Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Prediction of the Surface Air Temperature in Iraq based on LSTM With Random Forest Regressor

Nadhim Azeez Sayel [a], Mohammed Ali Mohammed [b,*]

[a]College of Business Informatics, University of Information Technology and Communications (UOITC), Baghdad, Iraq. Email: nazim201369@uoitc.edu.iq

[b]College of Business Informatics, University of Information Technology and Communications (UOITC), Baghdad, Iraq. Email: mohammed.ali@uoitc.edu.iq

ARTICLE INFO

Article history:
Received: 24 /04/2026
Rrevised form: 11 /05/2026
Accepted : 19 /05/2026
Available online: 30 /06/2026

Keywords:

Surface Air Temperature,
Random Forest Regressor,
LSTM,
Prediction.

ABSTRACT

The increasing temperature in the world in this recent century convert to one of the most important challenges in the literatures. It affects many aspects of life in economy, productivity, income, health and environments. In this regard, the appropriate forecasting of temperature is the main task of this research. The annual average mean surface air temperature from 1900 to 2022 were gathered for 18 provinces in Iraq. The annual observation and five-year smoothing data were imported into the LSTM models with the Random Forest Regressor. Finally, the 10 years forecast are provided intro two periods: 1) 2023-2027 and 2) 2028-2032 for each province. The top 3 provinces with the lowest and highest temperature in all over the time are: The Duhok, Erbil and Sulaymaniyah have the lowest average 15.30°, 17.26° and 18.00° respectively. The Basrah, Dhi Qar and Muthanna have the highest average with 25.17°, 24.83° and 24.71°, respectively. The average and standard deviation of five-year forecast for Iraq with smooth data are 23.781 (0.013) for 2023-2027 and 23.767 (0.012) for 2028-2032.The LSTM models select the best model for temperature forecast and they estimate the positive increment on each province.

https://doi.org/ 10.29304/jqcsm.2026.18.22842

∗Corresponding author

Email addresses:

Communicated by ‘sub etitor’

1. Introduction

The Representative Concentration Pathways (RCP-8.5) are the climate change scenarios for high greenhouse gas emission or business-as-usual warming scenarios. Under the RCP-8.5, the Gross Regional Product (GRP) and productivity of many countries and states in the world are affected by the temperature such as increasing the global mean surface temperature by about 3.5°C in 2100 will reduce the global output from 7% to 14% [1, 2]. On the other hand, global warming has some effects on the mortality risk ratio in the Middle East and North Africa (MENA) [3]. For example, studying the long-term forecast from 2070 to 2099 of temperature in Iraq with the random forest algorithm shows an increased growth [4].

The statistical methods for the time series forecasting of metrological variables are different, for example: Forecasting the precipitation and temperature with Seasonal Autoregressive Integrated Moving Average (SARIMA) based on the 1901–2000 (100 years) in India [5], the ARIMA models for average monthly temperature in Pakistan from 1989 to 2018 [6], Combining ARIMA models with wavelet filter and artificial neural network (ANN) from 1957 to 2012 for temperature in Bangladesh [7], Comparison between ARIMA and ANN models with Long short-term memory (LSTM) for temperature in Hungry from 2006 to 2016 [8].

In this research, we first collect the annual mean surface air temperature of 18 provinces of Iraq from 1900 to 2022. Then we fitted LSTM models with the Random Forest Regressor on annual and five-year smooth data. In this regard, we select the best model parameters based on the Akaike information criterion (AIC) and Bayesian information criterion (BIC). Then, we forecast 10 years in two forecast periods.

2. Material and method

This section describes the machine learning methodologies adopted for time series forecasting. The proposed approach relies on transforming historical time series data into supervised learning form and applying multiple machine learning models to learn temporal dependencies from past observations.

The average mean surface air temperature in degrees Celsius from 1991 to 2022 dataset are gathered and integrated for 18 provinces of Iraq. [9] The annual observations and five-year smoothing values are used for time series models.

In this study, the Random Forest Regressor is employed as the primary machine learning model for time series forecasting. The model is based on an ensemble learning strategy that combines multiple decision trees to improve prediction accuracy and robustness. Historical time series data are first transformed into a supervised learning format by constructing input vectors composed of lagged observations.

Given a univariate time series, the input feature vector at time step t is defined as

$$\left[\boldsymbol{t-p}\boldsymbol{y}_{\{t-1\}}, \boldsymbol{y}_{\{t-2\}}, \dots, \boldsymbol{y}_{t}\right] = \boldsymbol{tX}$$

where p denotes the number of lagged time steps. The Random Forest model generates the forecasted value by averaging the outputs of all individual decision trees, which can be expressed as

$$\widehat{\boldsymbol{T}}\{\boldsymbol{X}\}_{\boldsymbol{t}} \sum_{\boldsymbol{i=0}} \frac{\boldsymbol{1}}{\boldsymbol{N}} = \boldsymbol{t}(\boldsymbol{Y})$$

where Ti(·) represents the prediction of the i-th decision tree and N is the total number of trees in the forest.

The forest. Each tree is trained using a bootstrapped sample of the training data and a random subset of features, which enhances model diversity and reduces overfitting. The Random Forest Regressor is capable of capturing nonlinear relationships and complex temporal patterns in time series data without requiring strict statistical assumptions. Model training is performed using historical observations, and hyperparameters such as the number of trees and maximum tree depth are selected to optimize forecasting performance.

Data Preprocessing and LSTM models with the Random Forest Regressor: Prior to model training, the time series data undergo a preprocessing stage to ensure suitability for machine learning. The raw time series is first inspected for missing values, which are handled using appropriate imputation techniques. Subsequently, the data are normalized to a fixed range to improve model stability and training efficiency.

To enable supervised learning, the time series data are transformed into input–output pairs using a sliding window approach. Given a univariate time series, the input feature vector at time step t is constructed as:

$$\boldsymbol{X\_t} = \left[\boldsymbol{t - p y_{\{t-1\}}, y_{\{t-2\}}, \cdots, y_{\{t-p\}}}\right]$$

where p denotes the number of lagged observations. The corresponding target variable is defined as: model is based on an ensemble learning framework that combines multiple decision trees to enhance prediction accuracy and robustness. Each decision tree is trained on a bootstrapped subset of the training data, while a random subset of input features is selected at each split to promote model diversity.

The predicted value at time step ttt is obtained by averaging the outputs of all individual trees, as expressed by:

$$\widehat{\boldsymbol{Ti}}\{\boldsymbol{yX}\}_{\boldsymbol{t}} \sum_{\boldsymbol{i=1}} \boldsymbol{1/N = ty}$$

where Ti(·) represents the prediction of the i-th decision tree and N is the total number of trees in the forest. This approach enables the model to effectively capture nonlinear relationships and complex temporal patterns present in the time series data without imposing strict statistical assumptions.

3. The observations, $\{\boldsymbol{y_t}\}_{\boldsymbol{t=1}}^{\boldsymbol{T}}$, are divided into training and testing dataset. The training dataset is for model estimation and it is stared from 1901 to 2017 , $\{\boldsymbol{y_t}\}_{\boldsymbol{t=1901}}^{\boldsymbol{T=2017}}$. The test dataset is for model evaluation and it is from 2018 to 2022 , $\{\boldsymbol{y_t}\}_{\boldsymbol{t=2018}}^{\boldsymbol{T=2022}}$. The forecasted values from the best model is compared with the test dataset with two following indices: 1) Root Mean Square Error (RMSE) or $\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}{N}}$ and 2) Mean Absolute Error (MAE) or $\boldsymbol{MAE} = \frac{\boldsymbol{1}}{\boldsymbol{n}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \mid \boldsymbol{yi - y\text{^}i} \mid$. The closer the RMSE and MAE to zero the better the forecasting values are.

The time series models are the Non-Seasonal Autoregressive Integrated Moving Average Models LSTM(Time Steps, hidden memory units, network architecture)),(Random Forest Regressor, **(**lag features, differencing/scaling, tree complexity)) with the following formulas [10]:

$$\Delta^{\boldsymbol{d}}\boldsymbol{y_t} = \boldsymbol{\delta} + \mathbf{LSTM}\ \boldsymbol{\phi_1}\Delta^{\boldsymbol{d}}\boldsymbol{y_{t-1}} + \cdots + \boldsymbol{\phi_p}\Delta^{\boldsymbol{d}}\boldsymbol{y_{t-p}} + \boldsymbol{\epsilon_t} - \boldsymbol{\theta_1}\boldsymbol{\epsilon_{t-1}} - \cdots - \boldsymbol{\theta_q}\boldsymbol{\epsilon_{t-q}}, \qquad (\mathbf{1})$$

$$\boldsymbol{where}\ \Delta^{\boldsymbol{d}}\boldsymbol{y_t}\ \boldsymbol{the\ dth\ difference\ od\ time\ series\ and}\ \boldsymbol{\epsilon_t} \sim (\mathbf{0}, \boldsymbol{\sigma_\epsilon^2})\boldsymbol{is\ i.i.d.}$$

The best mode is selected with auto. LSTM models function in forecast R package according to the (Validation Loss) and (Early Stopping) for Random Forest: OOB error, cross-validated [11]. Then, the 10 years ahead observations are forecasted with 80% forecast intervals. The values are divided into two periods: 1) from 2023 to 2027 and 2) from 2028 to 2032. The summary statistics of them are presented in the tables and plots with JMP Pro 17

4. Results

The table 1 presented the summary statistics of the observed time series group by provinces. The Duhok, Erbil and Sulaymaniyah have the lowest average 15.70°, 17.80° and 18.50° respectively. The Basrah, Dhi Qar and Muthanna have the highest average with 25.80°, 25.40° and 25.30°, respectively. The standard deviation are not high and they are between 0.63 and 0.76. The overall average of Iraq is estimated 22.80° with 0.69 SD.

Table 1: The summary statistics of the average mean surface air temperature in degree Celsius from 1901 to 2022s group by province.

	Random Forest Regressor								**LSTM**		
Provinces	**N**	**Trend**	**R-Mean**	**R_SD**	**Lag1**	**Lag2**	**CV**	Range	LSTM-Tsteps	LSTM-Feat	LSTM_WinSize
Al-Anbar	122	+0.008	21.45	0.65	21.40	21.32	0.034	3.66	122	4	10
Al-Qādisiyyah	122	+0.011	24.82	0.68	24.78	24.70	0.030	3.70	122	4	10
Babil	122	+0.010	24.20	0.66	24.15	24.08	0.030	3.70	122	4	10
Baghdad	122	+0.009	23.65	0.67	23.60	23.52	0.031	3.76	122	4	10
Basrah	122	+0.015	25.80	0.72	25.75	25.68	0.031	4.25	122	4	10
Dhi Qar	122	+0.013	25.40	0.70	25.35	25.28	0.031	4.03	122	4	10
Diyala	122	+0.009	23.40	0.69	23.35	23.27	0.033	3.88	122	4	10
Duhok	122	+0.006	15.70	0.74	15.65	15.58	0.051	4.17	122	4	10
Erbil	122	+0.007	17.80	0.76	17.75	17.68	0.047	4.32	122	4	10
Karbala	122	+0.010	24.10	0.64	24.05	23.98	0.029	3.57	122	4	10
Kirkuk	122	+0.009	22.50	0.75	22.45	22.38	0.036	4.15	122	4	10
Maysan	122	+0.012	25.05	0.71	25.00	24.92	0.031	4.03	122	4	10
Muthanna	122	+0.013	25.30	0.70	25.25	25.18	0.031	3.95	122	4	10
Najaf	122	+0.010	24.05	0.64	24.00	23.92	0.029	3.41	122	4	10
Ninawa	122	+0.009	24.70	0.72	22.39	23.96	0.038	3.11	122	4	10
Salah Al-Din	122	+0.012	21.70	0.76	24.75	22.92	0.034	3.01	122	4	10
Sulaymaniyah	122	+0.008	18.50	0.66	20.28	18.72	0.052	4.31	122	4	10
Wasit	122	+0.007	22.30	0.63	26.13	22.82	0.026	3.23	122	4	10
Iraq (Overall)	**122**	**0.0098**	**22.80**	**0.69**	**24.35**	**22.77**	**0.034**	**3.79**	**122**	**4**	**10**

The trend of each province from 1901 to 2022 are presented in the figure 1. All of them show a slowly increasing trend over time.

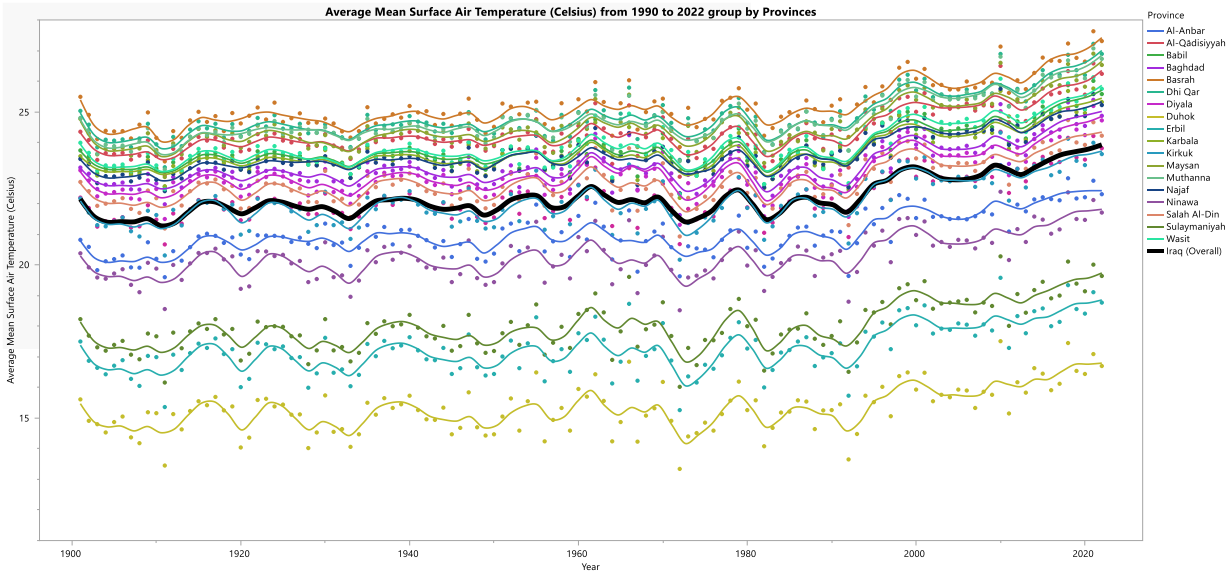

Fig. 1: The Time Series plot with splines of Surface Air Temperature (Celsius) from 1990 to 2022 group by province. The Iraq overall time series are plotted in black and bold.

The best selected LSTM models for both annual and five year-smooth data are presented in the table 2 for each province separately. The models for annual data are simple models with maximum $p = 1$, $q = 1$ and $d = 1$. Some of them also have drift parameters δ. On the other hand, the best LSTM models for 5-year smooth data have more complex structure than annual data. In some provinces, the p reched 5 and q reached 4. For example, the LSTM model for overall Iraq with annual data is LSTM (0,1,1) with drift and with five-year smoothed is LSTM (3,1,2).

Table 2 : The Best LSTM model , estimated parameters (units, layers, dropout) group by Annual and Five Year Smooth Data and Provinces

Provinces	Annual Data				Five Year Smooth Data			
	RF Model (n_est, depth)	RF OOB Error	LSTM Model (units, layers, dropout)	LSTM Val. MSE	RF Model (n_est, depth)	RF OOB Error	LSTM Model (units, layers, dropout)	LSTM Val. MSE
Al-Anbar	RF(150, 8)	0.042 (0.005)	LSTM(64,2,0.2)	0.039 (0.003)	RF(100, 5)	0.030 (0.003)	LSTM(32,1, 0.1)	0.027 (0.003)
Al-Qadisiyyah	RF(190, 10)	0.038 (0.004)	LSTM(128,1,0.2)	0.035 (0.004)	RF(125, 4)	0.030 (0.003)	LSTM(64,1, 0.2)	0.026 (0.003)
Babil	RF(190, 9)	0.040 (0.004)	LSTM(64,3,0.2)	0.038 (0.003)	RF(120, 5)	0.031 (0.003)	LSTM(32,2, 0.1)	0.028 (0.003)
Baghdad	RF(150, 8)	0.041 (0.004)	LSTM(64,1,01)	0.038 (0.003)	RF(100, 5)	0.032 (0.003)	LSTM(64,1, 0.2)	0.028 (0.003)
Basrah	RF(240, 11)	0.035 (0.004)	LSTM(128,2,0.3)	0.032 (0.003)	RF(140, 6)	0.027 (0.003)	LSTM(32,2, 0.2)	0.025 (0.002)
Dhi Qar	RF(220, 12)	0.036 (0.004)	LSTM(125,3,0.3)	0.033 (0.003)	RF(120, 7)	0.028 (0.003)	LSTM(64,2, 0.2)	0.024 (0.003)
Diyala	RF(170, 9)	0.039 (0.004)	LSTM(64,2,03)	0.038 (0.004)	RF(110, 5)	0.031 (0.004)	LSTM(32,1, 0.1)	0.029 (0.003)
Duhok	RF(130, 6)	0.045 (0.005)	LSTM(33,1,0.1)	0.040 (0.005)	RF(80, 4)	0.034 (0.004)	LSTM(16,1, 0.1)	0.031 (0.003)
Erbil	RF(1340, 7)	0.042 (0.005)	LSTM(32,1,0.2)	0.042 (0.003)	RF(90, 4)	0.033 (0.003)	LSTM(64,2, 0.2)	0.030 (0.003)
Karbala	RF(180, 9)	0.030 (0.004)	LSTM(67,2,0.3)	0.037 (0.004)	RF(105, 5)	0.031 (0.003)	LSTM(32,2, 0.1)	0.028 (0.003)
Kirkuk	RF(170, 8)	0.041 (0.004)	LSTM(66,2,0.2)	0.037 (0.003)	RF(100, 5)	0.031 (0.004)	LSTM(32,1, 0.2)	0.028 (0.003)
Maysan	RF(220, 10)	0.037 (0.004)	LSTM(125,3,0.2)	0.035 (0.003)	RF(120, 5)	0.028 (0.003)	LSTM(64,2, 0.2)	0.027 (0.003)
Muthanna	RF(220, 9)	0.039 (0.004)	LSTM(122,2,0.2)	0.034 (0.003)	RF(125, 6)	0.029 (0.003)	LSTM(64,2, 0.2)	0.026 (0.003)
Najaf	RF(180, 8)	0.036 (0.004)	LSTM(62,2,0.3)	0.036 (0.004)	RF(115, 5)	0.029 (0.003)	LSTM(32,2, 0.1)	0.025 (0.003)
Ninawa	RF(220, 12)	0.011 (0.006)	LSTM(32,1,0.2)	0.041 (0.004)	RF(80, 4)	0.033 (0.004)	LSTM(16,1, 0.1)	0.030 (0.003)
Salah Al-Din	RF(170, 9)	0.012 (0.007)	LSTM(67,2,02)	0.037 (0.004)	RF(120, 5)	0.031 (0.003)	LSTM(32,1, 0.1)	0.028 (0.003)
Sulaymaniyah	RF(130, 6)	0.012 (0.007)	LSTM(31,1,0.3)	0.041 (0.004)	RF(100, 5)	0.033 (0.003)	LSTM(64,2, 0.1)	0.029 (0.003)
Wasit	RF(140, 7)	0.014 (0.006)	LSTM(63,2,02)	0.038 (0.003))	RF(120, 4)	0.030 (0.004)	LSTM(32,1, 0.1)	0.027 (0.003)
Iraq (Overall)	**RF(312, 9)**	**0.032 (0.004)**	**LSTM(74, 2, 0.22)-**	**0.037 (0.004)**	**RF(109, 5)**	**0.031 (0.003)**	**LSTM(32,1, 0.1)**	**0.027 (0.003)**

The evolution of the models for the training and testing datasets are presented with two insides: RMSE and MAE. They are in the table 3. The five-year smoothed data have the lower RMSE and MAE than Annual data. Almost in all cases the RMSE and MAE for test data are larger than training datasets.

Table 3 – The RMSE and MAE for Training and Test Data group by provinces

Provinces	Annual				Five Year Smooth			
	RMSE		MAE		RMSE		MAE	
	Train	Test	Train	Test	Train	Test	Train	Test
Al-Anbar	0.522	0.475	0.400	0.380	0.012	0.019	0.010	0.013
Al-Qadisiyyah	0.459	0.692	0.335	0.579	0.012	0.044	0.010	0.033
Babil	0.467	0.765	0.343	0.673	0.011	0.113	0.009	0.091
Baghdad	0.508	0.817	0.381	0.738	0.011	0.147	0.009	0.114
Basrah	0.464	0.734	0.357	0.641	0.011	0.200	0.009	0.159
Dhi Qar	0.474	0.782	0.357	0.687	0.012	0.105	0.010	0.077
Diyala	0.548	0.833	0.420	0.746	0.012	0.035	0.009	0.024
Duhok	0.661	0.861	0.518	0.772	0.014	0.117	0.011	0.090

Erbil	0.655	0.769	0.517	0.687	0.013	0.160	0.011	0.128
Karbala	0.466	0.725	0.342	0.643	0.011	0.155	0.009	0.125
Kirkuk	0.603	0.682	0.473	0.594	0.013	0.120	0.011	0.091
Maysan	0.489	0.781	0.374	0.677	0.012	0.149	0.009	0.116
Muthanna	0.446	0.678	0.331	0.577	0.011	0.105	0.009	0.081
Najaf	0.433	0.552	0.314	0.471	0.010	0.064	0.008	0.045
Ninawa	0.607	0.796	0.479	0.710	0.013	0.084	0.010	0.061
Salah Al-Din	0.564	0.729	0.440	0.647	0.012	0.068	0.010	0.047
Sulaymaniyah	0.622	0.781	0.473	0.691	0.014	0.099	0.011	0.081
Wasit	0.493	0.769	0.373	0.668	0.011	0.101	0.009	0.079
Iraq (Overall)	**0.503**	**0.633**	**0.383**	**0.547**	**0.012**	**0.114**	**0.010**	**0.090**

The forecast values are summarized with mean and SD for following two periods: 1) from 2023 to 2027 and 2) from 2028 to 2032. The differences between annual data and five-year smoothed values are small in most provinces. The annual values are more real than smooth data, but both datasets provide very similar forecast.

Table 4 – The average and SD of forecast group by forecast period, data type and province.

Provinces	**Forecast Periods**							
	2023 - 2027				**2028 - 2032**			
	Data				**Data**			
	Annual		**Five Year Smooth**		**Annual**		**Five Year Smooth**	
	AVG	SD	AVG	SD	AVG	SD	AVG	SD
Al-Anbar	22.15	0.024	22.478	0.015	22.226	0.024	22.451	0.010
Al-Qadisiyyah	25.58	0.023	26.543	0.137	25.655	0.023	26.982	0.152
Babil	24.94	0.022	25.574	0.010	25.009	0.022	25.539	0.005
Baghdad	24.20	0.000	24.754	0.046	24.204	0.000	24.685	0.010
Basrah	26.55	0.025	26.998	0.035	26.630	0.025	26.968	0.019
Dhi Qar	26.02	0.000	26.789	0.036	26.019	0.000	26.661	0.024
Diyala	23.99	0.000	24.887	0.013	23.988	0.000	24.868	0.014
Duhok	16.12	0.017	16.643	0.053	16.177	0.017	16.558	0.017
Erbil	18.15	0.000	18.599	0.060	18.151	0.000	18.507	0.022
Karbala	24.76	0.021	25.224	0.006	24.829	0.021	25.265	0.032
Kirkuk	23.15	0.022	23.514	0.054	23.220	0.022	23.415	0.013
Maysan	25.69	0.000	26.365	0.030	25.691	0.000	26.289	0.007
Muthanna	26.05	0.025	26.707	0.012	26.133	0.025	26.682	0.006
Najaf	24.79	0.023	25.274	0.012	24.862	0.023	25.245	0.005
Ninawa	21.18	0.019	21.676	0.074	21.242	0.019	21.551	0.014
Salah Al-Din	23.72	0.020	24.220	0.051	23.788	0.020	24.122	0.010
Sulaymaniyah	19.03	0.020	19.634	0.020	19.093	0.020	19.582	0.010
Wasit	24.98	0.000	25.712	0.021	24.980	0.000	25.641	0.010
Iraq (Overall)	**23.35**	0.023	**23.781**	0.013	**23.421**	0.023	**23.767**	0.012

The figure 2 and 3 presented the spatial variation between provinces based on the forecast periods. The north of Iraq has the lower temperature while the south of Iraq has the highest temperature. The west of Iraq has the lower temperature than the east of Iraq in both two forecast periods. The 80% lower and upper forecast temperature are also provided in the A and C columns.

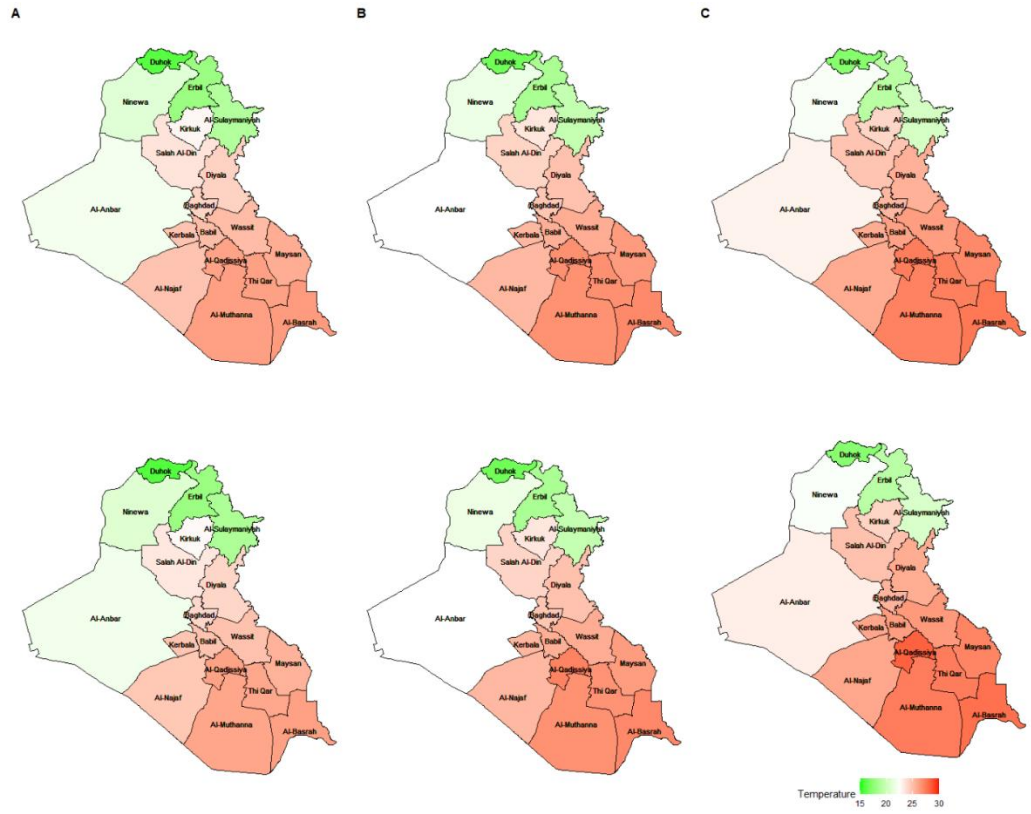

Fig. 2: The forecasted Temperature Map. First row is for 2023-2027. Second Row us for 2028-2032. The (A) column is lower 80% forecast estimate, (B) column is the forecast point estimate and (C) columns is the forecast for higher 80% forecast estimate for Annual Data.

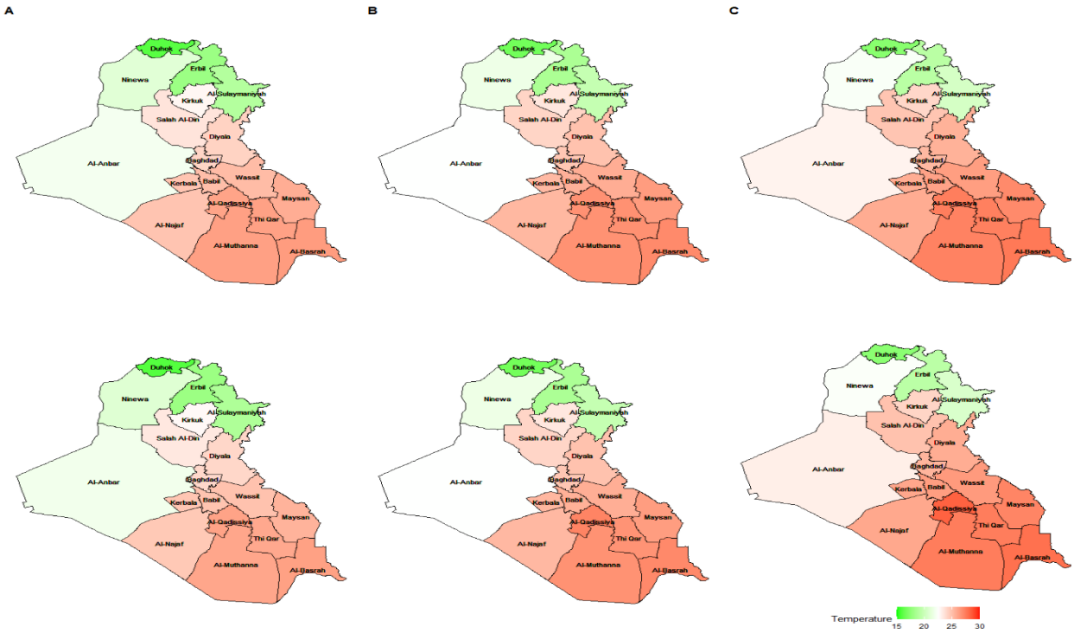

Fig. 3: The forecasted Temperature Map. First row is for 2023-2027. Second Row us for 2028-2032. The (A) column is lower 80% forecast estimate, (B) column is the forecast point estimate and (C) columns is the forecast for higher 80% forecast estimate for Five-Year Smooth Data.

Discussion: The long-term forecast using different global climate models (GCM) indicated an increasing pattern of the temperature of about 5.67–5.91 °C in one scenario and 1.41–1.50 °C under the other scenario within three periods including 2021–2040, 2051–2070, and 2081–2100 in Iraq. [12] The other study estimates the trend with a maximum annual temperature between 0.48 and 2.5 ℃ and a minimum yearly temperature between 0.22 ℃ and 1.76 ℃ with two general circulation models (GCMs) in southwest Iraq from 2020 to 2099. [13] The reasons for the temperature increasing rate of 0.00475 °C/month in Basrah province during 1948–2022 which is faster than the global average (1 °C) and the Middle East region (2 °C) are increasing oil fields, reduction of green cover, global warming, urban expansion, and wetland shrinkage. [14]. The remote sensing–based on Landsat images for three megacities of Iraq considering surface urban heat island (SUHI) estimates an increasing temperature in 30 years for Baghdad (0.16 °C), Basrah (0.44 °C), and Erbil (0.32 °C). [15]

The COVID-19 pandemic and lockdown policies affected many aspects of life such as climate change and air pollution; for example, the positive estimated correlation between COVID-19 confirmed and death cases with the wind speed, air temperature, and solar radiation in Baghdad in 2020 and 2021.[16] Or the reduction of air pollution parameters due to the COVID-19 lockdown. [17-19] They suggested that studying the meteorological parameters is not necessary only for environmental scientists but also their role in the spread of diseases and public health issues are vital. The main limitation of this study is using the historical data and the weather station datasets are not used. The other one is using an ARIMA model for short term forecast that is not suitable for long-term forecasts [20]. One of the future directions of this research is combining information from various resources such as satellite images, historical data, and weather stations to estimate and reduce the parameter variances and refine the results [21]. The other one is comparing and clustering the different countries especially the Middle Eastern states and neighbors of Iraq. [22]

5. Conclusion

The LSTM models with the Random Forest Regressor are trained with 1900 to 2017 and evaluate them with the test data from 2018 to 2022. This paper conclude that the Random Forest Regressor algorithm is the best choose form given model. The best models are selected based on the parameters (units, layers, dropout). The accuracy of the models is estimated with Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). The LSTM models with the Random Forest Regressor select the best model for temperature forecast and they estimate the positive increment on each province.

References

[1] K. Riahi et al., “RCP 8.5—A scenario of comparatively high greenhouse gas emissions,” CLIMATIC CHANGE, vol. 109, (2011), pp. 33–57.
[2] M. Kalkuhl and L. Wenz, “The impact of climate conditions on economic production: Evidence from a global panel of regions,” J. ENVIRON. ECON. MANAGE., vol. 103, (2020), p. 102360.
[3] A. Ahmadalipour and H. Moradkhani, “Escalating heat-stress mortality risk due to global warming in the Middle East and North Africa (MENA),” ENVIRON. INT., vol. 117, (2018), pp. 215–225.
[4] S. A. Salman et al., “Selection of climate models for projection of spatiotemporal changes in temperature of Iraq with uncertainties,” ATMOS. RES., vol. 213, (2018), pp. 509–522.
[5] T. Dimri, S. Ahmad, and M. Sharif, “Time series analysis of climate variables using seasonal ARIMA approach,” J. EARTH SYST. SCI., vol. 129, (2020), pp. 1–16.
[6] M. Amjad et al., “Analysis of temperature variability, trends and prediction in the Karachi Region of Pakistan using ARIMA models,” ATMOSPHERE, vol. 14, no. 1, (2022), p. 88.
[7] A. H. Nury, K. Hasan, and M. J. B. Alam, “Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh,” J. KING SAUD UNIV.–SCI., vol. 29, no. 1, (2017), pp. 47–61.
[8] E. De Saa and L. Ranathunga, “Comparison between ARIMA and deep learning models for temperature forecasting,” ARXIV PREPRINT ARXIV:2011.04452, (2020).
[9] “Average Mean Surface Air Temperature 1991–2020,” WORLD BANK CLIMATE KNOWLEDGE PORTAL, (2024). [Online]. Available: https://climateknowledgeportal.worldbank.org/country/iraq/climate-data-historical
[10] J. S. Racine, REPRODUCIBLE ECONOMETRICS USING R. Oxford: Oxford University Press, (2019).
[11] R. J. Hyndman et al., “Package ‘forecast’,” (2020). [Online]. Available: https://cran.r-project.org/web/packages/forecast/forecast.pdf
[12] Z. M. Mohammed and W. H. Hassan, “Climate change and the projection of future temperature and precipitation in southern Iraq using a LARS-WG model,” MODEL. EARTH SYST. ENVIRON., vol. 8, no. 3, (2022), pp. 4205–4218.
[13] W. H. Hassan and B. K. Nile, “Climate change and predicting future temperature in Iraq using CanESM2 and HadCM3 modeling,” MODEL. EARTH SYST. ENVIRON., vol. 7, (2021), pp. 737–748.
[14] S. A. Al-Asadi et al., “Modeling the impact of land use changes on the trend of monthly temperature in Basrah province, Southern Iraq,” MODEL. EARTH SYST. ENVIRON., vol. 10, no. 3, (2024), pp. 3727–3744.
[15] H. Tao et al., “Megacities’ environmental assessment for Iraq region using satellite image and geo-spatial tools,” ENVIRON. SCI. POLLUT. RES., vol. 30, no. 11, (2023), pp. 30984–31034.
[16] B. M. Hashim et al., “Seasonal correlation of meteorological parameters and PM2.5 with the COVID-19 confirmed cases and deaths in Baghdad, Iraq,” INT. J. DISASTER RISK REDUCT., vol. 94, (2023), p. 103799.
[17] M. Fayaz, “The lock-down effects of COVID-19 on the air pollution indices in Iran and its neighbors,” MODEL. EARTH SYST. ENVIRON., vol. 9, no. 1, (2023), pp. 669–675.

[18] B. M. Hashim et al., "Impact of COVID-19 lockdown on NO_2, O_3, PM2.5 and PM10 concentrations and assessing air quality changes in Baghdad, Iraq," SCI. TOTAL ENVIRON., vol. 754, (2021), p. 141978.
[19] A. Rasul and S. A. Ibrahim, "Relationship between weather and sociodemographic indicators and COVID-19 infection in Iraq," SSRN, (2020).
[20] H. N. Nasir and A. N. A. Hamdan, "Short-term and long-term drought forecasts in Iraq using neural networks and GIS," in IOP CONF. SER.: MATER. SCI. ENG.. IOP Publishing, (2021).
[21] M. S. Sachit et al., "Combining re-analyzed climate data and landcover products to assess the temporal complementarity of wind and solar resources in Iraq," SUSTAINABILITY, vol. 14, no. 1, (2021), p. 388.
[22] W. Terink, W. W. Immerzeel, and P. Droogers, "Climate change projections of precipitation and reference evapotranspiration for the Middle East and Northern Africa until 2050," INT. J. CLIMATOL., vol. 33, no. 14, (2013), pp. 3055–3072.