Earth Observation and Geomatics Engineering 4(2) (2020) 98-108

EARTH OBSERVATION AND GEOMATICS ENGINEERING
website: https://eoge.ut.ac.ir

Thermal anomaly detection using NARX neural network method to estimate the earthquake occurrence time

Mehran Nekoee[1], Reza Shah-Hosseini [1]*

[1] School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

Article history:
Received: 9 October 2019, Received in revised form: 15 July 2020, Accepted: 20 July 2020

ABSTRACT

The researchers' view toward earthquake has been vastly changed from the past until now in such a way that the field of earthquake prediction and assessing its pre-indicators have been receiving significant attention from researchers. By the development of remote sensing techniques and obtaining thermal data of the earth's surface and different layers above it, it has become possible to properly study the thermal changes before, during, and after the earthquake. In this study, the remotely sensed thermal data from the earth's surface of the center zone of the earthquake was used in order to predict the time of earthquake occurrence. To date, smart methods which include Artificial Neural Networks (ANN), Support Vector Machine (SVM), Genetic Algorithm (GA), etc., contain different uncertainties depending on their training algorithms in such a way that by defining an inappropriate threshold between the predicted value and the real ones, they are not able to isolate the variable but natural behavior of the under study area from the anomaly. For instance, they identify the natural increase in the environment temperature, which could be as a result of seasonal or climatic conditions, as a thermal abnormality which could enter high levels of errors in determining the time of earthquake occurrence. Considering the fact that a series of time dependent data should be used in studying earthquakes, the prediction of these time series can be done using Artificial Neural Networks. In order to make it more accurate, two different methods of dynamic NARX (Nonlinear Auto Regressive with eXternal input) neural network algorithm namely Levenberg Marquardt and Scaled conjugated gradient have been applied. After that, the responses of these two methods have been compared with the response derived from mean and variance. The important advantage of the NARX neural network is that it can detect and consider small thermal anomalies caused by natural climate change, which cannot be done by regular earthquake pre indicators. The results elucidate that the Saravan earthquake, 5 days before (Levenberg-Marquardt method response) and 11 days before (Scaled conjugated gradient method response), Goharan earthquake 13 days before and Borujerd earthquake 8 and 6 days before occur has been predicted. The thermal anomaly about 5, 7 and 7 degrees of Kelvin of earthquakes center zone respectively detected. Thus the thermal anomaly detected by this method can be a good pre indicator for earthquake prediction.

KEYWORDS

Thermal anomaly
NARX Neural Network
Land Surface Temperature

* Corresponding author
E-mail addresses: mehran.nekoee@ut.ac.ir (M. Nekoee); rshahosseini@ut.ac.ir (R. Shah-Hosseini).
DOI: 10.22059/eoge.2021.292253.1067

1. Introduction

Nowadays with the development of cities and aggregation of sensitive industrial centers and residential areas, an earthquake has more lethal and destructive consequences and therefore, preventing and minimizing its losses is of great importance. Hence the researchers and scientists of remote sensing have vastly studied this field but yet have not been able to provide an independent and comprehensive model for predicting an earthquake. In order to conduct a feasibility study of earthquake prediction, first, its process has to be evaluated. Earthquake is the result of deep ground movements in such a way that tectonic movements of the earth's ground layers lead to the creation of pressure between the rocks forming these layers and causes the earthquake to happen. These movements will increase the temperature of the layers' forming materials from the deep ground to their surface and therefore the earth's outer surface.

There are several ground-based methods to estimate the earth's parameters such as land surface temperature. In these methods, it is necessary to access every ground pixel in order that produces thermal time series. In addition, this limitation impels us to use remote sensing techniques in wide data ranges and high accuracy information of the earth parameters in continuous time series packages.

The remote sensing technology is able to measure different underground, surface and aerial physical parameters of the Earth using strong satellite information. Therefore, predicting parameters of the earthquake such as earth surface temperature, the temperature of layers above the earth surface, OLR, the amount of TEC existing in the atmosphere and several other parameters can be measured utilizing this technology. Several remote sensing satellites are present in this field that are able to measure the above parameters. For instance, for calculation of earth surface temperature, which is the subject of the present study, two thermal bands of Landsat 8 or/and a single thermal band in one of the Landsat set sensors can be used. Moreover, the MODIS sensor which is installed on Terra and Aqua platforms provides the users with the radiance amount and the thermal product independently.

Despite the capabilities expressed in remote sensing, this method also has limitations in the areas of modeling, data, and evaluation of the accuracy of methods. One of the challenges of modeling in such problems is the irregular and nonlinear behavior of temperature changes in the Earth's surface and atmosphere. Another problem is modeling noise elimination from the time series data, which are caused by multiple factors, such as seasonal effects or temporary effects of climate change. Furthermore, available satellite thermal data have a poor spatial resolution which affects the accuracy of the smart methods presented in all researches and has become a challenge for researchers. In addition, thermal data from satellite sensors do not record data in some days and nights due to the presence of clouds over the region, which in turn creates gaps in the time series of data that must be resolved with careful interpolation. It should be noted that earthquake-related thermal changes generally occur in a small local window around the earthquake center and often within the earthquake itself. Given this, there is a large gap between satellite measurements that have poor spatial resolution and ground measurements that are more accurate in many areas, and therefore the accuracy of the methods is also limited.

In the field of earthquake prediction based on the thermal anomaly, an increase of 3 to 12 Centigrade degrees approximately in 1 to 10 days prior to the earthquake occurrence has been reported, which would gradually degrade in several days after the earthquake (Qiang et al. 1991). Initially, the relation between thermal anomaly and the intensity as well as the time of earthquake occurrence was investigated to predict massive earthquakes in Russia, China and Japan (Tronin 2002), also after that numerous researchers in different studies provided reports on thermal anomalies before massive earthquake occurrences (Qiang 1999, Tronin et al 2000, Ouzounov & Freund 2004, Pulinets et al 2011, Akhoondzadeh 2014)

In order to predict and estimate earthquake parameters, many researchers proposed different approaches. In research, researchers studied about influences of multiple layers of air temperature differences on tidal forces and tectonic stress before, during and after the Jiujiang earthquake (Ma Weiyu, Zhang Xuedong, Jun Liu, Qi Yao, Bo Zhou, Chong Yue, Chunli Kang, Xian Lu). The deficiency of the proposed method is that the detection is not automatic and it is just applicable for areas that have no blast.

In automatic anomaly detection, many approaches have been proposed. Akhoondzadeh used an intelligent hybrid system, Kalman filtering, Artificial neural network (ANN), wavelet, Support Vector Machine (SVM) and Genetic Algorithm (GA) to detect the thermal anomalies (M. Akhoondzadeh 2013a; M. Akhoondzadeh 2013b; M. Akhoondzadeh 2013c). In another research's, the thermal anomaly, around 1-10 days prior to an earthquake has been detected (Ouzounov, D., Freund, T.,2004; Ouzounov, D., Bryant, N., Logan, T., Pulinets, S., Taylor, P.,2006; Pulinets, S., Ouzounov, D., 2011; Tronin, A.A, 2000; Saradjian, M. Akhoondzadeh, 2011).

In this study, has been tried to reduce recent intelligence anomaly detector's imperfection. As they say, the major defect of some approaches, which is ignoring the natural changes in LST, has been removed. In the proposed method, first, the data are processed. Generally, the data that are in the form of time series have noises caused by physical changes that happen due to time changes. Therefore, wavelet filters to reduce high frequencies of these time series, which are representative of very steep changes and are not in the domain defined for the anomaly, are used. The data are ready for the next step after the wavelet filtering. Then, in order to predict the thermal time series behavior, which is a non-linear series, a strong neural network algorithm must be utilized, which in this study the NARX neural network algorithm with two distinct training methods is suggested.

The advantages of the presented method are, the natural behavior and increase of the area temperature which is a result of seasonal conditions are considered in detecting thermal anomalies and the anomaly detection results are more reliable in comparison with other smart methods. This means that, for a specific under-study case with equal conditions and a similar threshold, the results obtained from the presented method are more precise in comparison with other smart methods. This is achievable by selecting a proper training method in such a way that the points which behave like an anomaly but are actually caused by the nature of the area are used in the training method and finally the time series prediction of the thermal data will be closer to reality.

Furthermore, good results have been obtained by this method using two training algorithms that are to be discussed in the following. The thermal anomaly was predicted 11 days and 5 days prior to the Saravan earthquake occurred in the first and second methods, respectively, which suggests that it has at least a 1-day improvement time-wise in comparison with other presented methods. In addition, in this study, the thermal anomaly in best result, 11 days prior to the Saravan earthquake has been detected which is the better detection relative to others (M. Akhoondzadeh, 2013; Andrew, A.T., Masashi, H., Oleg, A.M., 2002; Gou, G.M., Yang., j., 2013; Akhoondzadeh, 2012; Akhoondzadeh 2013; Khalili, M., Abdollahi Eskandar, S.S. & Alavi Panah, S.K. 2020). Moreover, the Goharan and Boroujerd earthquakes have also shown thermal anomalies 13 and 8 days before the occurrence of the earthquake, respectively, which is identified by the proposed algorithm.

2. The Studied Areas

2.1. Saravan Earthquake

Measured 7.8 on the Richter scale, happened at 15:14 of 16th of April 2013 at the depth of 82 Km in Sistan and Baluchistan Province (IRAN). The casualties were reported to be 35 deaths and 117 injuries. (Fig. 1)

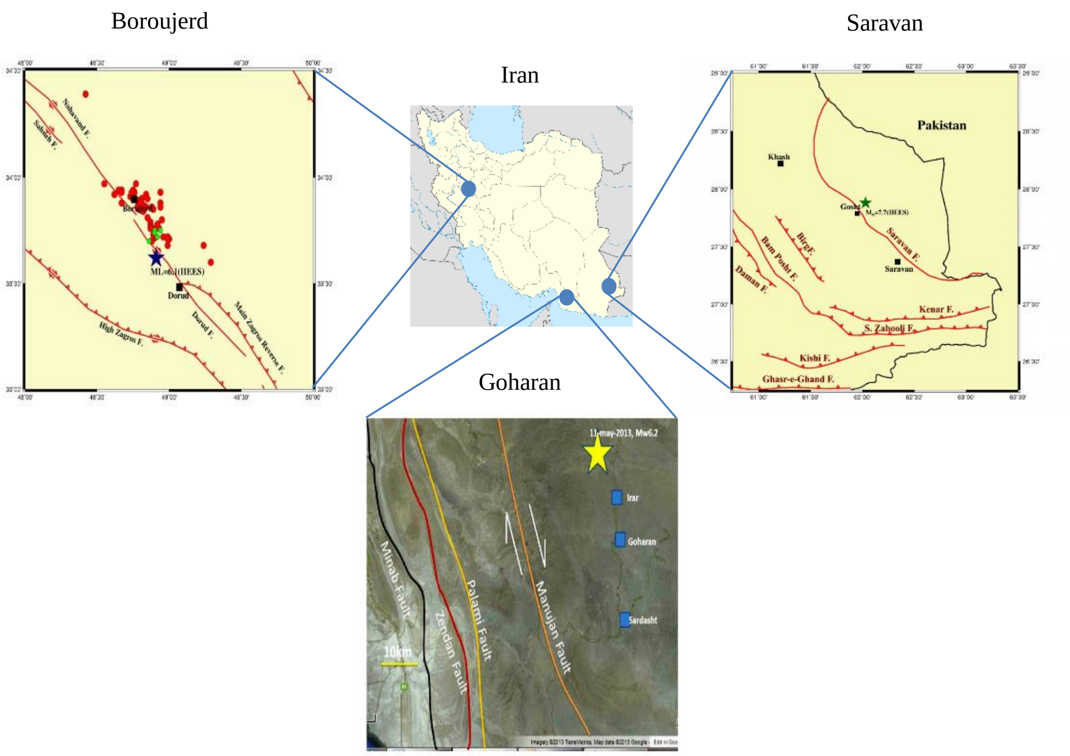


Figure 1. Location of the Saravan, Goharan and Borujerd cities provided by International Institute of Earthquake Engineering and Seismology and Google Earth

2.2. Goharan Earthquake

The Goharan earthquake occurred at 6:38 am on May 11, 2013, with a magnitude of 6.2 Richter. The center of the earthquake is 20 to 25 kilometers deep, 10 kilometers north of Goharan.

2.3. Boroujerd Earthquake

The Boroujerd earthquake, measuring 6.1 on the Richter

3. Data and methodology

3.1. Remote Sensing Data

Thermal raw data from earth's surface of the earthquake center zone in Saravan, in time series of 30 days, in Goharan and Borujerd in time series of 40 days before, during and after the earthquake derivative from MODIS thermal product have been studied. Land surface thermal time series data are obtained from satellite images and thermal outputs of the Modis sensor mounted on the Terra satellite. Given that these products are available on a daily basis; they are a reliable source for analyzing the thermal behavior of the Earth's surface. It should be noted that these satellite images have pixels of 0.05 degrees in latitude and longitude. In this research, the average surface temperature of the earth in a small local window around the center of the earthquake is calculated and entered into the proposed algorithm.

scale happened at 7 am local time on Friday, March 31, 2006, at a depth of 7 kilometers. It occurred in the east part of the Lorestan province in Iran and subjected the Boroujerd and Doroud cities to considerable destruction. The center of the earthquake was 30 kilometers south of Boroujerd on the Great Silakhor Plain Fault.

3.2. Methodology

As stated in the introduction section, the earthquake will indicate some signs in different forms before its happening. In this study, thermal changes pre-indicator was used for assessing the feasibility of earthquake prediction which is the following:

- Extraction of thermal information's from MODIS thermal product and pre-processing
- Prediction and anomaly detection by NARX neural network algorithm

Below, the diagram of the proposed algorithm has been shown in Fig. 2.

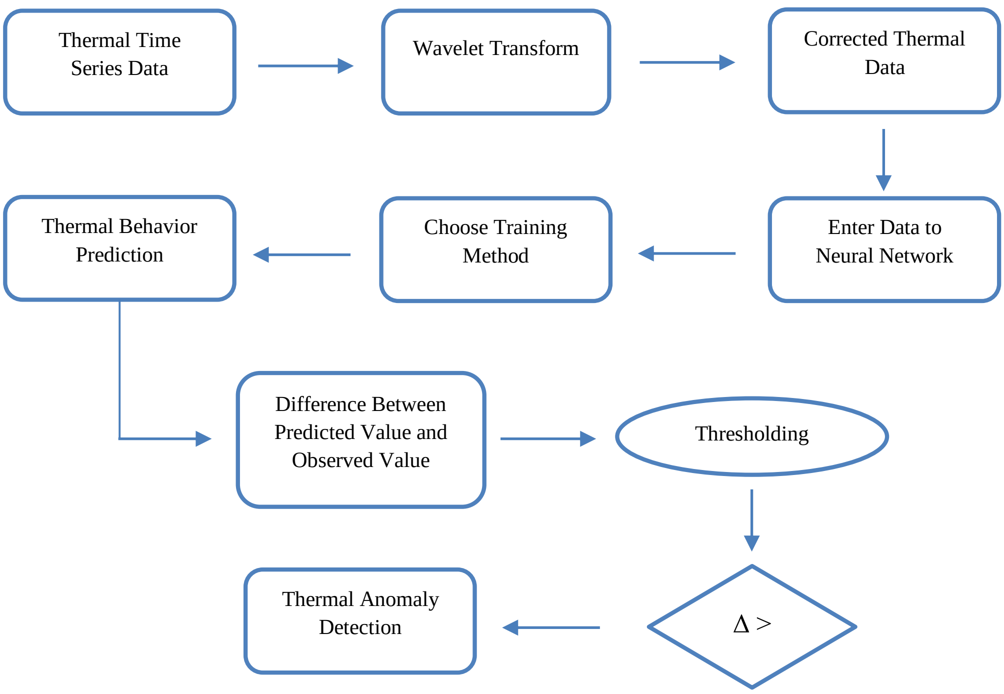


Figure 2. Flowchart of proposed anomaly detection algorithm

3.2.1. Extraction of thermal Information's from MODIS Thermal Product

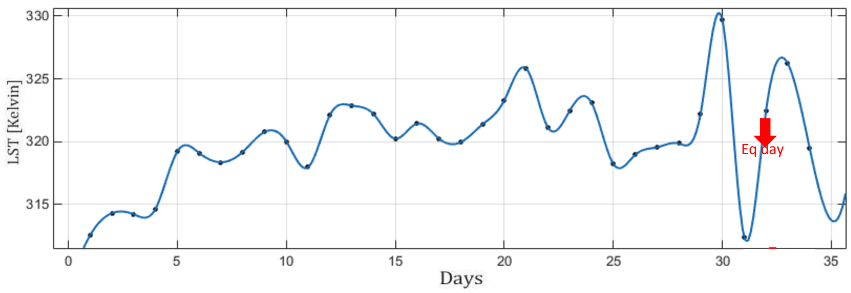
In this study, the thermal data of the MODIS sensor mounted on the TERA satellite was used. These data are measured by bands 31 and 32 of this sensor using split window and day and night methods, with a spatial resolution of 1 Km and a thermal resolution of 0.02 Kelvin degrees which is available on the NASA website. In order to reduce the effects of solar radiance, the daily night-time LST data have been used. The time series of thermal data before, during and after the earthquake are analyzed in this section.

3.2.2. Pre-processing: Wavelet Transform

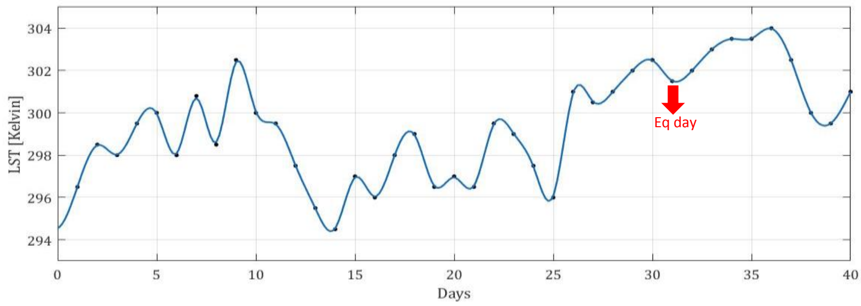
The wavelet transform, transform the signal into scale and time-spaces (Fugal 1994). In the below the mathematic formulation for continuous mode (CWT) has been written. In this equation (a) is the shift and (b) is the scale parameter (Equation 1).

$$CWT(a,b) = \Psi(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t)\Psi\left(\frac{t-b}{a}\right) dt \qquad (1)$$

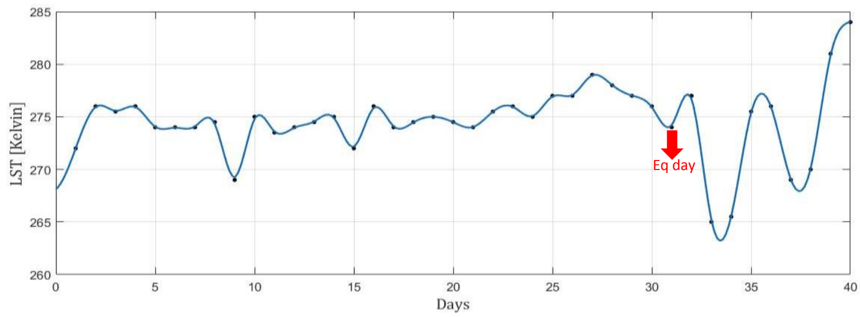
Initially, because of noise existence that results from the solar radiance and seasonal effects, a wavelet transform has been applied. In this section, by wavelet transform, the entrance continuous data divided into two parts. That part that has a higher frequency which is the result of the noise has been deleted. Thus, after that, the data will be nearest to reality (Fig. 3)



(a)



(b)



(c)

Figure 3. LST variations: (a)Saravan, (b)Goharan, (c)Borujerd

3.2.3. Prediction and Anomaly Detection by NARX Neural Network Algorithm

3.2.3.1. NARX Neural Network

The dynamic NARX neural network is a feed-forward algorithm that can predict nonlinear and irregular time series data with high accuracy. This method, predicts series y(t) given d past values of y(t) and another series x(t). The mathematical formulation is: (Equation. 2)

$$y(t) = f\big(x(t-1), \dots, x(t-d), y(t-1), \dots, y(t-d)\big) \quad (2)$$

3.2.3.2. NARX Neural Network Training Methods

In this research, in order to train the NARX neural network, two training methods have been applied. In this section, a summary of these two methods will be expressed.

- **Levenberg-Marquardt algorithm**

The Levenberg-Marquardt algorithm, which was independently developed by Kenneth Levenberg and Donald Marquardt, provides a numerical solution to the problem of minimizing a non-linear function (Kenneth Levenberg 1944, Donald Marquardt 1963). The Levenberg-Marquardt algorithm, also known as the damped least-squares method, has been designed to work specifically with loss functions which take the form of a sum of squared errors. It works without computing the exact Hessian matrix. Instead, it works with the gradient vector and the Jacobian matrix.

Consider a loss function which can be expressed as a sum of squared errors of the form 'f' as seen in the below: (Equation. 3)

$$f = \sum e_i^2 \ , i = 0, \dots, m \quad (3)$$

where 'm' is the number of instances in the data set.

We can define the Jacobian matrix of the loss function as that containing the derivatives of the errors with respect to the parameters: (Equation. 4)

$$J_{i,j}\, f(w) = \frac{de_i}{dw_j} \quad (i = 1, \dots, m \ \& \ j = 1, \dots, n) \quad (4)$$

where 'm' is the number of instances in the data set and 'n' is the number of parameters in the neural network. In the next step, the gradient vector of the loss function will be computed as (Equation. 5)

$$\Delta f = 2\, J^T . e \quad (5)$$

and here 'e' is the vector of all error terms.

Then, the Hessian matrix can be approximated with the following expression: (Equation. 6)

$$HF \approx 2\, J^T . J + \lambda I \quad (6)$$

where 'λ' is a damping factor that ensures the positiveness of the Hessian matrix and 'I' is the identity matrix. Finally, the next expression defines the parameters improvement process with the Levenberg-Marquardt algorithm: (Equation. 7)

$$W_{i+1} = W_i - (J_i^T . J_i + \lambda_i I)^{-1} . (2\, J_i^T e_i), \ \ i = 0,1, \dots \quad (7)$$

It is important to know that the Levenberg-Marquardt algorithm is a method tailored for functions of the type of sum-of-squared-error that makes it very fast when training neural networks measured on that kind of errors.

- **Scaled conjugated gradient algorithm**

The optimization problems on a large scale always have many challenges. One of the most important and useful methods proposed for resolving these problems is the Scaled Conjugate Gradient algorithm (SCG) (Magnus Hestenes and Eduard Stiefel 1952). These methods represent an important innovation for solving large scale unconstrained optimization problems. The conjugate gradient method is often implemented as an iterative algorithm, applicable to sparse systems that are too large to be handled by a direct implementation. When we want to solve an optimization problem, we are faced with these large sparse systems. This method can be implemented both directly and in a repeated way. Due to the non-linearity of the problem in this study and the use of the neural network, the iterative model is used. To solve the system problem, an initial answer is considered first. If we consider the unknown variable X, then this initial response is X_0. Given the nature of the problem, this initial answer must be entered into an iterative algorithm. Here we need to define a constraint function that tells us whether we are closer to the final answer. This constraint is expressed by minimizing the following quadratic function: (Equation. 8)

$$f(X) = \tfrac{1}{2} X^T A X - X^T b \quad , X \in R^n \quad (8)$$

And then its second derivative is given by a symmetric positive-definite matrix. This is the unique minimizer of the mentioned function: (Equation. 9)

$$\nabla f(X) = AX - b \quad (9)$$

This suggests taking the first basis vector P_0 to be the negative of the gradient of f(X) at $X = X_0$. Given the initial value x_0, for P_0, the following answer is also considered: (Equation. 10)

$$P_0 = b - AX_0 \quad (10)$$

The other vectors in the basis will be conjugate to the gradient, hence the name conjugate gradient method. P_0 is also the residual provided by this initial step of the algorithm.

Here, a parameter with the name of r_k is defined. This parameter has the following value for each iteration: (Equation. 11)

$$r_k = b - AX_k \tag{11}$$

As observed above, r_k is the negative gradient of f at $X = X_k$, so the gradient descent method would require moving in the direction r_k. Here, however, we insist that the directions P_k be conjugate to each other. A practical way to enforce this is by requiring that the next search direction be built out of the current residual and all previous search directions. This gives the following expression: (Equation. 12)

$$P_k = r_k - \sum \frac{P_i^T A\, r_k}{P_i^T A\, P_i} P_i \tag{12}$$

Following this direction, the next optimal location is given by (Equation. 13)

$$X_{k+1} = X_k - \alpha_k P_k \tag{13}$$

Where α_k is: (Equation. 14)

$$\alpha_k = \frac{P_k^T\,(b - AX_k)}{P_k^T\, A\, P_k} \tag{14}$$

And according to Equation (11), for α_k, the following equation is obtained: (Equation. 15)

$$\alpha_k = \frac{P_k^T\, r_k}{P_k^T\, A\, P_k} \tag{15}$$

The expression for α_k can be derived if one substitutes the expression for X_{k+1} into f(X) and minimizing it with respect to α_k. The final relation for f(X) is as follows, which establishes the relationship between two repetitions X_k and $X_{k+1:}$ (Equation. 16)

$$f(X_{k+1}) = f(X_k + \alpha_k P_k) \tag{16}$$

3.2.4. Referenced Method

The statistical parameters of the mean and standard deviation of the time series data are used in the reference method. Meaning that the whole thermal data are plotted daily and their mean values are calculated, after which a standard threshold value is used to detect the anomaly. Since the thermal data of an area are in the form of non-linear and irregular time series, the threshold values were considered twice the standard deviation values for the detection of the anomaly, which is proved to have high precision in anomaly detection.

The next step is about the prediction of the thermal behavior of the area. For this purpose, the NARX neural network will be useful, thus, the LST data has been used to train the neural network to extract the thermal model. And then by compare this model with observed daily LST and determine a threshold, the thermal anomaly has been detected. (Equation 17)

$$\text{If } \begin{cases} x > \text{upper bound} & x\text{: thermal anomaly} \\ \text{upper bound} < x < \text{lower bound} & x\text{: normal} \\ x < \text{lower bound} & x\text{: is not earthquake pre-indicator} \end{cases} \tag{17}$$

In this case, the neural network with 10 neurons and 4 delays for hidden layers have been applied (Fig. 4) and then 70 percent of data selected to train the algorithm and 15 percent for test and 15 percent for validation of an algorithm.

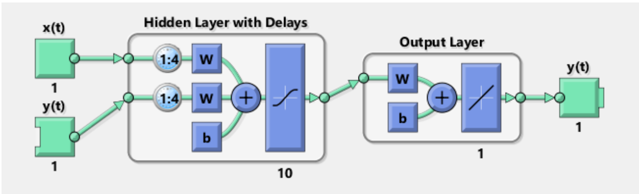


Figure 4. NARX neural network diagram

Finally, by 2 different methods of training, a neural network trained, thus different response derived. The thermal anomaly will be detected by comparing results from neural network prediction and real thermal data for each day. If the difference between them, is upper than 2σ (2*standard deviation = 5 degree of Kelvin), the thermal anomaly has occurred

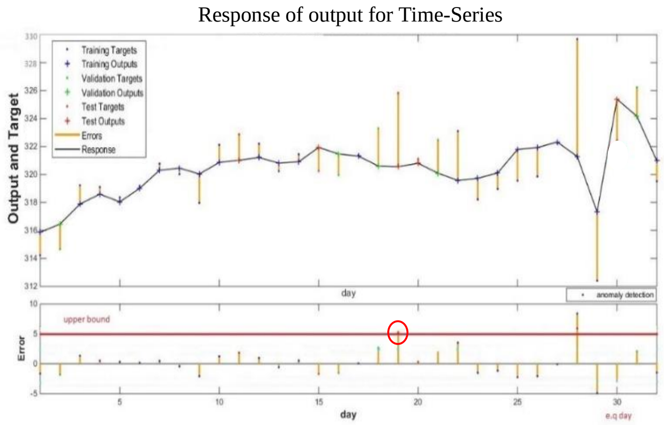
4. Experimental Results

4.1. Levenberg-Marquardt Training Method

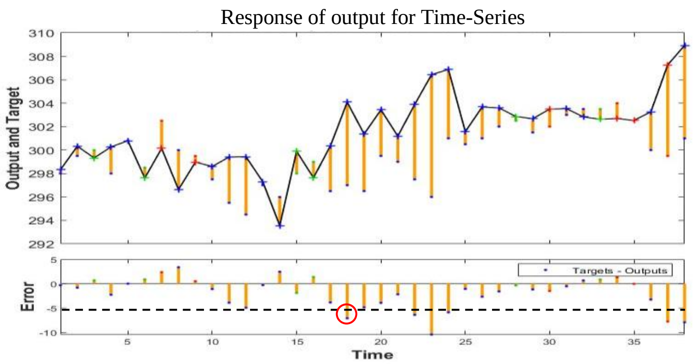
According to the results of implementation of this training method which showed in Fig. 5, 11 days' prior, the Saravan earthquake predicted. In this case in the time-series data the thermal degree softly increases but suddenly a decrease of LST observed. It could be a result of climate change due to global anomalies. Thus, in network analysis, it should be disregarded. Also, according to the results of Fig. 5, Goharan and Boroujerd earthquakes have shown thermal anomalies 13 days and 8 days before, respectively, which was detected by this network training method.

4.2. Scaled Conjugated Gradient Training Method

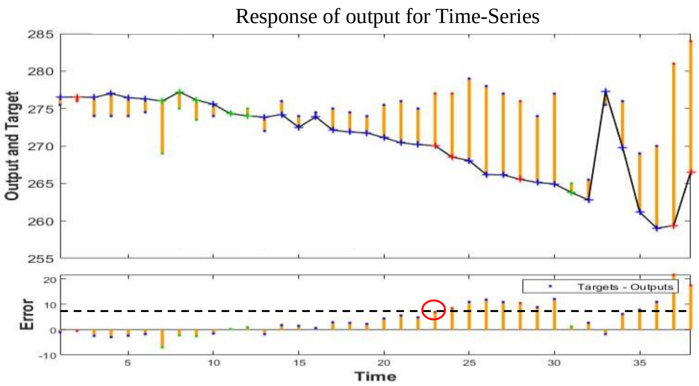
The results of this method which showed in Fig.6 have shown that the thermal anomaly 5 days prior to the Saravan earthquake occurrence has been detected. But in this neural network training method, the algorithm for detecting thermal anomalies caused by the Goharan earthquake has performed poorly. According to Figure 1, thermal anomalies have been discovered 24 days before the earthquake which is not definitely earthquake-related. It is further evidence that this training method has been able to detect very well a thermal anomaly in the Boroujerd earthquake 6 days prior to its occurrence which can be related to the motions of the underground plates due to its proximity to the time of the earthquake.



(a)

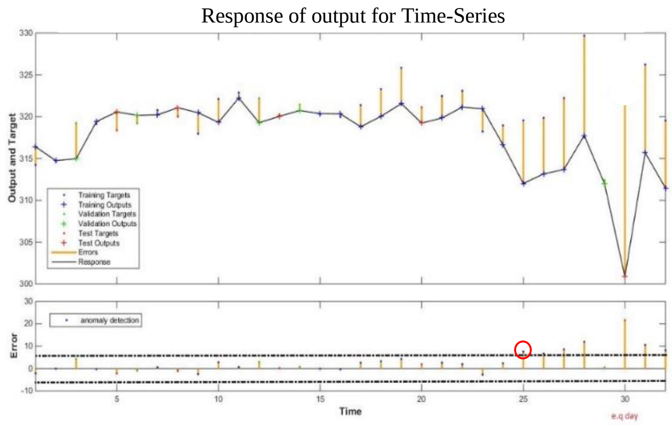


(b)

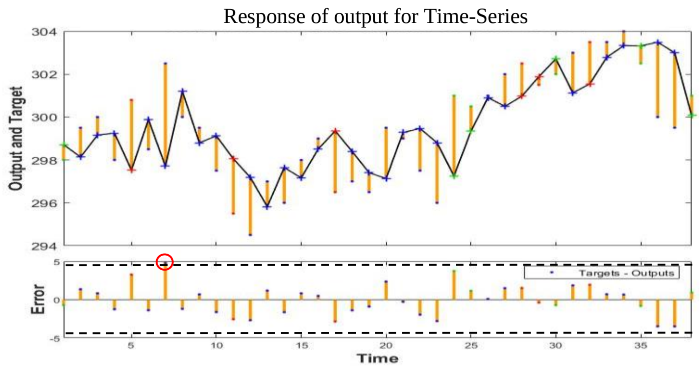


(c)

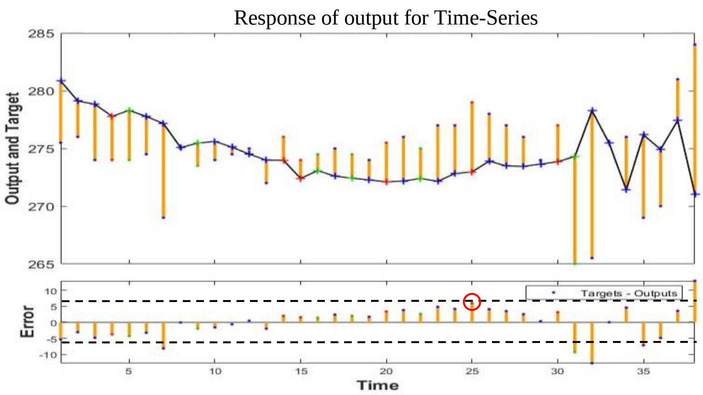
Figure 5. Levenberg-Marquardt method: (a)Saravan, (b)Goharan, (c)Borujerd



(a)



(b)



(c)

Figure 6. Scaled conjugated gradient method: (a)Saravan, (b)Goharan, (c)Borujerd

4.3. Anomaly Detection by Using Mean and Standard Deviation

In order to define a good response to predict the time of earthquake occurrence, the visual analysis and calculating mean and standard deviation can be useful. In this method, because of operator existence, the thermal anomaly has been detected with higher certainty. Now there is a reference model to validate our results. Mean values and deviations were extracted from numerical data of ground surface temperature. In the Saravan and the Goharan earthquakes, due to more noticeable fluctuations in the surface temperature before, during and after the earthquake, thermal anomalies can be extracted using Mean values and thresholds equal to twice the standard deviation. As it is evident, the Saravan earthquake 11 days before and the Goharan earthquake 17 days before have shown abnormal thermal behavior. But in Boroujerd earthquake due to very close fluctuations, we are not even able to detect any thermal abnormalities even with mean and standard deviation. The Boroujerd area has undergone almost normal thermal behavior before, during and after the earthquake.

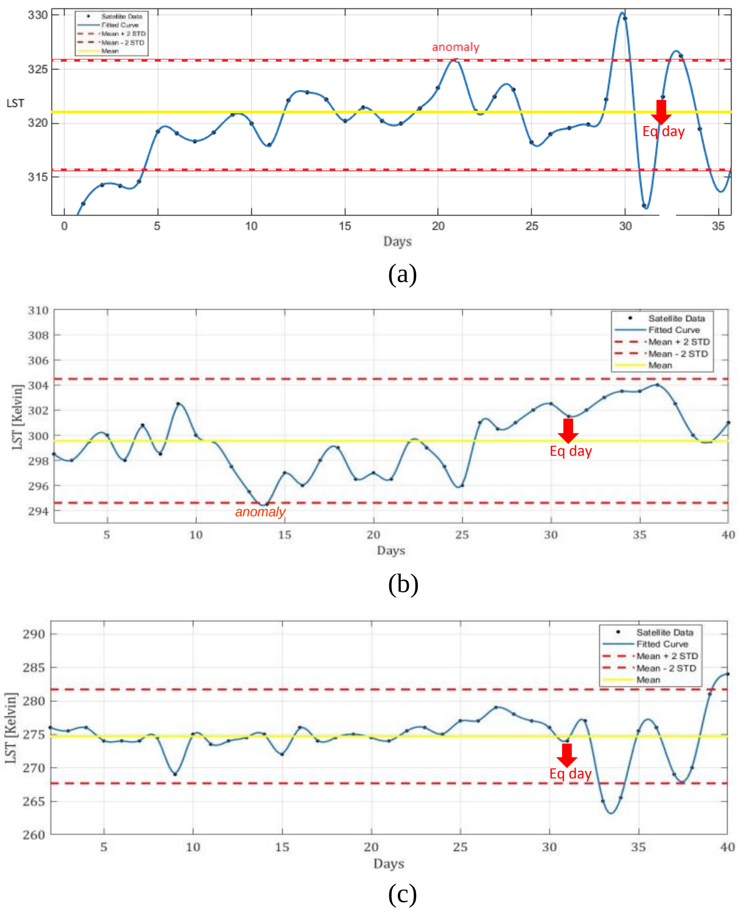


Figure 7. Mean and standard deviation result: (a)Saravan, (b)Goharan, (c)Borujerd

All results obtained from two different neural network training methods in this study for the three study areas are reported separately in Table 1. According to the results shown in the table, the algorithm was able to detect and report the Saravan earthquake 11 days in advance, the Goharan earthquake 13 days in advance and the Boroujerd earthquake 8 days in advance.

Table 1. Proposed algorithm results

Case study	NARX Neural Network training method	Result	LST anomaly (degree)
Saravan earthquake	Levenberg-Marquardt	11 days prior	About 5
	Scaled Conjugate Gradient	5 days prior	About 5
	Reference model	11 days prior	Between 2 and 8
Goharan earthquake	Levenberg-Marquardt	13 days prior	About 7
	Scaled Conjugate Gradient	–	–
	Reference model	17 days prior	About 5
Borujerd earthquake	Levenberg-Marquardt	8 days prior	About 7
	Scaled Conjugate Gradient	6 days prior	About 7
	Reference model	_	_

5. Discussion

The results obtained by the reference method are of high reliability and can be considered as the ideal state since they are fully non-automatic and are analyzed without the intervention of any human agent. Therefore, the reference method results are used to compare the ones attained by the proposed method. As can be seen in Table 1, the thermal anomaly of 2 to 8 degrees is detected about 11 days before the Saravan earthquake using the reference method. The NARX neural network with the Levenberg training method was able to predict the thermal anomaly of near 5 degrees in about 11 days before the earthquake.

Also according to Table 1, in the Goharan earthquake, the proposed algorithm with the first training method has been able to predict the surface thermal anomaly 13 days before, but with the second training method, the algorithm has failed. This may be due to the geographical location of the Gohran area. As shown in Figure 1, this area is located in the province of Hormozgan, near the Persian Gulf and the Strait of Hormuz, and therefore has a humidity of about 50%. This feature of Goharan City in addition to the seasonal thermal changes and the thermal changes caused by the underwater plate movements causes the city to be bereft of the winds and thermal flows born in the sea and poses challenges to any smart algorithm. Likewise, the surface thermal anomaly in Boroujerd has been well detected by both neural network training methods, and this diagnosis has been made in spite of very common and close fluctuations.

6. Conclusions

In order to evaluate the performance of the proposed method, its results were compared with other intelligence methods such as ANN, GA, SVM, ANN+PSO, Single Wavelet, and Kalman. In accordance with outcome results, the improvement of anomaly detection especially in occurrence time was reported. In other words, the results point out that the proposed method can be a good anomaly detector especially in nonlinear time series with natural changes in thermal behavior. As was stated in the introduction part, other smart and automatic methods like neural networks, Kalman Filter and Genetic Algorithm have been able to detect the thermal anomaly about 1 to 10 days before the earthquake occurrence, whereas in this study the time of Saravan earthquake occurrence was approximated 11 days before it, which suggests at least 1-day improvement in comparison with other smart methods (M. Akhoondzadeh, 2013; Andrew, A.T., Masashi, H., Oleg, A.M., 2002; Gou, G.M., Yang., j., 2013; Akhoondzadeh, 2012; Akhoondzadeh 2013; Khalili, M., Abdollahi Eskandar, S.S. & Alavi Panah, S.K. 2020).

The Goharan and Boroujerd earthquakes are also best predicted 13 and 8 days before they occur, respectively. It is evident from the results of this study that the Luneberg training method in the neural network in all three studied regions has achieved a satisfactory performance, but the Scaled method fails in the Goharan earthquake despite its success in the Boroujerd earthquake. Therefore, it can be concluded that the Levenberg and Scaled training methods in the neural network presented in the desert and mountainous regions have acceptable performance and the Scaled Network training methods in coastal and nearshore areas have led to unacceptable results.

As seen, the earthquake phenomena have pre-indicators which are in the range of thermal infrared wavelengths. Considering this factor, certainly analyzing the changes in other ranges of electromagnetic wavelengths can contribute to the improvement of methods and models, therefore it is suggested that radar remote sensing is implemented for future researches and its results are integrated with others to enhance the accuracy.

Acknowledgment

The authors would like to express their acknowledgment to NASA organization for providing us with satellite thermal data on Goharan and Boroujerd earthquakes.

References

Levenberg, Kenneth (1944). "A Method for the Solution of Certain Non-Linear Problems in Least Squares". Quarterly of Applied Mathematics. 2 (2): 164–168.

Marquardt, Donald (1963). "An Algorithm for Least-Squares Estimation of Nonlinear Parameters". SIAM Journal on Applied Mathematics. 11 (2): 431–441. doi:10.1137/0111030.

Hestenes, Magnus R.; Stiefel, Eduard (December 1952). "Methods of Conjugate Gradients for Solving Linear Systems". Journal of Research of the National Bureau of Standards. 49 (6): 409.

Rashki, A., Kaskaoutis, D.G., Eriksson, P.G. et al. Nat Hazards (2014) 71: 563.

Saraf, A.K., Rawat, V., Banerjee, P. et al. Nat Hazards (2008) 47: 119.

Asim, K.M., et al., Earthquake prediction model using support vector regressor and hybrid neural networks. PloS one, 2018. 13(7): p. e0199004..

Kancherla, V.K., V.R. Mandla, and C. Arrowsmith, Study of thermal IR phenomena associated with 27 February 2010 Chile Mw 8.8 earthquake using MODIS data. Geocarto International, 2018. 33(3): p. 293-309.

Mahmood, I., M.F. Iqbal, and M.I. Shahzad, Precursor anomalies prior to the 2006 Java and 2016 Yujing earthquakes. Journal of Geophysics and Engineering, 2018. 15(4): p. 1506.

Roessner, S., Wetzel, HU., Kaufmann, H. et al. Nat Hazards (2005) 35: 395.

Shah, M., et al., THERMAL ANOMALIES PRIOR TO THE 2015 GORKHA (NEPAL) EARTHQUAK FROM MODIS LAND SURFACE TEMPERATURE AND OUTGOING LONGWAVE RADIATIONS. GEODYNAMICS & TECTONOPHYSICS, 2018. 9(1): p. 123-138.

Ma Weiyu, Zhang Xuedong, Jun Liu, Qi Yao, Bo Zhou, Chong Yue, Chunli Kang, Xian Lu, Influences of multiple layers of air temperature differences on tidal forces and tectonic stress before, during and after the Jiujiang earthquake,Remote Sensing of Environment,Volume 210,2018,Pages 159-165,ISSN 0034-4257.

M. Akhoondzadeh, Thermal and TEC anomalies detection using an intelligent hybrid system around the time of the Saravan, Iran, (Mw=7.7) earthquake of 16 April 2013,Advances in Space Research,Volume53,Issue4,2014,Pages647-655,ISSN0273-1177.

Andrew, A.T., Masashi, H., Oleg, A.M., 2002. Thermal IR satellite data application for earthquake research in Japan and China. J. Geodyn. 33, 519–534. Arun,K.S., Vineeta,R.,Priyanka, B., Swapnamita,C.,Santosh, K.P.,Sudipta, D.,Das,J.D., 2008. Satellite detection of earthquake thermal infrared precursors in Iran. Nat. Hazards 47 (1), 119–135.

Guo, G.M., Yang, J., 2013. Three attempts of earthquake prediction with satellite cloud images. Nat. Hazards Earth Syst. Sci. 13 (1), 91–95.

Ma, W.M., Ma, W.Y., Zhao, H., Li, H., 2008. Temperature changing process of the Hokkaido (Japan) earthquake on 25 September 2003. Nat. Hazards Earth Syst. Sci. 8, 985–989.

Akhoondzadeh, M., 2012. Anomalous TEC variations associated with the powerful Tohoku earthquake of 11 March 2011. Nat. Hazards Earth Syst. Sci. 12, 1453–1462. http://dx.

Akhoondzadeh, M., 2013a. A MLP neural network as an investigator of TEC time series to detect seismo-ionospheric anomalies. Adv. Space Res. 51, 2048–2057.
Akhoondzadeh, M., 2013b. Support vector machines for TEC seismoionospheric anomalies detection. Ann. Geophys. 31, 173–186.
Akhoondzadeh, M., 2013c. Genetic algorithm for TEC seismo-ionospheric anomalies detection around the time of the Solomon (Mw = 8.0) earthquake of 06 February 2013. Adv. Space Res. 52, 581–590.
Ouzounov, D., Freund, T., 2004. Mid-infrared emission prior to strong earthquakes analyzed remote sensing data. Adv. Space Res. 33, 268– 273.
Ouzounov, D., Bryant, N., Logan, T., Pulinets, S., Taylor, P., 2006. Satellite thermal IR phenomena associated with some of the major earthquakes in 1999–2003. Phys. Chem. Earth 31, 154–163.
Pulinets, S., Ouzounov, D., 2011. Lithosphere–Atmosphere–Ionosphere Coupling (LAIC) model – an unified concept for earthquake precursors validation. J. Asian Earth Sci. 41, 371–382.
Qiang, Z.J., Xu, X.D., Dian, C.G., 1991. Thermal infrared anomaly precursor of impending earthquakes. Chin. Sci. Bull. 36, 319–323.
Qiang, Z.J., Dian, C.G., Li, L., Xu, M., Ge, F., Liu, T., Zhao, Y., Guo, M., 1999. Satellite thermal infrared brithness temperature anomaly image short-term and impending earthquake precursors. Sci. China 42, 313–324.
Tronin, A.A., 2000. Thermal IR Satellite sensor data application for earthquake research in China. Int. J. Remote Sens. 21, 3169–3177.
Tronin, A.A., Hayakawa, M., Molchanov, O.A., 2002. Thermal IR satellite data application for earthquake research in Japan and China. J. Geodyn. 33, 519–534.
Tronin, A.A., Biagi, P.F., Molchanov, O.A., Khatkevich, Y.M., Gordeev, E.I., 2004. Temperature variations related to earthquakes from simultaneous observation at the ground stations and by satellites in Kamchatka area. Phys. Chem. Earth 29, 501–504.
Akhoondzadeh, M. (2013). A comparison of classical and intelligent methods to detect potential thermal anomalies before the 11 August 2012 Varzeghan, Iran, earthquake (M w= 6.4). Natural Hazards and Earth System Science, 13(4), p. 1077-1083.
Natarajan, V. & Philipoff, P. Nat Hazards (2018) 92: 1097.
Saradjian, M. & M. Akhoondzadeh (2011). Thermal anomalies detection before strong earthquakes (M> 6.0) using interquartile, wavelet and Kalman filter methods. Natural Hazards and Earth System Science, 11(4), p. 1099-1108.
Khalili, M., Abdollahi Eskandar, S.S. & Alavi Panah, S.K. Thermal anomalies detection before Saravan earthquake (April 16th, 2013, M_W = 7.8) using time series method, satellite, and metrological data. J Earth Syst Sci 129, 5 (2020).