



OPEN

Reliability prediction and evaluation of communication base stations in earthquake prone areas

Xueming Li¹, Yao Wei^{2,3}, Zheng Ming¹, Hao Cong¹, Xuanyu Zheng¹ & Qihai Chang¹✉

One of the primary tasks for effective disaster relief after a catastrophic earthquake is robust communication. In this paper, we propose a simple logistic method based on two-parameter sets of geology and building structure for the failure prediction of the base stations in post-earthquake. Using the post-earthquake base station data in Sichuan, China, the prediction results are 96.7% and 90% for the two-parameter sets and all parameter sets, respectively, and 93.3% for the neural network method sets. The results show that the two-parameter method outweighs the whole parameter set logistic method and the neural network prediction and can effectively improve the accuracy of the prediction. The weight parameters of two-parameter set by the actual field data significantly show that the failure of base stations after earthquake is mainly due to the geological differences where the base stations are located. It can be envisioned that if the geological distribution between the earthquake source and the base station is parameterized, the multi-parameter sets logistic method can not only effectively solve the failure prediction after earthquakes and the evaluation of communication base stations under complex conditions, but also provide site selection evaluation for the construction of civil buildings and power grid towers in earthquake-prone areas.

Earthquake disasters can cause collapse of houses, damage to communication base stations towers and transmission lines, resulting in the disruption of communication services over a large area. The interruption of communication services prevents effective information exchange in the affected areas, which seriously hinders rescue and causes huge loss of life and property to local people. How to ensure communication after the earthquake is a key concern for disaster prevention and mitigation. In order to grasp the operation condition of post-earthquake communication base stations, Liu et al.¹ from China Earthquake Administration conducted a study and analysis of typical seismic damage of Wenchuan post-earthquake ground base stations and floor base stations to obtain the operational factors affecting post-earthquake communication base stations, respectively. Bai et al.² used a fault tree model to analyze the causes of service withdrawal of base station due to earthquake damage and obtained the failure probability of post-earthquake communication base stations. Li et al.³ derived the failure probability of post-earthquake communication base station rooms and towers and gave an assessment by conducting seismic vulnerability analysis of ground base station equipment and facilities. Thakur et al.⁴ selected rooftop towers of buildings with different slopes for dynamic analysis and comparison. Masoud et al.⁵ analyzed the seismic damage probability and vulnerability assessment of combined steel–concrete structures using finite element software. Agrawal et al.⁶ combined stochastic factors and real data to simulate earthquake occurrence and proposed a stochastic model to estimate the impact of seismic hazard on backbone optical network. To provide communication services to post-earthquake disaster areas, Peer et al.⁷ proposed a new multi-hop device-to-device (D2D) communication framework that connects devices which are not in the coverage area of the communication network to active base stations through relaying. Ibrah et al.⁸ designed a scheme to build an emergency relay network using multiple UAVs. Yao et al.⁹ designed an emergency wireless communication resource allocation scheme based on 5G-UAV. Li et al.¹⁰ proposed an air-ground integrated “Super-BTS” to form a dual routed transmission of local ring network + satellite transmission.

The above studies mainly analyzed the causes of failures based on the working conditions of post-earthquake communication base stations or propose a new emergency communication scheme. These new solutions are

¹College of Information Engineering, Xizang Minzu University, Xianyang 712082, China. ²College of Mathematics and Physics, Chengdu University of Technology, Chengdu 610059, China. ³Geomathematics Key Laboratory of Sichuan Province, Chengdu University of Technology, Chengdu 610059, China. ✉email: chsh720201@126.com

difficult and costly to implement in practice and are not conducive to post-earthquake disaster relief work. The post-earthquake communication base station condition analysis is limited to the relationship between the tower type of the base station¹¹, building structure¹², etc. and the earthquake. While ignoring that the damage of the post-earthquake communication base station is also related to many factors such as the geographical location of the base station, the distance from the earthquake source, the geography and geology between the earthquake source and the communication base station. Through the analysis of the causal relationship between the post-earthquake communication base station working conditions and these factors, a reasonable model is selected to make the siting of communication base stations in catastrophic earthquake-prone areas reliable, thus effectively reducing the failure rate of post-earthquake communication base stations. At present, there is no reliable method for siting and constructing communication base stations in earthquake-prone areas, and they are implemented according to standards, such as selecting seismic intensity according to geographical areas, there will be problems such as increased investment and failure of communication base stations after earthquakes. In fact, the siting of communication base stations requires a specific analysis process, so that the causes of post-earthquake failure problems can be known, and relevant preparations can be made at the communication base station siting stage.

Based on the real operation data of post-earthquake communication base stations, this paper proposes a logistic method of parameter grouping, which can effectively evaluate the failure probability of post-earthquake communication base stations. And it can be used to judge whether the location of communication base stations is proper, to reduce the probability of failure of post-earthquake communication base stations, so that failure-free base stations can form the continuous coverage of emergency communication networks, and thus reducing the special investment in post-earthquake emergency networks.

Conclusion

In the past, it was often thought that complex nonlinear methods such as artificial intelligence were needed to solve the problem of accurate prediction and evaluation of the working conditions of communication base stations after earthquakes, but in fact, traditional simple methods can improve the accuracy of prediction and evaluation just as well when more factors such as geology and geography and building structure are taken all into account. The advantage of traditional Logistic method compared to the method of innovative technology such as artificial intelligence is that the whole physical process is clear. Through the judgment of weighed-parameters of two-parameter set, it is derived clearly that geological factors have the greatest impact on communication base stations after earthquakes. In the construction of base stations, we can effectively reduce the construction cost of base stations by selecting the location of base stations reasonably rather than seismic intensity level.

Data processing and discussion. *Data sources.* To avoid the accidental source of post-earthquake communication base station data in a single region, it is necessary to select post-earthquake communication base station data from different earthquake magnitudes, different earthquake fault zones, different geological and geographical environments, and other composite factors when selecting experimental data. In this paper, the post-earthquake communication base station statistics of three earthquakes, Wenchuan Ms8.0 earthquake, Lushan Ms7.0 earthquake and Changning Ms6.0 earthquake in Sichuan Province, are used as the object of analysis and research. The geographical contour maps of the seismic regions are shown in Figs. 1, 2 and 3, respectively. The pentagram is the epicenter, and the quadrangular star is the cities and counties where the base station is distributed.

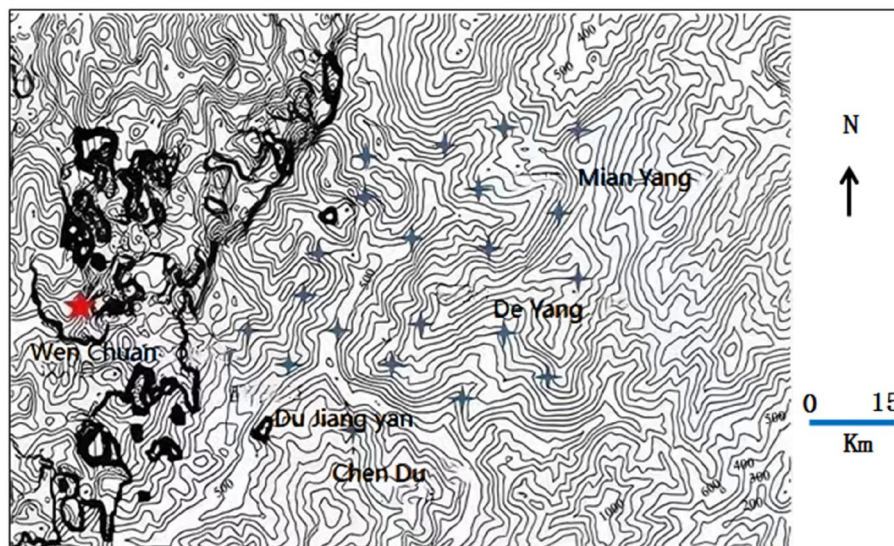


Figure 1. Wenchuan Ms8.0 earthquake.

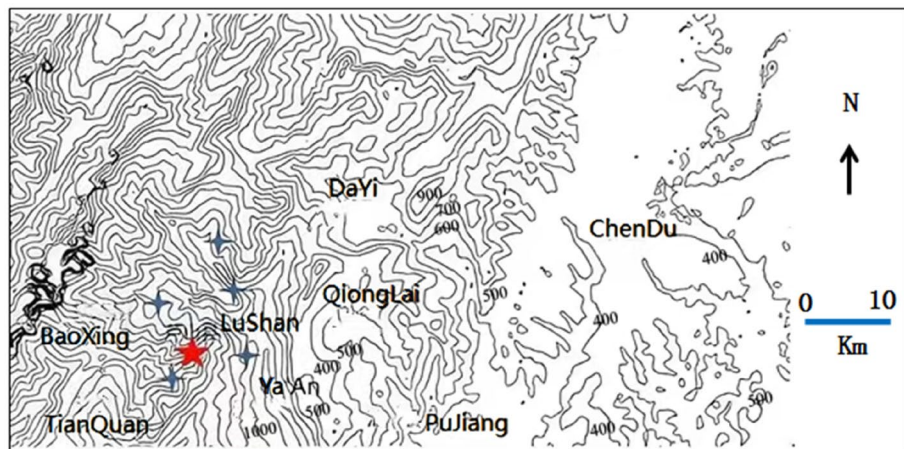


Figure 2. Lushan Ms7.0 earthquake.



Figure 3. Changning Ms6.0 earthquake.

Figure 1 shows that the selected base station data are mainly distributed in Dujiangyan City, Deyang City, Mianyang City, and Wenchuan County. These areas have dense geographical contours, are located in the northeast of Chengdu city, and are distributed along the Longmenshan earthquake fault zone; the fault zone is northeast-southwest trending, the topography is mainly mountainous and hilly, the geological structure is more complex, and a variety of rock formations coexist, and the communication base stations are mostly mountain base stations dominated by ground stations.

Lushan County is a district and county under Ya'an City. As can be seen in Fig. 2, Lushan County is next to Chengdu City in the east and Wenchuan County in the north and is in the southern part of the Longmenshan earthquake fault zone, with the whole territory being a basic intensity 7 degrees or higher intensity zone. The topography is high in the north and low in the south, there are more mountain valleys in the county, and the geological structure is mainly sedimentary rock, and communication base stations are mainly mountain base stations.

Changning is a county of Yibin, located in the southeast of Yibin City and belongs to Huayingshan fault zone. As can be seen from Fig. 3, the terrain is high in the South and low in the north, mainly mountainous and hilly, with complex geological structure and mainly sedimentary rocks.

Parameter selection. Although all three earthquakes occurred in Sichuan Province, China, they were in different areas with different magnitudes and seismic intensities. In the post-earthquake survey, it was found that the communication base stations could maintain basic operation if the main equipment such as power supply system, wireless equipment and transmission equipment were not damaged after the earthquake, and the towers played the role of antenna support and information transmission without changes; if any of them were damaged, the communication base stations could not operate effectively³. Therefore, scenarios can be classified based on

the above description of the post-earthquake base station room condition and tower condition, normal or damaged is selected as the measurement criteria.

The impact of earthquakes on communication base stations can be divided into geological factors underground and environmental factors above ground. The data in this article is from Sichuan Province, China. Sichuan Province is a mountainous area, and the lithology of the area is mainly sedimentary, magmatic and metamorphic rocks, so we use these three rocks as a description of geological structure. The above-ground environmental factors include the geographical location of the communication base station construction and the construction type of the communication base station^{3,4,13,14}. The occurrence of earthquakes is on seismic fault zones, which are directional in nature, and the communication base stations are more severely damaged in the area towards the seismic fault zone, and the seismic intensity is higher in this direction, which requires the selection of the Base station relative epicentral position in the selection of parameters.

According to the above definition of variable parameters that affect the working conditions of the base station, the parameters can be divided into local variable parameters and global variable parameters. Global variables have global scope, which means that all communication base stations in the affected area are affected by the same objective factors at the time of an earthquake, such as magnitude, seismic intensity, geological structure, and the relative orientation of the base station to the epicenter. Local variables only have a local scope, which are affected by the geological and geographical characteristics of the communication base station location, the structure of the base station and other factors, such as magnitude, seismic intensity, geological structure, geographical location of the base station, base station category, computer room situation, tower type, etc. Earthquake magnitude and seismic intensity are the direct factors affecting the working conditions of post-earthquake communication base stations, so the two divisions of local variable parameters and global variable parameters overlap in magnitude and seismic intensity.

Due to the enormous number of influencing factors, the small amount of data, and the significant differences between local factors such as geological changes and seismic grade, a two-parameter set consisting of global and local parameters was modeled, and the results were derived from the weighted sum of the local and global parameters estimated through binary regression. The local parameters and global parameters are shown in Tables 1 and 2, respectively. The factors affecting the base station operation are set as input variables X , and whether the base station operates after the earthquake is set as output variable Y .

Data pre-processing. If collinearity occurs in the quantified data, it will seriously affect the inference and prediction in regression analysis. In this paper, the variance inflation factor VIF is used to detect covariance in the data¹⁵. When $0 < VIF < 10$, there is no collinearity. The VIF value between the actual data is less than 2, there is no collinearity in the data set. The results are shown in Table 3.

Variable	Variable parameters
Magnitude X_1	6.0, 7.0, 8.0
Seismic intensity X_2	6, 7, 8, 9, 10, 11
Geological structure X_3	Sedimentary rock, magmatic rock, metamorphic rock
Base station location X_4	Flat ground, Half-hillside, mountaintop
Base station category X_5	Ground base station, rooftop base station
Base station room condition X_6	Normal, damaged
Tower condition X_7	Normal, damaged

Table 1. Local variable parameters.

Variable	Variable parameters
Magnitude X_1	6.0, 7.0, 8.0
Seismic intensity X_2	6, 7, 8, 9, 10, 11
Geological structure X_3	Sedimentary rock, magmatic rock, metamorphic rock
Base station relative epicentral position X_8	Due east, northeast, due north, northwest, due west, southwest, due south, southeast

Table 2. Global variable parameters.

Variable	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
VIF	1.57	1.59	1.71	1.63	1.82	1.25	1.33	1.25

Table 3. VIF value.

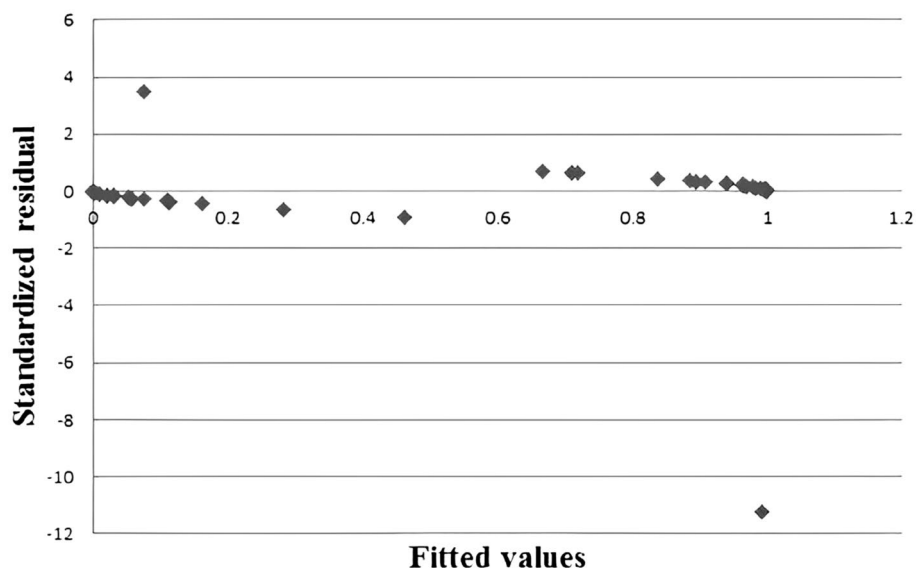


Figure 4. Outlier detection results.

To avoid outliers in the data set from affecting the results of regression analysis, standardized residuals were used for testing¹⁵. There are two sets of experimental data outliers on the 95% confidence interval, and the outlier data were removed. The results are shown in Fig. 4.

The problem of different nature and comparability of data is solved by using data normalization to unify the comprehensive variables to the same standard through homogenization and dimensionless processing of multiple variables parameters in the data set after eliminating outliers. In this paper, 80 sets of experimental data were finally selected and randomly divided into training and test sets in the ratio of 7:3 for experiments.

Experimental results and discussion. In this subsection, the logistic model and the two-parameter set model are used to predict the working condition of the post-earthquake base station respectively. In view of the dynamic changes in the location of the earthquake source and the operating working condition of the communication base station, a neural network model^{16–20} with an autonomous learning function that can find the optimal solution at high speed by continuous feature learning in the face of complex environmental problems is also introduced for comparison. The neural network model in this paper is a three-layer structure, with the first layer as the input layer, taking the variable parameters as the input; the second layer as the hidden layer, responsible for weighting the data; and the third layer as the binary output layer, deciding the base station operating conditions based on the output. The coefficients of the hidden layer of the neural network are adjusted by the backward propagation algorithm to complete the model training.

In the experiment, the default classification threshold of the binary classification model is 0.5, which is not reasonable in specific scenarios. In this paper, the method is applied to practical engineering, and the threshold value is selected according to the maximum prediction correct rate in the training set results. The distribution of the prediction probability of post-earthquake communication base station conditions in the training set is analyzed; the correct prediction rate is the maximum when the threshold is set to 0.4. According to the training set training results, the result is judged as 1 when the output prediction probability value $P > 0.4$ and 0 when $P < 0.4$.

Table 4 shows the correct prediction rate of the logistic model with a threshold value of 0.4 compared with the default threshold value of 0.5, which is used to demonstrate that a threshold value of 0.4 is more suitable for the

Sample	Threshold	Actual	Predicted results		Accuracy rate (%)
			0	1	
Training	0.5	0	19	2	86.00
		1	5	24	
	0.4	0	19	2	88.00
		1	4	25	
Test	0.5	0	13	1	86.70
		1	3	13	
	0.4	0	13	1	90.00
		1	2	14	

Table 4. Threshold selection.

Models	Actual	Predicted results		Accuracy rate (%)
		0	1	
Logistic	0	13	2	90
	1	1	14	
Two-parameter	0	14	0	96.7
	1	1	15	
Neural network	0	12	1	93.3
	1	1	16	

Table 5. Experimental results.

model in this paper than a threshold value of 0.5. The correct prediction rate in the training set is 86.00% when the threshold is 0.5 and 88.00% when the threshold is 0.4. The correct prediction rate in the test set is 86.70% when the threshold is 0.5 and 90.00% when the threshold is 0.4. With reasonable threshold choice, the correct prediction rates of the training and test sets are improved by 2% and 3.3%, respectively. Therefore, in the sense of probability distribution, it is more reasonable for the model to set the threshold value to 0.4. Subsequent experiments on the two-parameter set model and the neural network model are based on the threshold value of 0.4.

Table 5 gives a comparison of the correct prediction rates for three groups of models including the two-parameter set model. Among them, the logistic method and the neural network method use non-group parameter experiments, the two-parameter set model based on the logistic method uses grouping parameter experiments with grouping parameter weights of a and b , where a is the local parameter weight and b is the global parameter weight.

Several factors affecting the post-earthquake communication base station operating conditions are modeled with non-grouping and grouping parameters, respectively. Based on the grouping results, the probability weighted accumulation obtained by conducting evaluation prediction using the two classification parameters obtained from regression modeling is jointly evaluated to predict the post-earthquake base station operating conditions. As shown in Table 5, among the three groups of models, the prediction correctness of the two-parameter model using grouping is perfectly accurate and much higher than that of the non-grouping logistic method and neural network model.

The method of parameter grouping can clearly present the relationship between the grouped parameters. The weight of the global parameter obtained from the field data processing is 2.2 times of the weight of the local parameter, showing that the influence of the geological parameter and the base station relative epicentral position in the global parameter is huge, showing that the choice of the base station construction related to geological factors is important. The non-grouping parameters do not reflect the role played by a specific parameter among the influencing factors, while in the grouping parameters we can indicate which parameter is more important by the weights. By this method it is possible to consider geology as the first factor in the failure prediction of post-earthquake communication base stations or in the choice of future base station construction sites.

In this study, we found that in addition to the traditionally perceived factors such as magnitude and seismic intensity, base station relative epicentral position and geological geography are also factors that affect the post-earthquake communication base station working conditions, among which geological factors are the most important. Unlike the previous methods for predicting the post-earthquake base station working conditions, the two-parameter method of grouping these influencing factors can extract useful information of the influencing factors to a greater extent and improve the correct prediction rate in a complex geographical background environment using a smaller number of base stations.

Methods

The operating condition of a post-earthquake base station can be represented by a binary classification, i.e., 1 is operation and 0 is fault, so the prediction and evaluation of its operating condition is a binary classification problem. Therefore, a binary classification (0 or 1) method is needed to estimate the operating condition of the post-earthquake base station. Logistic regression is a method based on linear regression theory, which can effectively deal with the binary classification (0 or 1) problem with distinct characteristics by introducing nonlinear factors through the sigmoid function. It is often used in data mining, automatic disease diagnosis, economic forecasting and other fields^{21–25}.

The essence of logistic regression is to divide the probability of occurrence by the probability of non-occurrence, and then take logarithm, so that a linear relationship is formed between dependent variables and independent variables. The logistic model is set up as follows:

$$\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_i X_i \quad (1)$$

The above formula can visually present the probability of dependent variable after transformation:

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \cdots + \beta_i X_i)}} \quad (2)$$

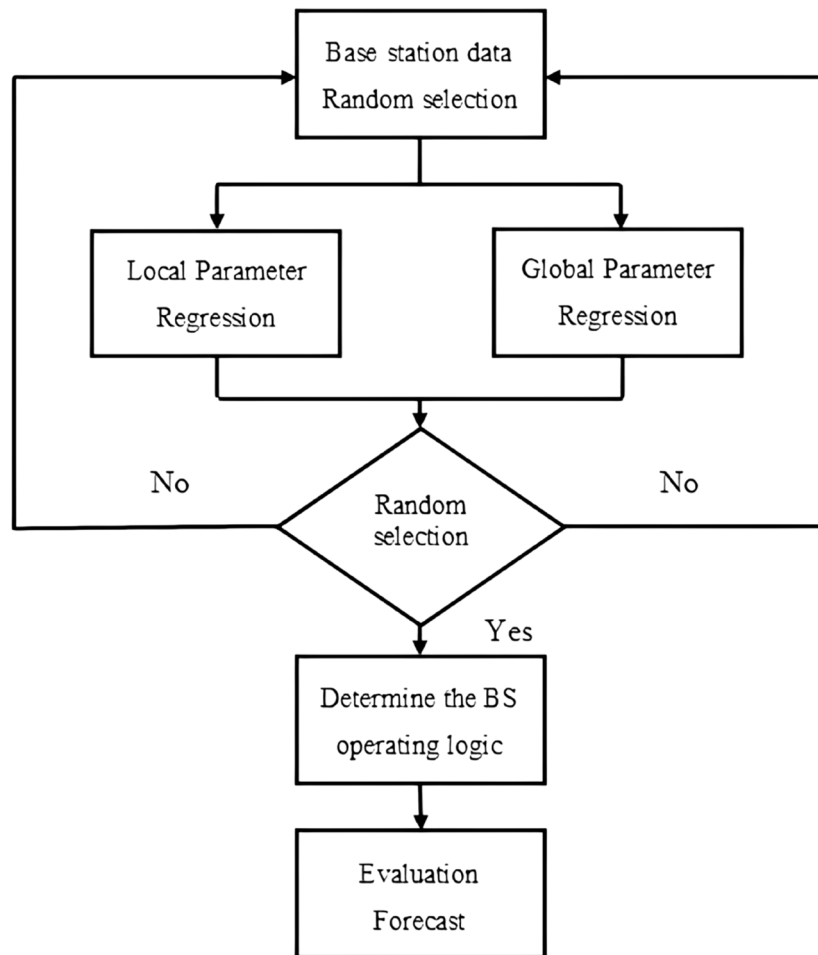


Figure 5. Logic diagram for determining the working conditions of communication base stations.

where P is the probability of post-earthquake communication base stations; β_0 is a constant term, $\beta_1, \beta_2, \dots, \beta_i (i \geq 1, i \in \gamma)$ is an independent variable parameter, $X_1, X_2, \dots, X_i (i \geq 1, i \in \delta)$ is an independent variable.

Due to the limited actual data of post-earthquake communication base stations, the amount of data in this paper has just reached the requirement of logistic regression analysis, i.e., the amount of data is 10 times of the variable parameters. It is unreasonable to process the regression parameters by data from only a single regional earthquake source, and the accuracy of the obtained results may not be high, so it is necessary to analyze and figure out the distribution of regression parameters using random combinations of data from different earthquake sources. To avoid the low prediction accuracy caused by the small amount of data, this study proposes a two-parameter set method based on logistic regression to improve the prediction accuracy, and the processing flow of the post-earthquake base station failure probability with two-parameter set is given in Fig. 5.

Data grouping. The communication base station data from different seismic sources are randomly combined and randomly divided into training set and test set according to the ratio of 7:3. 70% of the training set data are used for learning and 30% of the test set data are used for testing.

Two-parameter regression. Step 1, the first set of base station data is randomly selected in the training set, and the base station data is divided into the set of global variable parameters and local variable parameters. Perform logistic regression tests on these two parameters sets to obtain the prediction probabilities of the global variable parameters and the prediction probabilities of the local variable parameters.

Step 2, the second set of base station data is randomly selected in the training set, repeat the first step, ..., until the logistic parameter regression test is completed for all finite data combinations, and obtain the prediction probabilities of all global variable parameters and local variable parameters in the training set.

Two-parameter weights. The weights are obtained by using the post-earthquake base station conditions of the test data, giving the post-earthquake base station conditions as Z , and the post-earthquake base station conditions of normal and failure as 1 or 0, respectively, with the weights set to a and b , and the prediction prob-

abilities obtained from the local and global parameters of the model are set to x and y , respectively, c as constants. One communication base station corresponds to a binary equation, and the equation for post-earthquake communication base station condition prediction is set up according to the number of actual test data, the weights a and b can be obtained by the principal component analysis method¹⁵, see Eq. (3).

$$Z = ax + by + c \quad (3)$$

The weights a and b were obtained using principal component analysis in three steps:

The first step is to determine the composite score of local parameters. Firstly, the selected local variable parameters are subjected to principal component analysis to obtain the corresponding principal components. Then, the composite score model is calculated from the variable coefficients, component contribution rates and cumulative contribution rates of the principal components, and finally, the selected local parameter data are brought into the above local composite score model to obtain the composite score of the base station data, which is denoted as k .

The second step determines the composite score of global parameters. In the same way as the local parameter weights are decided; the composite score of the base station data is calculated and denoted as g .

The third step is to determine the weight. Using the k derived in the first step and the g derived in the second step, the weights a and b are calculated for the local and global parameters, respectively.

Based on the grouped prediction probabilities from the two-parameter group regression, a weighted accumulation of grouped probability values is performed to obtain the final prediction evaluation probabilities²⁴.

Through the above learning process, the weight parameters of the model can be obtained, and the weight parameters of the model can be used to predict the post-earthquake conditions of 30% of the test set of base station data once they are determined.

In this method, the geological structure, geographic location of the base station, and the category of the base station in the parameter variables are objectively available when evaluating the siting of communication base stations in earthquake-prone areas. Base station relative epicentral position can be determined according to the preset epicenter, the earthquake magnitude and seismic intensity are preset according to human needs.

However, the condition of the base station room and the tower, which are also input parameters, are unknown. The joint solution of seismic magnitude and seismic intensity is used to judge the seismic capability of the communication base station infrastructure. The logic given to the post-earthquake condition of the communication base station infrastructure is H , and its normal and failure are 1 or 0, respectively. The weights of seismic intensity and magnitude are set to m and n , respectively, and seismic intensity and magnitude are set to α and β , respectively, with e as a constant.

A binary linear system of equations is set up, with H_1 as the fortification value and H_2 as the preset value. Calculate the fortification value and the preset value according to Eq. (4), and then compare the two values. If H_1 is greater than H_2 , the input is 1, which means that the base station room and tower are in normal condition after the earthquake; if H_1 is less than H_2 , the input is 0, which means that the base station room and tower fail after the earthquake. By comparing H_1 and H_2 , we can decide the condition of the existing base stations and towers after the earthquake.

$$\begin{cases} H_1 = m\alpha_1 + n\beta_1 + e_1 \\ H_2 = m\alpha_2 + n\beta_2 + e_2 \end{cases} \quad (4)$$

Data availability

The datasets used during the current study available from the corresponding author on reasonable request.

Received: 18 November 2022; Accepted: 24 May 2023

Published online: 02 June 2023

References

- Liu, A. W. *et al.* Typical damage analysis for mobile communication base stations in the extremely damage area of Wenchuan earthquake. *Telecom Eng. Tech. Stand.* **25**, 82–86. <https://doi.org/10.13992/j.cnki.tetas.2012.12.017>(InChinese) (2012).
- Bai, P. F. *et al.* Fault tree modeling of mobile communication infrastructure damage under earthquake disaster. *Math. Pract. Theory.* **47**, 1–7 (2017).
- Li, S. Y. *Seismic Vulnerability and Functional Failure Probability Assessment of Landing Communication Station* (Institute of Engineering Mechanics, China Earthquake Administration, 2017) (In Chinese).
- Thakur, A. *et al.* Influence of rooftop telecommunication tower on set back-step back building resting on different ground slopes. *Earthq. Eng. Eng. Vib.* **18**, 351–362. <https://doi.org/10.1007/s11803-019-0508-7> (2019).
- Masoud, A. *et al.* Seismic failure probability and vulnerability assessment of steel-concrete composite structures. *Periodica Polytech. Civ. Eng.* **61**, 939–950. <https://doi.org/10.3311/PPci.10548> (2017).
- Agrawal, A. *et al.* Network and risk modeling for disaster survivability analysis of backbone optical communication networks. *J. Lightwave Technol.* **37**, 1–1. <https://doi.org/10.1109/JLT.2019.2904328> (2019).
- Peer, M. *et al.* Enabling disaster-resilient communication using multi-hop device-to-device framework. *Wirel. Netw.* **27**(12), 1–13. <https://doi.org/10.1007/s11276-020-02481-2> (2021).
- Ibrah, A. D. *et al.* Optimization method of relay network deployment using multi-UAV for emergency communication. *J. Phys.* **1060**, 012042–012048. <https://doi.org/10.1088/1742-6596/1060/1/012042> (2018).
- Yao, Z. *et al.* Resource allocation for 5G-UAV based emergency wireless communications. *IEEE J. Sel. Areas Commun.* **39**(11), 3395–3410. <https://doi.org/10.1109/jsac.2021.3088684> (2021).
- Li, Y. B. *et al.* Satellite communication application in emergency communication. In *IEEE 2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS)* <https://doi.org/10.1109/ursigass.2014.6929338> (2014) (In Chinese).
- Sag, O. & McClure, G. A novel approach to evaluate the equivalent dynamic stiffness of guy clusters in telecommunication masts under ground excitation. *Eng. Struct.* **33**, 1764–1772. <https://doi.org/10.1016/j.engstruct.2011.02.015> (2011).

12. Kobayashi, M. Experience of infrastructure damage caused by the Great East Japan Earthquake and countermeasures against future disasters. *IEEE Commun. Mag.* **52**, 23–29. <https://doi.org/10.1109/MCOM.2014.6766080> (2014).
13. Liu, J. *et al.* Seismic damage investigation and analysis of the communication system in Lushan M_s7.0 earthquake. *Nat. Disasters* **22**, 91–97. <https://doi.org/10.13577/j.jnd.2013.0513> (2013) (In Chinese).
14. Liu, J. *et al.* Investigation and analysis of seismic damage to communication system in Nepal M_s 8.1 earthquake. *Nat. Disasters* **25**, 146–151. <https://doi.org/10.13577/j.jnd.2016.0217> (2016) (In Chinese).
15. Chatterjee, S. & Hadi, A. S. In *Regression Analysis by Example* 5th edn (ed. Chatterjee, S.) 189–199 (Wiley, 1977).
16. He, H. *et al.* Evaluation modeling of highway collapse hazard based on rough set and support vector machine. *Sci. Rep.* **12**, 18723. <https://doi.org/10.1038/s41598-022-23567-1> (2022).
17. Zhang, Q. W. & Wang, C. Using genetic algorithm to optimize artificial neural network: A case study on earthquake prediction. In *2008 Second International Conference on Genetic and Evolutionary Computing* 128–131. <https://doi.org/10.1109/WGEC.2008.96> (2008) (In Chinese).
18. Yerlikaya-Özkurt, F. & Askan, A. Prediction of potential seismic damage using classification and regression trees: A case study on earthquake damage databases from Turkey. *Nat. Hazards* **103**, 3163–3180. <https://doi.org/10.1007/s11069-020-04125-2> (2020).
19. Lin, J. W. & Chiou, J. S. Active probability backpropagation neural network model for monthly prediction of probabilistic seismic hazard analysis in Taiwan. *IEEE Access* **7**, 108990–109014. <https://doi.org/10.1109/access.2019.2932991> (2019).
20. McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biol.* **52**(1–2), 99–115. [https://doi.org/10.1016/s0092-8240\(05\)80006-0](https://doi.org/10.1016/s0092-8240(05)80006-0) (1990).
21. Luo, X. *et al.* Coupling logistic model tree and random subspace to predict the landslide susceptibility areas with considering the uncertainty of environmental features. *Sci. Rep.* **9**, 15369. <https://doi.org/10.1038/s41598-019-51941-z> (2019).
22. Xu, C. *et al.* Application of an incomplete landslide inventory, logistic regression model and its validation for landslide susceptibility mapping related to the May 12, 2008 Wenchuan earthquake of China. *Nat. Hazards* **68**, 883–900. <https://doi.org/10.1007/s11069-013-0661-7> (2013).
23. Saputra, A. *et al.* Seismic vulnerability assessment of residential buildings using logistic regression and geographic information system (GIS) in Pleret Sub District (Yogyakarta, Indonesia). *Geoenviron. Disasters* **4**, 11–33. <https://doi.org/10.1186/s40677-017-0075-z> (2017).
24. Bao, X. *et al.* Prediction of tunnel earthquake damage based on a combination weighting analysis method. *Symmetry* **14**, 1922. <https://doi.org/10.3390/sym14091922> (2022).
25. Cramer, J. S. The Origins of Logistic Regression. *Tinbergen Institute Discussion Papers*. <https://papers.tinbergen.nl/02119.pdf> (2022)

Author contributions

Q.C. and X.L. devised the study; Y.W. and X.L. performed the experiments; Z.M., H.C. and X.Z. analyzed the data; X.L. wrote the manuscript with contributions from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Q.C.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2023