

Article

A Pre-Seismic Anomaly Detection Approach Based on Earthquake Cross Partial Multi-View Data Fusion

Yongming Huang ^{1,*} , Kun'ao Zhu ¹, Wen Shi ¹, Yong Lu ², Gaochuan Liu ³, Guobao Zhang ¹ and Yuntian Teng ³¹ School of Automation, Southeast University, Nanjing 210096, China² Seismological Bureau of Jiangsu Province, Nanjing 210006, China³ Institute of Geophysics, China Earthquake Administration, Beijing 100081, China

* Correspondence: huang_ym@seu.edu.cn

Abstract: It is a challenge to detect pre-seismic anomalies by using only one dataset due to the complexity of earthquakes. Therefore, it is a promising direction to use multiparametric data. The earthquake cross partial multi-view data fusion approach (EQ-CPM) is proposed in this paper. By using this method, electromagnetic data and seismicity indicators are fused. This approach tolerates the absence of data and complements the missing part in fusion. First, the effectiveness of seismicity indicators and electromagnetic data was validated through two earthquake case studies. Then, four machine learning algorithms were applied to detect pre-seismic anomalies by using the fused data and two original datasets. The results show that the fused data provided better performance than the single-modal data. In the Matthews correlation coefficient index, the results of our method showed an 8% improvement compared with the latest study.

Keywords: electromagnetic data; seismicity indicator; pre-seismic anomaly detection; multi-view data fusion



Citation: Huang, Y.; Zhu, K.; Shi, W.; Lu, Y.; Liu, G.; Zhang, G.; Teng, Y. A Pre-Seismic Anomaly Detection Approach Based on Earthquake Cross Partial Multi-View Data Fusion. *Magnetochemistry* **2023**, *9*, 48. <https://doi.org/10.3390/magnetochemistry9020048>

Academic Editors: Octavio Castillo-Reyes and Guang Li

Received: 21 November 2022

Revised: 26 January 2023

Accepted: 1 February 2023

Published: 3 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Earthquakes are profound natural disasters that threaten economic development and people's lives and property. Pre-seismic anomalies have been found before almost all major earthquakes [1], such as the 1995 Kobe $M_s 7.3$ [2,3] and the 2008 Wenchuan $M_s 8.0$ [4]. Therefore, the research for extracting and detecting pre-seismic anomalies has been a top priority in recent studies.

In the past 20 years, seismologists have conducted ample research on seismic-related data, including electromagnetic data (EM data) [5–7], geoacoustic data [8,9], radon [10–12], stress [13], and seismicity indicators [14–22]. Among them, EM data and seismicity indicators are often employed for pre-seismic anomaly detection.

Seismicity indicators are usually calculated from earthquake catalogs by using Gutenberg Richter's law [16], which describes the relationship between the magnitude and frequency of earthquakes. The parameter b in this law is considered a statistical parameter and is related to the area's physical characteristics [18,19]. Reyes et al. [14] calculated seven seismicity indicators based on the value of b as input into artificial neuronal networks (ANN) for pre-seismic anomaly detection. Asencio-Cortés [20] implemented these seven indicators, but he modified T , τ and c . Fault density (FD) [21] was proposed to describe the density of active faults in the study zone. The authors thought that there is a higher probability of a large earthquake occurring along active fault zones or in their proximity. Salam et al. [22] calculated seven seismicity indicators based on an earthquake catalog for California. They predicted a specific magnitude in the following fifteen days. However, the reliability of the results is yet to be examined, and taking pre-seismic anomaly detection as a binary classification is a choice.

The EM data are considered to be closely related to earthquakes [23–25]. Chen et al. [26] summarized seismo-electromagnetic phenomena in the lithosphere, the lower atmosphere

and the upper atmosphere. Hayakawa [3] found a seismo-ionospheric perturbation before the 1995 Kobe earthquake ($M_s7.2$), where the amplitude or phase of the very-low-frequency (VLF) reached a minimum at sunrise and sunset. It is worth mentioning that the same phenomenon was found in [27]. Hayakawa analyzed the statistical correlation between VLF anomalies and earthquakes. He discovered that earthquakes with a source depth of less than 40 km are usually more correlated with VLF anomalies. Moustra et al. [28] utilized magnitude data and seismic electric signal (SES) to predict earthquakes. Li et al. [24] discovered that the anomalies of earthquakes appeared 20 days before the earthquake by analyzing the ultra-low-frequency (ULF) and VLF bands of electrical data.

Geographic activities are complex, which increases the difficulty of pre-seismic anomaly detection. A survey of available research is listed in Table 1. Some seismicity indicators calculated from earthquake catalogs are treated as inputs of pre-seismic detection models [14,20–22,29]. However, these studies mainly focus on the choice of seismicity indicators, but some small earthquakes may be a mistake due to the lack of sensor accuracy. It is necessary to calculate the minimum completeness magnitude and then ignore earthquakes smaller than the minimum completeness magnitude [30]. In addition, some research on EM data ignores the effects of magnetic storms, which usually yield anomalous fluctuations [31]. Sometimes the EM data during a magnetic storm is excluded, which leads to missing data. Therefore, how to deal with missing data is still a problem. Moreover, the above studies detected pre-seismic anomalies by using one kind of data. It is a promising direction to use multiparameteric data. It can make full use of the complementarity between different related data and improve the accuracy of pre-seismic anomaly detection.

Table 1. Survey of pre-seismic anomaly detection research. \checkmark means covered and \times means not covered.

Research	Seismicity Indicators	EM Data	Consider Data Fusion	Consider Missing Data
[14,20–22,29]	\checkmark	\times	\times	\times
[11,28]	\times	\checkmark	\times	\times
This Paper	\checkmark	\checkmark	\checkmark	\checkmark

A novel approach based on the fusion of seismicity indicators and EM data is proposed in this paper. Considering that there may exist missing data, the earthquake cross partial multi-view data fusion approach (EQ-CPM) is proposed based on a partial multi-view fusion method [23]. This approach tolerates the absence of data and complements the missing part in fusion. Firstly, the effectiveness of EM data and seismicity indicators was analyzed based on two earthquake case studies. Then, EQ-CPM was applied to acquire fusion data from seismicity indicators and EM data. Finally, a comparison of pre-seismic anomaly detection between fused data and two original data was conducted based on four machine learning methods: logistic regression, AdaBoost based on decision trees, stochastic gradient descent (SGD) and support vector machines (SVM). The main contributions of this paper are summarized as follows:

1. The effectiveness of EM data and seismicity indicators was further validated in pre-seismic anomaly detection based on the 2017 Jiuzhaigou $M_s7.0$ earthquake and the 2019 Changning $M_s6.0$ earthquake in China.
2. Propose an earthquake data fusion approach that leverages the temporal characteristic of earthquake data by convolution. The proposed approach tolerates the absence of data and complements the missing part in fusion.

The following content is organized as follows: Section 2 introduces the datasets used in this paper, including EM data and seismicity indicators; Section 3 specifies the data fusion algorithm; Section 4 analyzes EM data, seismicity indicators and the fused data separately and compares the pre-seismic anomaly detection of fused data and two original data; Section 5 presents a summary of this paper.

2. EM Data and Seismicity Indicators

This section introduces the data used to detect pre-seismic anomalies, including EM data and seismicity indicators. The details of the data are presented in each subsection.

2.1. EM Data

The Key Laboratory of Integrated Microsystems (IMS) of Peking University Shenzhen Graduate School deployed the acoustic and electromagnetics on an artificial intelligence (AETA) system. AETA provides the EM data used in our study. The EM data are collected at 0.1 Hz. There are 159 AETA stations in Sichuan and Yunnan (22° N–34° N, 98° E–107° E). Unfortunately, the datasets are incomplete due to weather or equipment failure. Therefore, stations with more than 30% missing data were filtered out. As a result, 74 stations were selected for this study, which is shown in Figure 1.

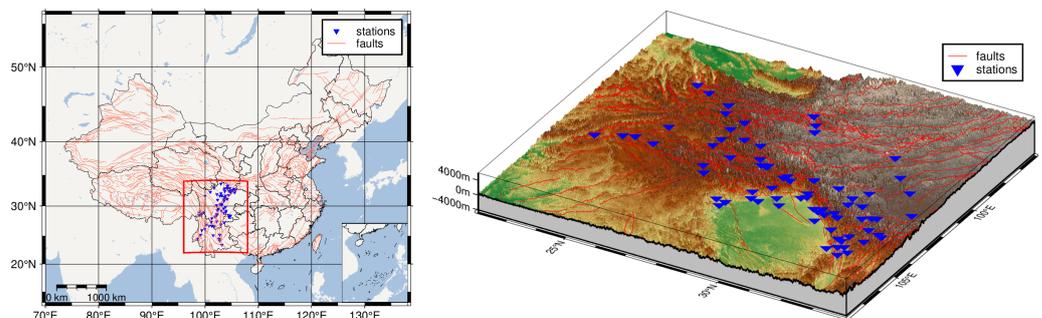


Figure 1. AETA stations' distribution. The blue inverted triangle indicates stations, and the red solid line plots the fault contours in the area.

There are 51 types of EM data features provided by AETA, which are shown in Figure 2. The features are divided into three categories: time-domain-related features, frequency-domain-related features and wavelet-transform-related features.

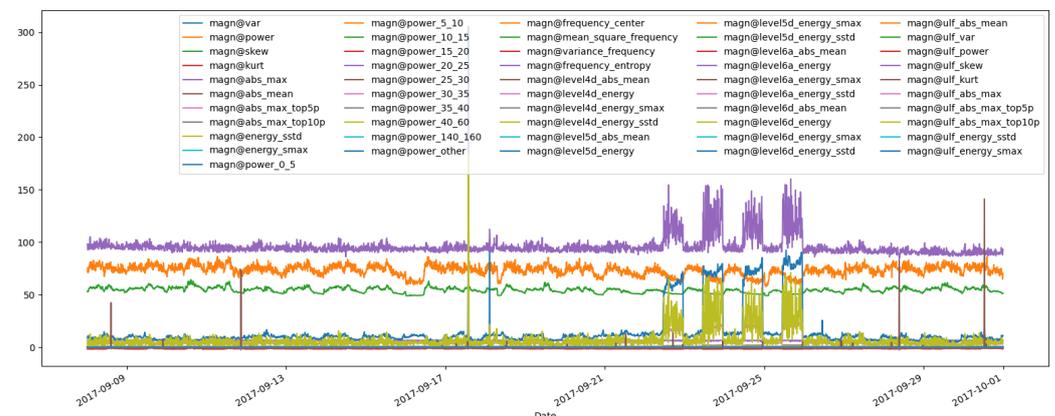


Figure 2. Waveform of EM data from AETA.

The following is some description of these features, and more details can be found on the homepage of AETA.

Some basic statistical features, such as mean value and maximum value, were extracted. The variance is the expectation of the squared deviation of the original signal from its

population mean or sample mean. The power means the energy of the signal, and in most cases, the power of the time series is reckoned in terms of the square of the signal. The skewness indicates the asymmetry of the observed signal. The kurtosis represents how often outliers occur and reflects anomalies of the signal to a certain extent. The energy of the signal of different frequency bands is expressed by `power_rate_atob`. `levelx` means the reconstruction of the x th level while wavelet transforming. The features with `ulf` as a prefix represent relative features in ultra-low frequencies (300 Hz to 3 kHz).

Solar activity can cause disturbances in the magnetosphere of Earth, which are generally known as magnetic storms. Strong storms can cause significant disruptions in EM data, leading to many false alarms. The magnetic storm can be measured through the disturbance storm time (Dst) index. Magnetic storms can be divided into five categories: weak magnetic storms ($Dst \in [-50, -30]$), medium magnetic storms ($Dst \in [-100, -50]$), strong magnetic storms ($Dst \in [-200, -100]$), intense magnetic storms ($Dst \in [-350, -200]$) and giant magnetic storms ($Dst \leq -350$ nT). In our study, the data were excluded when $Dst \leq -30$ nT, even if the data are viewed as anomalies.

2.2. Seismicity Indicators

Earthquake catalogs typically contain the time, latitude and longitude of the epicenter, magnitude and depth of the source. The earthquake catalog is available from the National Earthquake Data Center (NEDC). The Gutenberg–Richter law [16] is used when calculating seismicity indicators from the earthquake catalog, which states the relationship between magnitude and frequency of earthquakes:

$$\log N = a - bM, \quad (1)$$

where a and b is constant and N is the number of earthquakes with a magnitude greater or equal to M . The b value is a vital seismicity indicator. Reyes et al. [14] calculated seven seismicity indicators for pre-seismic anomaly detection based on the b value. Maximum likelihood method is used to estimate the b value [32]:

$$b = \frac{\log e}{M_{\text{mean}} - (M_c - \Delta M/2)}, \quad (2)$$

where M_{mean} represents the average of magnitude, M_c is the completeness magnitude and ΔM is the magnitude accuracy (equals 0.1 in most cases). The calculation is affected by the completeness of the earthquake catalog [30]. Therefore, we have to check the completeness of the earthquake catalog and get the completeness magnitude. Then, the earthquake events smaller than the completeness magnitude should be ignored.

The maximum curvature method (MAXC) was used to calculate the completeness magnitude of the earthquake catalog as $M_{s0.3}$, as shown in Figure 3. The blue box represents the number of earthquakes larger than the current magnitude; the orange triangle is the number of earthquakes equal to the current magnitude. The green curve is the fitted curve. The red pentagram is the maximum curvature point, and the magnitude $M_{s0.3}$ is the completeness magnitude of this earthquake catalog.

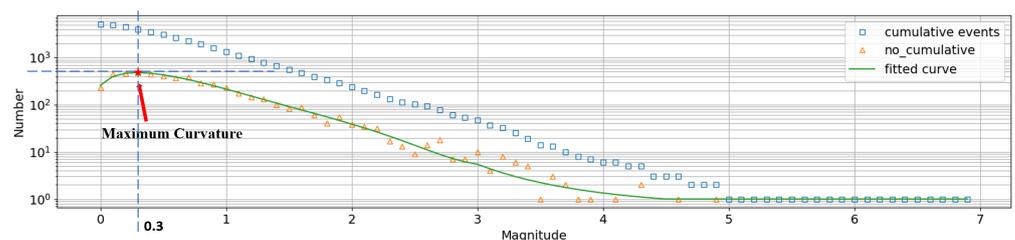


Figure 3. The completeness magnitude of the earthquake catalog is $M_{s0.3}$.

A sliding window was used to generate seismicity indicators for each AETA station, and the length of the window was one day. Seismicity indicators calculated from the earthquake catalog are listed in Table 2.

Table 2. Seismicity indicators [14,20–22,29].

Seismicity Indicator	Expression
b	$\frac{\lg e}{M_{\text{mean}} - (M_c - \Delta M/2)}$
a	$\lg N + b$
T	$t_n - t_1$
M_{mean}	$\frac{1}{n} \sum_{i=1}^n M_i$
$dE^{\frac{1}{2}}$	$\frac{1}{T} \sum_{i=1}^n \sqrt{10^{11.8+1.5M_i}}$
x_6	$\max\{M_i \mid i \in [1, n]\}, \text{ when } t \in [-7, 0]$
x_7	$P(M_s \geq 6.0) = e^{-3b/\log(e)} = 10^{-3b}$
δ_b	$2.3b^2 \sqrt{\frac{\sum_{i=1}^n (M_i - M_{\text{mean}})^2}{n(n-1)}}$
η	$\frac{\sum (\lg N - a - bM)^2}{n-1}$
M_{def}	$M_{\text{max, actual}} - M_{\text{max, expected}} = M_{\text{max, actual}} - \frac{a}{b}$
$T_{\text{recurrence}}$	$\frac{T}{10^{a-bM_0}}, \text{ for } M_0 \text{ in } [4.0, 4.1, 4.2, \dots, 6.0]$

b and a are the constants of the Gutenberg–Richter Law; they imply the activity of seismic events to some extent. T represents the time interval between N earthquakes, and this indicator partially reflects the frequency of earthquakes. M_{mean} is the average magnitude of n seismic events. $dE^{\frac{1}{2}}$ describes the release of seismic energy; seismic energy is continuously released from the ground through small earthquakes; and a seismic quiet period is reached when the energy stops being released. x_6 means the maximum magnitude in the past seven days. x_7 is the probability of an earthquake with magnitude six or greater. δ_b represents the deviation value of b . η indicates the correspondence of the real data distribution and the Gutenberg–Richter law. M_{def} is the difference between the expected maximum magnitude and the observed maximum magnitude. $T_{\text{recurrence}}$ represents the time between two earthquakes with magnitude greater or equal to M_0 .

3. Method

This section introduces the fusion method, including the network design and loss function design, as shown in Figure 4. The network is a generative model and converts latent representations into seismic-related data. Reconstruction and classification loss functions are applied to update the network parameters and the latent representations.

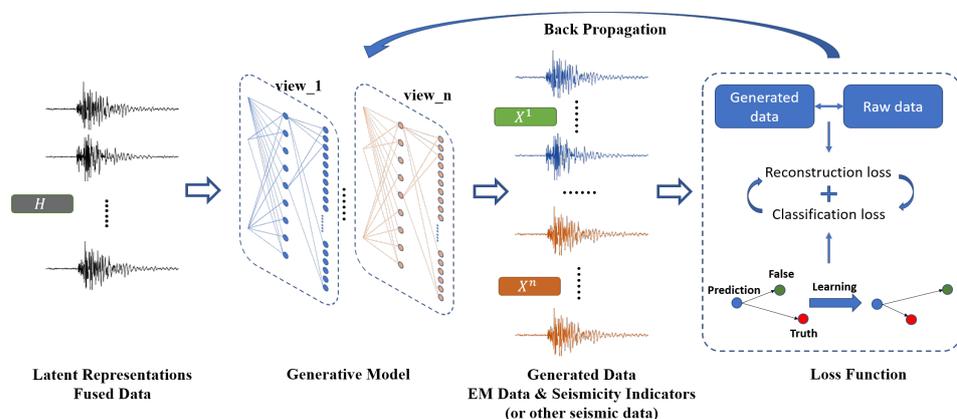


Figure 4. Framework of the fusion method.

3.1. Network

The network is designed as a generative model with two layers: the 1D convolution layer and the densely-connected layer. Each view (dataset) corresponds to a model with the same structure but different parameters, as shown in Figure 5.

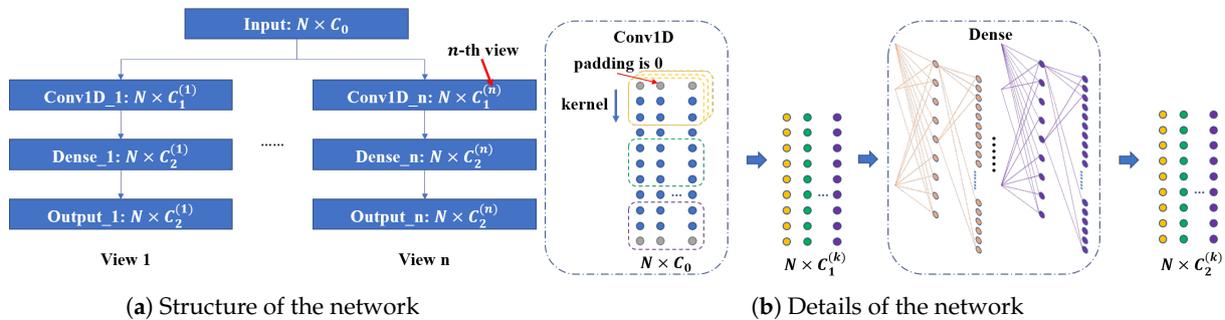


Figure 5. Network design.

The 1D convolution layer can extract deep features of latent representations, which takes advantage of the temporal characteristic of the sequence. The size of the convolution kernel is 3. The features of the adjacent three days are converged when the convolution kernel slides. Zeros are padded evenly into the upstream and downstream parts of the input such that the output has the same length as the input. The convolution layer is followed by a densely-connected layer, whose main roles are feature aggregation and mapping. The output is the generated seismic-related data, including EM data and seismicity indicators. The network parameters and the latent representations are updated by comparing output with true data.

The network can fuse missing data because the data are not used directly but in the process of backpropagation. In other words, the problem of missing is put into the part of the loss function.

3.2. Loss Function

Given a dataset X with multiple views:

$$X = \left\{ x_i^{(v)} \right\}_{v=1}^V, i \in [1, N], \quad (3)$$

where V is the number of views and N is the number of samples. The latent representations of X are expressed as $H = \{h_i\}$, where $i \in [1, N]$ and $h_i \in \mathbb{R}^{D \times 1}$.

Reconstruction and classification loss functions are applied to update the network parameters and the latent representations, as shown in Figure 4. The reconstruction loss function is mainly used to update network parameters. Both loss functions are used to update the latent representations.

Matrix $\Omega \in \mathbb{R}^{N \times V}$ describes the completeness of X , and $\Omega(i, j) = 1$ denotes the j -th view of the i -th sample data that are available or missing. For the convenience of expression, $\Omega^{(v)} \in \mathbb{R}^{N \times 1}$ is used to describe the completeness of the v -th view for all samples. Missing parts of the data are ignored by multiplying $\Omega^{(v)}$ when calculating the reconstruction loss (see Equation (4) for details).

$$L_r(H, \Theta) = \min_{H, \Theta} \sum_{i=1}^N \sum_{v=1}^V \Omega_i^{(v)} \left\| x_i^{(v)} - f_v(h_i; \Theta_v) \right\|_F^2, \quad (4)$$

where $f_v(h_i; \Theta_v)$ is the forward propagation of the network and Θ_v are the parameters of the v -th view.

Classification loss is derived from triplet loss [33], which increases the distance between different classes by clustering.

$$G(h_i, h_j) = \frac{h_i^T \cdot h_j^T}{|h_i^T| \cdot |h_j^T|};$$

$$\hat{y}_i = \underset{y_i}{\operatorname{argmax}} E_{h^* \sim T(y_i)} G(h_i, h^*); \quad (5)$$

$$L_c(\hat{y}_i, y_i, h_i) = \max \left(E_{h^* \sim T(\hat{y}_i)} G(h_i, h^*) - E_{h^* \sim T(y_i)} G(h_i, h^*) + \Delta(\hat{y}_i, y_i), 0 \right),$$

where $G(h_i, h_j)$ is cosine similarity indicating the similarity of h_i and h_j , and the higher the similarity, the greater the value. \hat{y}_i is the prediction of the i -th sample. The expression of $h^* \sim T(y_i)$ represents all the samples that belong to class y_i . A margin distance denoted as $\Delta(\hat{y}_i, y_i)$ in L_c is aimed at avoiding the training process stopping before convergence.

Classification loss is controlled by λ , and the combined loss function is:

$$L = \min_{H, \Theta} \left(\sum_{i=1}^N \sum_{v=1}^V \Omega_i^{(v)} \left\| x_i^{(v)} - f_v(h_i; \Theta_v) \right\|_F^2 + \lambda \frac{1}{N} \sum_{i=1}^N L_c(\hat{y}_i, y_i, h_i) \right) \quad (6)$$

L_r is used to update the parameters of the network, and L is used to update the latent representations. During training, update H and Θ alternately.

4. Results

This section presents the results and analysis of seismicity indicators, EM data and the fused data. Firstly, EM data and seismicity indicators are assessed, which is a prerequisite for fusion. Then, a comparison between the fused data with two original data is presented.

4.1. Assessment of Seismicity Indicators

The earthquake catalog from 2009 to 2017 was selected to analyze the changes in b before the 8th of August 2017, Jiuzhaigou Ms7.0 earthquake in the study area of 32.5° N to 34° N, 103° E to 104.5° E, as shown in Figure 6a.

The epicenter of the Jiuzhaigou Ms7.0 earthquake was located at the intersection of the Kunlun Fault Zone and the Minshan Tectonic Zone. In this area, the most recent earthquake of magnitude seven or greater occurred in 1976, well before the occurrence of Jiuzhaigou Ms7.0. In other words, there was a seismic quiet period between 1976 and 2017. It is worth mentioning that no earthquakes with a magnitude greater than four occurred from 2009 to 2017. In the region's surrounding areas, the Wenchuan Ms8.0 earthquake occurred on 12 May 2008, and the Lushan Ms7.0 earthquake occurred on 20 April 2013. The strong earthquakes in the surrounding areas increased the coulomb stress in this region and promoted the occurrence of Jiuzhaigou Ms7.0. This geotectonic situation can be reflected to some extent by the trend in the b value in Figure 6b.

Changning Ms6.0 occurred in June 2019. Earthquakes that occurred between 2018 and 2019 (in Figure 6c) were selected to study the value of b before Changning Ms6.0. The value of b increased before Changning Ms6.0 and decreased after the earthquake, as shown in Figure 6d. The trend matches the change pattern of b presented in [17]. However, the change trend was subtle and not as obvious as before Jiuzhaigou Ms7.0. The reasons are complex. Unlike the Jiuzhaigou area, there are more earthquakes near Changning area, including earthquakes of magnitude 5 or higher. There were more than ten thousand earthquakes within two years after the 2018 event around Changning. On the contrary, only 5561 earthquakes occurred between 2009 and 2017 around Jiuzhaigou. Namely, the coulomb stress accumulated around Jiuzhaigou, and this contributed in part to the Ms7.0 earthquake. Contrary to this, the coulomb stress was released through continuous earthquakes before Changning Ms6.0, such as Ms5.7 in December 2018 and Ms5.1 in January 2019, and these two earthquakes may have caused the change in b to not be obvious.

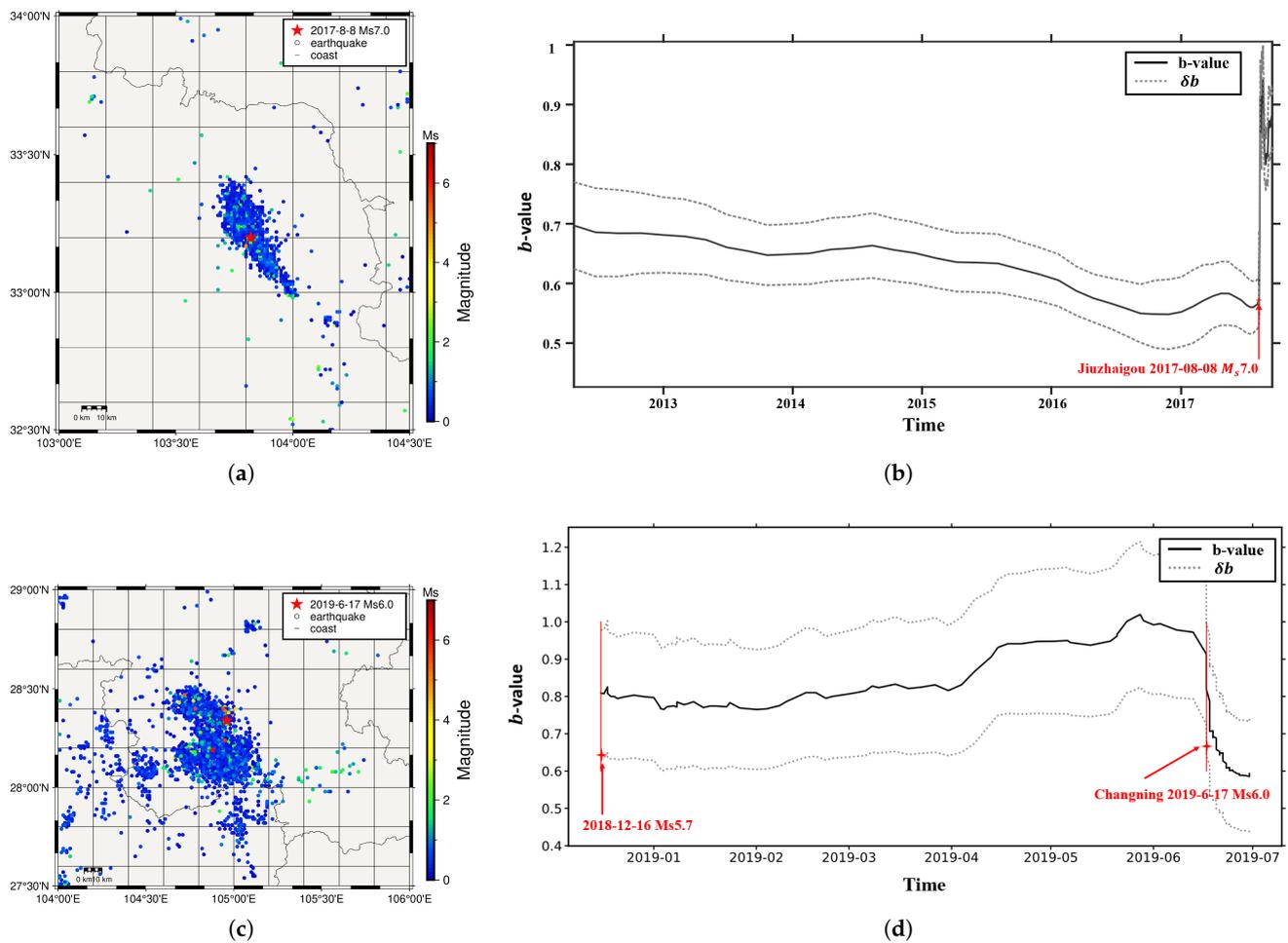


Figure 6. Analysis of b based on Jiuzhaigou $M_s 7.0$ and Changning $M_s 6.0$. (a) The distribution of earthquakes near Jiuzhaigou $M_s 7.0$; the redder the color the larger the magnitude. (b) Trend of b before the Jiuzhaigou $M_s 7.0$ earthquake. (c) The distribution of earthquakes near Changning $M_s 6.0$. (d) Trend of b before the Changning $M_s 6.0$ earthquake.

4.2. Assessment of EM Data

The Dst index shows no apparent magnetic storms before the occurrence of Jiuzhaigou $M_s 7.0$, 2017. We found that anomalies in EM data existed both before and after the earthquake. The types of anomalies varied from station to station, which can be roughly classified as period anomalies, amplitude anomalies and waveform-disturbance anomalies. The different anomalies may be related to the distance between each station and the epicenter. As shown in Figure 7:

1. The periodic disturbance anomalies were found when the distance between the station and the epicenter was closer than 200 km; see Figure 7(ai).
2. The amplitude anomalies appeared when the distance between the station and the epicenter was about 400 km; see Figure 7(aii,aiii).
3. The waveform-disturbance anomalies appeared when the distance between the station and the epicenter was farther than 1000 km; see Figure 7(aiv,av).

To analyze the statistical characteristics of the EM data, firstly, we generated the subseries of the EM data for each station by sliding windows. Then, the fractal dimension of each subseries was calculated using the Higuchi fractal algorithm [34]. Finally, the fractal dimensions were analyzed by using the interquartile range (IQR) method. We took Changning $M_s 6.0$ as the earthquake case.

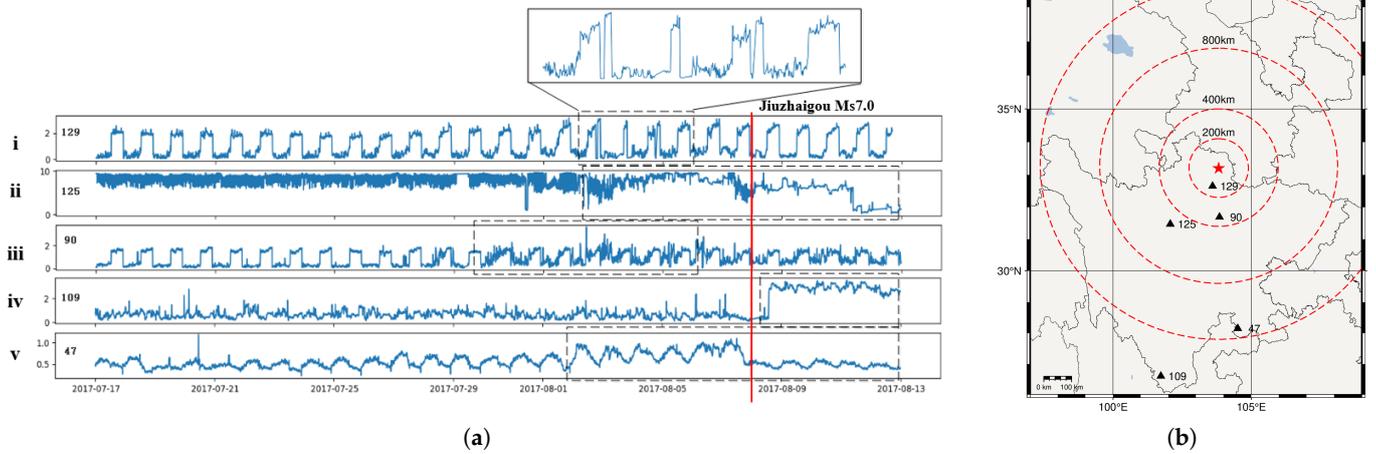


Figure 7. EM data anomalies based on Jiuzhaigou $M_s7.0$. (a) EM data waveforms from different stations. (b) Distribution of stations in relation to the Jiuzhaigou $M_s7.0$ earthquake.

Stations 48 and 77 were selected for analysis because their epicenter distances were lower than 100 km. The fractal dimension is shown in Figure 8a. It can be seen that before the Changning $M_s6.0$ earthquake, simultaneous fractal dimension anomalies occurred at the two stations on 20 April 2019 and 15 May 2019. It is worth mentioning that almost the values are negative, which means the fractal dimension decreased before the earthquake [35]. In addition, Station 48 had more anomalies from June 17 to early July, and one of possible reasons is that it is closer to the epicenter, as shown in Figure 8b. In contrast, Station 77 had only one anomaly of a smaller magnitude from June 17 to early July. A clear hypothesis is that the closer the epicenter, the more the EM data can reflect the associated anomalies.

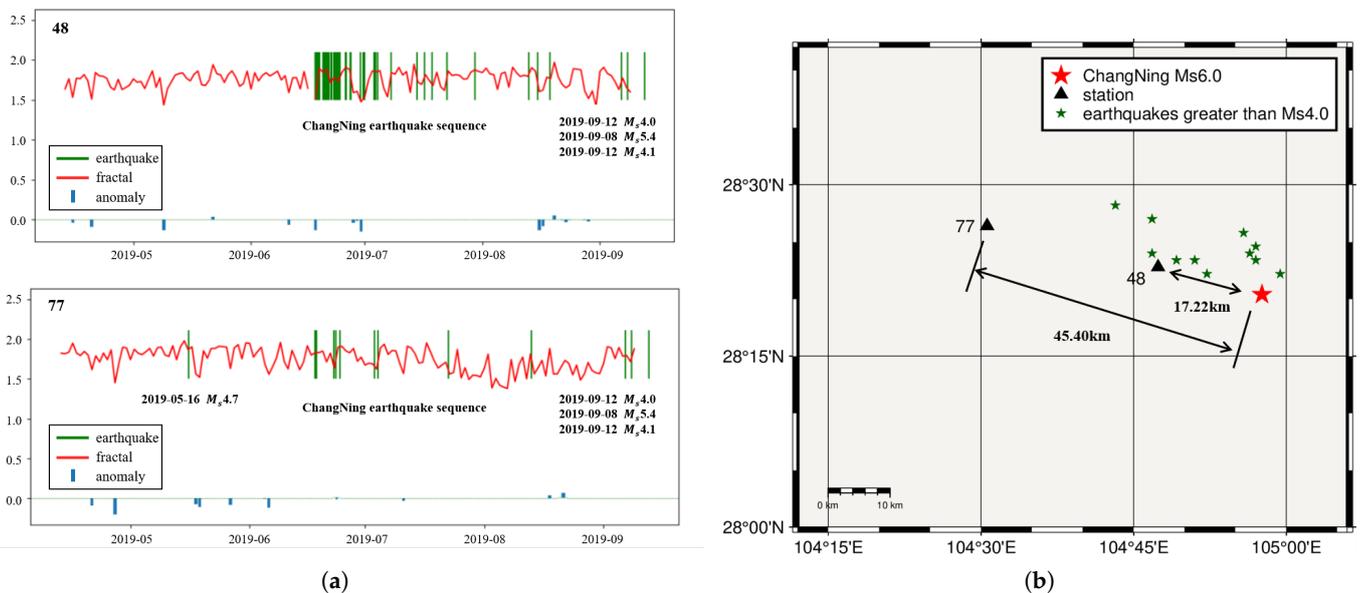


Figure 8. Fractal dimension anomalies of EM data based on Changning $M_s6.0$. (a) Fractal dimension of EM data. The green vertical lines are earthquakes, the red line is the fractal dimension of EM data and the blue vertical line represents anomalies detected by the IQR method. (b) Distribution of stations and the Changning $M_s6.0$ earthquake.

4.3. Comparison between Fused Data and Single-Modal Data

The ultimate goal of data fusion is to improve the accuracy of detection tasks. In this paper, pre-seismic anomaly detection is viewed as a binary classification problem, where positive samples are pre-seismic anomalies and negative samples are background data. The fused data are high-dimensional data and challenging to visualize. The t-SNE method [36] reduced the data to two dimensions for data visualization.

After dimensionality reduction, the data are only a projection of a two-dimensional plane, without any physical meaning. However, we can visualize the two-dimensional data to analyze the distribution of positive samples and negative samples. When there is an overlap of positive and negative samples, the negative samples will cover the positive samples. The more red points (positive samples), the easier it is to distinguish between positive and negative samples.

It can be seen in Figure 9 that it is easier to detect positive samples from fused data, because there are more red points. As a pre-processing step for pre-seismic anomaly detection, the data fusion is meaningful. The overlap of positive and negative samples may lead to important information being hidden. In contrast, the distribution of fused data is more dispersed, probably because in the process of reconstruction, the loss function L_c increases the distance of different classes and influences all samples meanwhile.

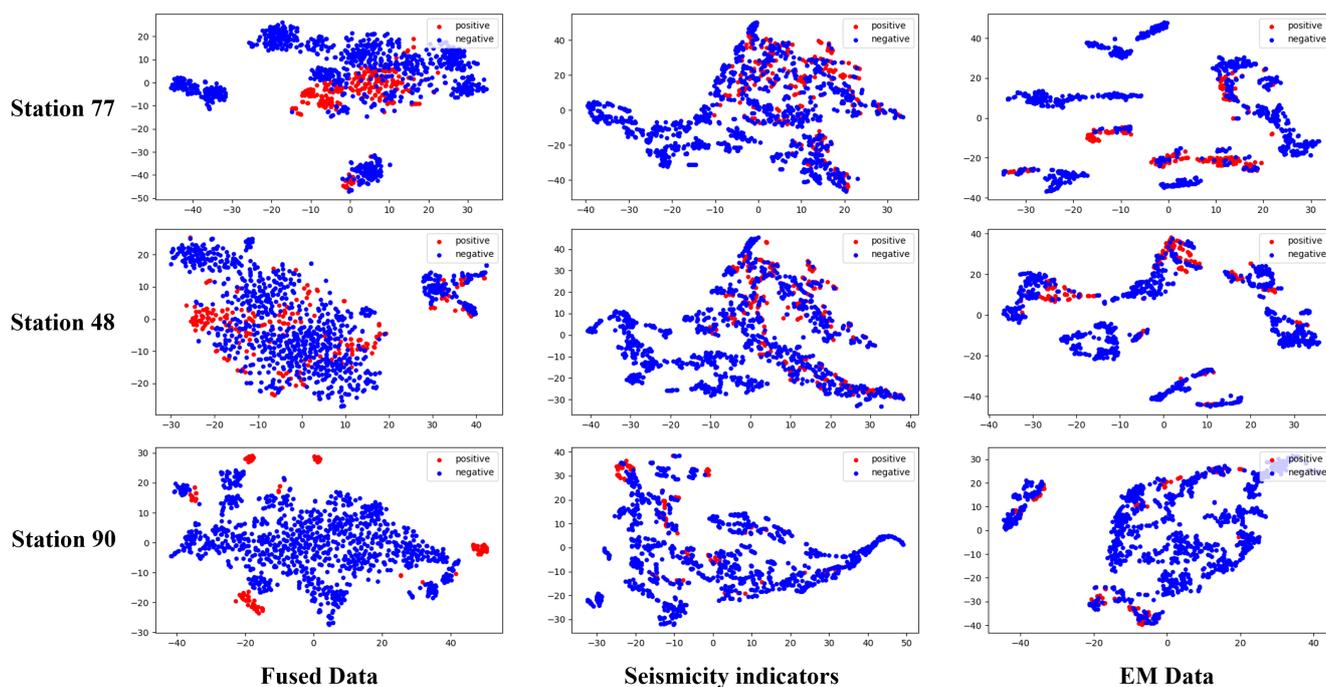


Figure 9. Visualization of Stations 77, 48 and 90. The blue points are background data and the red points are anomalies. After the dimensionality reduction, the horizontal and vertical coordinates only represent the scale.

Four machine learning methods were applied to compare the fused data with EM data and seismicity indicators. The dataset was split into a training set and a test set in the ratio of 8:2. The SMOTE method was used to alleviate the data imbalance problem. Five metrics were used to measure the detection result: PPV (positive predictive value), NPV (negative predictive value), recall, TNR (true negative rate), MCC (Matthews correlation coefficient) and AVG, where AVG is the average of PPV, NPV, recall and TNR.

Table 3 shows the detection performance of the fused data, EM data and seismicity indicators of station 47. As Table 3 depicts, the values of six metrics of fused data are large except *recall* based on decision tree method. The good performance mainly resulted from increasing the distances of different classes through L_c . The bad result of the decision

tree method was probably caused by the imbalance of the dataset. The proposed method, EQ-CPM, complements the missing part of data through L_r , and fully utilizes the complementary information between different views (different dataset). Therefore, in two overall metrics, *AVG* and *MCC*, the value of the fused far exceeds those of the two original data.

Table 3. Comparison of detection results among fused data, EM data and seismicity indicators.

Method	Dataset	PPV	NPV	Recall	TNR	AVG	MCC
Logistic Regression	EM data	0.4000	0.9179	0.5926	0.8367	0.6868	0.3694
	seismicity indicators	0.1522	0.7706	0.4565	0.4179	0.4493	−0.0984
	fused data	1.0000	0.9213	0.6222	1.0000	0.8859	0.7571
Decision Tree	EM data	0.0483	0.3103	0.2593	0.0612	0.1698	−0.6602
	seismicity indicators	0.1959	0.8200	0.4130	0.6119	0.5102	0.0199
	fused data	0.3636	0.8550	0.3556	0.8593	0.6084	0.2167
SGD	EM data	0.1145	0.7209	0.5556	0.2109	0.4005	−0.1961
	seismicity indicators	0.1250	0.7748	0.2609	0.5821	0.4357	−0.1254
	fused data	1.0000	0.9387	0.7111	1.0000	0.9124	0.8170
SVM	EM data	0.5200	0.9060	0.4815	0.9184	0.7065	0.4127
	seismicity indicators	0.1364	0.7862	0.2609	0.6219	0.4513	−0.0953
	fused data	1.0000	0.9522	0.7778	1.0000	0.9325	0.8606

There were 74 stations selected for the study, and the data fusion results of these 74 stations are given in Table 4. For the imbalance dataset, the recall metric is more important, and a lower recall value in pre-seismic anomaly detection means a lower leakage rate and a better detection effect. With the recall rate as the main index, the effect of fusion is considered better when the recall rate of the fused data is higher than that of the EM data and the seismicity indicators. All four methods had better fusion effects for more than 50% of the stations, indicating that the method proposed in this paper has better generalization for seismic data.

Table 4. The detection results for all stations. Effective stations represent stations whose recall of fused data is higher than that of EM data and seismicity indicators.

Method	Effective Stations ¹	Total Number of Stations	Ratio
Logistic Regression	65	74	87.84%
Decision Tree	51	74	68.92%
SGD	57	74	77.03%
SVM	73	74	98.65%

¹ The number of effective stations.

The results in this paper are compared with those of existing studies, as shown in Table 5. The fused data are better than EM data in all indexes but inferior to the seismicity indicators in some indexes, such as NPV and recall. A possible reason is that earthquake catalogs are different in different regions. However, the performance of the fused data in both *AVG* and *MCC* is much better than that of the EM data and seismicity indicators. That is to say, the fused data has a better overall performance. Since the distances of some classes are larger, the false positives and false negatives, which are related to the metrics PPV, NPV, recall and TNR, can be significantly reduced.

Table 5. Comparing our fusion dataset with the datasets from other research.

	Used Dataset	Year	PPV	NPV	Recall	TNR	AVG	MCC
ANN [23]	Seismic Electric Signals	2011	0.61	-	-	-	-	-
ANN [14]	seismicity indicators	2013	0.63	0.81	0.48	0.88	0.70	0.58
EQP-ANN [20]	seismicity indicators	2017	1.00	0.69	0.18	1.00	0.72	0.35
DNN [21]	seismicity indicators	2021	0.88	0.97	-	0.96	-	-
K-means [37]	seismicity indicators	2021	0.90	0.85	0.99	0.07	0.70	0.22
AMA [29]	seismicity indicators	2022	0.73	0.96	0.98	0.82	0.87	0.78
LSTM [8]	EM data	2022	0.64	0.74	0.81	0.55	0.69	-
This paper(SVM)	Fusion data	2023	1.00	0.95	0.78	1.00	0.93	0.86

5. Conclusions

In this paper, we leveraged the advantage of multimodal data for pre-seismic anomaly detection. EM data and seismicity indicators were considered. First, the effectiveness of EM data and seismicity indicators with related earthquake cases was investigated. The anomalies were further confirmed by focusing on the b-value change curves before the occurrence of the Jiuzhaigou M_s 7.0, 2017 earthquake. At the same time, different EM data anomalies were discovered from different AETA stations before the Jiuzhaigou M_s 7.0, 2017 earthquake. Multiple anomalies in the fractal dimension of the EM data were found in relation to the Changning M_s 6.0, 2019 earthquake.

There are few relevant studies that have detected pre-seismic anomalies using multi-parameteric data. We proposed a method called EQ-CPM to fuse EM data and seismicity indicators. The proposed method tolerates the absence of data and complements the missing data by fusion. Therefore, we can improve the use of a dataset with missing parts via EQ-CPM. Meanwhile, the proposed method brings together complementary information between different datasets during the reconstruction and leverages the advantage of multimodality. The fused data surpassed the original data in performance when used by four classical machine learning algorithms. Most importantly, more relevant datasets can be fused by EQ-CPM if possible.

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

Funding: This research was financially supported by the the Jiangsu Provincial Key R&D Programme (BE2020116, BE2022154).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The earthquake catalog data can be found here: https://data.earthquake.cn/datashare/report.shtml?PAGEID=earthquake_zhengshi (accessed on 20 January 2022). The Dst index can be found here: <https://wdc.kugi.kyoto-u.ac.jp/dstae/index.html> (accessed on 20 January 2022). The EM data presented in this study are available from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Akhoondzadeh, M.; Parrot, M.; Saradjian, M.R. Electron and ion density variations before strong earthquakes ($M > 6.0$) using DEMETER and GPS data. *Nat. Hazards Earth Syst. Sci.* **2010**, *10*, 7–18. [CrossRef]
2. Tsunogai, U.; Wakita, H. Precursory Chemical Changes in Ground Water: Kobe Earthquake, Japan. *Science* **1995**, *269*, 61–63. [CrossRef] [PubMed]
3. Hayakawa, M. Earthquake prediction with electromagnetic phenomena. *AIP Conf. Proc.* **2016**, *1709*, 020002.

4. Li, M.; Lu, J.; Parrot, M.; Tan, H.; Chang, Y.; Zhang, X.; Wang, Y. Review of unprecedented ULF electromagnetic anomalous emissions possibly related to the Wenchuan M_S = 8.0 earthquake, on 12 May 2008. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 279–286. [[CrossRef](#)]
5. Hayakawa, M. Seismo Electromagnetics and Earthquake Prediction: History and New directions. *Int. J. Electron. Appl. Res.* **2019**, *6*, 1–23. [[CrossRef](#)]
6. Hayakawa, M. Earthquake precursor studies in Japan. In *Pre-Earthquake Processes: A Multidisciplinary Approach to Earthquake Prediction Studies*; Advancing Earth and Space Science: Washington, DC, USA, 2018; pp. 7–18.
7. Schekotov, A.; Izutsu, J.; Asano, T.; Potirakis, S.M.; Hayakawa, M. Electromagnetic Precursors to the 2016 Kumamoto Earthquakes. *Open J. Earthq. Res.* **2017**, *6*, 168–179. [[CrossRef](#)]
8. Wang, C.; Li, C.; Yong, S.; Wang, X.; Yang, C. Time Series and Non-Time Series Models of Earthquake Prediction Based on AETA Data: 16-Week Real Case Study. *Appl. Sci.* **2022**, *12*, 8536. [[CrossRef](#)]
9. Rouet-Leduc, B.; Hulbert, C.; Lubbers, N.; Barros, K.; Humphreys, C.J.; Johnson, P.A. Machine Learning Predicts Laboratory Earthquakes. *Geophys. Res. Lett.* **2017**, *44*, 9276–9282. [[CrossRef](#)]
10. Barkat, A.; Ali, A.; Siddique, N.; Alam, A.; Wasim, M.; Iqbal, T. Radon as an earthquake precursor in and around northern Pakistan: A case study. *Geochem. J.* **2017**, *51*, 337–346. [[CrossRef](#)]
11. Barkat, A.; Ali, A.; Hayat, U.; Crowley, Q.G.; Rehman, K.; Siddique, N.; Haidar, T.; Iqbal, T. Time series analysis of soil radon in Northern Pakistan: Implications for earthquake forecasting. *Appl. Geochem.* **2018**, *97*, 197–208. [[CrossRef](#)]
12. Alam, A.; Wang, N.; Zhao, G.; Barkat, A. Implication of Radon Monitoring for Earthquake Surveillance Using Statistical Techniques: A Case Study of Wenchuan Earthquake. *Geofluids* **2020**, *2020*, 1–14. [[CrossRef](#)]
13. Ren, Y.; Ma, J.; Liu, P.; Chen, S. Experimental Study of Thermal Field Evolution in the Short-Impending Stage Before Earthquakes. *Pure Appl. Geophys.* **2017**, *175*, 2527–2539. [[CrossRef](#)]
14. Reyes, J.; Morales-Esteban, A.; Martínez-Álvarez, F. Neural networks to predict earthquakes in Chile. *Appl. Soft Comput.* **2013**, *13*, 1314–1328. [[CrossRef](#)]
15. Panakkat, A.; Adeli, H. Neural Network Models for Earthquake Magnitude Prediction Using Multiple Seismicity Indicators. *Int. J. Neural Syst.* **2007**, *17*, 13–33. [[CrossRef](#)] [[PubMed](#)]
16. Gutenberg, B.; Richter, C. *Seismicity of the Earth*; Geological Society of America: Boulder, CO, USA, 1941.
17. Wang, J.H. A mechanism causing b-value anomalies prior to a mainshock. *Bull. Seismol. Soc. Am.* **2016**, *106*, 1663–1671. [[CrossRef](#)]
18. Lee, K.; Yang, W.S. Historical seismicity of Korea. *Bull. Seismol. Soc. Am.* **2006**, *96*, 846–855. [[CrossRef](#)]
19. Marzocchi, W.; Spassiani, I.; Stallone, A.; Taroni, M. How to be fooled searching for significant variations of the b-value. *Geophys. J. Int.* **2020**, *220*, 1845–1856.
20. Asencio-Cortés, G.; Martínez-Álvarez, F.; Troncoso, A.; Morales-Esteban, A. Medium–large earthquake magnitude prediction in Tokyo with artificial neural networks. *Neural Comput. Appl.* **2015**, *28*, 1043–1055. [[CrossRef](#)]
21. Yousefzadeh, M.; Hosseini, S.A.; Farnaghi, M. Spatiotemporally explicit earthquake prediction using deep neural network. *Soil Dyn. Earthq. Eng.* **2021**, *144*, 106663. [[CrossRef](#)]
22. Salam, M.A.; Ibrahim, L.; Abdelminaam, D.S. Earthquake Prediction using Hybrid Machine Learning Techniques. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 654–665. [[CrossRef](#)]
23. Zhang, C.; Han, Z.; Fu, H.; Zhou, J.T.; Hu, Q. CPM-Nets: Cross Partial Multi-View Networks. In *Advances in Neural Information Processing Systems*; Neural Information Processing Systems Foundation, Inc. (NeurIPS): La Jolla, CA, USA, 2019; p. 32.
24. Li, Z.; Yang, B.; Huang, J.; Yin, H.; Yang, X.; Liu, H.; Zhang, F.; Lu, H. Analysis of Pre-Earthquake Space Electric Field Disturbance Observed by CSES. *Atmosphere* **2022**, *13*, 934. [[CrossRef](#)]
25. Zhao, G.; Bi, Y.; Wang, L.; Han, B.; Wang, X.; Xiao, Q.; Cai, J.; Zhan, Y.; Chen, X.; Tang, J.; et al. Advances in alternating electromagnetic field data processing for earthquake monitoring in China. *Sci. China Earth Sci.* **2015**, *58*, 172–182. [[CrossRef](#)]
26. Chen, H.; Han, P.; Hattori, K. Recent Advances and Challenges in the Seismo-Electromagnetic Study: A Brief Review. *Remote Sens.* **2022**, *14*, 5893. [[CrossRef](#)]
27. Chakrabarti, S.K. *Propagation Effects of Very Low Frequency RadioWaves*; American Institute of Physics: New York, NY, USA, 2010.
28. Moustra, M.; Avraamides, M.; Christodoulou, C. Artificial neural networks for earthquake prediction using time series magnitude data or Seismic Electric Signals. *Expert Syst. Appl.* **2011**, *38*, 15032–15039. [[CrossRef](#)]
29. Zhou, W.; Liang, Y.; Wang, X.; Ming, Z.; Xiao, Z.; Fan, X. Introducing macrophages to artificial immune systems for earthquake prediction. *Appl. Soft Comput.* **2022**, *122*, 108822. [[CrossRef](#)]
30. Wiemer, S.; Wyss, M. Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. *Bull. Seismol. Soc. Am.* **2000**, *90*, 859–869. [[CrossRef](#)]
31. Parrot, M.; Buzzi, A.; Santolik, O.; Berthelier, J.J.; Sauvaud, J.A.; Lebreton, J.P. New observations of electromagnetic harmonic ELF emissions in the ionosphere by the DEMETER satellite during large magnetic storms. *J. Geophys. Res. Atmos.* **2006**, *111*. <https://doi.org/10.1029/2005JA011583>. [[CrossRef](#)]
32. Utsu, T. A method for determining the value of “b” in a formula $\log n = a - bM$ showing the magnitude-frequency relation for earthquakes. *Geophys. Bull. Hokkaido Univ.* **1965**, *13*, 99–103.
33. Schroff, F.; Kalenichenko, D.; Philbin, J. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 7–12 June 2015; pp. 815–823. [[CrossRef](#)]

34. Gotoh, K.; Hayakawa, M.; Smirnova, N.A.; Hattori, K. Fractal analysis of seismogenic ULF emissions. *Phys. Chem. Earth Parts A/B/C* **2004**, *29*, 419–424. [[CrossRef](#)]
35. Han, P.; Hattori, K.; Huang, Q.; Hirooka, S.; Yoshino, C. Spatiotemporal characteristics of the geomagnetic diurnal variation anomalies prior to the 2011 Tohoku earthquake (Mw 9.0) and the possible coupling of multiple pre-earthquake phenomena. *J. Asian Earth Sci.* **2016**, *129*, 13–21. [[CrossRef](#)]
36. Van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
37. Yuan, R. An improved K-means clustering algorithm for global earthquake catalogs and earthquake magnitude prediction. *J. Seism.* **2021**, *25*, 1005–1020. [[CrossRef](#)]

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