EARTHQUAKE PREDICTION MODEL BASED ON DANGER THEORY IN ARTIFICIAL IMMUNITY

W. Zhou, Y.W. Liang*, Z. Ming*, H.B. Dong†*

Abstract: Earthquake prediction is an extraordinarily stochastic process. Determining the occurrence time, location of epicenter and magnitude of a coming earthquake in the following month is an extremely difficult task. Nowadays, some geophysical, statistical and machine learning methods are adopted to predict earthquakes, however, for the insufficient medium-large seismic data, their results are not satisfactory. Due to there is no obvious empirical relationship between seismicity features, magnitude and location of a coming earthquake in a particular time window, an earthquake prediction approach based on danger theory is proposed in this paper. It extracts eight indicators calculated from earthquake data for recent years in Sichuan and surroundings by Gutenberg-Richter(GR) inverse power-law, and predicts quakes with magnitude lager than 4.5 during the following month by numerical differential based Dendritic Cell Algorithm (ndDCA). We compare this approach with six state-of-art earthquake prediction algorithms. Overall our algorithm yields the encouraging results in all the qualified parameters assessed, and it provides technical support for the application of earthquake prediction.

Key words: *danger theory, earthquake prediction, Gutenberg-Richter inverse power-law, ndDCA*

Received: July 4, 2019 **DOI:** 10.14311/NNW.2020.30.016
Revised and accepted: August 30, 2020

1. Introduction

Earthquake prediction is designed to detect anomalous changes in earthquake indicators and find the relationship between seismic indicators and occurrence of earthquakes, meanwhile predict the magnitude, location of epicenter, and occurrence time of future earthquakes [1]. For impermeability, the rarity of earthquakes and the complexity of physical processes, earthquake magnitude is difficult to predict, and then resulting in vast loss of life and property. Many scholars of geology, mathematics and computer have proposed various methods for earthquake prediction. However, due to the differences of region, earthquake frequency and quantity,

*Wen Zhou; Yiwen Liang – Corresponding author; Zhe Ming; School of Computer Science of Wuhan University, Bayi road 299, Wuchang district, Wuhan city, Hubei province, China, E-mail: zw_mmwh@163.com, ywliang@whu.edu.cn, mzh116@126.com

†Hongbin Dong; School of Cyber Science and Engineering of Wuhan University, Bayi road 299, Wuchang district, Wuhan city, Hubei province, China, E-mail: hbdong@whu.edu.cn

the specificity of prediction methods and insufficient medium-large seismic data, these methods have poor performance.

The Sichuan province of China is one of the most seismically active cities, with more than 200 earthquakes with magnitude greater than 3 occurring per year. The most recent major earthquakes include the 2008 Wenchuan earthquake, 2013 Ya'an earthquake and 2017 Jiuzhaigou earthquake [2]. The plate movement results in seismic activity of Sichuan, and it locates at Longmenshan fault zone which is formed by the active motion of India plate to Eurasian plate. Since the Sichuan province's special geological structure, earthquake prediction of Sichuan and surroundings may be a problem worth studying.

With the development of intelligent technology, seismologists begin to introduce intelligent methods to build earthquake prediction models, Shi and Liu [3] were pioneers in introducing neural network into earthquake prediction, and the relationship between magnitude and epicentral intensity of an earthquake is established in the paper, nevertheless, the method has poor performance. In 2015, Asencio-Cortés et al. [4] adopted the Principal Component Analysis (PCA) for data dimensionality reduction and new datasets generated, so as to generalize the exist prediction model, however, the difference in geological structure hinders its universality. In [5], an expert system was proposed for earthquake prediction based on historical seismic events. Panakkat and Adeli [6] conducted a large number of reviews in earthquake prediction in the past 15 years, and divided the existing methods into two categories: (1) researches on the seismic precursors analysis; (2) researches based on analyzing historical seismic data. In 2018, DeVries et al. [7] introduced a deep learning scheme to discern criteria for predicting the location of aftershocks based on static stress. Jia et al. [8] adopted Information Entropy Principle (IEP) to predict mine earthquake. However, due to the insufficient medium-large seismic data, these methods have poor performance, for example, the neural network and deep learning approaches require more training sets, otherwise, the prediction results will be unsatisfactory.

In an earthquake prediction system, various anomalies will appear at different levels when the system enters an anomalous status, and this process is similar to human immune system. The danger theory is a method that distinguishes danger from non-danger, it does not require training process, and is skilled in finding anomalies through indicator changes. Hence, it can be used to predict earthquake which is a task that lacking large and medium earthquake data [9]. In [10] and [11], Dendritic Cell Algorithm (DCA) inspired approaches were introduced to predict earthquakes. However, the signals extraction process in these methods depended on artificial experiments, which influences the performance of DCA in earthquake prediction. In [12], the concept of numerical differentiation based DCA was proposed and implemented to four UCI datasets. However, since the first-order numerical differentiation is used to directly calculate changes, the normal changes will also be extracted as danger signal in [12], which may affect the performance of DCA. Therefore, we present a second-order numerical differentiation approach to extract signals of DCA, and then, the new version of ndDCA is adopted to do earthquake prediction in Sichuan and surroundings in our research. In the first place, the earthquake indicator system is collected from different level of features in seismic silence and activity period. Moreover, the earthquake indicators are converted into the input signals in ndDCA through numerical differentiation. In the end, we predict whether there will be an earthquake with a magnitude higher than 4.5 in Sichuan and surroundings during the following month.

The experiment in the paper is consistent with the earthquake prediction conditions proposed by Allen [13]: (1) a definite location (Sichuan and surroundings), (2) a definite time span (the flowing month), (3) a definite range of magnitude ($M \geq 4.5$). The problems are modeled using seven different algorithms: ndDCA, DCA [10], k-Nearest Neighbor (KNN), Haskell based deterministic DCA (hDCA) [11], Negative Selection Algorithm (NSA) [14], Back Propagation Neural Network (BPNN) [15] and Support Vector Machine (SVM). The performance of these algorithms are evaluated by various machine learning and earthquake prediction evaluation indicators [16], and the results depict that our approach yields better prediction accuracies.

The main contributions of our study are described as below:

- We present an architecture that combines a numerical differentiation for signal extraction with DCA that performs anomaly detection, then, it is implemented to earthquake prediction.
- When the first-order numerical differentiation is adopted to directly calculate changes, the normal changes will also be extracted as danger signal in [12], which may affect the performance of DCA. Therefore, we present a second-order numerical differentiation to extract signals of DCA.
- The seismic data for recent years in Sichuan and surroundings are chosen to extract eight indicators by Gutenberg-Ritcher inverse power-law and seismic magnitude distribution for earthquake prediction.
- We apply ndDCA to earthquake prediction, as far as the author knows, it has not been implemented to do earthquake magnitudes prediction. Overall our algorithm yields the encouraging results in all qualified parameters assessed, and it provides technical support for the application of earthquake prediction.

The remainder of the paper is arranged as follows. The latest and most related work of earthquake magnitude prediction and DCA are summarized in Section 2. Section 3 outlines the proposed approach to implement earthquake magnitude prediction. The data preprocess and prediction verification methods are demonstrated in Section 4. Section 5 addresses the experimental results and analysis. Section 6 describes the conclusions and future work.

2. Related work

The artificial intelligence inspired earthquake prediction methods mainly depend on historical seismic catalog or time series earthquake indicators, so as to establish learning model to predict a future earthquake. In 2007, the artificial neural network was used to predict earthquakes in southern California [17]. In 2007, a multilayer perceptron neural network was introduced to predict earthquakes based on the total electron content time series, and it can detect changes in ionospheric variables that are considered as earthquake precursors [18]. The earthquake magnitude was predicted by adaptive neuro-fuzzy model using fuzzy subtractive clustering [19]. An adaptive neuro-fuzzy inference system was proposed for earthquake magnitude prediction in Iran [20]. The regression algorithm and cloud-based big data infrastructure were used for earthquake prediction in California within the next 7 days [21]. In 2017, the imbalanced classifiers and ensemble learning were used to do large earthquake magnitude prediction in Chile [22]. In 2015, for the expert system can provide the characteristics of flexible and effective solutions to different problems, the association rules were used for historical seismic data processing, and the expert systems based on association rules were applied to forecast the earthquake probability over the next 12 hours [5]. Huang et al. [23] proposed a deep learning approach to predict continuous earthquake. For current neural networks are prone to local minimum problems during the training phase, some scholars use optimization algorithms such as genetic algorithms to optimize neural networks [6, 21, 22].

Since Sichuan is an earthquake-prone area, scholars have also been involved in the study of Sichuan earthquake. A ground temperature sensor that implemented in the detection of earthquake precursors was designed by Wang et al. [24]. The Kohonen neural network was adopted to do distribution of aftershock prediction, and implemented to 2008 Sichuan earthquake [25]. In [14], a negative selection algorithm was proposed for earthquake prediction to reduce the impact of lacking large earthquake data on the training performance.

When danger theory is applied to the field of data mining, the study of danger and non-danger discrimination is closely bound up with concentration of indicators. In most data mining systems, it is important to define and measure the correspondence between indicators and danger theory signals. According to status transition process of dendritic cells in dangerous patterns, Greensmith [26] proposed DCA and implemented it to anomaly detection, SYN scanning, intrusion detection and internet of things security. DCA belongs to the category of innate immune algorithm, without the need of pattern matching and dynamic learning, the current researches focus on the optimization of DCA signal acquisition process and its application. Chelly and Elouedi [27] have summarized that the DCA data preprocessing methods including advantages and disadvantages of different versions of DCA and their application areas. This research team uses rough set theory as a data preprocess method of DCA and applies it to solve the problem of binary classification. The DCA is used to do spam detection in [28], and it is also applied to load balancing in cloud computing in [29]. Since DCA does not require a training process when performing anomaly detection, it can solve the problem that the artificial neural network requires a large number of training sets. Therefore, Gan et al. [10] introduced the DCA to predict an earthquake of magnitude greater than 4.5 in Sichuan next month. In [11], a Haskell based deterministic DCA is proposed to do earthquake prediction of Sichuan and surroundings with magnitude greater than 4.5 in the next month. In [12], since the changes described by numerical differential can better reflect the signal generation process in danger theory, it is recommended to do signal mapping for DCA, and then the proposed ndDCA was implemented to classify standard UCI Wisconsin breast cancer dataset.

Generally speaking, an earthquake prediction method includes three steps: acquisition of seismic activity indicators, establishment of prediction models, and evaluation of prediction results. Nevertheless, since the medium and large earthquake data is insufficient, the prediction accuracies are not ideal. Danger theory is an adaptive dynamic intelligent approach to do anomaly detection without was implemented data. According to a large number of research activities in the last decade, as far as the authors know, the ndDCA has not been used to predict earthquakes. Hence, this paper uses ndDCA to predict earthquakes in Sichuan and surroundings.

3. Methodology

An earthquake prediction model based on ndDCA (EQP-ndDCA) is outlined in this section, which transforms earthquake prediction into a binary classification problem based on machine learning approaches.

3.1 Seismic indicators

The proposed earthquake prediction method is according to GR law and earthquake magnitude distribution [30]. This paper uses eight earthquake indicators to predict earthquake, the indicator matrix F is depicted as Eq. (1), and the Tab. I shows the indicators computed from catalogs for earthquake prediction, where M_s is earthquake magnitude threshold, in this paper, it is set to 4.5.

$$F = \{b, \eta, \triangle M, T, \mu, c, \mathrm{d}E^{1/2}, M_{mean}\}. \tag{1}$$

Feature	Description
b	The b value in GR law
η	Sum of mean value of regression curve according to GR law
ΔM	Difference between the largest observed value and the largest expected value according to the GR law
T	Elapsed time of the most recent N events with magnitude $\geq M_s$
μ	Average time among the characteristic events
c	Variation coefficient in mean time between characteristic events (μ)
$\mathrm{d}E^{1/2}$	Square root rate of the released seismic energy
M_{mean}	Mean Richter magnitude of the most recent N events

Tab. I *The indicators computed from catalogs for earthquake prediction.*

The calculation methods of the eight indicators are shown in Tab. II [15, 31], where a refers to the cumulative frequency of earthquakes above zero, N is the number of seismic event, t_n denotes the occurrence time of the nth seismic event (t_1 shows the occurrence time of the first seismic event), M denotes Richter magnitude, n is the total number of earthquake events, $t_{\text{icharacteristic}}$ denotes the observed elapsed time between characteristic events (M_i), and $n_{\text{characteristic}}$ is the total number of characteristic events.

Attribute	Calculation
b	$(n\Sigma(M_i \log N_i) - \Sigma M_i(\Sigma \log N_i))/((\Sigma M_i)^2 - n\Sigma {M_i}^2)$
η	$(\Sigma(\log_{10} N_i - (a - bM_i)^2))/(n-1)$
	$a = \Sigma(\log_{10} N_i + bM_i)/n$
ΔM	$M_{\mathrm{max,observed}} - M_{\mathrm{max,expected}}$
T	$t_n - t_1$
μ	$\Sigma(t_{\mathrm{icharacteristic}})/n_{\mathrm{characteristic}}$
c	standard deviation of the observed times $/u$
$\mathrm{d}E^{1/2}$	$E = 10^{11.8+1.5M} ergs$
	$\mathrm{d}E^{1/2} = \Sigma E^{1/2}/T$
M_{mean}	$\Sigma M_i/n$

Tab. II *Calculation method of the earthquake indicators.*

3.2 Overall prediction procedure

Fig. 1 shows the process diagram of EQP-ndDCA. Firstly, Sichuan (China) and surroundings are chosen as experimental data (red circle), which is acquired from the earthquake catalog of China Earthquake Network Center (CENC), and it provides historical seismic data for the proposed EQP-ndDCA. As can be seen from Fig. 1, the historical seismic data includes five features: event, data time, latitude, longitude and magnitude, and each data is denoted as $E = \{e_1, \ldots, e_i, \ldots, e_m\}$. Then, the GR law is adopted to calculate seismic indicator matrix F, which is described in Section 3.1, four indicators are according to GR law (blue rectangle), and others are not related to the assumed temporal distribution of earthquake magnitude (pink rectangle). The class C for each event is denoted as $C = \{0, 1\}$, and it is defined as Eq. (2) according to [31], where C is equal to 1 if the largest magnitude of an event occurring within the prediction range is greater than or equal to M_s, otherwise $C = 0$. Ω is the event number that occurred in the catalog within the prediction range. According to [15], since the prediction horizon is related to the calculation of seismicity indicators, the prediction horizon is set as one month. To evaluate the seismic indicators for a specific event e_i, the previous N events are calculated, and the N is set to 100 [15]. In this paper, the proposed EQP-ndDCA has been compared to six state-of-art approaches: DCA, hDCA, NSA, SVM, KNN and BPNN, and some statistical indicators are introduced to evaluate the performance of these algorithms.

$$C_i = \begin{cases} 1 & \text{if } \max\{M_j : i < j < i + \Omega\} \geq M_s, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

3.3 Danger theory based earthquake prediction model

The DCA is a danger theory inspired algorithm. Immature DC is to deal with the collection of antigens in tissue, pathogen associated molecular pattern (PAMP), safe signal (SS), danger signal (DS) and inflammatory cytokines (IC), and then the concentrations of costimulatory molecules (CSM), semi-mature dendritic cy-

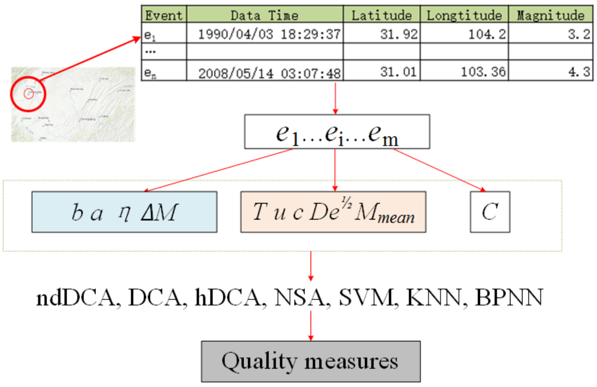

Fig. 1 *The program diagram of the proposed earthquake prediction approach.*

tokines (semi), and mature dendritic cytokines (mature) are exported after antigen presentation (signal fusion) [26]. The status transformation of DC is according to CSM value and migration condition. The DC will migrate when CSM concentration satisfies the migration condition, and the DC context is estimated. The antigen is labeled as 'mature' when mature cytokines concentration is higher than semi-mature cytokines concentration. The mature context indicates 'danger', and semi-mature context indicates 'safe'. DC context determines safe or danger status of all antigens which are presented by the DC. For each antigen, the times it has presented in 'mature' or 'semi-mature' context can be used to decide whether the antigen is normal or anomalous, and this value is defined as mature context antigen value (MCAV). A higher MCAV represents that the antigen is anomalous, whereas the antigen is safe, and the anomalous threshold is usually determined by the application environment. For the signal acquisition of the original DCA does not realize the dynamics of DC, the ndDCA is proposed to realize the purpose of signals acquisition dynamically [12]. Fig. 2 describes the architecture of the

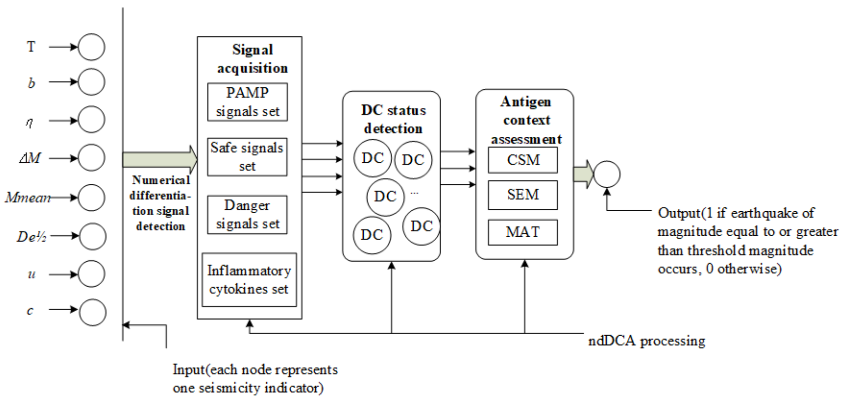

Fig. 2 *The proposed EQP-ndDCA architecture for earthquake magnitude prediction.*

ndDCA model, which is used to predict earthquakes of magnitude greater than or equal to M_s that will occur in the next month. The proposed model involves four phases, signal collection, DC status detection, antigen context evaluation and classification.

(1) Signal collection The core of DCA application is signals collection, and it is various for different application environment. In 2010, packet loss rate, TCP reset packet per second and other four characteristics are selected as indicators [32]. The number of packets per second and packet conversion rate per second are used as DCA signals [33]. In this paper, the indicators described in Tab. I are selected as signal source of ndDCA, and these indicators will be mapped to PAMP, DS, SS and IC signals by the second-order numerical differentiation on the basis of the method ndDCA. DS denotes high possibility of earthquake occurring. PAMP indicates that there are abnormal indicators and earthquake activity is strong. SS indicates that the probability of normal is relatively high, namely earthquake activity is weak. IC is to amplify other signals. The first-order numerical differentiation represents the change between discrete data [12], and the second-order numerical differentiation introduced in this paper is based on the first-order numerical differentiation. If only the first-order numerical differentiation is used to directly calculate the change, the normal change will also be extracted as a danger signal. Therefore, it is necessary to introduce the second-order numerical differentiation to distinguish whether the change represented by the first-order numerical differentiation is anomalous, as the basis for signal extraction. Suppose the indicator set is defined as $I = \{x_1, \ldots, x_i, \ldots, x_n\}$, the first-order and second-order numerical differentiation are described as Eq. (3) and (4), respectively, where t is the unit of equidistance between discrete data points, that is, $t = x_{i+1} - x_i$, which is expressed as a time unit in this paper.

$$f'(x_i) \approx \frac{f(x_{i+1}) - f(x_{i-1})}{2t}, \tag{3}$$

$$f''(x_i) \approx \frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1})}{t^2}. \tag{4}$$

The variation coefficient of first-order and second-order numerical differentiation of indicators are expressed as Eq. (5) and (6), respectively, where m is the number of data instance, and it normalizes each dimension to compare the dispersion degree of each indicator. In this paper, it serves as an indicator for judging whether the change is anomalous.

$$fond_i = \frac{\frac{1}{m}\sum_{j=1}^{m}(f'_j(x_i) - \frac{1}{m}\sum_{j=1}^{m} f'_j(x_i))^2}{\frac{1}{m}\sum_{j=1}^{m} f'_j(x_i)}, \tag{5}$$

$$sond_i = \frac{\frac{1}{m}\sum_{j=1}^{m}(f''_j(x_i) - \frac{1}{m}\sum_{j=1}^{m} f''_j(x_i))^2}{\frac{1}{m}\sum_{j=1}^{m} f''_j(x_i)}. \tag{6}$$

When the variation coefficient of the second-order numerical differential is greater than a danger threshold α, it indicates that the first-order numerical differential is an anomalous change and may be dangerous. Then, the indicator with the largest variation coefficient of the first-order numerical differential is extracted as a danger signal, the safe signal is 0, the PAMP is the variation coefficient of other indicators, IC is the sum of the variation coefficient of all indicator values; on the contrary, the first-order numerical differentiation is a normal change, and the indicator with the largest variation coefficient of the first-order numerical differential is extracted as a safe signal, the danger signal is 0, the PAMP is the variation coefficient of other indicators, IC is the sum of the variation coefficient of all indicator values. The extraction method of each signal of a single data instance is shown in Eq. (7).

$$\begin{cases} ds_i = \max(fond_i), ss_i = 0, \quad pamp_i = \sum_{i=1}^{n} fond_i - \max(fond_i), \\ \qquad ic_i = \sum_{i=1}^{n} fond_i, \text{ if } sond_i > \alpha, \\ ds_i = 0, ss_i = \max(fond_i), \quad pamp_i = \sum_{i=1}^{n} fond_i - \max(fond_i), \\ \qquad ic_i = \sum_{i=1}^{n} fond_i, \text{ otherwise.} \end{cases} \tag{7}$$

(2) DC status detection DC is used to collect antigens in tissue, and the output CSM, SEMI and MAT signals are calculated using a weighted average formula as Eq. (8). The W_P, W_S, W_D denote weight values of input signals, respectively. The C_P, C_S, C_D and IC are input PAMP, SS, DS and IC, and the signal weight values are shown as Tab. III [26].

$$\begin{aligned} C_{[\text{CSM, SEMI, MAT}]} &= ((W_P * C_P) + (W_S * C_S) + W_D * C_D)/ \\ &\quad (|W_P| + |W_S| + |W_D|) * ((1 + IC)/2). \end{aligned} \tag{8}$$

	PAMP	DS	SS
CSM	2	1	2
SEMI	0	0	-3
MAT	2	1	-2

Tab. III *Signal weight values.*

The concentrations of CSM, SEMI, MAT are calculated by antigens and signals in the antigen poll which are sampled by DC randomly. If CSM concentration is larger than the predefined migration threshold, the DC begins to migrate.

(3) Antigen context assessment For a DC, while its accumulated semi value is higher than accumulated mature value, it is in a semi-mature context and vice versa. Then, for each antigen, the times it has presented by a 'mature' or 'semi-mature' DC can be used to decide whether the antigen is normal or anomalous.

(4) Classification The antigens presented by a semi-mature DC are labeled as normal, otherwise they are anomalous. Then, we calculate the normal number and anomalous number of each antigen, and assess the abnormality of the antigen through MCAV.

MCAV is used to assess antigen environment, as shown in Eq. (9), where A_m is the number of mature antigen, and A_s is immature antigen number. The antigen is anomalous when the MCAV is larger than the anomalous threshold, and vice versa. The closer the MCAV value is to 1, the higher antigenic abnormality and the threshold is determined by experimental data.

$$\mathrm{MCAV} = A_m/(A_m + A_s). \tag{9}$$

The earthquake prediction model by ndDCA is described as Algorithm 1.

Algorithm 1 ndDCA.

```
Input: antigens and signals.
Initialize: DC poll = 100, Antigen poll = 10.
Output: antigen context vectors.
for each antigen poll do
  ICs, PAMPs, SSs, DSs calculation by Eq. (7);
end for
for each DC do
  antigen acquisition;
  concentration calculation by Eq. (8) and Tab. III;
  if C_CSM > migration threshold then
    antigen acquisition;
  end if
  for the DC do
    if SEMI > MAT then
      DCContext = SEMI;
    else
      DCContext = MAT;
    end if
  end for
end for
for each antigen do
  if MCAV > anomaly threshold then
    antigen = anomalous;
  else
    antigen = normal;
  end if
end for
```

Danger theory is an adaptive and dynamic intelligent approach to do anomaly detection. The earthquake prediction system and the human immune system are similar in complexity, so this paper uses the danger theory to construct a bionic system for earthquake prediction. We uses the ndDCA which is inspired from danger theory to predict whether an earthquake equal to or greater than 4.5 magnitude will occur in Sichuan and surroundings during the next month. By collecting historical earthquake data in Sichuan and surroundings, the decisive characteristics of earthquakes are collected as an indicator system. Then, the mapping of characteristic vector PAMP, SS, DS and IC signals are established, and the weight matrix of ndDCA is defined, finally, we use ndDCA to do earthquake prediction.

3.4 Model parameters

The EQP-ndDCA has been compared to three artificial immune algorithms: DCA, hDCA, NSA, and three state-of-art machine learning methods: BPNN, SVM and KNN. The proposed approach, DCA and hDCA use PCA as the signal extraction method. In DCA inspired algorithms, DC poll number is 100, and the anomalous threshold is set as a random number between 0 and 1. The danger threshold α in our EQP-ndDCA is set to 0.5. In NSA, the detector radius is set to 0.2. In KNN, k is 1, the Euclidean distance and linear search of neighbors are adopted. The SVM employs a sequential minimal optimization method. The complexity parameter is 1 and it uses a polynomial kernel of exponent 1. The BPNN includes three layers, therein, the learning rate is set to 0.0003, the number of training epochs is 100, and the standard gradient descent is used as weight learning method.

4. Example applications

4.1 Data preprocess

The seismic information of Sichuan and surroundings is intercepted from January 1, 1990 to February 28, 2020. In our research, we select 12,306 samples with the magnitude higher than 3.0 for experiments. Each data instance contains a list of records: time of earthquake, epicentral location, depth and Richter scale. The purpose of data preprocessing is to map these data to earthquake indicators according to GR law, earthquake magnitude distribution and the recent related researches.

In this paper, Yunnan, Gansu, Sichuan, Qinghai are denoted by DS1, DS2, DS3 and DS4, respectively. The first column of Tab. IV describes the symbol of each earthquake zone. The second to third columns show the date of experiment data. The column Total numbers indicates the events number of experiment instance. The column Number +/− shows the number of positive and negative events in experiment dataset. The positive and negative criteria is defined in Eq. (2) with M_s = 4.5. Finally, column Mean depicts the average magnitude of each dataset, meanwhile, column SD describes the standard deviation of magnitudes of all datasets.

DS	Date begin	Date end	Total numbers	Number+/−	Mean	SD
DS1	19900116	20200228	3439	176/3263	3.47	0.49
DS2	19900214	20191231	1170	68/1102	3.51	0.50
DS3	19900109	20200225	5573	292/5281	3.52	0.49
DS4	19900114	20191231	2124	221/1903	3.67	0.61

Tab. IV *Analysis of the four datasets.*

According to the calculation process of earthquake indicators in Tab. II, the eight seismicity indicators are calculated, and the N is set to 100, the initially magnitude threshold is set to 3.0 [31]. The eight indicators and the labels (M_s = 4.5) of Sichuan province are presented in Tab. V.

Date	b	η	ΔM	T	μ	c	d$E^{1/2}$($\times 10^{10} ergs$)	M_{mean}	C
201710	0.80	0.002	−0.31	271	21	1.38	0.019	3.47	0
201505	0.87	0.006	0.62	175	18	0.73	0.033	3.401	1
...	...	...	...	...	...	...	...	...	...
201009	1.05	0.005	−0.24	99	11	0.67	0.036	3.39	1
200805	0.80	0.004	−0.02	3	0	0	3.955	3.907	0

.

Tab. V *Partial earthquake data of Sichuan province after preprocess.*

For the corresponding relationship between seismic indicators and PAMP, DS, SS and IC signal, the signal vectors {PAMP, DS, SS, IC} of DS1, DS2, DS3 and DS4 are calculated by the first-order numerical differential and the second-order numerical differential. The partial signals of Sichuan province (DS3) are shown in Tab. IV.

ID	DS	PAMP	SS	IC
3	0.0087	0.2933	0	1
4	0.0045	0.1145	0	0.90
...	...	...	...	...
45	0	0.0121	0.0042	1
46	0	0.0378	0.0012	0.83

Tab. VI *Partial signals of Sichuan province (DS3).*

4.2 Prediction verification

The performance of the seven earthquake magnitude prediction models are compared by various statistical indicators: R score (R), negative prediction value (NPV), recall (R_n), specificity (S), predictive positive value (PPV), false acceptance rate (FAR), Matthews Correlation Coefficient (MCC), Area Under Curve (AUC). In addition, the average of PPV, NPV, recall and S, named Avg is calculated to provide overall quality measures. Including, the meaning and calculation methods of each seismic evaluation indicators refer to the literature [15].

5. Experimental results and analysis

This section demonstrates the results of four earthquake regions based on the evaluation indicators of each earthquake prediction experiment. The experimental results of each area will be analyzed separately. Tab. VII, VIII, IX and X respectively refer to the statistical analysis of experiment in the four areas, and Fig. 3 is a comprehensive analysis of the experimental results.

	EQP -ndDCA	DCA	hDCA	NSA	SVM	KNN	BPNN
PPV	0.73	0.49	0.32	0.23	0.11	0.11	0.54
NPV	0.88	0.78	0.74	0.62	0.64	0.64	0.69
R_n	0.90	0.76	0.70	0.59	0.10	0.10	0.35
S	0.70	0.53	0.37	0.32	0.10	0.10	0.15
FAR	0.22	0.25	0.26	0.38	0.46	0.46	0.37
MCC	0.31	0.18	0.06	0.12	0.10	0.10	0.23
AUC	0.73	0.60	0.53	0.41	0.56	0.56	0.61
Avg	0.80	0.64	0.53	0.44	0.24	0.24	0.43
R	0.78	0.57	0.44	0.33	0.36	0.36	0.02

Tab. VII *Comparison of experimental results in DS1.*

	EQP -ndDCA	DCA	hDCA	NSA	SVM	KNN	BPNN
PPV	0.79	0.47	0.61	0.57	0.27	0.00	0.67
NPV	0.92	0.82	0.78	0.53	0.73	0.83	0.96
R_n	0.92	0.50	0.71	0.61	0.81	0.00	0.67
S	0.81	0.52	0.65	0.49	0.76	0.00	0.04
FAR	0.02	0.18	0.10	0.31	0.27	0.17	0.04
MCC	0.01	−0.01	0.66	0.43	0.21	0.00	0.63
AUC	0.90	0.46	0.79	0.47	0.45	0.46	0.82
Avg	0.86	0.58	0.69	0.55	0.64	0.20	0.59
R	0.93	0.32	0.87	0.53	0.67	−0.17	0.63

Tab. VIII *Comparison of experimental results in DS2.*

Referring to DS1 (see Tab. VII), we can find that for the R score, the EQP-ndDCA is the highest among the seven classifiers, followed by DCA. The highest AUC of EQP-ndDCA indicates that it is the best classification algorithm for DS1. Meanwhile, the MCC has obvious advantages compared with NSA, hDCA, SVM, KNN and BPNN, which also shows that the compared methods are not suitable for this dataset.

The experimental results in Tab. VIII show that the proposed approach has the highest R score, and it is a very satisfactory evaluation indicator in a seismic prediction system. Meanwhile, the evaluation indicators of hDCA are also not bad. However, the R score of KNN is −0.17, which depicts that it is not suitable for the dataset.

Tab. IX demonstrates the experimental results of DS3, from which it can be seen that the R score and Avg of the proposed algorithm provide better results.

	EQP -ndDCA	DCA	hDCA	NSA	SVM	KNN	BPNN
PPV	0.73	0.33	0.46	0.31	0.00	0.00	0.56
NPV	0.87	0.78	0.83	0.43	0.76	0.24	0.90
R_n	0.75	0.24	0.86	0.47	0.00	0.00	0.50
S	0.69	0.15	0.63	0.51	0.00	0.00	0.08
FAR	0.16	0.22	0.12	0.33	0.24	1.00	0.10
MCC	0.59	0.10	0.46	0.56	0.00	0.00	0.44
AUC	0.69	0.62	0.71	0.53	0.59	0.59	0.71
Avg	0.76	0.38	0.70	0.43	0.19	0.56	0.51
R	0.85	0.01	0.73	0.43	−0.24	−1.00	0.40

Tab. IX *Comparison of experimental results in DS3.*

For SVM and KNN, neither of them passed any true positive, but also do not find any false positive, namely, all the data are labeled as positive, therefore, the result cannot be considered.

The experimental results in Tab. X depicts that the MCC difference between EQP-ndDCA and the compared approaches is obvious, producing a difference greater than 0.13 units with the sub-optimal classifier (hDCA). However, the R scores of EQP-ndDCA and BPNN are the same, which indicates that BPNN is also suitable for predicting earthquakes in DS4.

	EQP -ndDCA	DCA	hDCA	NSA	SVM	KNN	BPNN
PPV	0.91	0.43	0.83	0.53	0.34	0.00	0.78
NPV	0.89	0.52	0.78	0.61	0.53	0.63	0.79
R_n	0.88	0.48	0.81	0.49	0.69	0.00	0.95
S	0.75	0.59	0.75	0.57	0.80	0.00	0.89
FAR	0.01	0.31	0.10	0.33	0.47	0.37	0.90
MCC	0.76	0.09	0.63	0.41	−0.12	0.00	0.62
AUC	0.87	0.56	0.83	0.43	0.42	0.48	0.59
Avg	0.86	0.51	0.79	0.55	0.59	0.16	0.85
R	0.92	0.48	0.83	0.33	0.22	−0.37	0.92

Tab. X *Comparison of experimental results in DS4.*

Through the combination analysis of Tab. VII–X, we can find that for most datasets, EQP-ndDCA has the highest R score and Avg value, therefore, we can conclude that EQP-ndDCA is the most suitable classifier for these datasets. Nevertheless, since almost all forecasts have too much true positive, we still have a lot of work to be improved. Since Avg is a comprehensive manifestation of PPV, NPV, Rn and S, our paper concludes by observing the Avg of datasets DS1-DS4 in Fig. 3, and it shows that the Avg of EQP-ndDCA are better. The second best algorithm is hDCA, but the result has high variability, making this algorithm an unstable method, and may not be an estimated method for future work. In the

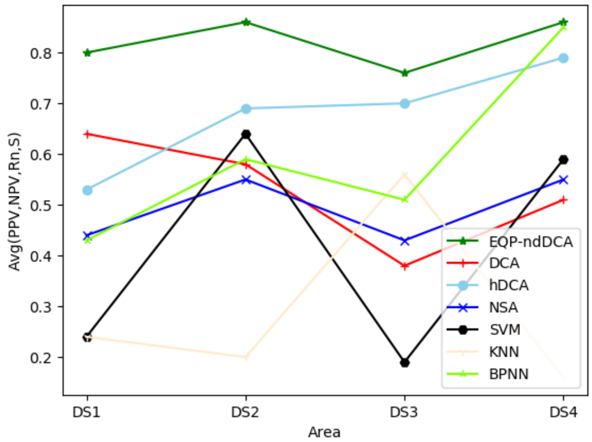

Fig. 3 *The Avg between EQP-ndDCA and other algorithms.*

end, the methods with worst results are SVM and KNN, which have an average accuracy less than 50 %.

From Tab. VII–X and Fig. 3, we can find that the performance of EQP-ndDCA in dataset DS2 and DS4 are better than DS1 and DS3, the reason is that the standard deviation of magnitudes of DS2 and DS4 are higher than that of DS1 and DS3, which can be easily find in Tab. IV, and it makes numerical differentiation has better results in signal extraction, so as to improve the prediction accuracies of EQP-ndDCA.

6. Conclusion and future work

Since the influence of many variables in earthquake magnitude prediction has not been fully understood, it is also a very complicated problem. To the best of author's knowledge, the danger theory inspired ndDCA has not been introduced to do earthquake prediction although it is an efficacious classification method in computer safety, image recognition, and etc.

In our study, the ndDCA in the danger theory is firstly adopted to do earthquake prediction, and the earthquake indicators are acquired by GR law and earthquake magnitude distribution. For danger theory has characteristics of dynamic, and requires no training samples, which guarantee the optimization of the danger theory model and make it suitable to predict earthquakes.

The 12306 samples with a magnitude greater than 3.0 in Sichuan and surroundings are selected as experimental dataset, according to the experiments in DS1, DS2, DS3 and DS4, since the most R score and Avg values for EQP-ndDCA are better than that for other compared algorithms, we conclude that the proposed approach has the best prediction accuracy compared to the other six machine learning algorithms: DCA, hDCA, NSA, BPNN, KNN and SVM. Nevertheless, the proposed EQP-ndDCA has a lack of self-adaptive characteristics. Regarding the impact of seismic indicators on the EQP-ndDCA, the future work is to start from this point.

Acknowledgement

The authors want to thank NSFC http://www.nsfc.gov.cn/ for the support through Grants Number 61877045, and Fundamental Research Project of Shenzhen Science and Technology Program for the support through Grants Number JCYJ2016042815- 3956266, and Scientific research project of Hubei Provincial Department of Education for the support through Grants Number D20191406.

References

[1] CHENG Y.T. Earthquake prediction: retrospect and prospect. *Science in China (Series D): Earth Science*, 2009, 39(12), pp. 1633–1658.

[2] China Earthquake Networks Center. http://news.ceic.ac.cn/index.html,2017.

[3] SHI C.G., LIU X.L. Application of neural network to earthquake engineering. *Earthquake engineering and engineering vibration*, 1991, 11(2), pp. 39–46.

[4] ASENCIO-CORTÉS G., MARTÍNEZ-ÁLVAREZ F. , MORALES-ESTEBAN A. Improving Earthquake Prediction with Principal Component Analysis: Application to Chile. *Lectúre Notes in Computer Science*. 2015, In press. doi: 10.1007/978-3-319-19644-2_33.

[5] IKRAM A., Qamar U. Developing an expert system based on association rules and predicate logic for earthquake prediction. *Knowledge-Based Systems*, 2015, 75(C), pp. 87–103, doi: 10.1016/j.knosys.2014.11.024.

[6] PANAKKAT A. , ADELI H. Recent efforts in earthquake prediction (1990–2007). *Natural Hazards Review*, 2008, 9(2), pp. 70–80, doi: 10.1061/(ASCE)1527-6988(2008)9:2(70).

[7] DEVRIES P.M.R., VIÉGAS F., WATTENBERG M., MEDE B.J. Deep learning of aftershock patterns following large earthquakes. *Nature*, 2018, 560(7720), pp. 632, doi: 10.1038/s41586-018-0438-y.

[8] JIA B.X., ZHOU L.L., PAN Y.S., YANG C.F., LI Z.Y. The Research on Method of Prediction Mine Earthquake Based on the Information Entropy Principle. *International Conference on Information and Software Technologies*. Springer, Cham, 2018, pp. 303–311, doi: 10.1007/978-3-319-99972-2_24.

[9] STRIKWERDA C. The Danger Theory and Its Application to Artificial Immune Systems. *University of Kent at Canterbury*, 2008, pp. 141–148.

[10] GAN Y., LIANG Y.W., TAN C.Y., ZHOU W., WU J.J. Earthquake prediction method based on Danger Theory. *Computer Engineering*, 2019, pp. 40–46.

[11] ZHOU W., DONG H.B., LIANG Y.W. The deterministic dendritic cell algorithm with Haskell in earthquake magnitude prediction. *Earth Science Informatics*, 2020, 13(2), pp. 447–457, doi: 10.1007/s12145-020-00442-z.

[12] ZHOU W., LIANG Y.W., DONG H.B., TAN C.Y., LIU W.W. A Numerical Differentiation based Dendritic Cell Model]. *International Conference on Tools with Artificial Intelligence (ICTAI 2017)* [C], Boston, MA, USA. November 06-08, doi: 10.1109/ICTAI.2017.00167.

[13] ALLEN C.R. Responsibilities in earthquake prediction. *Bulletin of the Seismological Society of America*, 1976, 66(6), pp. 2069–2074.

[14] WU J.J., LIANG Y.W., TAN C.Y., ZHOU W. Method of earthquake prediction based on negative selection. *Application Research of Computers*, 2019, 36(4), pp. 1097–1100.

[15] PANAKKAT A., ADELI H. Neural network models for earthquake magnitude prediction using multiple seismicity indicators. *International journal of neural systems*, 2007, 17(1), pp. 13–33, doi: 10.1142/S0129065707000890.

[16] XU S.X. Estimate of earthquake prediction [A]. State Seismological Bureau. *Beijing: Earthquake Press*, 1989, pp. 586–589.

[17] KAGAN Y.Y., JACKSON D., RONG Y. A testable five-year forecast of moderate and large earthquakes in southern California based on smoothed seismicity. *Seismological Research Letters*, 2007, 78(1), pp. 94–98, doi: 10.1785/gssrl.78.1.94.

[18] SHEN Z.K., JACKSON D., KAGAN Y.Y. Implications of geodetic strain rate for future earthquakes, with a five-year forecast of M5 earthquakes in southern California. *Seismological Research Letters*, 2007, 78(1), pp. 116–120, doi: 10.1785/gssrl.78.1.116.

[19] KAMATH R.S., KAMAT R.K. Earthquake Magnitude Prediction for Andaman-Nicobar Islands: Adaptive Neuro Fuzzy Modeling with Fuzzy Subtractive Clustering Approach. *Journal of Chemical and Pharmaceutical Sciences*, 2017, 10(3), pp. 1225–1230.

[20] MIRRASHID M. Earthquake magnitude prediction by adaptive neuro-fuzzy inference system (ANFIS) based on fuzzy C-means algorithm. *Natural Hazards*, 2014, 74(3), pp. 1577–1593, doi: 10.1007/s11069-014-1264-7.

[21] ASENCIO-CORTÉS G., MORALES-ESTEBAN A., SHANG X., MARTINES-ALVAREZ F. Earthquake Prediction in California Using Regression Algorithms and Cloud-based Big Data Infrastructure. *Computers & Geosciences*, 2017, In press, doi: 10.1016/j.cageo.2017.10.011.

[22] FERNÁNDEZ-GÓMEZ M., ASENCIO-CORTÉS G., TRONCOSO A., MARTINES-ALVAREZ F. Large Earthquake Magnitude Prediction in Chile with Imbalanced Classifiers and Ensemble Learning. *Applied Sciences*, 2017, 7(6), pp. 625, doi: 10.3390/app7060625.

[23] HUANG J.P., WANG X.A., ZHAO Y., XIN C., XIANG H. Large earthquake magnitude prediction in Taiwan based on deep learning neural network. *Neural Network World*, 2018, 28(2), pp. 149–160, doi: 10.14311/NNW.2018.28.009.

[24] WANG T., WANG X.A., YONG S.S., XIE Z., GUO Z.Y., LIN K. A sensor of ground temperature and its application in big earthquake monitoring. *South china journal of seismology*, 2014, 34(01), pp. 33–37.

[25] MADAHIZADEH R., ALLAMEHZADEH M. Prediction of Aftershocks Distribution Using Artificial Neural Networks and Its Application on the May 12, 2008 Sichuan Earthquake. *Ocean Dynamics*, 2009, 11(3), pp. 111–120.

[26] GREENSMITH J. The Dendritic Cell Algorithm. *University of Nottingham*, 2007, doi: 10.2139/ssrn.2831280.

[27] CHELLY Z., ELOUEDI Z. A survey of the dendritic cell algorithm. *Knowledge and Information Systems*, 2016, 48(3), pp. 505–535, doi: 10.1007/s10115-015-0891-y.

[28] TAN Y., DENG C., RUAN G. Concentration based feature construction approach for spam detection. *International Joint Conference on Neural Networks (IJCNN 2009)* [C]. Atlanta, Ga, USA, 2009, pp. 3088–3093, doi: 10.1109/IJCNN.2009.5178651.

[29] DUBEY I., SAURABH P. Danger Theory Based Load Balancing (DTLB) Algorithm for Cloud Computing. *International Journal of Computer Science Issues (IJCSI 2015)*, 2015, 12(2), pp. 301–306.

[30] GUTENBERG B., RICHTER C.F. Seismicity of the Earth. *Princeton University*, Princeton, 1954.

[31] ASENCIO-CORTÉS G., MARTÍNEZ-ÁLVAREZ F., TRONCOSO A., MORALES-ESTEBAN A. Medium large earthquake magnitude prediction in Tokyo with artificial neural networks. *Neural Computing & Applications*, 2015, pp. 1–13, doi: 10.1007/s00521-015-2121-7.

[32] GREENSMITH J., AICKELIN U. Dendritic cells for SYN scan detection, 2010, pp. 49–56.

[33] GREENSMITH J., AICKELIN U. Dendritic Cells for Real-Time Anomaly Detection. *Computer Science*, 2006, doi: 10.2139/ssrn.2831378.