DOI: 10.24996/ijs.2024.65.10(SI).5

ISSN: 0067-2904

Estimation of some Climatological Parameters by WEKA Software for Selective Regions in Iraq

Dher I. Bakr[1], Jasim Al-Khalidi[1], Abas Wisam Mahdi Abas[2], Layla A. ALHak[1]

[1] *Department of Physics , College of Sciences, University of Diyala, Diyala, Iraq.*

[2] *Presidency of Diyala University, University of Diyala, Diyala, Iraq*

Received: 1/10/2023 Accepted: 26/6/2024 Published: 15/11/2024

Abstract

An Estimation is a direct summation that can produce new data from past measurements. In this study, we confirm the possibility of using the (WEKA) program in estimating the monthly values of some climatological parameters and investigate the influence of the time series' length parameters on the accuracy of estimation for selected regions of Iraq. Satellite data were used, which represent the monthly values for each of the minimum and maximum temperature, wind velocity, and relative humidity for the 1981 – 2021 time period, the absolute error rate (MAE), and the square root of the error rate (RMSE) with the correlation (R^2) are also identified to test the confidence of the prediction. Using (WEKA) software, which depends only on the time series of the parameters in the estimation, 12 months were estimated for the mentioned parameters. The estimated values by WEKA were close to the satellite data, thus we can depend on the software as a good source of meteorological data. Also, the study indicates that the time series' length strongly affects the accuracy of the estimation, as the increase in the time series of weather parameters increases the estimate's accuracy, and this increase fluctuates among parameters and varies from one region to another.

Keywords: Estimation, temperature, WEKA, Mean Absolute Error.

تخمين بعض المتغيرات المناخية باستخدام برنامج WEKA لمناطق مختارة في العراق

ذر انتصار بكر[1] ، جاسم محمد خليل[1]، وسام مهدي عباس[2] ، ليلى عبد الحق اسماعيل[1]

[1] قسم علوم الفيزياء ، كلية العلوم، جامعة ديالى ، ديالى العراق

[2] رئاسة جامعة ديالى، جامعة ديالى، ديالى، العراق

الخلاصة

التخمين هو محصلة مباشرة يمكن انتاجها من البيانات السابقة المقاسة. تهدف هذه الدراسة الى بيان إمكانية استخدام برنامج (WEKA) في تخمين القيم الشهرية لبعض العناصر الجوية لمناطق مختارة من العراق مع دراسات تأثير طول السلسلة الزمنية للعنصر على دقة تخمين العناصر الجوية. في هذه الدراسة تم استخدام بيانات الأقمار الاصطناعية والتي تمثل القيم الشهرية لكل من درجة الحرارة الصغرى والعظمى ، سرعة الرياح و الرطوبة النسبية للفترة (1981 – 2021) بالإضافة الى معايير الخطأ وهما معدل الخطأ المطلق(MAE) والجذر التربيعي لمعدل الخطأ(RMSE) مع الارتباط (R^2) لاختبار دقة التنبؤ. باستخدام برنامج (WEKA) الذي يعتمد فقط على السلاسل الزمنية للعناصر الجوية في التخمين ، تم تخمين 12 شهرًا للعناصر المذكورة.

*Email: dharentisar78@gmail.com

باستخدام برنامج (WEKA) كانت القيم المخمنة مقاربة من القيم للقمر الاصطناعي، وبالتالي يمكننا الاعتماد على البرنامج كمصدر جيد لبيانات الأرصاد الجوية. كما أشارت الدراسة إلى أن طول السلسلة الزمنية له تأثير قوي على دقة التخمين، حيث تؤدي زيادة السلاسل الزمنية للعناصر الجوية إلى زيادة دقة التخمين، وهذه الزيادة متذبذبة بين العناصر الجوية وتختلف من منطقة إلى أخرى.

1.Introduction

Estimating in common can be defined as a value assessment process of some variable in future. Estimating is a significant difficult process used by the management and business for planning and choice building, to protect life and property, and by every individual to carry out life activities [1].

Time series Estimating is a crucial instrument for meteorological and environmental applications, including stream flow, temperature, humidity, and rainfall. This method involves using historical data to create the most accurate model possible for predicting future values; in other words, it's a way to determine future values by comprehending historical data [2]. The most important indicators of climate change are changes in temperature, precipitation, relative humidity, and wind speed, Therefore, an accurate forecast of future climate parameters is necessary for preventing the country from natural hazards and managing natural resources [3]. Given the determinant impact of the environment on biological systems, evaluations of the spatial circulation of climatic factors are required like never before for the reasonable administration of regular assets. Deciding spatial environment conditions isn't difficult because drawn-out typical climate perceptions come from scanty, discrete and unpredictable dispersed meteorological stations [4] The world's climate is changing due to several elements coming from indoors and outdoors of the climate system [5]. Although climate factors can be detected practically everywhere, climate location networks are still inadequate and/or restricted in some regions of the world because of subjects relevant to natural features and at times economics [6].

A forecast or estimating model is based on the presumption that an underlying (known or unknown) association prevail between the outputs and the inputs. Due to the complexity of the actual estimation system, where inputs have the past values of the time series and/or other related variables [7].

Measurement of Air temperature at 2 m over the earth's surface in metropolitan scenes is firmly connected with different urban problems like the metropolitan intensity island impact [8], air pollution [9], and human mortality [10]. Because they are directly linked to fatal disasters like heat waves and tropical nights, monitoring and forecasting urban maximum and minimum air temperatures (T_{max} and T_{min}, respectively) are especially important in densely populated areas. [11]. The future energy landscape will be significantly influenced by wind energy. By 2020, wind energy could provide approximately 10% of the country's electricity when the wind's kinetic energy is converted into electrical energy by wind turbines [12,13,14]. This type of energy is considered one of the most rapidly expanding and environmentally friendly renewable energy sources.

The amount of water in the air is referred to as air humidity. Air humidity is significant for the advancement of plants and keeps the vast majority of the radiation from the sun and reflected sunbeams from arriving at the ground, hence forestalling unreasonable warming or cooling [15]

2. Methodology

2.1 Data source

The data utilized in this study is a historical series of monthly values from 1981 to 2021 for the atmospheric parameters, which are the minimum and maximum temperature, wind velocity and relative humidity.

The meteorological data used in this study is from a historical series of four stations indicating four provinces distributed throughout Iraq, as seen in Figure (1). The Data was obtained from NASA satellite.

2.2 Study area

In this research, The study regions consist of six stations, representing six provinces, namely Baghdad, Nineveh, Anbar, and Basra.

These stations differ among themselves in terms of geological characteristics (mountainous, plains, plateau and desert), vegetation cover, soil nature and climate (Mediterranean, steppe, desert and semi-desert), and their latitude, longitude, and elevation above sea level differ, as clarified in Table 1.

Table 1: Shows the geography of the study stations.

Stations	Altitude (m)	Longitude (deg)	Latitude (deg)
Baghdad	34	44.3661	33.3152
Nineveh	223	43.1577	36.3489
Basra	2.4	47.8164	30.5008
Anbar	375	43.3	33.4167

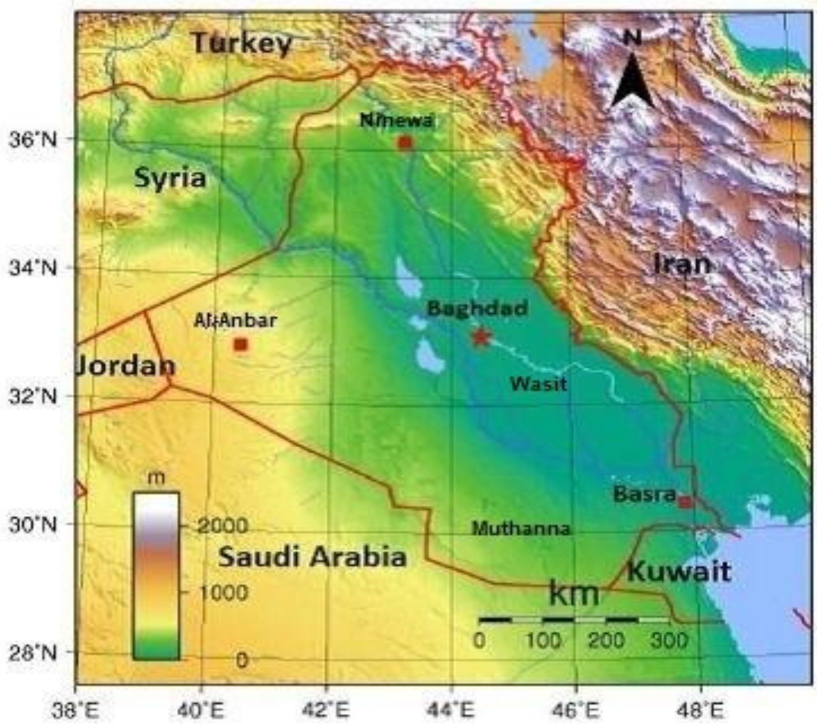


Figure 1: Iraq map, Baghdad, Mosul, Anbar, and Basra weather stations' locations [16].

2.3 WEKA program

WEKA is an innovative program of computer developed at the University of Waikato in New Zealand. Its primary aim is to extricate important experiences and data from crude information gathered across different domains. WEKA offers a wide range of capabilities for data mining tasks like clustering, regression, classification, feature selection, and visualization. The basic idea driving WEKA is to prepare a PC application to use AI methods and find significant examples and patterns. Because of this, scientists and researchers can gain a more in-depth understanding and base their decisions on accurate data. Through WEKA, you can smooth out your information examination process and reveal important information concealed inside your information. Regardless of whether you are a scientist or work in another field, WEKA offers powerful tools and a user-friendly interface to assist you in effective use to analyse and interpret your data [17]. It has accomplished far-reaching acknowledgement inside business and scholarly community circles and has become a generally involved device for data mining research [18].

Weather Forecasting is a high-level framework that uses different machine learning and the WEKA tool for data mining. It utilizes modelling of predictive and analysis of data to estimate atmospheric conditions precisely
WEKA is an open-source software that involves a collection of machine learning algorithms for analyzing data and modelling estimating (predictive). These algorithms can be applied to different datasets [19].

2.4 The Method

In this study, to forecast monthly values of climatic elements using the program (WEKA), it was as follows
1- Enter the monthly values of the climatic elements for the period (1991-2020) for each parameter at each station.
2- Increase the inputs by entering the monthly values of the climatic elements for the duration (1986-2020) for each parameter at each station.
3- Increase the inputs by entering the monthly values of the climatic elements for the duration (1981-2020) for all parameters of each station.
4- In each stage, the same monthly values of weather variables are estimated (12 months) or equivalent to one year.
5- Calculate the statistics (correlation and absolute error) between the estimated values using the WEKA software, and the measured values at each stage help us to distinguish which stage is better and to know the effect of increasing data for atmospheric parameters on the accuracy of parameter estimating.

The statistics used in this study are:

Mean Absolute Error (MAE) is an error difference calculation between two observations for the same phenomenon. Comparison of predicted versus observed results, subsequent time versus initial time, and one measurement method versus an alternative measurement method are all examples of ŷ versus yi. The MAE is calculated by dividing the sample size by the total number of absolute errors [20].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y} - y_i}{y_i}\right| \tag{1}$$

A typical indicator of Estimate inaccuracy in time series analysis is the mean absolute error, Pearson type coefficient of determination index, R^2 [21]:-

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (2)$$

Root-mean-square error (RMSE) is a commonly used metric for comparing observed values to those predicted by a model or an estimator. When the calculations are carried out over the estimation data sample, these deviations are referred to as residuals. However, the computations out-of-sample represent errors (or prediction errors). These deviations are used to compare the forecasting errors of various models for a specific dataset [22].

$$RMSE = \frac{1}{\mathrm{n}}\left[\sum_{i=1}^{n}(y_i - \hat{y})^2\right]^{\frac{1}{2}} \quad (3)$$

yi: is the observed value, ŷ: is the estimated value, and n is the number of sample size

3. Results and Discussion

Using (WEKA) to estimate atmospheric parameters requires entering a time series for the parameter and does not require any addition. Therefore, the accuracy of the estimate relies upon the length of the time series entered into the software in addition to the nature of the parameter within the time series, and this methodology can be summarized as follows:

3.1 Estimating the atmospheric parameters

The time series is inserted for a period of 480 months, amounting to (40) years, to estimate 12 months, and it was:

For the four stations, the minimum temperature is seen in Table 2. The estimated minimum temperature for Basra Station accomplished great qualities regarding mistake measures, indicating the accuracy of the prediction. Compared to the Basra station, there was a weaker correlation between the estimated and measured values for the other stations. The incorrect criterion also had values when the MAE and RMSE more or equal 1. Nevertheless, it is essential to remember that all of these remain within the acceptable estimate range.

Similarly, the maximum temperature estimation in error criteria term produced good results and revealed a significant correlation between the estimated and observed values. Anbar Station, in particular, stood out for having the lowest error values and the greatest correlation values, making it the most precise for determining the maximum temperature.

The wind velocity in Iraq fluctuates continuously and is influenced by numerous factors, including weather factors, geological factors, and the nature of the surface. Compared to the other parameters, this parameter's estimate has higher error criteria and lower correlation. Hence the error criteria had the best values. Nineveh Station has the lowest value of error and the highest correlation between the estimated and measured data.

Relative humidity estimation achieved good values in some stations, so the best values of those standards were in the stations of Baghdad, Basra, and Anbar, where the correlation values were large and there were few error values between the measured and the estimated values. In contrast for the station of Nineveh, the correlation value was low and the error criteria values were large.

Table 2: Clarifies the statistics values to estimate the minimum and Maximum Temperature, Wind Velocity, and Relative Humidity for different time series in the study stations.

Parameter	Station	MAE	RMSE	R^2
Minimum Temperature	**Baghdad**	0.2963	0.0135	0.98719
	Nineveh	0.4426	0.0157	0.98003
	Anbar	0.4415	0.0136	0.98944
	Basra	0.1475	0.014	0.98729
Maximum Temperature	**Baghdad**	0.0509	0.0197	0.98041
	Nineveh	0.0737	0.0235	0.96779
	Anbar	0.0503	0.0178	0.98078
	Basra	0.0522	0.021	0.96993
Wind Velocity	**Baghdad**	0.0768	0.002701	0.811356
	Nineveh	0.1048	0.001864	0.859679
	Anbar	0.0905	0.003419	0.786455
	Basra	0.1361	0.005426	0.488868
Relative Humidity	**Baghdad**	0.218	0.0632	0.930186
	Nineveh	0.4166	0.1146	0.82613
	Anbar	0.1937	0.0778	0.927447
	Basra	0.2337	0.0678	0.965306

3.2 Test the effect of time series on estimating accuracy

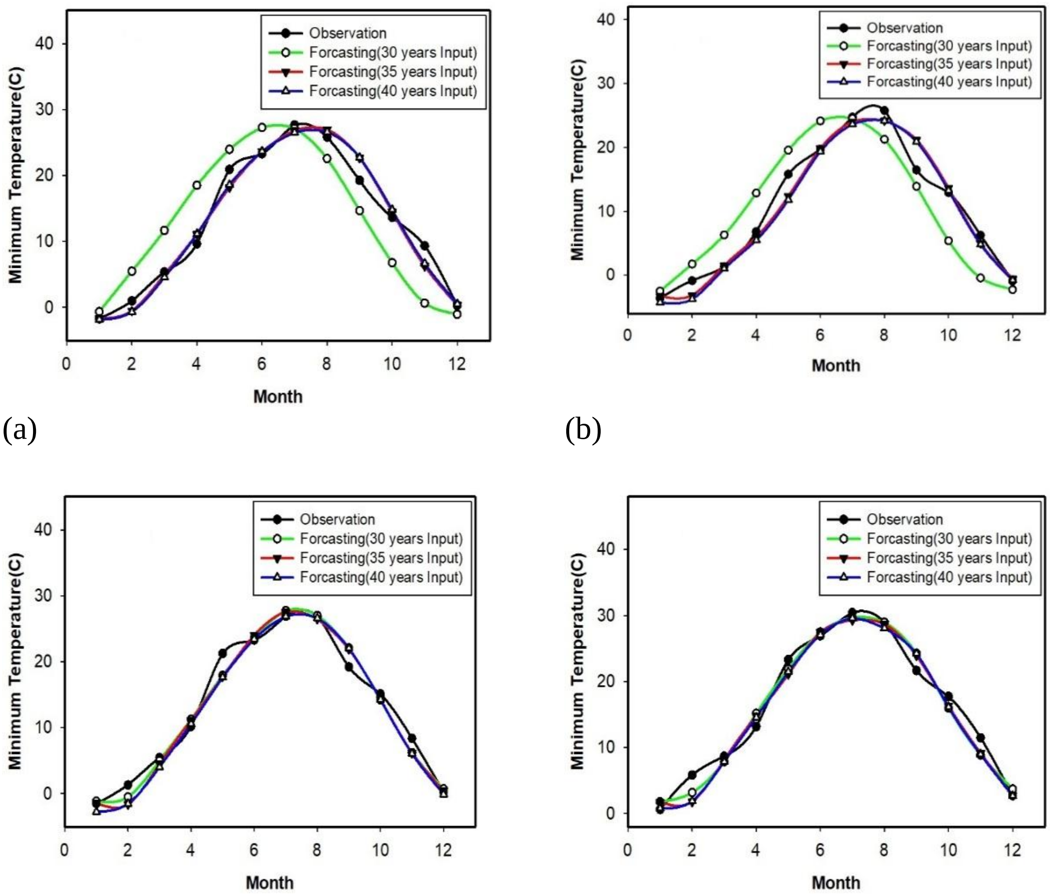
This test is based on entering the time series data for the parameter and different periods, as it was found that the length of the time series greatly affects the accuracy of the estimated value in the (WEKA) software, and it was as follows;

The estimation of minimum temperature shows a relationship between the observed values and the estimation values for all periods in the time series (30-35-40) years. The ratio of the correlation between the two values and the value of the error criteria between the observed values and the estimated values of the minimum temperature is shown in Table 3. It can be noted that there is a significant decrease in the error standards for the Nineveh station and

a clear increase in the correlation, as well as for the Baghdad station. It is possible to notice a significant convergence between the observed and the estimated values for all cases. As for the rest of the stations, we can notice a decrease in the correlation and divergence between the estimated and the observed values, especially when the input (30) year, as shown in Figure 2.

Table 3: Shows the statistics values to estimate the minimum temperature for different time series in the study stations.

	30 years Input			**35 years Input**			**40 years Input**		
Station	**MAE**	**RMSE**	R^2	**MAE**	**RMSE**	R^2	**MAE**	**RMSE**	R^2
Baghdad	1.15	0.044	0.870	0.22	0.0132	0.986	0.2963	0.0135	0.9871
Nineveh	0.98	0.038	0.900	0.35	0.0138	0.981	0.4426	0.0157	0.9800
Anbar	0.33	0.012	0.988	0.32	0.0134	0.987	0.4415	0.0136	0.9894
Basra	0.29	0.014	0.986	0.27	0.0146	0.985	0.1475	0.014	0.9872



(a) (b)

(c) (d)

Figure 2: The observation and estimation minimum temperature for different time series (a) Baghdad (b) Nineveh, (c) Anbar, and (d) Basra station.

The estimation of the maximum temperature shows a relationship between the estimated and measured values for all periods in the time series (30-35-40) years, It can be observed that the convergence between the later values is significant for Anbar and Basra stations and in almost all cases. For Mosul and Baghdad stations, one can notice a decrease in the correlation and divergence between the estimated and the observed values, especially when the input is (30) years. This indicates that the model will require data.from Figure 3 The best case was in the Baghdad station, where there was a significant decrease in the error standards and an increase in the relationship between the estimated and the observed values, as clarified in Table 4.

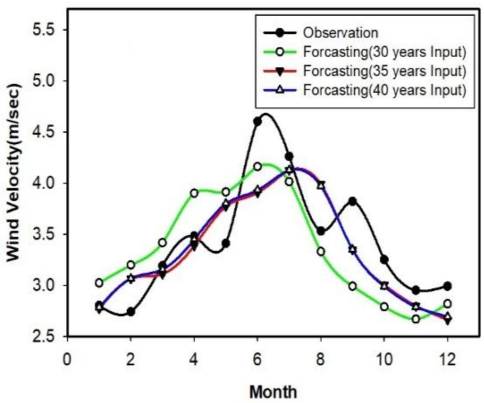
Table 4: Illustrates the statistics values to estimate the maximum temperature for different time series in the study stations.

	30 years Input			**35 years Input**			**40 years Input**		
Station	**MAE**	**RMSE**	**R^2**	**MAE**	**RMSE**	**R^2**	**MAE**	**RMSE**	**R^2**
Baghdad	0.104	0.039	0.86	0.0587	0.021	0.983	0.0509	0.019	0.9804
Nineveh	0.106	0.042	0.92	0.0749	0.025	0.972	0.0737	0.023	0.9677
Anbar	0.052	0.018	0.97	0.0582	0.020	0.982	0.0503	0.017	0.9807
Basra	0.058	0.022	0.96	0.0603	0.023	0.971	0.0522	0.021	0.9699

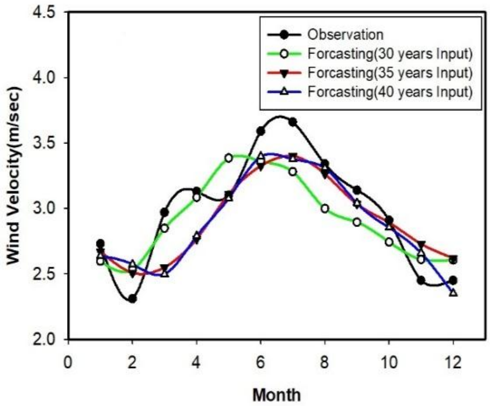
For the wind velocity estimation, the wind velocity fluctuates for all four stations, It is possible to notice the divergence between the observed and estimated values, especially in June because the wind fluctuation is the highest in this month. The highest discrepancy between the observed and estimated values was at the Basra station in June due to the high activity of the eastern winds, see Figure 3. Hence, the increase in the time series accomplished fluctuating qualities concerning the upsides of the blunder standards. This is obvious in the Anbar station qualities. The Nineveh station and all stations had the highest accuracy values. Better values for estimating wind velocity were obtained with greater accuracy as the time series expanded, see Table 5.

Table 5: Displays the statistics values to estimate the Wind Velocity for different time series in the study stations.

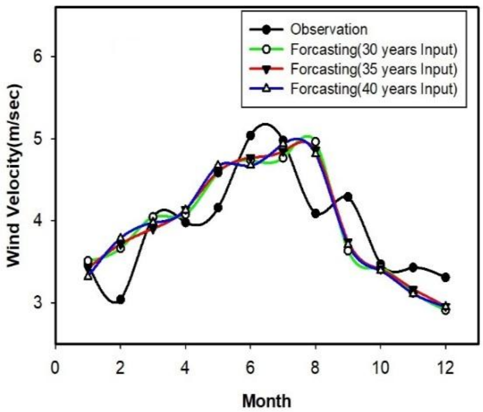
	30 years Input			35 years Input			40 years Input		
Station	**MAE**	**RMSE**	**R^2**	**MAE**	**RMSE**	**R^2**	**MAE**	**RMSE**	**R^2**
Baghdad	0.10	0.0037	0.703	0.080	0.0028	0.80	0.0768	0.0027	0.811
Nineveh	0.068	0.002	0.872	0.063	0.0018	0.87	0.1048	0.0018	0.859
Anbar	0.086	0.0033	0.776	0.084	0.003	0.80	0.0905	0.0034	0.786
Basra	0.144	0.0056	0.498	0.144	0.0056	0.48	0.1361	0.0054	0.488



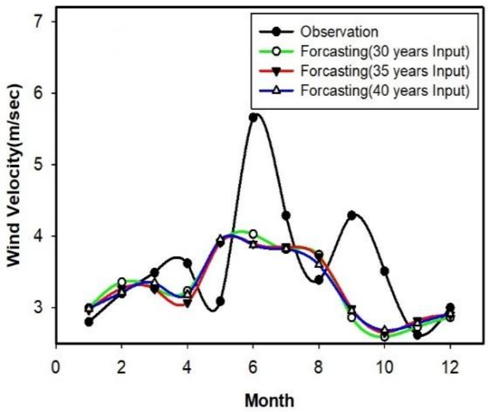
(a)



(b)



(c)



(d)

Figure 3: The observation and estimation of Wind Velocity for different time series (a) Baghdad(b) Nineveh, (c) Anbar, and (d) Basra station.

Regarding the relative humidity Estimation, Figure 4 shows the assembly between the measured values and outputs of the (WEKA) software, which represents the estimated values at various times. It can be observed that the discrepancy between the estimated and observed values is the largest at Basra station for all three estimation cases because it is near the Western Bay, which leads to no fluctuation in the relative humidity. At Anbar station, one can notice the difference in months due to the interplay of reasons group for the humidity fluctuation, including altitude and desert nature. The difference between the estimated and observed values is the largest possible in the Baghdad and Nineveh stations. It has been discovered that as the duration of time series increases the error standards decrease, whilst the correlation between the estimated values and measured values increases the change in the values is not similar by the same mechanism. Sometimes, the length increase of the time series results in the error standards increasing with a decrease in correlation, as in the Anbar station because the area is desert, which increases the disturbance of moisture consumption and leads to inaccurate estimates, see Table 6.

Table 6: illustrates the statistics values to estimate the Relative Humidity for different time series in the study stations

	30 years Input			**35 years Input**			**40 years Input**		
Station	**MAE**	**RMSE**	**R^2**	**MAE**	**RMSE**	**R^2**	**MAE**	**RMSE**	**R^2**
Baghdad	0.17	0.052	0.951	0.266	0.071	0.925	0.218	0.063	0.930
Nineveh	0.26	0.074	0.938	0.338	0.099	0.832	0.416	0.114	0.826
Anbar	0.23	0.076	0.910	0.246	0.077	0.914	0.193	0.077	0.927
Basra	0.27	0.074	0.970	0.263	0.072	0.965	0.233	0.067	0.965

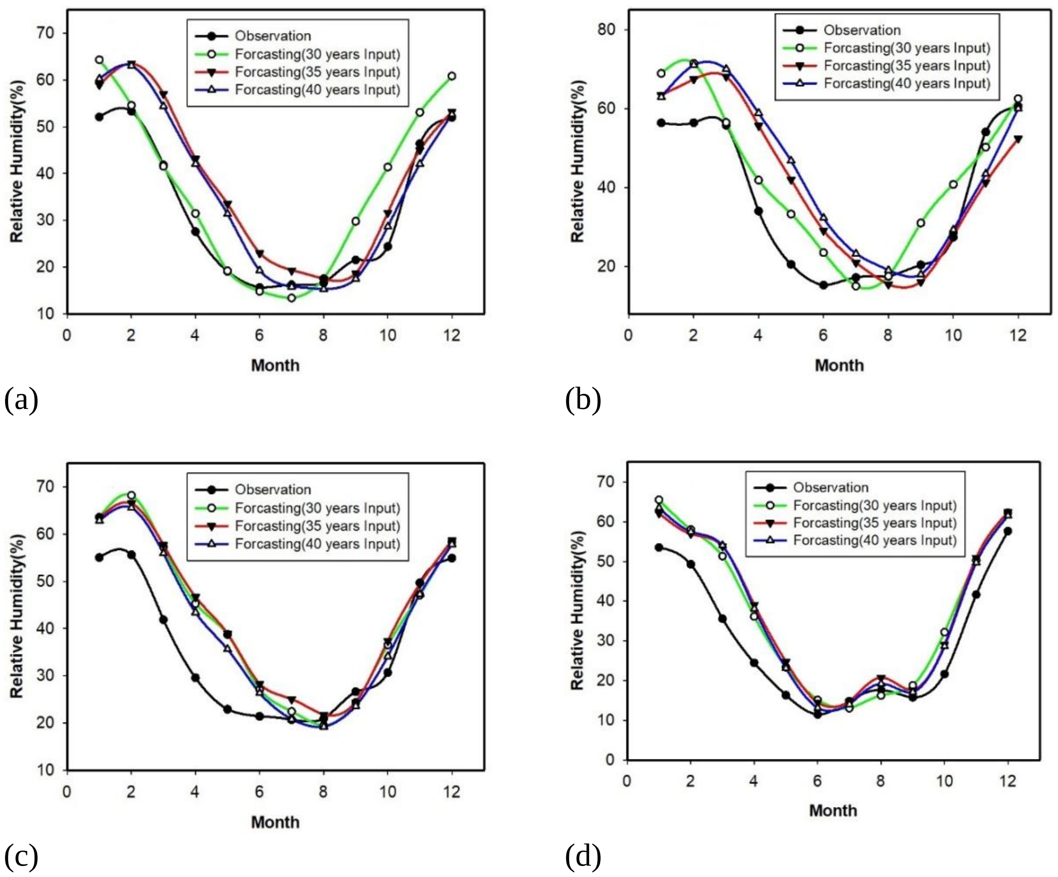


Figure 4: The observation and estimation of Relative Humidity for different time series(a) Baghdad, (b) Nineveh, (c) Anbar, and (d) Basra station.

Here, it must be pointed out that there is uncertainty around projections resulting from the fact that the climate elements of a particular region are affected by global climate changes, as well as noise and natural climate fluctuations that affect for a small period including the passing of cold air masses. In addition to the increase in carbon dioxide emissions and greenhouse gases. The accuracy of any model used to estimate climate elements depends on the quality of the data included in the model, so uncertainty may sometimes be due to the inaccuracy of the data**.**

Conclusions

In this study, time series data for estimation parameters were used using the (WEKA) software, where the size of the data entering the software was changed to determine its effect on the accuracy of the estimation. It was found that the estimation value of the same parameter differs from one station to the latter in terms of accuracy and error standards, as well as the correlation between the estimated and the measured values.

Also, the accuracy of the estimation is not equal and similar in terms of error and correlation criteria for all parameters, even if they are for the same station and region. As the estimate accuracy depends mainly on the time series entering the parameter, it is directly affected by the length of the time series of the parameter entering the (WEKA). The length increase of the time series often increases the accuracy of the estimate, and this increase differs from one parameter to another and from one time series to another in addition to differing from one station to another. This increase in accuracy is fluctuating and uneven in terms of amount depending on the time series increase.

Acknowledgements

The authors express their thanks to the General Authority for Meteorology and Seismic Monitoring in Baghdad Governorate for providing the data used in this study.

Conflict of interest

The authors have no conflicts of interest.

References

[1] K. Parul, A. Tripti, and S. Nikita Mitra, "Climate change forecasting using data mining algorithm," *AQUA — Water Infrastructure, Ecosystems and Society*, vol. 00, no. 2, 2023.

[2] A. K. Shathir, A. S. Layla, and A. A. Sumayah, "Forecasting Monthly Maximum Temperatures in Kerbala Using Seasonal ARIMA Models," *Journal of University of Babylon for Engineering Sciences*, vol. 27, no. 2, 2019.

[3] A. O. Abe, Q. A. Adeniji, J. A. Rabiu, O. Adegboyega, I. O. Raheem, M. G. Rasaki, S. M. Sada, and L. Fidelis, "Statistical Analysis and Forecasting of Rainfall Patterns and Trends in Gombe North-Eastern Nigeria," *Iraqi Journal of Physics*, vol. 20, no. 2, pp. 64-77, 2022.

[4] H. M. Berivan, A. M. Jwan, and I. M. Amera, "Development and Assessment of Feed Forward Back Propagation Neural Network Models to Predict Sunshine Duration," *Iraqi Journal of Physics*, vol. 20, no. 3, pp. 64-75, 2022.

[5] J. Fatima and S. M. Alaa, "Testing the MODIS Thermal Modes for Dust Storms Monitoring," *Iraqi Journal of Physics*, vol. 21, no. 2, pp. 17-24, 2023.

[6] A. Rahimikhoob, S. M. R. Behbahani, and M. E. Banihabib, "Comparative study of statistical and artificial neural network's methodologies for deriving global solar radiation from NOAA satellite images," *International Journal of Climatology*, vol. 33, no. 2, pp. 480-486, 2013.

[7] P. F. Yunus and S. M. Shaahid, "Forecasting Wind Speed using Artificial Neural Networks – A Case Study of a Potential Location of Saudi Arabia," *E3S Web of Conferences*, vol. 173, 01004, 2020. [Online]. Available: https://doi.org/10.1051/e3sconf/202017301004.

[8] E. Krüger and R. Emmanuel, "Accounting for atmospheric stability conditions in urban heat island studies: the case of Glasgow, UK," *Landscape Urban Plann*., vol. 117, pp. 112-121, 2013.

[9] J. Aw and M. J. Kleeman, "Evaluating the first-order effect of intra-annual temperature variability on urban air pollution," J. Geophys. Res.: Atmos., vol. 108, 2003.

[10] L. Liu, S. Breitner, X. Pan, U. Franck, A. M. Leitte, A. Wiedensohler, S. von Klot, H.-E. Wichmann, A. Peters, and A. Schneider, "Associations between air temperature and cardio-respiratory mortality in the urban area of Beijing, China: a time-series analysis," *Environ. Health*, vol. 10, no. 51, 2011.

[11] P. Romero, H. Qin, and K. Dickinson, "Urban vulnerability to temperature-related hazards: A meta-analysis and meta-knowledge approach," *Global Environmental Change*, vol. 22, no. 3, pp. 670-683, 2012.

[12] G. M. Joselin Herbert et al., "A review of wind energy technologies," *Renewable and Sustainable Energy Reviews*, vol. 11, pp. 1117-1145, 2007.

[13] A. K. Sharma, "Study of wind turbine based SEIG under balanced/unbalanced loads and excitation," *International Journal of Electrical and Computer Engineering*, vol. 2, pp. 353-370, 2012.

[14] Y. Zheng et al., "Mode analysis of horizontal axis wind turbine blades," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 12, pp. 1212-1216, 2014.

[15] E. Kuzugudenli, "Relative humidity modeling with artificial neural networks," *Applied Ecology and Environmental Research*, vol. 16, no. 4, pp. 5227-5235, 2018.

[16] H. K. Jasim, S. H. Ahmed, and J. M. Ali, "Phase Lag of Temperature behind Solar Radiation in Iraq," *Journal of Mechanical Engineering Research and Developments*, vol. 44, no. 7, pp. 100-108, 2021.

[17] E. Witten and I. H. Frank, Data Mining: Practical Machine Learning Tools and Techniques with Java Implementation, Morgan Kaufmann Publishers, 2000.

[18] I. H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, 2nd ed., Morgan Kaufmann, San Francisco, 2005.

[19] G. Holmes, A. Donkin, and I. H. Witten, "WEKA: A Machine Learning Workbench," *in Proc. Second Australia and New Zealand Conference on Intelligent Information Systems, Brisbane, Australia*, 1994.

[20] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research*, vol. 30, pp. 79-82, 2005.

[21] S. Mondal and H. Mondal, "Value for (2) in Statistical Analysis by Pearson Correlation Coefficient," *Journal of Clinical and Diagnostic Research*, vol. 11, no. 11, pp. 1-2, 2017.

[22] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679-688, 2006.