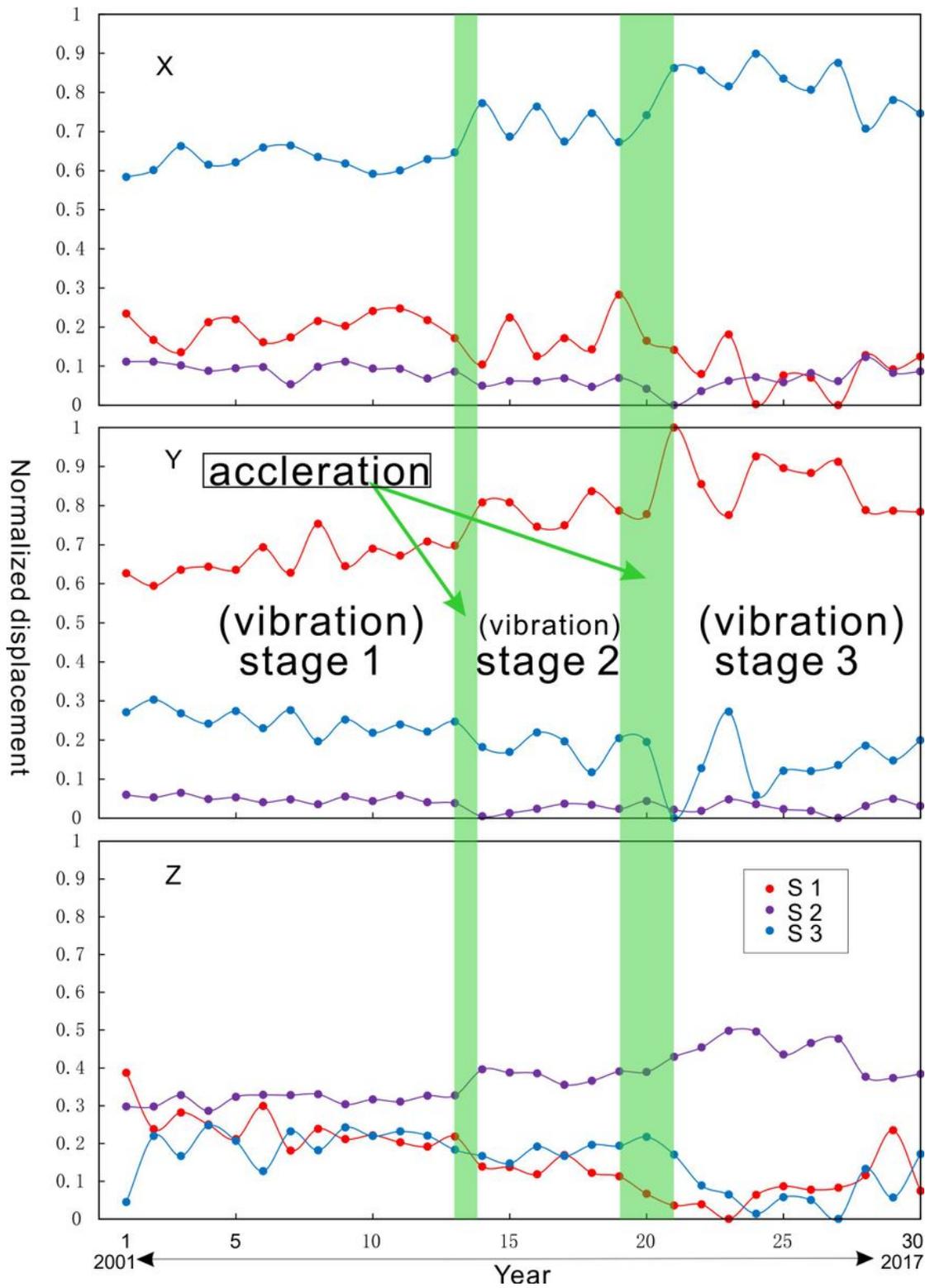


Figure 8

The evolution of ground deformation consisting of three stages of X, Y, and Z displacement. Y axis represents the normalized displacement. Dot represents the normalized displacement in a monitoring campaign.

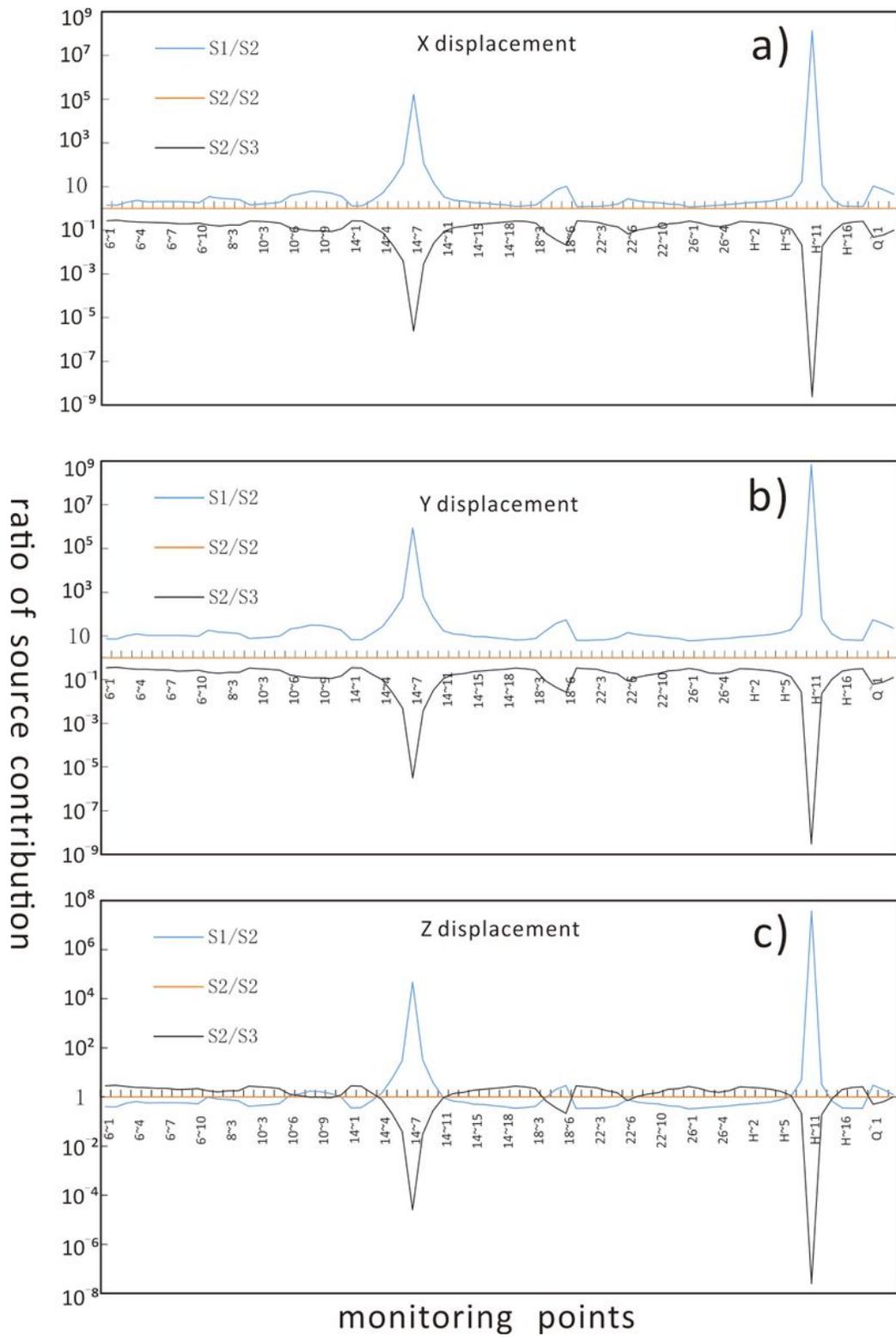


Figure 9

Dynamic of source contribution of S1, S2, and S3 in X, Y, Z displacement, respectively.

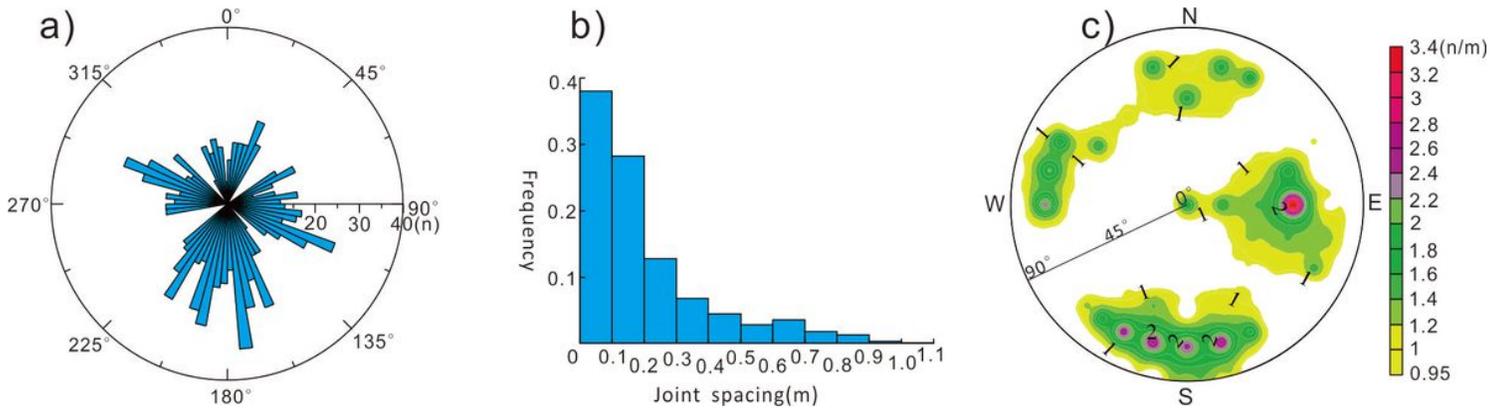


Figure 10

Statistical graphs of joints from $n=1134$ measured data (modified from Bibo et al. 2016). a) rose diagram of dip direction, b) histogram of joint spacing, c) joint density map.

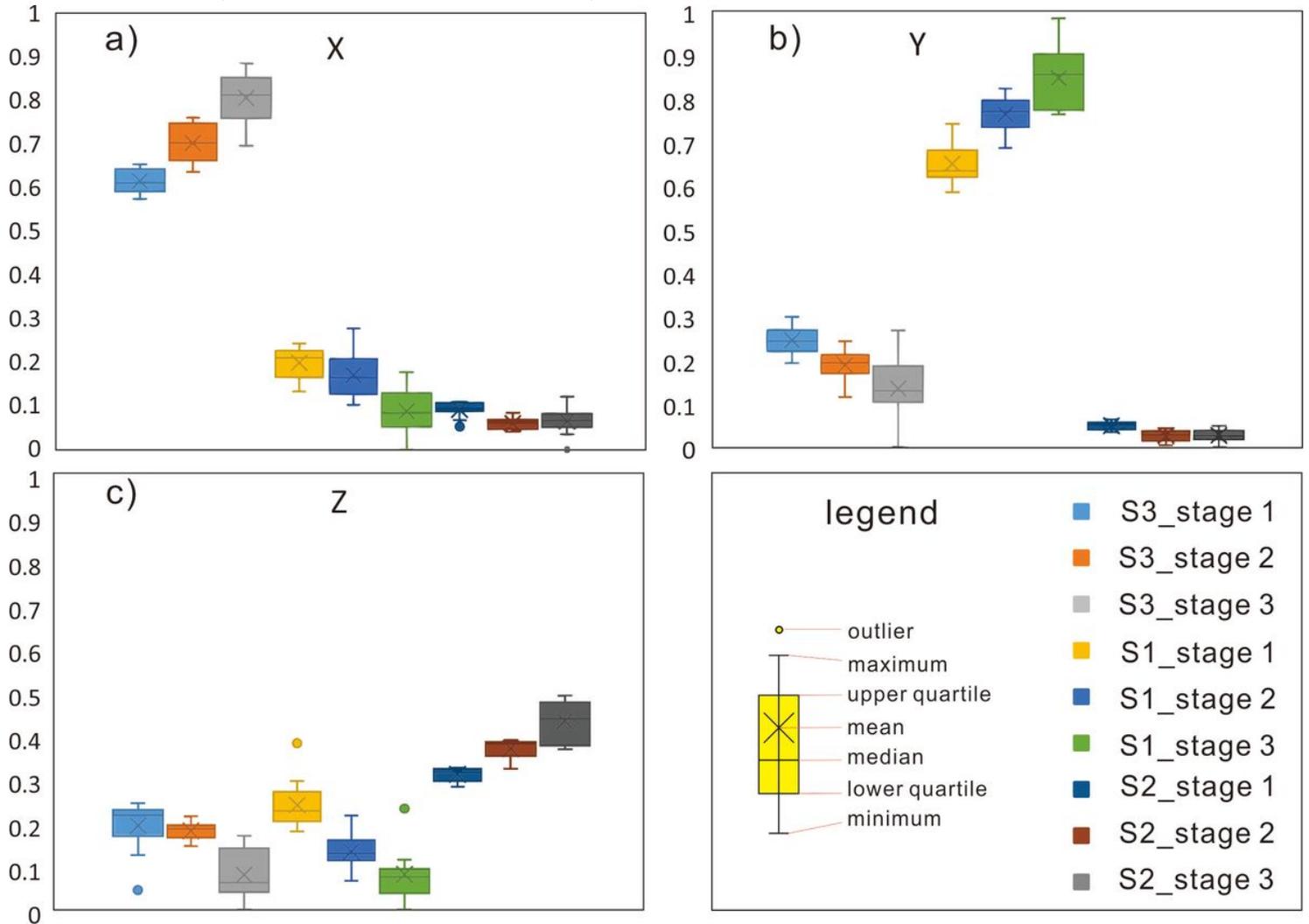


Figure 11

Boxplots of normalized oscillation amplitude of S1, S2, and S3 in three stages.

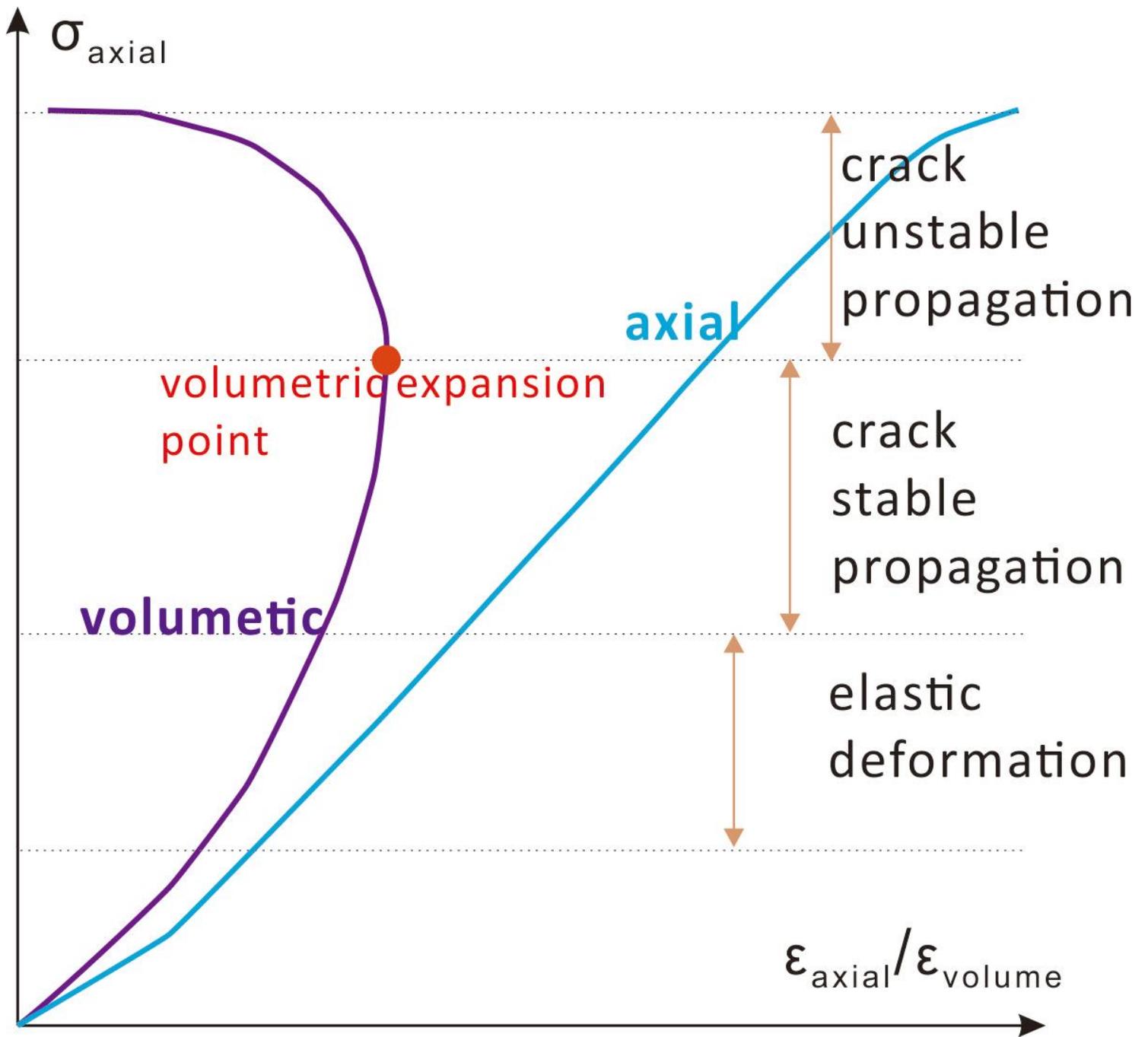


Figure 12

Schematic diagram of uniaxial compression test (axial stress vs ratio of axial strain to volume strain).