A Comparative Performance Model of Machine Learning Classifiers on Time Series Prediction for Weather Forecasting

Sudhir Sharma, Kaushal Kishor Bhatt, Rimmy Chabra, and Nagender Aneja

Abstract Machine learning is a booming technical term in every domain of research. The majority of the technical concepts sounds to accomplish classification task in a real-life scenario. In the literature, the huge number of classification tools, it becomes very necessary to justify the performance of machine learning classifiers. This paper describes four classification techniques that are successfully applied for the prediction of the two most significant features for weather forecasting temperature and relative humidity (RH). A brief introduction of the proposed model with four prediction methodologies—ARMA, MLP, SVM and ELANFIS—follows the discriminate ideas that can create the space for such research. The techniques are then compared on a public data set containing the time series of the two parameters: temperature and relative humidity. As per the data statistics, the parameters are registered on an hourly basis and recorded over a field in an Italian city. An elaborating analysis of the results is performed to provide insights into the satisfactory performance of the models.

Keywords Time series prediction · Weather forecasting · Support vector machine (SVM) · Multi-layer perceptron (MLP)

S. Sharma (✉)
Department of Information Technology, Manipal University, Jaipur, Rajasthan 302004, India
e-mail: sudhir33.04@bitmesra.ac.in

K. K. Bhatt
Birla Institute of Applied Sciences, Nainital 263136, India
e-mail: kaushalbhatt@birlainstitute.co.in

R. Chabra
College of Engineering Roorkee, Roorkee 247667, India
e-mail: rimmy@coer.ac.in

N. Aneja
School of Digital Science, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei Darussalam
e-mail: nagender.aneja@ubd.edu.bn

V. Goar et al. (eds.), *Advances in Information Communication Technology and Computing*, Lecture Notes in Networks and Systems 392,
https://doi.org/10.1007/978-981-19-0619-0_50

1 Introduction

Weather forecasting is one field which has emerged greatly in the past few years. It finds huge application in the lives of most people who look to plan their activities accordingly. Two of the most important parameters of weather prediction are temperature and RH which contain the majority of the information required for the prediction. The relevance of these features is evident from the fact that most forecasting stations provide the future values of only temperature and RH values. Thus, considering their practical application, temperature and RH data set is selected to test the four prediction models. Time series prediction (TSP) has a great demand in a number of fields which can be mainly attributed to its increasing importance in practical applications there. This has led to a lot of research in TSP [1], and as a result, various methods have evolved over the years. This paper deals with some of the popular techniques present in literature. The first model used is ARMA [16] which consists of two interlocked parts that operate simultaneously: first an autoregressive (AR) part and second a moving average (MA) part. The function of the first part is to regress the variable on its own lagged values, i.e. express the variable as a linear function of its own past values. The MA part performs the same role for the error term. It models the error term as a linear combination of various past values of itself. The second model used is artificial neural networks (ANNs) [7]-based multilayer perception (MLP). Their inherent capability of nonlinear modelling without any knowledge of statistics is the main reason for their popularity. Another powerful tool which has been employed is support vector machine (SVM). An amazing feature of SVM [6, 11] is that it only yields good classification results but also generalizes well to new data. The final method used for modelling is extreme learning adaptive neuro-fuzzy inference system (ELANFIS) [10] which represents a class of systems in which the learning potential of neural networks is combined with the expressiveness of the fuzzy logic and implemented using the time-saving approach of extreme learning.

The rest of the paper is organized as follows. Section 2 provides a brief overview of some of the prior works in this field. Section 3 describes the four models which have been compared over the acquired data set. It contains a brief introduction to the algorithms of those models. The next section illustrates the results of the time series prediction on the test data and provides a comparative discussion on the four aforementioned methods in terms of various errors. The last section concludes the paper with the inferences drawn from the analyses carried out.

2 Related Work

There has been a large body of work in the field of weather forecasting and modelling the time series of some of the primary metrics related to weather. Most of the research efforts in this domain have been to try and forecast different temperatures [4, 15, 19] such as minimum temperature, daily dew point temperature, ground temperature and indoor temperature in different parts of the world with the help of different input data variables. Some of them look to perform spectral analysis of climate indices [4], while others use infrared sounder observations [2] to make their predictions. There are other research works that look to model and predict other metrics such as humidity [2, 14, 18], or global solar radiation [17], or reference evapotranspiration [15]. The popular classification techniques used in their work have been considered major state of the arts in the comparative studies of machine learning domain. They used variants of the SVM such as LS-SVM and fuzzy LS-SVM, extreme learning machine, ELANFIS, different modifications of ANNs and other adaptive neuro-fuzzy inference systems.

3 Proposed Work

3.1 Auto-regressive Moving Average (ARMA)

The theory presented in [16] states that any linear stationary process can be modelled as follows:

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_m Y_{t-m} + \epsilon_t - \beta_1 \epsilon_{t-1} - \beta_2 \epsilon_{t-2} - \cdots - \beta_n \epsilon_{t-n} \tag{1}$$

where Y_t is the processed output of the forecasting system, Y_{t-i} are uncorrelated input variables, α and β are model coefficients, and ϵ_{t-j} is random error term associated with the forecasting process [8].

Input variables are basically past values of the same process; this means to say the system works on feedback policy. The total number of nonzero terms to be included in the model and the value of the model coefficients are determined in the following algorithmic steps:

1. Calculate the auto-correlation function (ACF) and partial auto-correlation (PACF) [13] for the process.
2. Try to match the above estimated functions with ACFs and PACFs of standard theoretical functions available. Some of the standard theoretical models are AR(1), AR(2), MA(1), MA(2), ARMA(1, 1), etc. Here (m, n) indicates number of input variables and random error terms used.
3. Take the best matching model. Thus, the number of terms in series has been determined.
4. To determine the coefficients α and β, we equate estimated ACF to theoretical ACF. This will give us one equation corresponding to each coefficient. Determine the coefficients from the equations.

After the successful follow-up of the mentioned algorithmic steps, the proposed model is ready to predict the time series values.

3.2 *Multi-layer Perceptron (MLP)*

Perceptrons are analogous to neurons in human brain. A network of more than one layer of neurons is MLP [13]. MLP learns from the experience obtained by analysing input data. The processed features from input data are fed to the network, based on which the desired model is learnt in adaptive fashion. In most of the practical time series, the past and the future values are highly correlated. Thus, the MLP can be trained on some selective time lagged samples of the same time series for which the prediction is to be done. The number of lags to be used for a prediction is decided from ACF and PACF plots of data with different time lags. Those time lags, at which ACF and PACF values are above a predecided threshold, are taken for prediction. Input samples to be considered are determined accordingly, and MLP is trained on that data.

3.3 *Support Vector Machine (SVM)*

SVM is one of the most recent forecasting methods with its foundation in statistical learning. It is based on the principle of structural risk minimization (SRM). The input and output format is same as in MLP, only the training method is different. Training SVM [6] involves optimization of a quadratic cost function. We can increase the performance of our SVM by transforming input data to higher dimension. We have used LIBSVM library [3] for prediction using SVM.

3.4 *Extreme Learning Adaptive Neuro-Fuzzy Inference System (ELANIS)*

Adaptive neuro-fuzzy inference system (ANFIS) [10] is a very popular learning algorithm derived from a combination of two of the most basic learning architectures. It leverages the capability of the fuzzy inference systems to store knowledge and unites it with the power of neural networks to adapt resulting in a much more powerful learning system. The initial parameters of the fuzzy inference system are trained using back propagation which makes ANFIS highly dependent on learning through gradient-based methods. This leads to slower training and thus more cost computationally. With the idea of making the training of the network faster, Hunag introduced the extreme learning machine (ELM) [9] which modifies the learning method in a traditional single-layer feed forward neural network. In order to learn, ELM looks to find a solution for the weights from the hidden layer to the output layer after randomly projecting the input variables. It uses Moore–Penrose pseudo-inverse to find a solution that is not only minimum norm but also minimum error. Extreme learning ANFIS (ELANFIS) is a hybrid of ANFIS and ELM. It combines the fast learning method of ELM with the architecture of ANFIS to tune the parameters of a Sugeno-type fuzzy system. ELANFIS involves computing the consequent parameters with the help of pseudo-inverse method as in ELM while the premise parameters are assumed randomly.

4 Results and Discussion

The hourly data of Temperature and RH is obtained from the UCI repository [5]: Air Quality Data Set. This data was recorded by a multi-sensor device deployed on the field in an Italian city from November 2004 to April 2005. This provided a data set of 10,000 instances. However, only the first 2200 values were utilized for further processing because of the constraint of computation time. The first 2000 points have been used for training, including parameter calculation, and the rest 200 for testing in all the models. The prediction results for the same have been discussed ahead. It is clear from the plots that only two lags are significant for prediction, namely 1 and 4. Hence, out of the various ARMA models, the one which gave the best results is AR(2) which involves only the α coefficients and not the β coefficients. The values of the parameters α_1 and α_2 are obtained for lags 1 and 4, respectively, which are 0.77 and 0.32. Figure 1a shows the prediction results for instances 2001–2200 of the temperature data using the above model. Figure 1b shows the same results but for humidity data. The results from the MLP model are illustrated in Fig. 2 which contain the results when first four lags, i.e. previous four values of the time series, were randomly chosen for prediction of temperature and RH, respectively. Again, Fig. 2a, b shows the same results by choosing only first and fourth lags with marked improvement in accuracy.

The prediction using SVM is shown ahead in Fig. 3 which shows the result for temperature when last seven values were taken as input data for predicting the eighth value. The number of lags was chosen randomly. Figure 3a, b demonstrates the same with lags chosen as 1 and 4 specifically. This resulted in better prediction, as is evident from the figures, with the same argument as above. Figure 4 contains the result obtained using ELANFIS which are the best so far in terms of error. This can be mainly attributed to the capabilities of fuzzy logic combined with the learning ability of a neural network together in an inference system. This combination yields a very powerful tool for time series prediction as is proven in Fig. 4a containing the temperature prediction and Fig. 4b containing the RH forecasting result (Tables 1 and 2).

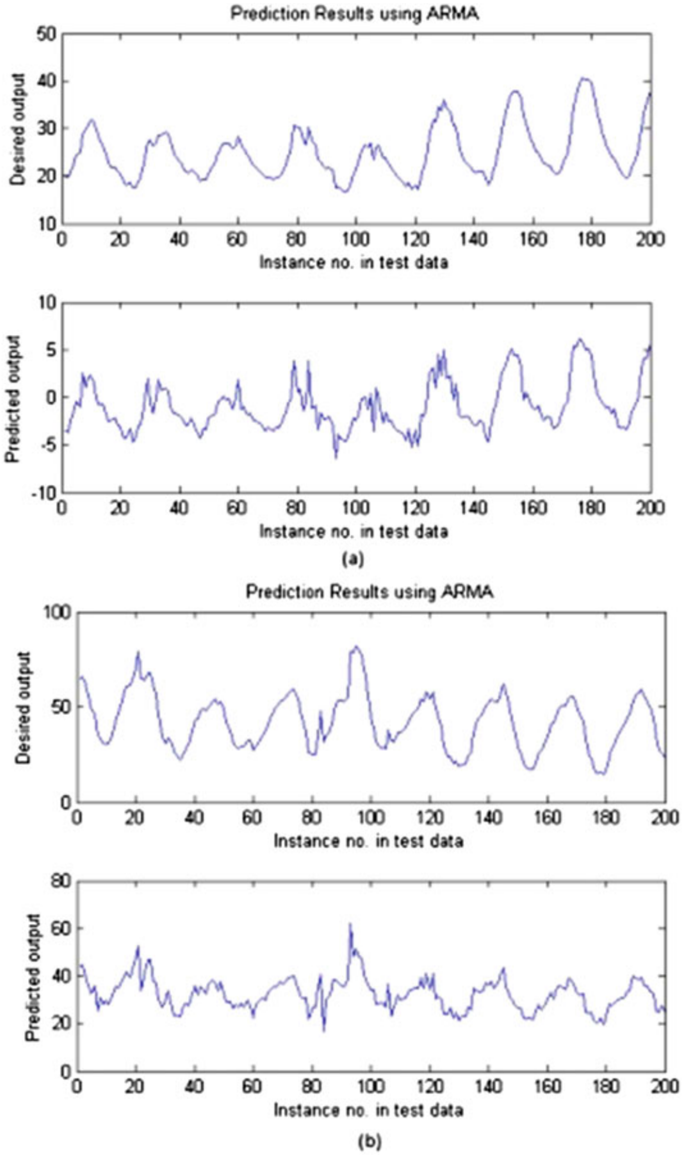

Fig. 1 Prediction results using ARMA model for **a** temperature, **b** RH

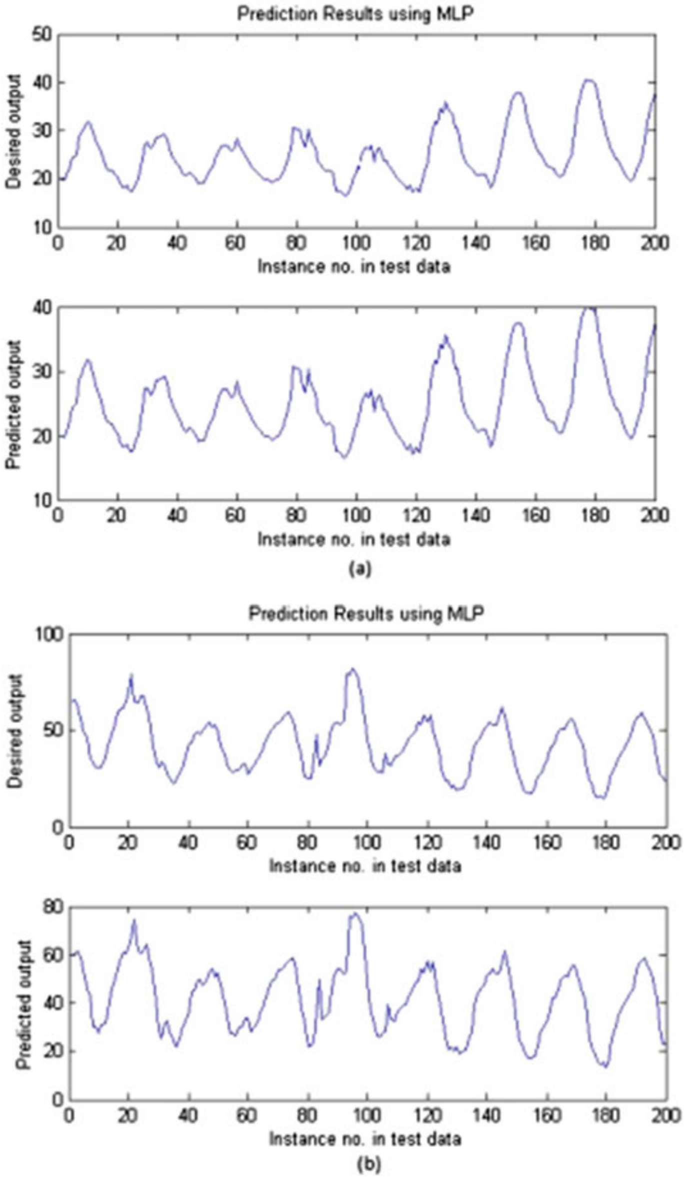

Fig. 2 Prediction results using MLP with first and fourth lags for **a** temperature, **b** RH

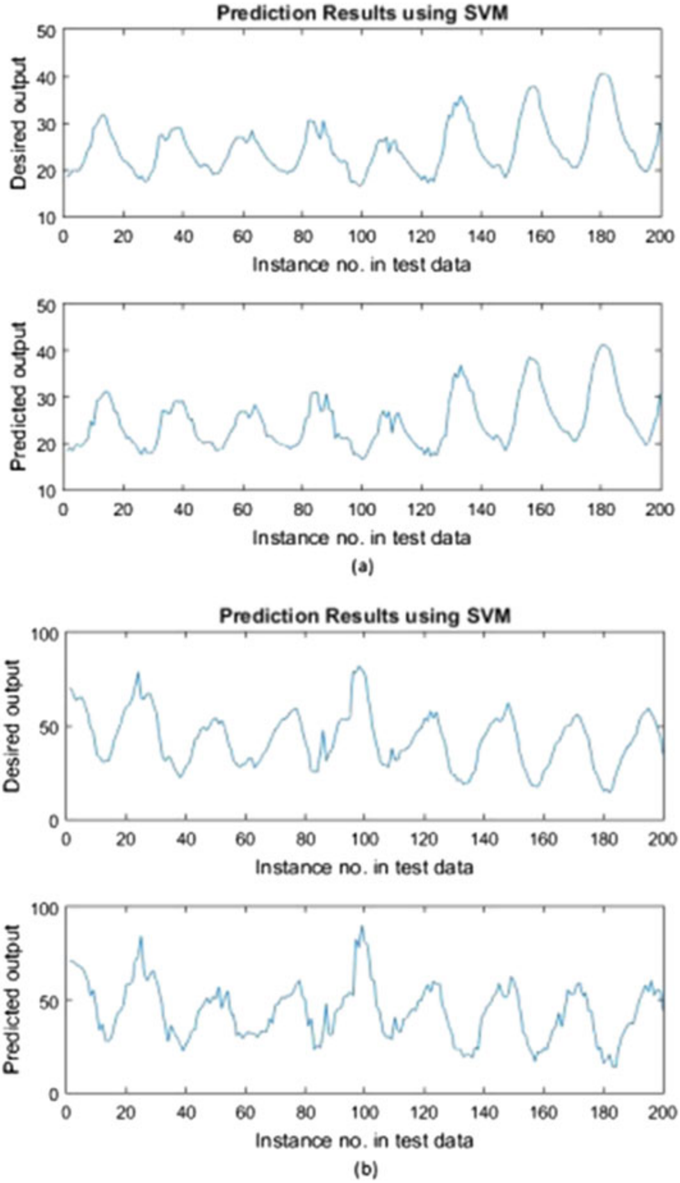

Fig. 3 Prediction results using SVM with first and fourth lags for **a** temperature, **b** RH

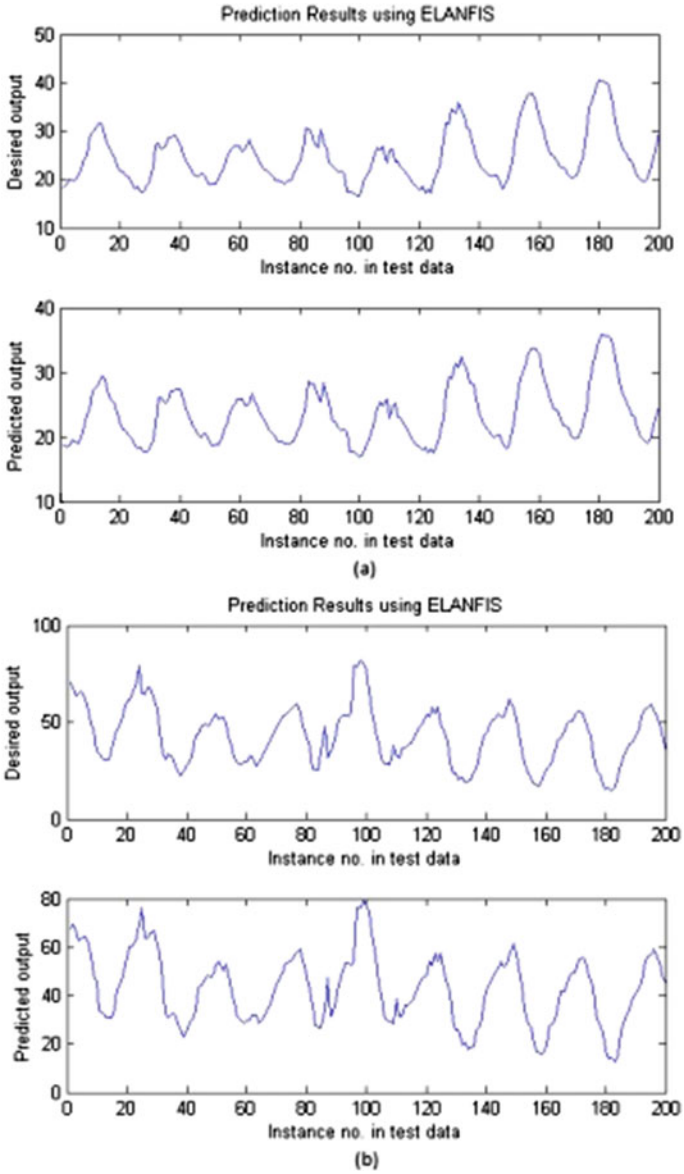

Fig. 4 Prediction results using ELANFIS with first and fourth lags for **a** temperature, **b** RH

Table 1 Comparison of the four classification models over three types of error for temperature

	ARMA	MLP	SVM	ELANFIS
RMSE	25.03	0.70	13.01	0.20
MAE	27.92	0.69	8.79	0.31
MAP	0.37	0.07	0.23	0.003

Table 2 Comparison of the four classification models over three types of error for humidity

	ARMA	MLP	SVM	ELANFIS
RMSE	19.58	4.44	6.39	1.63
MAE	23.96	5.02	5.86	2.39
MAP	0.38	0.12	0.16	0.09

5 Conclusion and Future Recommendations

The paper successfully achieves the objective of providing a comparative study amongst the four popular time series prediction models. The models were trained and tested on small and separate subsets of the complete data set acquired from UCI repository. Temperature and relative humidity were chosen for data selection because of their practical application in weather forecasting. Three different types of errors were evaluated to convincingly prove the superiority of some methods over others. The prediction results along with the original time series were also plotted for better visualization of the prediction accuracy.

The future recommendation of this work is looked for developing a better classifier which may effectively tune with deep data analytics [12] of spatiotemporal data statistics on several parameters such as snow, solar radiation soil moisture and different atmospheric pressure.

References

1. Adhikari, R., Agrawal, R.K.: An introductory study on time series modeling and forecasting (2013). arXiv:1302.6613
2. Ajil, K.S., Thapliyal, P.K., Shukla, M.V., Pal, P.K., Joshi, P.C., Navalgund, R.R.: A new technique for temperature and humidity profile retrieval from infrared-sounder observations using the adaptive neuro-fuzzy inference system. IEEE Trans. Geosci. Remote Sens. **48**(4), 1650–1659 (2010)
3. Chang, C.C., Lin, C.J.: Libsvm: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) **2**(3), 1–27 (2011)
4. Daneshmand, H., Tavousi, T., Khosravi, M., Tavakoli, S.: Modeling minimum temperature using adaptive neuro-fuzzy inference system based on spectral analysis of climate indices: A case study in iran. J. Saudi Society Agric. Sci. **14**(1), 33–40 (2015)
5. Dua, D., Graff, C., et al.: Uci machine learning repository (2017)

6. Fan, Y., Li, P., Song, Z.: Dynamic least squares support vector machine. In: 2006 6th World Congress on Intelligent Control and Automation, vol. 1, pp. 4886–4889. IEEE (2006)
7. Hamzaçebi, C.: Improving artificial neural networks-performance in seasonal time series forecasting. Inf. Sci. **178**(23), 4550–4559 (2008)
8. Hipel, K.W., McLeod, A.I.: Time series modelling of water resources and environmental systems. Elsevier (1994)
9. Huang, G.B., Zhu, Q.Y., Siew, C.K.: Extreme learning machine: theory and applications. Neurocomputing **70**(1–3), 489–501 (2006)
10. Jang, J.S.: Anfis: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. **23**(3), 665–685 (1993)
11. Kumar, N.: Better performance in human action recognition from spatiotemporal depth information features classification. In: Computational Network Application Tools for Performance Management, pp. 39–51. Springer (2020)
12. Kumar, N.: Recent issues with machine vision applications for deep network architectures. In: Cognitive Computing Systems, pp. 267–284. Apple Academic Press (2021)
13. Li, G., Shi, J.: On comparing three artificial neural networks for wind speed forecasting. Appl. Energy **87**(7), 2313–2320 (2010)
14. Martínez-Martínez, V., Baladrón, C., Gomez-Gil, J., Ruiz-Ruiz, G., Navas-Gracia, L.M., Aguiar, J.M., Carro, B.: Temperature and relative humidity estimation and prediction in the tobacco drying process using artificial neural networks. Sensors **12**(10), 14004–14021 (2012)
15. Mohammadi, K., Shamshirband, S., Motamedi, S., Petković, D., Hashim, R., Gocic, M.: Extreme learning machine based prediction of daily dew point temperature. Comput. Electron. Agric. **117**, 214–225 (2015)
16. Rojas, I., Valenzuela, O., Rojas, F., Guillén, A., Herrera, L.J., Pomares, H., Marquez, L., Pasadas, M.: Soft-computing techniques and ARMA model for time series prediction. Neurocomputing **71**(4–6), 519–537 (2008)
17. Shamshirband, S., Mohammadi, K., Chen, H.L., Samy, G.N., Petković, D., Ma, C.: Daily global solar radiation prediction from air temperatures using kernel extreme learning machine: a case study for Iran. J. Atmos. Solar-Terrestrial Phys. **134**, 109–117 (2015)
18. Suryono, S., Saputra, R., Surarso, B., Sukri, H.: Web-based fuzzy time series for environmental temperature and relative humidity prediction. In: 2017 IEEE International Conference on Communication, Networks and Satellite (Comnetsat), pp. 36–41. IEEE (2017)
19. Zhou, S., Chu, X., Cao, S., Liu, X., Zhou, Y.: Prediction of the ground temperature with ann, ls-svm and fuzzy ls-svm for gshp application. Geothermics **84**, 101757 (2020)