

Comparison of different artificial neural network (ANN) training algorithms to predict the atmospheric temperature in Tabuk, Saudi Arabia

ANUSHKA PERERA, HAZI MD. AZAMATHULLA* and UPAKA RATHNAYAKE

Department of Civil Engineering, Faculty of Engineering,

Sri Lanka Institute of Information Technology, Malabe, Sri Lanka

**Department of Civil and Environmental Engineering, The Faculty of Engineering,*

The University of West Indies, St. Augustine, Trinidad

(Received 3 September 2019, Accepted 13 February 2020)

email : anushkaminipro@gmail.com

सार – पिछले अनेक वर्षों से मौसम प्राचलों का अनुमान लगाने के लिए कृत्रिम तंत्रिका संजाल (ए एन एन) मॉडलों का उपयोग महत्वपूर्ण हो गया है। मौसम और जलवायु के पूर्वानुमान में अनेक अन्य पद्धतियों में ए एन एन मॉडल से अधिक सटीक परिणाम प्राप्त होते हैं। हालांकि अलग-अलग मॉडलों के लिए अलग-अलग आँकड़ों की आवश्यकता होती है, और तदनुसार इन आँकड़ों का सावधानीपूर्वक उपयोग किया जाना होता है। साथ ही, इनमें से अधिकांश आँकड़े - गैर रेखीय प्रक्रियाओं से प्राप्त होते हैं, इसलिए प्रागुक्ति के मॉडल आमतौर पर जटिल होते हैं। फिर भी तंत्रिका संजाल गैर रेखीय आँकड़ों पर भी बेहतर कार्य करता है और अच्छे अनुकूल परिणाम देता है। इसलिए सऊदी अरब के तबूक में न्यूनतम वायुमंडलीय तापमान और अधिकतम वायुमंडलीय तापमान की प्रागुक्ति के लिए विभिन्न ए एन एन मॉडलों की तुलना के लिए यह अध्ययन किया गया। सऊदी अरब के तबूक में जलवायु आँकड़ों के लिए आठ अलग-अलग प्रशिक्षण एल्गोरिदम, बी एफ जी एस क्वासी न्यूटन (बी एफ जी) कॉन्जुगेट ग्रेडिएंट विद पॉवेल- बी एल रिस्टार्टस (सी जी बी) लैवेन्बर्ग- मारक्वेड (एल एम), स्केल्ड कॉन्जुगेट ग्रेडिएंट (एस सी जी), फ्लेचर-रीव्स अपडेट कॉन्जुगेट ग्रेडिएन्ट एल्गोरिदम (सी जी एफ), वन स्टेप सीकेन्ट (ओ एस एस), पोलक रिबेयर अपडेट कॉन्जुगेट ग्रेडिएन्ट (सी जी पी) और रैसिलिएन्ट बैक प्रोपेगेशन (आर पी) प्रशिक्षण एल्गोरिदम डाले गए। ए एन एन मॉडलों को प्रशिक्षित करने के लिए अलग-अलग आँकड़ों के प्रदर्शन के मूल्यांकन के लिए माध्य वर्ग त्रुटि मीन स्क्वायर एरर (एम एस ई) और परस्पर संबंध गुणांक (आर) का उपयोग किया गया। मूल्यांकन से सिद्ध हुआ कि प्रशिक्षण एल्गोरिदम बी एफ जी, एल एम और एस सी जी का प्रदर्शन सबसे बेहतर था जबकि ओ एस एस प्रशिक्षण एल्गोरिदम का प्रदर्शन, उपयोग किए गए शेष सभी एल्गोरिदम की तुलना में सबसे कम था।

ABSTRACT. Use of Artificial neural network (ANN) models to predict weather parameters has become important over the years. ANN models give more accurate results in weather and climate forecasting among many other methods. However, different models require different data and these data have to be handled accordingly, but carefully. In addition, most of these data are from non-linear processes and therefore, the prediction models are usually complex. Nevertheless, neural networks perform well for non-linear data and produce well acceptable results. Therefore, this study was carried out to compare different ANN models to predict the minimum atmospheric temperature and maximum atmospheric temperature in Tabuk, Saudi Arabia. ANN models were trained using eight different training algorithms. BFGS Quasi Newton (BFG), Conjugate gradient with Powell-Beale restarts (CGB), Levenberg-Marquadt (LM), Scaled Conjugate Gradient (SCG), Fletcher-Reeves update Conjugate Gradient algorithm (CGF), One Step Secant (OSS), Polak-Ribiere update Conjugate Gradient (CGP) and Resilient Back-Propagation (RP) training algorithms were fed to the climatic data in Tabuk, Saudi Arabia. The performance of the different training algorithms to train ANN models were evaluated using Mean Squared Error (MSE) and correlation coefficient (R). The evaluation shows that training algorithms BFG, LM and SCG have outperformed others while OSS training algorithm has the lowest performance in comparison to other algorithms used.

Key words – Artificial neural network, Atmospheric temperature, Prediction, Tabuk, Training algorithms.

1. Introduction

Interest in use of Artificial Neural Networks (ANNs) for developing climate change prediction models has increased in recent years due to ever changing climate

patterns in the world (Yadav and Chandel, 2013; Acharya *et al.*, 2014; Belayneh *et al.*, 2014; Hashim *et al.*, 2017; Moghim and Bras, 2017; Mishra *et al.*, 2018). ANNs are computer systems inspired by biological neural networks to model relationships between independent and

dependent variables. They are capable of modelling complex non-linear relationships from raw data sets to find relationships among variables (Betiku *et al.*, 2015; Emeko *et al.*, 2015; Ebrahimi and Rajaei, 2017; Ravansalar *et al.*, 2017). Unlike traditional statistical techniques, ANNs do not require the transformation of raw data prior to model generation. Furthermore, pre-assumption of the nature of relationship between input and output variables is not required in ANNs (Agatonovic-Kustrin and Beresford, 2000; Ibric *et al.*, 2003). Literature gives different ANN algorithms and applied to many real-world events predicting future scenarios. In addition, these different ANN training algorithms were tested for their performance while comparing them each other (Kisi, 2004; Nayak *et al.*, 2004; Wang *et al.*, 2018). Climate change is such a real-world scenario which ANN was heavily used by many researchers.

The climate change, which is heavily influenced by the human activities has become a vital topic for discussion in the present world (Field *et al.*, 2015). Emissions from burning fossil fuel is one of the major factors for the climate change and climate variability, where gases like carbon dioxide (CO₂), methane (CH₄), ozone (O₃), etc., are emitted through these fossil fuels (Karl, 2003; Quadrelli and Peterson, 2007). Not only human life, but also other living species like plants and animals have been badly affected through the climate change and climate variability (Vorbsmarty *et al.*, 2000; Hughes *et al.*, 2003; Barnett and Adger, 2007; Harvell *et al.*, 2014; Patz *et al.*, 2016; Azamathulla *et al.*, 2018; Friedrich *et al.*, 2018).

Among the other climatic factors, atmospheric temperature is a key factor in defining the climate change and a slightest change in temperature could trigger the changes in people's daily routines (Kalkstein and Smoyer, 1993). Thus, prediction of atmospheric temperature using various research methods can be increasingly found in literature (Rotstayn *et al.*, 2014; Simmons *et al.*, 2014; Mears and Wentz, 2017).

Kisi and Uncuoglu (2005) studied on the use of three back propagation training algorithms, Levenberg-Marquadt (LM), resilient Back-Propagation (BP) and conjugate gradient for two case studies, stream-flow forecasting and to determine the lateral stress in cohesionless soils. Based on their comparisons of convergence velocities in training and testing performance, they found that LM algorithm is faster and has better performance than other algorithms in training. However, results showed that Resilient Back-Propagation algorithm has the best accuracy during training. Ghaffari *et al.* (2006) tested five training algorithms, Incremental Back-Propagation (IBP) and Batch Back-Propagation (BBP) under gradient

descent, Levenberg-Marquadt, Quick Propagation (QP) and Genetic Algorithm (GA) for their ability in predicting the effect of coating weight gain and pectin-chitosan amount in the coating solution for drug delivery. The performance was tested using the effect of two factors, coating weight gain and amount of pectin-chitosan in the coating solution on the in-vitro release profile of theophylline for biomedical drug delivery. No significant difference was observed between the performances of IBP and BBP, although, the convergence speed of BBP was found to be three-to-four-fold higher than IBP. In addition, they found that the predicting ability precision based on the performance was in the order of IBP, BBP > LM > QP > GA. However, Pham and Sagiroglu (2001) showed that the BP is the best training algorithm out of tested algorithms, (BP), QP, Delta-Bar-Delta (DBD) and Extended-Delta-bar-Delta (EBDB). These training algorithms were used to learn ANN to recognize control chart patterns and classify wood veneer defects. Kisi (2007) studied on the use of four different ANN training algorithms, BP, conjugate gradient (CG), cascade correlation (CC) and LM to forecast streamflow in the North Platte river in the United States. He found that LM gave the best flow forecasts and faster results compared to other training algorithms. In addition, his results showed that CG and CC models produced more satisfactory predictions than BP. Therefore, several studies illustrate different algorithms to reach the best prediction.

However, further studies have used different training algorithms in ANN to forecast various climatic parameters such as rainfall (Lee *et al.*, 1998; Hall *et al.*, 1999), evaporation (Shiri *et al.*, 2014a), dew point temperature (Shiri *et al.*, 2014b), solar radiation (Landeras *et al.*, 2012), daily reference evapotranspiration (Güven and Günel, 2008; Izadifar and Elshorbagy, 2010), etc. Azamathulla *et al.* (2018) presented a study to predict the atmospheric temperature in Tabuk, Saudi Arabia using gene expression techniques and compared that to the ANN model. However, no proper study has carried out to predict the minimum and maximum atmospheric temperature in Tabuk using different ANN algorithms and then to compare their performance. Therefore, identifying that research gap, we developed eight different ANN models to predict the atmospheric temperature of Tabuk, Saudi Arabia using other climatic factors. The results from eight different training algorithms including, Levenberg-Marquadt (LM), BFGS Quasi Newton, Resilient Back-Propagation, scaled conjugate gradient, Conjugate gradient with Powell-Beale restarts, Fletcher-Powell conjugate gradient, Polak-Ribiere conjugate gradient and One step secant are promising. Prediction of minimum and maximum atmospheric temperatures which are two most important climatic parameters to the dwellers in Tabuk is the major novelty of the presented paper.

2. Artificial Neural Networks (ANNs) training algorithms

As it was stated in the introduction section, artificial neural networks are popular among researchers and planners these days to predict real world scenarios. These ANN algorithms use local or global non-linear optimization methods for optimizing the feed-forward neural networks weights. The local searches are limited to local solutions whereas, global searches avoid this limitation (Ilonen and Kamarainen, 2003). The training performance varies based on the objective function of optimization process and the underlying error surface for a given problem and the network configuration. The most popular optimization methods are variants of gradient based back-propagation algorithms. This is because the gradient information of error surface is available in these algorithms. The widely used methods are Levenberg-Marquadt (LM), BFGS Quasi Newton, Resilient Back-Propagation, Scaled conjugate gradient, Conjugate gradient with Powell-Beale restarts, Fletcher-Powell conjugate gradient, Polak-Ribiere conjugate gradient, One step secant (Hagan and Menhaj, 1994; Liang *et al.*, 1994; Martin, 1997; Japkowicz and Hanson, 1999) The objective of the training process in an ANN is to reduce the global error (E) and the error is defined as follows in Equation (1) (Kişi, 2007).

$$E = \frac{1}{P} \sum_{p=1}^P E_p \quad (1)$$

where, P is the total no of training patterns and E_p is the error for training pattern P . E_p is calculated using the following Equation (2).

$$E_p = \frac{1}{2} \sum_{k=1}^N (O_k - t_k) \quad (2)$$

where, N is the total number of output nodes, O_k is the network output at the k^{th} output node and t_k is the target output at the k^{th} output node. The global error is reduced by adjusting the weights and biases in the training algorithm (Kişi, 2007). The following subsections present the widely used algorithms for training the neural networks.

2.1. Levenberg-Marquadt (LM) algorithm

The Levenberg-Marquadt (LM) optimization algorithm is identified to be more powerful than the conventional gradient descent techniques (Lera and Pinzolas, 2002; Lourakis, 2005; Kişi, 2007). It is the most widely used optimization algorithm and designed to

approach the second-order training speed without computing the Hessian Matrix (Moré, 1978). The Hessian Matrix can be approximated when the performance function is in the form of sum of squares and given in Equation (3).

$$H = J^T J \quad (3)$$

where, J is the Jacobian matrix, containing the first derivatives of the network errors with respect to the weights and biases.

The standard back-propagation technique is used to compute the Jacobian matrix, which is less complex than computing the Hessian matrix. Equation 4 gives the Newton-like update used in the LM algorithm. When $\mu = 0$, this is the Newton's method, using the approximate Hessian matrix.

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \quad (4)$$

where, e is a vector of network errors, μ is a scalar quantity, x_{k+1} is the predicted minimizer and x_k is the current point.

2.2. BFGS Quasi Newton Back-Propagation (BFG) algorithm

The BFGS Quasi Newton algorithm was independently developed by C. G. Broyden, D. Goldfarb, R. Fletcher and D. F. Shanno (Nocedal and Wright, 1999). The basic step in the Newton's method given in equation (5), where, H^{-1} is the Hessian matrix of the performance index at the current values of the weights and biases (usual notations are used here).

$$x_{k+1} = x_k - H^{-1} g \quad (5)$$

BFGS method does not calculate the 2nd derivatives. However, an approximate Hessian matrix is updated in each iteration of the algorithm. This update is calculated as a function of the gradient. Super linear convergence rate is observed in BFGS method on most practical problems, even though the algorithm requires more computations and storage in each of the iterations performed (Nocedal, 1980; Nocedal and Wright, 1999; Schraudolph *et al.*, 2018).

2.3. Resilient Back-Propagation (RP) algorithm

Resilient Back-Propagation is an algorithm which directly adapts weights based on local gradient. In this learning scheme, adaptation is not blurred by the behaviour of the gradient. For this to happen, an

individual update value Δ_{ij} is given for each weight to determine the size of the weight update. Δ_{ij} evolves during the learning process depending on its sight on local error function F , governed by the learning rule given in the equation (6) :

$$\Delta_{ij}^{(t)} \begin{cases} \eta^+ \times \Delta_{ij}^{(t-1)}, \text{if } \frac{\partial F^{(t-1)}}{\partial \omega_{ij}} \times \frac{\partial F^{(t)}}{\partial \omega_{ij}} > 0 \\ \eta^+ \times \Delta_{ij}^{(t-1)}, \text{if } \frac{\partial F^{(t-1)}}{\partial \omega_{ij}} \times \frac{\partial F^{(t)}}{\partial \omega_{ij}} < 0 \\ \Delta_{ij}^{(t-1)}, \text{else} \end{cases} \quad (6)$$

where, $0 < \eta^- < 1 < \eta^+$; η is update value factor and ω is corresponding weight. More details on this algorithm can be found in Saini (2008).

2.4. Scaled Conjugate Gradient Back-Propagation (SCG) algorithm

Usually, line search is required at each iteration in conjugate gradient algorithms. However, scaled conjugate gradient algorithm developed by Moller (1993) avoids the time consumption in line search. The model trust region approach used in the LM algorithm is combined with conjugate gradient approach in this algorithm. Scaled Conjugate Gradient Back-Propagation algorithm requires more iterations to converge than other algorithms. But, computations in each iteration are significantly less compared to others since line search is not performed (Moller, 1993; Andrei, 2007; Cetisli and Barkana, 2010).

2.5. Conjugate Gradient with Powell-Beale Restarts (CGB) algorithm

The search direction is periodically reset to the negative gradient in conjugate gradient algorithms. The standard reset point has occurred when number of iterations is equal to the number of network parameters. Powell (1977) proposed a reset method based on the earlier method proposed by Beale (1967) to improve the efficiency of the training. According to this technique, restart is set if a little orthogonality is left between current and previous gradients and this is tested using the inequality given in the equation (7) (Colaco and Orlande, 1999; Saini and Soni, 2002).

$$\left| \mathbf{g}_{k-1}^T \mathbf{g}_k \right| \geq 0.2 \mathbf{g}_k^2 \quad (7)$$

where, \mathbf{g}_{k-1}^T is the norm squared of the previous gradient and \mathbf{g}_k is the current gradient.

2.6. Fletcher-Reeves (CGF) and Polak-Ribiere (CGP) conjugate gradient algorithm

Fletcher-Reeves and Polak-Ribiere introduced two equations to calculate β_k positive scalar to find the P_0 , search vector in this algorithm. Equation 8 presents the β_k for Fletcher-Reeves method whereas Equation 9 presents it for the Polak-Ribiere method (Colaco and Orlande, 1999; Saini and Soni, 2002; Shaheed, 2004).

$$\beta_k = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}} \quad (8)$$

$$\beta_k = \frac{\Delta \mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}} \quad (9)$$

where, β_k is the ratio of the norm squared of the current gradient to the norm squared of the previous gradient and it is a positive scalar.

2.7. One step secant (OSS) algorithm

The one-step-secant (OSS) algorithm is an approach in bridging the gap between the quasi-Newton approach and the conjugate gradient algorithm. This approach does not store the complete Hessian matrix, instead it was assumed at each iteration. OSS also has the advantage of calculating the new search direction without computing an inverse matrix. However, OSS requires more computation in each iteration and more storage than conjugate gradient methods (Constantinescu *et al.*, 2008; Upadhyay, 2013).

3. Methodology

MATLAB installed in a personal computer (*Intel(R) Core (TM) i7-7700HQ CPU @ 2.80 GHz, 16 GB RAM*) was used to develop the above stated algorithms to predict the monthly minimum and maximum atmospheric temperature in Tabuk, Saudi Arabia. The input variables of the algorithms were monthly rainfalls, minimum monthly relative humidity, maximum monthly relative humidity, minimum monthly air pressure, maximum monthly air pressure and monthly average wind speed. In addition, minimum monthly atmospheric temperatures and maximum monthly atmospheric temperatures were also fed to the algorithms to calibrate and test the ANN algorithms. The developed ANN models are used to predict $y(t)$ (are dependent variables of the model; the minimum and maximum atmospheric temperatures) with d past values of $y(t)$ and $x(t)$, where $y(t)$ is a parameter depending on different $x(t)$ parameters for set of time steps, $x(t)$ is the independent parameter for different time steps (are the monthly rainfalls, minimum monthly

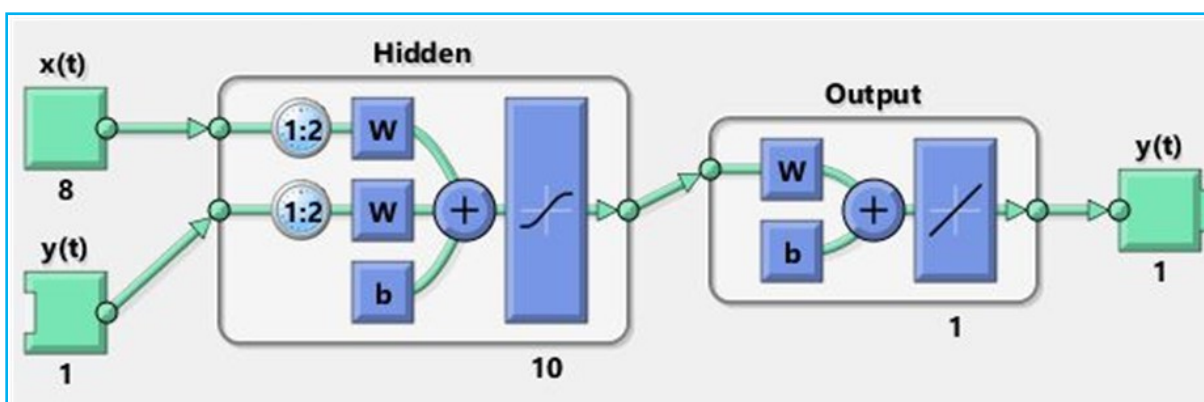


Fig. 1. Conceptual diagram of ANN

relative humidity, maximum monthly relative humidity, minimum monthly air pressure, maximum monthly air pressure and monthly average wind speed) and d is the available past values for $x(t)$ parameters.

The ANN was trained with 70% of target time steps while 15% each of target time steps were used to validate and test. In addition, 10 hidden neurons and 2 delays were used in the network. The details are shown in the Fig. 1. Performance of each training algorithm in predicting the atmospheric temperatures was evaluated using the Mean Squared Error (MSE) and the correlation coefficient (R). The simulation times for all algorithms were around 1-3 seconds from the above stated personal computer.

4. Case study application – Tabuk, Saudi Arabia

Tabuk city is in Saudi Arabia in its north western part as shown in the Fig. 2 and it has an area of 139,000 km². Tabuk province is bound by Saudi-Jordan country boundary from north, Red sea from south and west and Hufa depression from the eastern side. The city is located at an average altitude of 770 m from the mean sea level. Tabuk province is classified as a hyper-arid catchment and experiences a shorter winter season and a longer summer season (Abushandi and Alatawi, 2015). Monthly weather data were collected for 30 years from 1986 to 2015 from the Saudi General Authority of Meteorology and Environment Protection and fed to the developed ANN algorithms to predict the minimum and maximum atmospheric temperatures in Tabuk.

5. Results and discussion

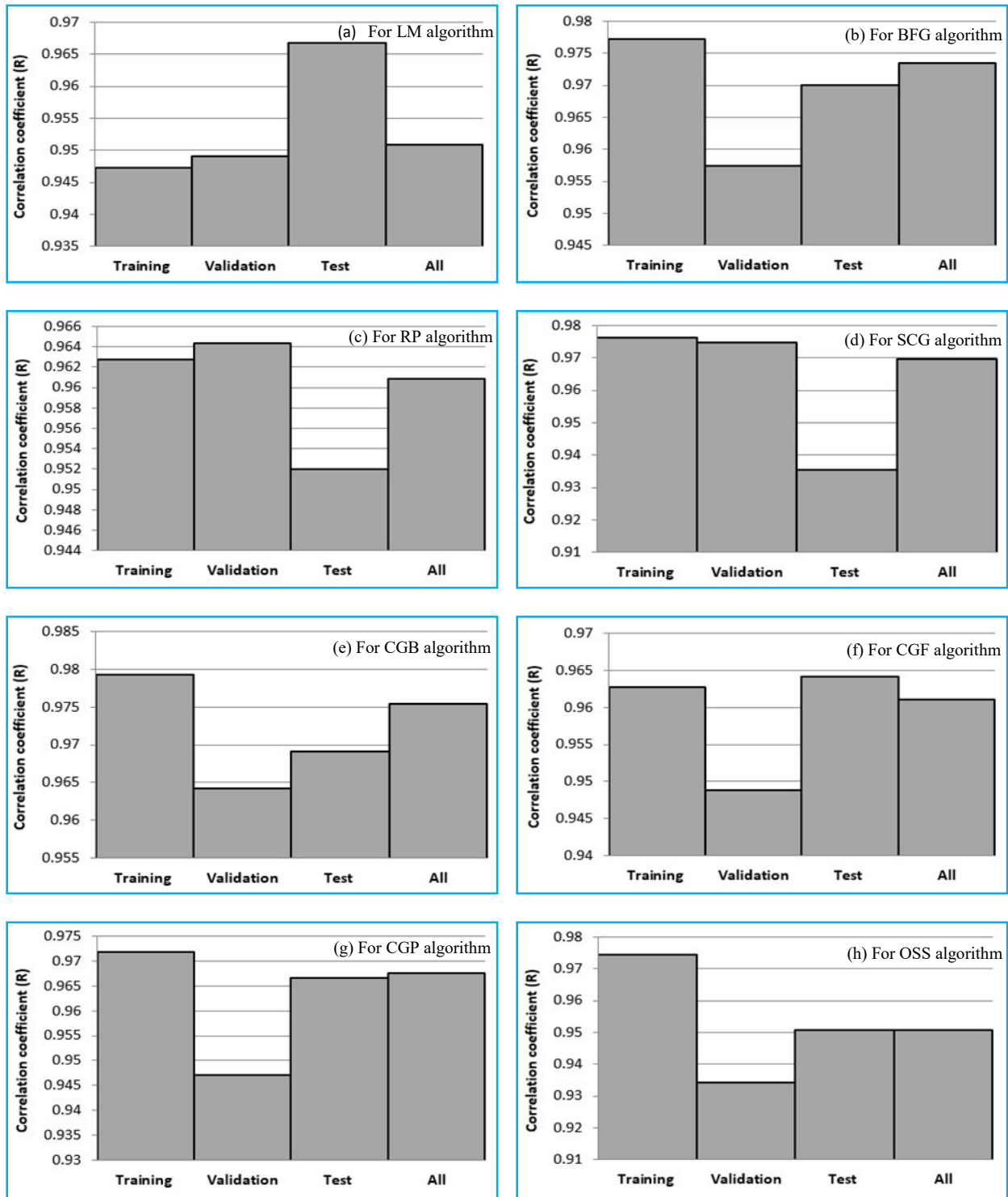
Figs. 3(a-h) show the correlation coefficients (R values) for the performance of the neural networks used for predicting minimum atmospheric temperature. R values are separately shown for training, testing and validation of the neural network models produced using



Fig. 2. Map of Tabuk (Al-Harbi, 2010)

different training algorithms. Figs. 3(a-h) clearly shows that training algorithms BFGS Quasi Newton (BFG) and Conjugate gradient with Powell-Beale restarts (CGB) perform better with R values closer to 1 (>0.97). However, in general, other algorithms also result higher R values.

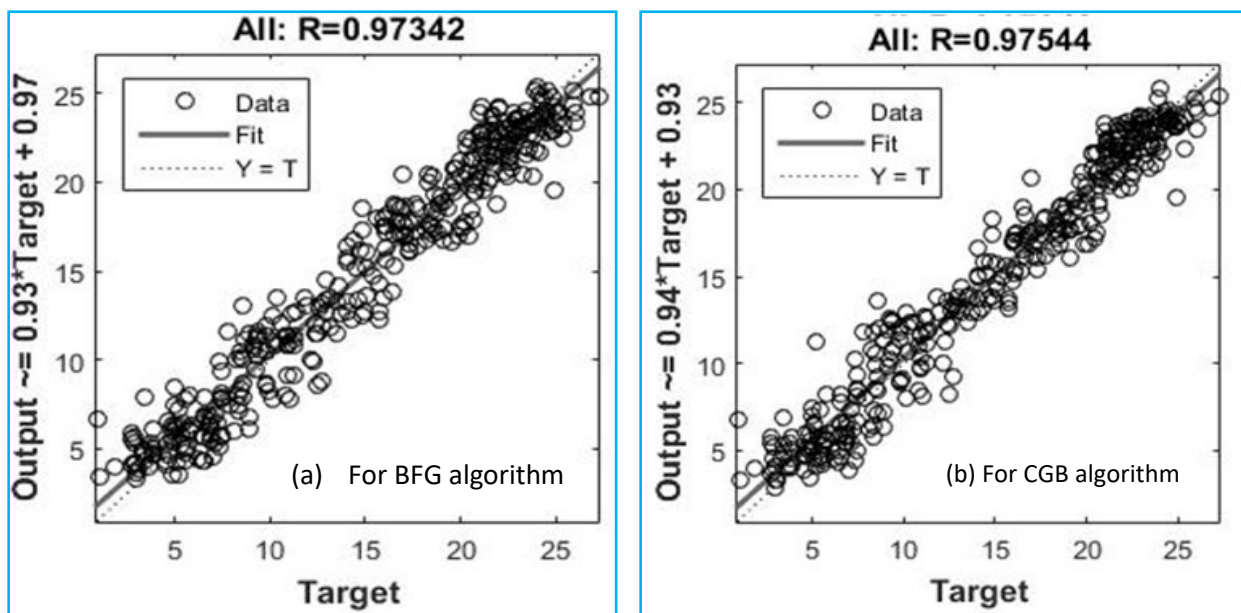
Figs. 4(a&b) present the results from the ANN models for BFG and CGB algorithms against the target. Even though BFG and CGB give the better



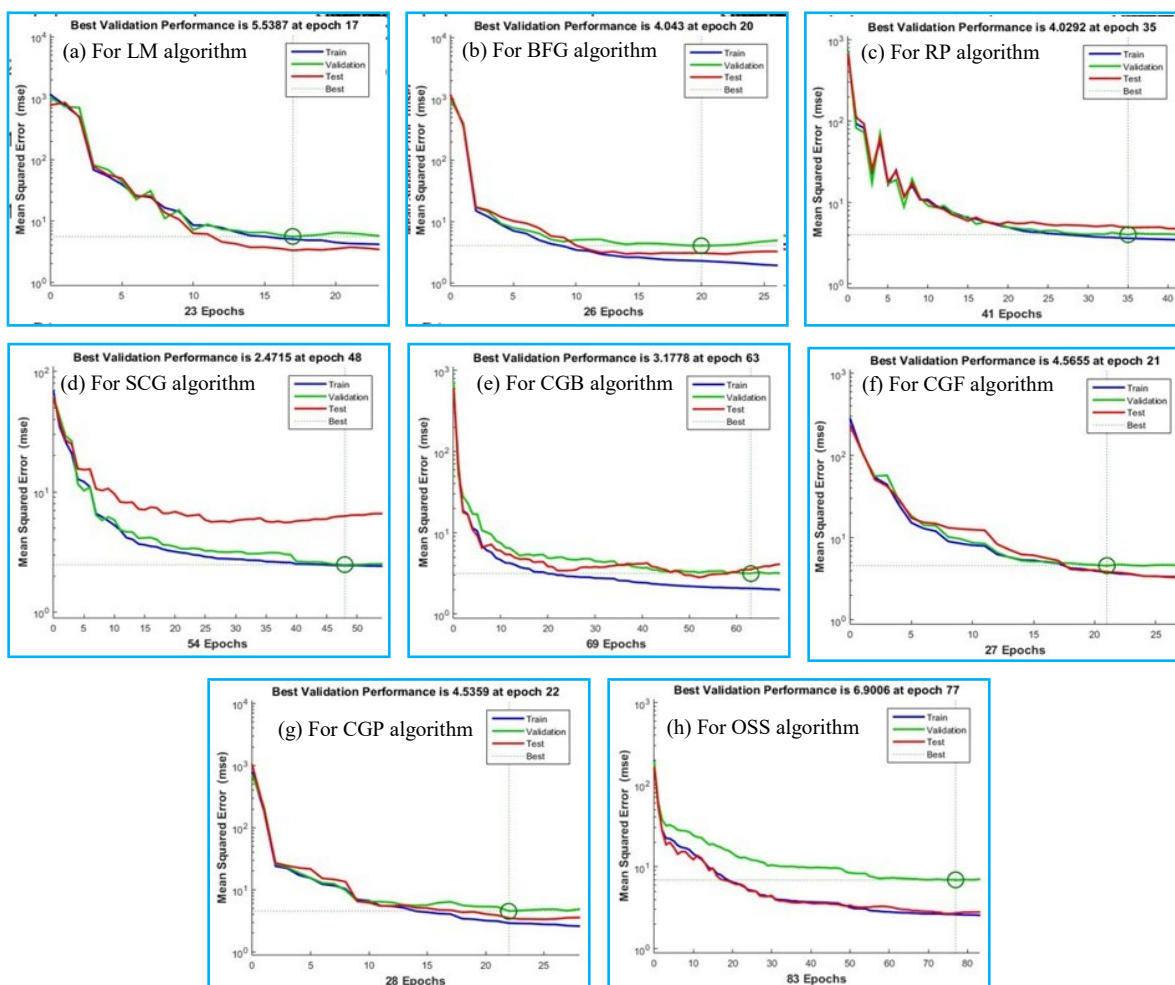
Figs. 3(a-h). Correlation coefficients for minimum atmospheric temperature

performance in R values [Figs. 3(a-h)], they have slightly lower gradients compared to the target [Figs. 4(a&b)]. Therefore, the minimum atmospheric temperature values for Tabuk, Saudi Arabia are under-

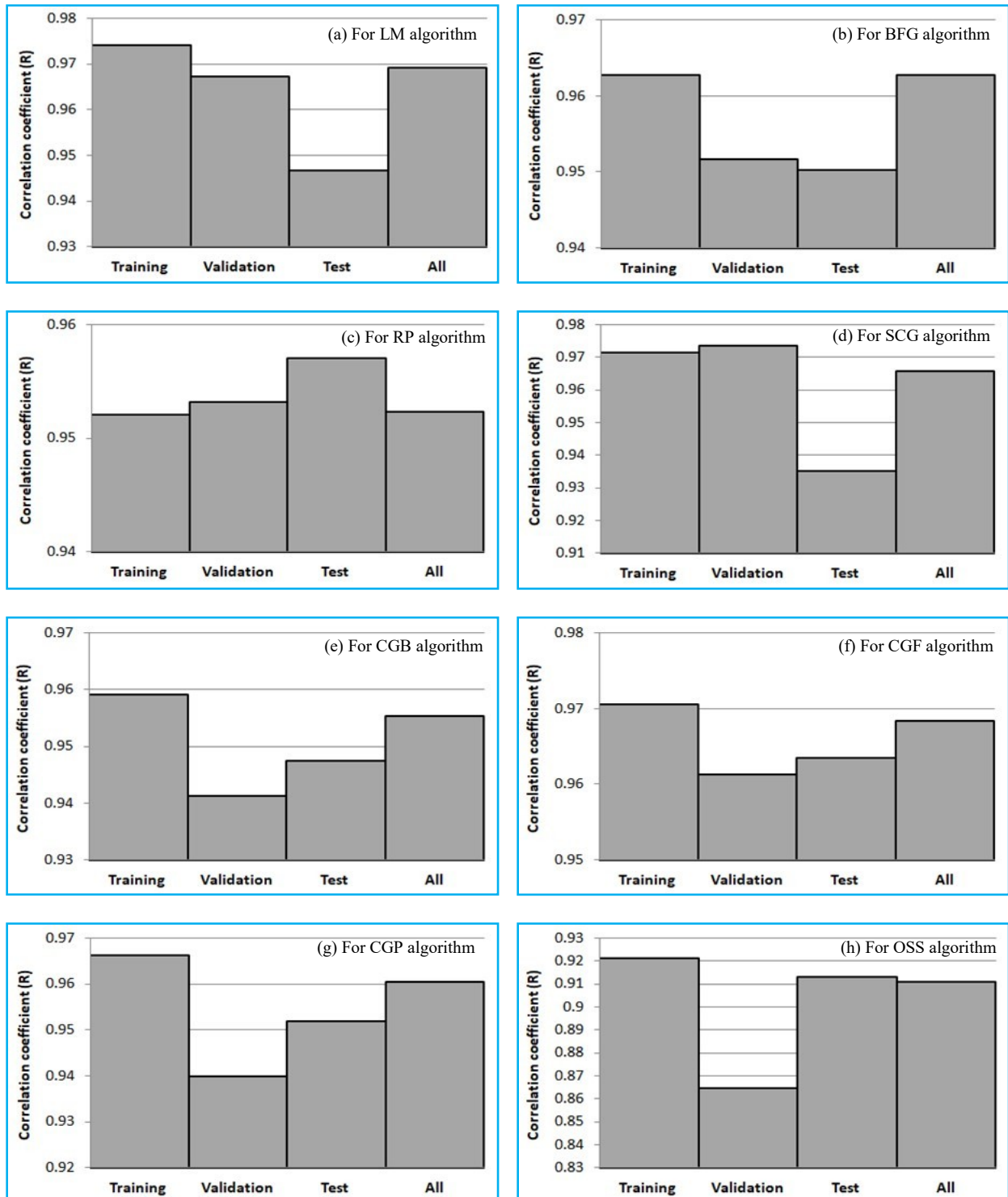
predicting as per training algorithms BFG and CGB. Nevertheless, it is not a significant reduction as the gradients are 0.93 and 0.94, respectively to the two algorithms.



Figs. 4(a&b). Best fits to the target for the ANN results for minimum atmospheric temperature



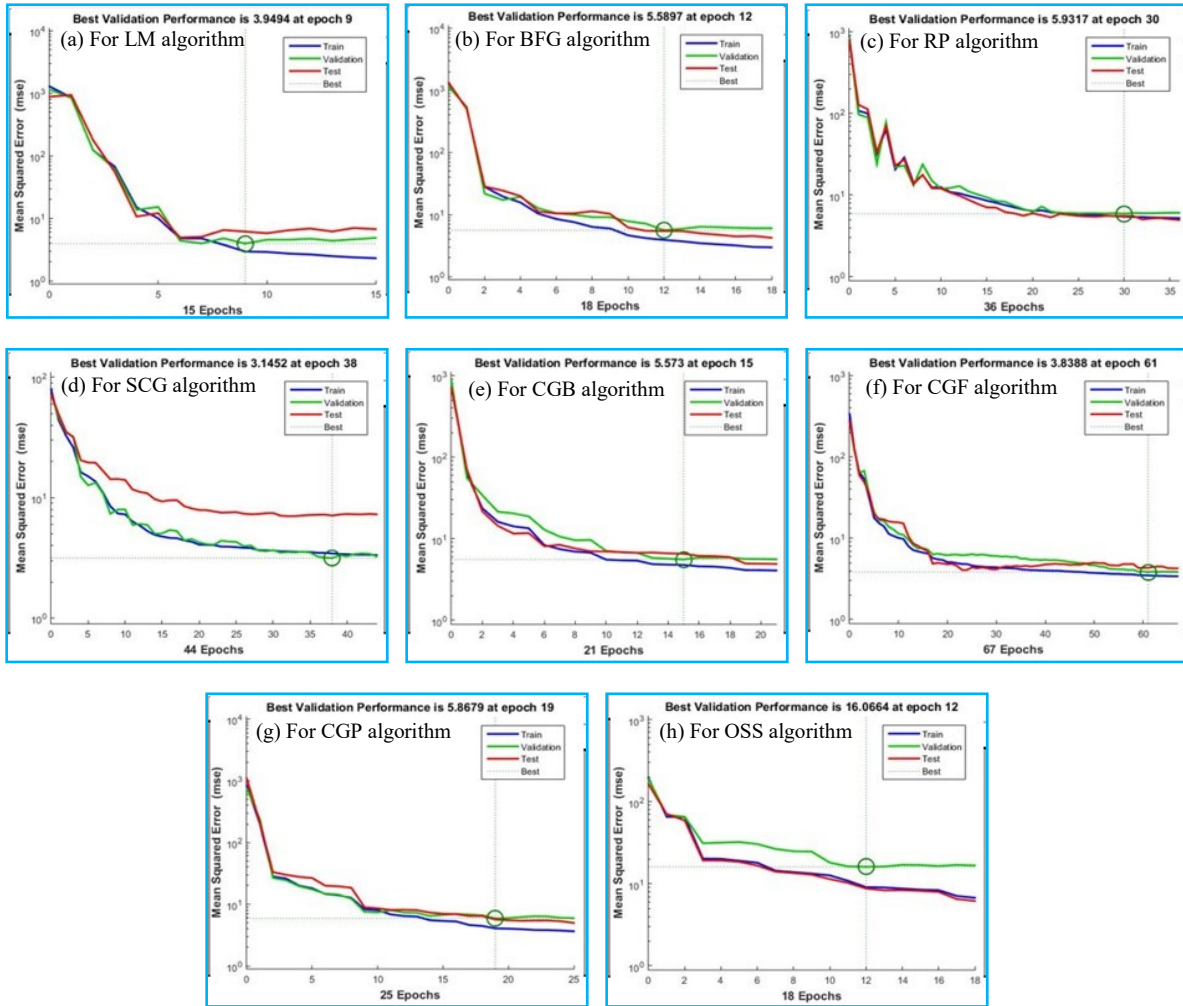
Figs. 5(a-h). Validation performance for different ANN algorithms for minimum temperature



Figs. 6(a-h). Correlation coefficients for maximum atmospheric temperature

Figs. 5(a-h) gives the Mean Squared Error (MSE) values for different algorithms in ANN. CGB and SCG algorithms have shown the lowest MSE values at convergence point. However, these two algorithms have

taken more epochs to converge to the least MSE when compared to other training algorithms. Nevertheless, algorithm LM has only taken 17 epochs to reach 5.5387 of MSE. The CGB algorithm has also performed averagely



Figs. 7(a-h). Validation performances for different ANN algorithms for maximum temperature

well in the validation process. It reached the convergence at 20 epochs while reaching a considerable low MSE (~ 4). However, the OSS algorithm has taken the greatest number of epochs (77) and it has the highest MSE (~ 7) at the convergence. Therefore, OSS was outperformed by the other algorithms. Therefore, the results show that the BFG algorithm has a better approach when compared to the others in predicting the minimum atmospheric temperature in Tabuk, Saudi Arabia. Nevertheless, CGB and SCG algorithms can also be considered alternative better approaches.

6. Maximum atmospheric temperature

Correlation coefficients for performance of ANN models in predicting the maximum temperature in Tabuk are shown in Figs. 6(a-h). However, slight reductions can be seen in the correlation coefficients of Figs. 6(a-h) when compared to those of at Figs. 3(a-h) (for minimum

atmospheric temperature). The maximum value of R value in minimum temperature prediction ANN is around 0.98; however, it is lower in the maximum temperature prediction. A similar observation can be seen for the minimum R values. The minimum atmospheric temperature prediction algorithms have the minimum R value of 0.935 while it is 0.865 in the maximum atmospheric temperature prediction algorithms. Therefore, this observation indicates that the regression plot of ANN models for predicting minimum atmospheric temperature are less scattered compared to that of in maximum atmospheric temperature. It can be clearly seen that the Figs. 6(a-h) give similar R values except in the OSS algorithm. OSS algorithm has shown lowered R values. However, among the others, LM and CGF algorithms have good performances.

Figs. 7(a-h) present the MSE values for the validation process of the different algorithms in maximum

atmospheric temperature prediction. Similar to the Fig. 5(a), LM algorithm shows the lowest number of epochs in convergence (9 epochs) in the maximum atmospheric temperature prediction. However, it also has the lower MSE value (~4) among the other algorithms. In addition, SCG (~3) and CGF (~4) algorithms have lower MSE values; however, they used greater number of epochs for the convergences (38 and 61, respectively). OSS algorithm converged faster compared to the other algorithms; however, it has the highest MSE value. Therefore, similar conclusions can be drawn in maximum atmospheric temperature prediction for OSS algorithm (similar to the minimum atmospheric temperature prediction). The OSS algorithm was outperformed by the other algorithms. Therefore, in general LM and CGF algorithms have performed well to predict the maximum atmospheric temperature for Tabuk, Saudi Arabia.

7. Conclusions

Eight different ANN algorithms were developed to predict the atmospheric temperature in Tabuk, Saudi Arabia. The algorithms used several weather parameters including, rainfall, relative humidity, wind speed, air pressure and atmospheric temperature. Results revealed that, in general, two different ANN algorithms have performed better for predicting minimum and maximum atmospheric temperature in Tabuk. They are BFG for minimum temperature and LM for maximum temperature. However, one common algorithm also performed better in minimum and maximum temperature prediction and that is SCG algorithm. Nevertheless, results show that all eight training algorithms have performed to an acceptable level because, all the correlation coefficients are greater than 0.85 and have acceptable MSE values for the validation processes. Therefore, it can be concluded herein that the atmospheric temperature forecasting process in Tabuk can be reliably done using the artificial neural networks. In addition, the weather forecasters have a choice of different algorithms to use in prediction. Among the other algorithms, LM, BFG and SCG algorithms are proposed as the preferred algorithms for the prediction.

Furthermore, the models are based on the real measured data. Therefore, they are applicable in the real-world cases. Not only applicable in real world cases, but also the models can be used to predict the future minimum and maximum atmospheric temperatures based on the various climate models' driven independent variables.

Acknowledgements

Authors would like to thank The University of Tabuk, Saudi Arabia for the support it has provided in obtaining the weather data. The contents and views

expressed in this research paper/article are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

- Abushandi, E. and Alatawi, S., 2015, "Dam Site Selection Using Remote Sensing Techniques and Geographical Information System to Control Flood Events in Tabuk City", *Hydrol. Current Res.*, **6**, 1, 1-13, doi: 10.4172/2157-7587.1000189.
- Acharya, N., Shrivastava, N. A., Panigrahi, B. K. and Mohanty, U. C., 2014, "Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India : An application of extreme learning machine", *Clim. Dynam.*, **43**, 5-6, 303-1310, doi: 10.1007/s00382-013-1942-2.
- Agatonovic-Kustrin, S. and Beresford, R., 2000, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research", *J. Pharmaceut. Biomed.*, **22**, 5, 717-727, doi: 10.1016/S0731-7085(99)00272-1.
- Al-Harbi, K. M., 2010, "Monitoring of agricultural area trend in Tabuk region - Saudi Arabia using Landsat TM and SPOT data", *Egyptian J. Rem. Sens. Spa. Sci. Elsevier B.V.*, **13**, 1, 37-42, doi: 10.1016/j.ejrs.2010.07.005.
- Andrei, N., 2007, "Scaled conjugate gradient algorithms for unconstrained optimization", *Comput. Optim. Appl.*, **38**, 3, 401-416, doi: 10.1007/s10589-007-9055-7.
- Azamathulla, H. M., Rathnayake, U. and Shatnawi, A., 2018, "Gene expression programming and artificial neural network to estimate atmospheric temperature in Tabuk, Saudi Arabia", *Appl. Water Sci.*, **8**, 6, 184-190, doi: 10.1007/s13201-018-0831-6.
- Barnett, J. and Adger, W. N., 2007, "Climate change, human security and violent conflict", *Polit. Geogr.*, **26**, 6, 639-655, doi: 10.1016/j.polgeo.2007.03.003.
- Beale, E. M. L., 1967, "Numerical Methods in Nonlinear Programming", J. Abadie (ed.), North-Holl and, Amsterdam, Netherland.
- Belayneh, A., Adamowski, J., Khalil, B. and Ozga-zielinski, B., 2014, "Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural network and wavelet support vector regression models", *J. Hydrol.*, **508**, 418-429, doi: 10.1016/j.jhydrol.2013.10.052.
- Betiku, E., Okunsolawo, S. S., Ajala, S. O. and Odedele, O. S., 2015, "Performance evaluation of artificial neural network coupled with generic algorithm and response surface methodology in modeling and optimization of biodiesel production process parameters from shea tree (*Vitellaria paradoxa*) nut butter", *Renew Energ.*, **76**, 408-417, doi: 10.1016/j.renene.2014.11.049.
- Cetisli, B. and Barkana, A., 2010, "Speeding up the scaled conjugate gradient algorithm and its application in neuro-fuzzy classifier training", *Soft Computing*, **14**, 4, 365-378, doi: 10.1007/s00500-009-0410-8.
- Colaco, M. J. and Orlande, H. R. B., 1999, "Comparison of different versions of the conjugate gradient method of function estimation", *Numer. Heat Tr. A-Appl.*, **36**, 2, 229-249, doi: 10.1080/104077899274859.
- Constantinescu, R., Lazarescu, V. and Tahboub, R., 2008, "Geometrical Form Recognition Using "One-Step-Secant" Algorithm in Case of Neural Network", *U.P.B. Sci. Bull., Series C*, **70**, 2, 15-28.

- Ebrahimi, H. and Rajaei, T., 2017, "Simulation of groundwater level variations using wavelet combined with neural network, linear regression and support vector machine", *Global Planet Change*, **148**, 181-191, doi: 10.1016/j.gloplacha.2016.11.014.
- Emeko, H. A., Olugbogi, A. O. and Betiku, E., 2015, "Appraisal of artificial neural network and response surface methodology in modeling and process variable optimization of oxalic acid production from cashew apple juice: A case of surface fermentation", *Bio. Resources*, **10**, 2, 2067-2082, doi: 10.15376/biores.10.2.2067-2082.
- Field, R. D., Luo, M., Kim, D., Del Genio, A. D., Voulgarakis, A. and Worden, J., 2015, "Sensitivity of simulated tropospheric CO to subgrid physics parameterization: A case study of Indonesian biomass burning emissions in 2006", *J. Geophys. Res. Atmos.*, **120**, 22, 11743-11759, doi:10.1002/2015JD023402.
- Friedrich, K., Theiler, G. and Klein, P., 2018, "Polymer composites for tribological applications in a range between liquid helium and room temperature", *Handbook of Polymer Tribology (Tier 2)*, 307-343, doi: 10.1038/nature01309.1.
- Ghaffari, A., Abdollahi, H., Khoshayand, M. R., Bozchalooi, I. S., Dadgar, A. and Rafiee-Tehrani, M., 2006, "Performance comparison of neural network training algorithms in modeling of bimodal drug delivery", *Int. J. Pharm.*, **327**, 1-2, 126-138, doi: 10.1016/j.ijpharm.2006.07.056.
- Güven, A. and Günel, M., 2008, "Prediction of scour downstream of grade-control structures using neural networks", *J. Hydraul. Engg.*, **134**, 11, 1656-1660, doi: 10.1061/(ASCE)0733-9429(2008)134:11(1656).
- Hagan, M. T. and Menhaj, M. B., 1994, "Training Feedforward Networks with the Marquardt Algorithm", *IEEE T. Neural Networks*, **5**, 6, 989-993, doi: 10.1006/brng.1996.0066.
- Hall, T., Brooks, H. E. and Doswell, C. A., 1999, "Precipitation Forecasting Using a Neural Network", *Weather Forecast*, **14**, 338-345, doi: 10.3389/fnagi.2018.00255.
- Harvell, D., Altizer, S., Cattadori, I. M., Harrington, L. and Weil, E., 2014, "Climate change and wildlife diseases: When does the host matter the most?", *Ecology*, **87**, 11, 2871-2881, doi: 10.1007/s10670-005-4007-z.
- Hashim, F. R., Daud, N. G. N., Ahmad, K. A., Adnan, J., Rizman, Z. I. and Camp, S. B., 2017, "Prediction of rainfall based on weather parameter using artificial neural network", *J. Fundam. Appl. Sci.*, **9**, 3S, 493-502.
- Hughes, T. P., Baird, A. H., Bellwood, D. R., Card, M., Connolly, S. R., Folke, C., Grosberg, R., Hoegh-Guldberg, O., Jackson, J. B. C., Kleypas, J., Lough, J. M., Marshall, P., Nyström, M., Palumbi, S. R., Pandolfi, J. M., Rosen, B. and Roughgarden, J., 2003, "Climate Change, Human Impacts and the Resilience of Coral Reefs", *Science*, **301**, 5635, 929-934, Available at: <http://www.uq.edu.au/news/index.phtml?article=4759>.
- Ibrić, S., Jovanović, M., Djurić, Z., Parojčić, J., Petrović, S. D., Solomun, L. and Stupar, B., 2003, "Artificial neural networks in the modeling and optimization of aspirin extended release tablets with Eudragit L100 as matrix substance", *AAPS Pharm. Sci. Tech.*, **4**, 1, 62-70, doi: 10.1208/pt040109.
- Ilonen, J. and Kamarainen, J., 2003, "Differential Evolution Training Algorithm for Feed-Forward Neural Networks", *Neural Process Lett.*, **17**, 1, 93-105.
- Izadifar, Z. and Elshorbagy, A., 2010, "Prediction of hourly actual evapotranspiration using neural networks, genetic programming and statistical models", *Hydrol. Process*, **24**, 23, 3413-3425, doi: 10.1002/hyp.7771.
- Japkowicz, N. and Hanson, S. J., 1999, "Adaptability of the backpropagation procedure; in IJCNN'99", International Joint Conference on Neural Networks, Proceedings (Cat. No. 99CH36339), Vol. 3, 1710-1715, doi: 10.1109/IJCNN.1999.832633.
- Kalkstein, L. S. and Smoyer, K. E., 1993, "The impact of climate change on human health: Some international implications", *Experientia*, **49**, 11, 969-979, doi: 10.1007/BF02125644.
- Karl, T. R., 2003, "Modern Global Climate Change", *Science*, **302**, 5651, 1719-1723, doi: 10.1126/science.1090228.
- Kisi, O. and Uncuoglu, E., 2005, "Comparison of three back-propagation training algorithms for two case studies", *Indian J. Engg. Mater. Sci.*, **12**, 5, 434-442.
- Kisi, O., 2004, "River Flow Modeling Using Artificial Neural Networks", *J. Hydrol. Eng.*, **9**, 1, 60-63, doi: 10.1061/(ASCE)1084-0699(2004)9:1(60).
- Kisi, Ö., 2007, "Streamflow Forecasting Using Different Artificial Neural Network Algorithms", *J. Hydrol. Engg.*, **12**, 5, 532-539.
- Landeras, G., López, J. J., Kisi, O. and Shiri, J., 2012, "Comparison of Gene Expression Programming with neuro-fuzzy and neural network computing techniques in estimating daily incoming solar radiation in the Basque Country (Northern Spain)", *Energ. Convers. Manage.*, **62**, 1-13, doi: 10.1016/j.enconman.2012.03.025.
- Lee, S., Cho, S. and Wong, P. M., 1998, "Rainfall Prediction Using Artificial Neural Networks", *J. Geogr. Inf. Dec. Ana.*, **46**, 2, 233-242, doi: 10.1002/jso.20181.
- Lera, G. and Pinzolas, M., 2002, "Neighborhood based Levenberg-Marquardt algorithm for neural network training", *IEEE T. Neural Networ.*, **13**, 5, 1200-1203, doi: 10.1109/TNN.2002.1031951.
- Liang, M., Wang, S. X. and Luo, Y. H., 1994, "Fast learning algorithms for multi-layered feed-forward neural network", Proceedings of National Aerospace and Electronics Conference (NAECON'94), 787-790.
- Lourakis, M. I. A., 2005, "A Brief Description of the Levenberg-Marquardt Algorithm Implemented by levmar", *Technical Report-Institute of Computer Science*, 1-5, doi: 10.1016/j.ijinfomgt.2009.10.001.
- Martin, M., 1997, "Efficient Training of Feed-Forward Neural Networks", Ph. D. thesis, Aarhus University, Denmark.
- Mears, C. A. and Wentz, F. J., 2017, "A satellite-derived lower-tropospheric atmospheric temperature dataset using an optimized adjustment for diurnal effects", *J. Climate*, **30**, 19, 7695-7718, doi: 10.1175/JCLI-D-16-0768.1.
- Mishra, N., Soni, H. K., Sharma, S. and Upadhyay, A. K., 2018, "Development and Analysis of Artificial Neural Network Models for Rainfall Prediction by Using Time-Series Data", *Int. J. Intell. Syst. Appl.*, **1**, 16-23, doi: 10.5815/ijisa.2018.01.03.
- Moghimi, S. and Bras, R. L., 2017, "Bias Correction of Climate Modeled Temperature and Precipitation Using Artificial Neural Networks", *J. Hydrometeorol.*, **18**, 1867-1884, doi:10.1175/JHM-D-16-0247.1.
- Moller, M. F., 1993, "A scaled conjugate gradient algorithm for fast supervised learning", *Neural Networks*, **6**, 4, 525-533, doi: 10.1016/S0893-6080(05)80056-5.
- Moré, J. J., 1978, "The Levenberg-Marquardt algorithm: Implementation and theory", In: Watson G.A. (eds) Numerical Analysis. Lecture Notes in Mathematics, **630**, 105-116, Springer, Berlin, Heidelberg.

- Nayak, P. C., Sudheer, K. P., Rangan, D. M. and Ramasastri, K. S., 2004, "A neuro-fuzzy computing technique for modeling hydrological time series", *J. Hydrol.*, **291**, 1-2, 52-66, doi: 10.1016/j.jhydrol.2003.12.010.
- Nocedal, J. and Wright, S. J., 1999, "Numerical optimization", Edited by P. Glynn and S. M. Robinson, Springer-Verlag, New York, doi: 10.1007/BF01068601.
- Nocedal, J., 1980, "Updating quasi-Newton matrices with limited storage", *Math. Comput.*, **35**, 151, 773-773, doi: 10.2307/2006193.
- Patz, J. A., Campbell-Lendrum, D., Holloway, T. and Foley, J. A., 2016, "Impact of climate change on human health", *Nature*, **438**, 17, 310-317, doi: 10.1007/978-3-319-16751-0_53.
- Pham, D. T. and Sagioglu, S., 2001, "Training multilayered perceptrons for pattern recognition: A comparative study of four training algorithms", *Int. J. Mach. Tool. Manu.*, **41**, 3, 419-430, doi: 10.1016/S0890-6955(00)00073-0.
- Powell, M., 1977, "Restart procedures for the conjugate gradient method", *Math Program*, **12**, 1, 241-254.
- Quadrelli, R. and Peterson, S., 2007, "The global climate challenge: Recent trends in CO₂ emissions from fuel combustion", *Energ. Policy*, **35**, 5938-5952, doi: 10.1016/j.enpol.2007.07.001.
- Ravansalar, M., Rajaei, T. and Kisi, O., 2017, "Wavelet-linear genetic programming: A new approach for modeling monthly streamflow", *J. Hydrol.*, **549**, 461-475, doi: 10.1016/j.jhydrol.2017.04.018.
- Rotstayn, L. D., Plymin, E. L., Collier, M. A., Boucher, O., Dufresne, J. L., Luo, J. J., Von Salzen, K., Jeffrey, S. J., Foujols, M. A., Ming, Y. and Horowitz, L. W., 2014, "Declining aerosols in CMIP5 projections: Effects on atmospheric temperature structure and midlatitude jets", *J. Climate*, **27**, 18, 6960-6977, doi: 10.1175/JCLI-D-14-00258.1.
- Saini, L. M. and Soni, M. K., 2002, "Artificial Neural Network-Based Peak Load Forecasting Using Conjugate Gradient Methods", *IEEE T. Power Syst.*, **17**, 3, 907-912.
- Saini, L. M., 2008, "Peak load forecasting using Bayesian regularization, Resilient and adaptive backpropagation learning based artificial neural networks", *Electr. Pow. Syst. Res.*, **78**, 7, 1302-1310, doi: 10.1016/j.eprsr.2007.11.003.
- Schraudolph, N. N., Yu, J. and Gunter, S., 2018, "A Stochastic Quasi-Newton Method for Online Convex Optimization", *Eur. J. Radiol.*, **108**, 7-12, doi: 10.1016/j.ejrad.2018.09.011.
- Shaheed, M. H., 2004, "Performance analysis of 4 types of conjugate gradient algorithms in the nonlinear dynamic modelling of a TRMS using feedforward neural networks", Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics, **6**, 5985-5990, doi: 10.1109/ICSMC.2004.1401153.
- Shiri, J., Kim, S. and Kisi, O., 2014b, "Estimation of daily dew point temperature using genetic programming and neural networks approaches", *Hydrol. Res.*, **45**, 2, 165, doi: 10.2166/nh.2013.229.
- Shiri, J., Marti, P. and Singh, V. P., 2014a, "Evaluation of gene expression programming approaches for estimating daily evaporation through spatial and temporal data scanning", *Hydrol. Process*, **28**, 3, 1215-1225, doi: 10.1002/hyp.9669.
- Simmons, A. J., Poli, P., Dee, D. P., Berrisford, P., Hersbach, H., Kobayashi, S. and Peubey, C., 2014, "Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim", *Q. J. Roy. Meteor. Soc.*, **140**, 679, 329-353, doi: 10.1002/qj.2317.
- Upadhyay, D., 2013, "Classification of EEG signals under different mental tasks using wavelet transform and neural network with one step secant algorithm", *Int. J. Sci. Engg. Tech.*, **2**, 4, 256-259, Available at: <http://ijset.com/ijset/publication/v2s4/paper14.pdf>.
- Vorbismarty, C. J., Green, P., Salisbury, J. and Lammers, R. B., 2000, "Global Water", *Science*, **289**, 284-288, doi:10.1126/science.289.5477.284 Resources : Vulnerability from Climate Change and Population Growth.
- Wang, X., Liu, F., Gao, Y., Xue, C., Li, Robert W. T. and Qing, J., 2018, "Transcriptome analysis revealed anti-obesity effects of the Sodium Alginate in high-fat diet -induced obese mice", *Int. J. Biol. Macromol.*, **115**, 861-870, doi: 10.1016/j.ijbiomac.2018.04.042.
- Yadav, A. K. and Chandel, S. S., 2013, "Solar radiation prediction using Artificial Neural Network techniques : A review", *Ren. Sus. Energ. Rev.*, **33**, 772-781, doi: 10.1016/j.rser.2013.08.055.