International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 14, Issue 5 [May. 2025] PP: 186-194

Study on Correction of Temperature Forecast Errors Based on Attention Mechanism LSTM Neural Network

Yao Li[a], Luoluo Shi[a,c,*], Landi Zhong[b], Zijian Zhao[a], Changqing Liu[a]

[a.] *NanchangHangkong University, Science and Technology College, Jiu Jiang, China*
[b.] *Nanchang Meteorological Bureau, Nanchang, China*
[c.] *Meteorological Service Center, Jiangxi Provincial Meteorological Bureau, China*
** Corresponding author*

Abstract: *In this paper, an attention mechanisim based Long Short-Term Memory (LSTM) neural network model cooperating with ECMWF products (EA_LSTM) had been proposed to correct the error of the site temperature forecast in Nanchang City and its districts. The effectiveness of the proposed method and the importance of real data were explored by using the key element of maximum temperature. Firstly, the LSTM was introduced to deeply dig out the temporal feature. Then, combined with attention mechanism, crucial information was strengthened to build an efficient model. Finally, generate temperature predictions over a fully connected network. Byintroducing the real data for the correction forecast test of the next day's maximum temperatures at Nanchang City and its district forecast test stations, the results showed that the maximum temperature values of the mean absolute error (MAE) proposed by EA_LSTM correction method were 14.183% in Nanchang City and 21.649% in its district area less than the forecast method using ECMWF interpolation, whichproved the effectiveness of the method and the importance of real data.*

Keywords: *temperature forecast; deep learning; numerical weather prediction; attention mechanism*

Date of Submission: 12-05-2025 Date of acceptance: 24-05-2025

I. Introduction

Temperature observation techniques have undergone continuous innovation in recent year. Significant advancements have been made in numerical forecasting, observation systems, and synchronization technology, leading to a notable improvement in the accuracy of numerical temperature forecasting technology. Therefore, these advancements have played a crucial role in the field of temperature forecasting [1]. However, the variability of temperature, and factors such as aerosols, numerical forecasts still suffer from certain errorsdue to imperfections in physical models [1-3]. As is well know that the analysis and correction of numerical forecast errors can generally improve forecast accuracy by exploring the relationship model between observed values and numerical forecast products. Therefore, extensive post-processing analysis have been conducted on numerical forecast results, likely statistical correction techniques including perfect prognosis, Kalman filtering, analog ensemble, etc.[4]. However, with the continuous innovation of observation technology and the storage device, the amount of data generated every day far exceeds the capability of traditional data processing methods to extract, proceed, and apply information [5].

The machine learning technology, such as neural networks and support vector machines had been largely applied on the bias correction in numerical weather prediction models[6, 7]. Zhou [8] developed the QPFNet deep learning model to predict precipitation, finding that its performance was superior to ECMWF HRES, especially with increasing precipitation. Moreover, the Model Residual Machine Learning (MRML) method was proposed for generating high-resolution medium-range forecasts, which outperformed traditional downscaling methods[9]. As a whole, these studies hadbeen fiercely demonstrated that machine learning technology was superior than traditional prediction methods.Convolutional Neural Networks (CNN), a special type of neural network, excels in processing image data, primarily extracting and learning features from raw image data. According to former research, Its application in forecast research had also yielded promising results. Atsushi Kudo [10] used an encoder-decoder CNN to predict ground temperature in the Kanto region of Japan, enhancing operational guidance capabilities and correcting NWP model biases. Zhao [11] combined CNN feature extraction with a random forest regression model to predict ECMWF feature factors and precipitation predictors, achieving a lower false alarm rate than the ECMWF interpolation method. As we know, long Short-Term Memory (LSTM) is a variant of the standard Recurrent Neural Network (RNN) [12] and its core lies in three gating units that mine the sequential relationships in data. This neural network effectively models temporal data such as temperature and precipitation changes. Geng [13] proposed the ED-ConvLSTM model to correct numerical weather forecasts, collaboratively extracting spatiotemporal features and fitting complex nonlinear relationships. Zhang [14] introduced the PCS-LSTM model, analyzing the periodicity and closure of time series for more accurate predictions. Hou [15] built a CNN–LSTM model to predict hourly temperatures. Compared to RNN models and traditional models, the excavation of temporal features through the LSTM model makes the model more practical.Hybrid models play a crucial role in temperature forecasting, combining the advantages of multiple models for better performance and wider applicability. Mehdi Neshat [16] established a hybrid model to predict wind power, enhancing the accuracy of wind power forecasting. Arif Ozbek [17] utilized LSTM and the Adaptive Neuro-Fuzzy Inference System (ANFIS) of the FCM machine learning model to predict relative humidity in Turkey, achieving satisfactory results. Zhang [18] leveraged AutoML technology to propose a deep learning framework that models the spatiotemporal dynamics of multimodal meteorological data.

In summary, the application of neural network systems in the field of meteorology has greatly facilitated the development of temperature forecasting systems and yielded promising research results. However, there is a lack of sufficient research on the impact of temporal features in real data on temperature correction. This paper proposed an LSTM algorithm based on the attention mechanism to explore the temporal characteristics of meteorological data and constructs a temperature forecast error correction model for ECMWF products [19]. Moreover, the performance of traditional LSTM and GRU modelshad been compared. By forecasting the maximum temperature at observation stations in Nanchang city (Capital of Jiangxi province, China), we investigate the influence of temporal features on temperature and evaluate the advantages, disadvantages, and feasibility of deep learning in numerical forecast bias correction. As a result, this research contributed to providing a method to improve the accuracy of numerical forecasts andcarried new ideas and methods for the field of meteorological prediction.

II. Research Methodology

Considering the complex combination of real and Times data and the variability of temperature data, an LSTM neural network prediction method based on the attention mechanism was proposed, i.e., using the information selection mechanism to aggregate the input information, and utilizing the attention mechanism to filter the information and the invalid information to improve the accuracy and effectiveness of the temperature prediction.

2.1 Long Short-Term Memory Neural Network (LSTM)

LongShort-TermMemory (LSTM) networks are widely utilized for data sequential relation mining, especially for long time series, which is well suited for temperature prediction. Compared with traditional recurrent neural networks, LSTM can easily memorize long-term information, which is attributed to its internal gating structure. Its specific unit structure is shown in Fig. 1.

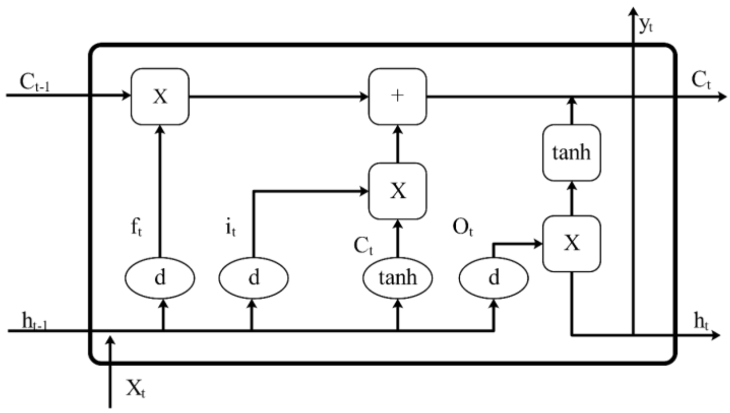

Fig. 1 Structure of LSTM network

Where X_t is the input and $h_t$represents the output, the core of which lies in the cell state, running like a bond through the entire network, Which allows partial information to pass through without loss. As a common rule, LSTM controls the flow of information through three gate structures - forgetting gate, input gate and output gate, all of which are realized by *Sigmoid* function and point-by-point multiplication operation. The Sigmoid function can decide whether the information passes or not and the output value is between [0,1], where '0' means the information has been discarded and '1' represent thatthe information has been retained and passed. As a whole, all LSTM cells follow the following set of equations for the operation.

$$f_t = \delta(w_f \cdot \left(h_{t-1}, x_t + b_f\right) \quad (1)$$

$$i_t = \delta(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\widetilde{C_t} = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$
$$C_t = f_t * C_{t-1} + i_t * \widetilde{C_t} \quad (4)$$
$$\delta_t = \delta(W_0[h_{t-1}, x_t] + b_0) \quad (5)$$
$$h_t = o_t * tanh(C_t) \quad (6)$$

From the above equation, It can be seen that the forgetting gate decides how much of the previous moment's cell state C_{t-1} is preserved to the current moment C_t, while the input gate decides how much of the current moment's input X_t of the network is preserved to the cell's state C_t, and its output gate: controls how much of the cell's current moment's state C_t is outputted to the LSTM's current output value h_t

2.2 The Attention Mechanism

To solve the problem of information overload that results in unsatisfactory network outcomes, the attention mechanism in the neural network was invoked in this paper to assign different weights to the data to obtain the optimization effect of the prediction curve. The structure of this mechanism is shown in Fig. 2. The model follows the equations (7)-(9) for operation.

$$\alpha_n = softmax\big(s(x_n, q)\big) = \frac{exp(s(x_n,q))}{\sum_{j=1}^{N} \exp(s(x_n,q))} \quad (7)$$
$$\mathrm{s}(x, q) = x^T q \quad (8)$$
$$\mathrm{attention}(X, q) = \sum_{n=1}^{N} \alpha_n x_n \quad (9)$$

In the equation, the $\boldsymbol{\alpha_n}$ represents the attention distribution, $\boldsymbol{s(x,q)}$ is the attention scoring function, while⊗ is the production of weights and inputs. Then, the correlation between the query vector q and x is calculated by dot product method, and different weight coefficients are also obtained. After calculating the weight coefficients, the Softmax activation function is used to normalize the output weights. The output of α_n and LSTM layer$\boldsymbol{h_t}$are weighted and summed to obtain a new output vector $\boldsymbol{\alpha}$. As a result, the attention mechanism can reduce the interference of invalid information to the model and thus improve the prediction accuracy.

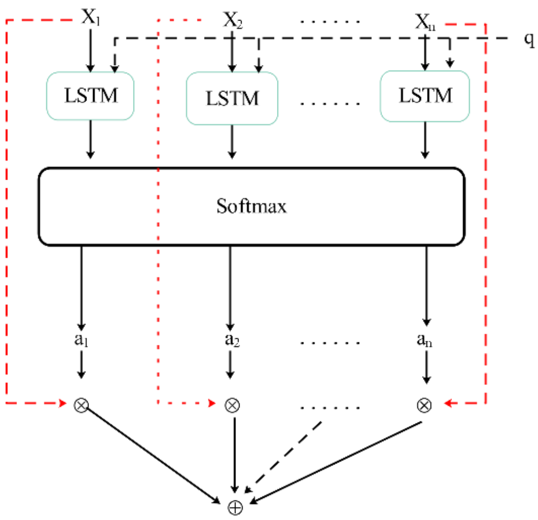

Fig. 2. Structure of the attention mechanism

2.3 Temperature prediction methods

In this paper, a deep learning method for temperature prediction had beenfigured out. The convolved input data had put into an LSTM network, and the main factors affecting the temperature were later enhanced using the attention mechanism. The detailed structure is shown in Fig. 3, which were proved to improve the accuracy of the model.In addition, the study invoked the attention mechanism to LSTM networks can learn sufficiently. First of all, the normalized data through a one-dimensional convolutional layer for convolutional processing, the use of convolutional neural networks of dimensional reduction ability to reduce the input data to achieve the purpose of reducing the amount of subsequent calculations. The output data after convolution is input to the LSTM layer to mine the time series features in the data. When the output of the LSTM network passes through the attention mechanism layer, the correlation between it and the output vector is calculated by the query vector, and the normalized attention distribution is obtained from the Softmax layer, and the distribution enhances the influence of the important time-series features on the model, and suppresses the interference of the unimportant features in order to achieve the purpose of improving the accuracy of the model. Finally, the output results are linearly fitted in a two-layer fully connected network to obtain the temperature prediction curve.

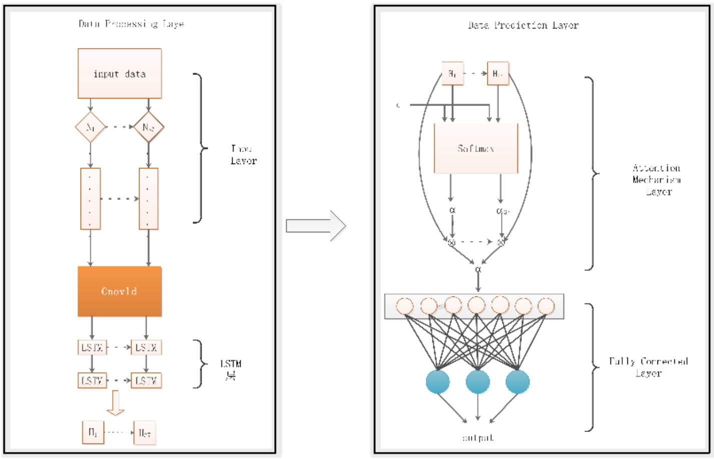

Fig. 3 LSTM neural network approach based on an attention mechanism

III. Data analysis and testing

3.1 Data source

In this paper, ECMWF data and real data are used to verify the validity of the proposed model by comparing the assessment indexes of the model and the temperature prediction curves. The experimental data are obtained from the European Center for Numerical Forecasting (ECMWF) numerical forecasts of Nanchang City and its jurisdictions as well as the local real data (as shown in Table 1). Among them, the time span of Nanchang City is from January 1, 2018 to December 31, 2021 (a total of 1461days), and the time span of the rest of the stations was from January 1, 2018 to August 27, 2022 (a total of 1701days), the ECMWF model starts at 20 o'clock every day, and the time span was 120 hours, in the first 72 hours, the data was forecast at a frequency of every 3 hours, and in the subsequent time period, the forecast frequency is reduced to once every 8 hours. The real data elements consisted of daily maximum temperature, minimum temperature, average temperature and daily precipitation, for single forecast valid time and single forecast validity period, the data sample size included 1461 days and 1701 days.

Table 1. Description of station data

Serial Number	Station location	Time span
Site 1	Nanchang City	2018-01-01 ~ 2021-12-31
Site 2	Anyi County	2018-01-01 ~ 2022-08-27
Site 3	Jinxian County	2018-01-01 ~ 2022-08-27
Site 4	Xinjian District	2018-01-01 ~ 2022-08-27

3.2 Data Preprocessing

To fully explore the feature information in the data so as to get more accurate results, it is necessary to process the numerical forecast data and real data. For the data pre-processing, the following steps are listed: data cleaning and data normalization. Normally, there were a large number of singular values and missing values in the original data-set. To begin with, the singular values of barometric pressure and the precipitation data were artificially cleaned in order to comply with the experimental requirements. Then, for the large number of missing values in the air pressure and part of the precipitation data, this study chose to delete this kind of feature data. Finally, for the missing values in the air temperature data and daily precipitation data, this paper chose the mean interpolation method to make up the time series data to ensure the integrity of its features. Given that the data-set was time series, the time series problem was required to be seriously notified. Firstly, the live data and numerical forecast data were integrated in chronological order to ensure the time consistency, and then the data-set was divided into a training set (80%) and a test set (20%), in which January 1, 2018 to March 13, 2021 was the training set, and March 14, 2021 to the end of the year was set as the test set.

Considering the normalization step, as the data used were combined by multiple categories with different units, different types of values had different impacts on the model training effect. To eliminate the impact of units and ensure the comparability of numerical features, this paper adopted the Z-Score standardization method, which could transform the different levels of data into a unified measure of the Z-Score score, enhance data comparability and reduce the difficulty of parsing. The standardization formula is written as followed:

$$\delta = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - \omega)^2} \tag{10}$$

$$Z = \frac{X-\omega}{\sigma} \tag{11}$$

where δ is the standard deviation of the overall data and ω is the mean of the overall data, and *X* is the individual observation. The standardized data were constructed as a three-dimensional matrix of *(S, T, F),* where S is the number of time series, *T* is the amount of meteorological elements (37), and *F* is the hyper-parameter (uniformly 1). In the case of a single meteorological site, the input samples would be a three-dimensional matrix of (3, 37, 1), if three consecutive days of complete data were used.

3.3 Testing Methods

Choosing the appropriate evaluation function is crucial to the experiment. It is obviously that the root mean square error (RMSE) can effectively reflect the error between the predicted value and the true value, and reduce the impact of large errors on the overall fit. As a comparison, the mean absolute error (MAE) are used to indicate the absolute magnitude of prediction error, which is less sensitive to outliers and is rarely affected by the positive or negative direction of the error. To develop both of the methods mentioned above and create innovative analysis method, the combination of the two method is used to comprehensively assess the performance of the model.

(1) Root Mean Squared Error (RMSE)

Root Mean Square Error (RMSE) is considered to be a commonly used performance evaluation index in regression problems, which can indicate the stability of the forecast level. Generally, a smaller value of RMSE indicates that the model's forecast levels of each region are relatively close, and vice versa when the value of RMSE is large, the model's forecast level varies greatly in different regions. In summary, the formula of RMSE can be written as followed:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widetilde{y_i} - y_i)^2} \tag{12}$$

（2）Mean Absolute Error (MAE)。

The MAE is the average of the absolute values of the deviations of all individual observations from the arithmetic mean, and the MAE better reflects the reality of the prediction value error. Its formula is as follows:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\widetilde{y_i} - y_i| \tag{13}$$

IV. Analysis of the revised effects of the model

4.1 Experimental setup

To explore the feasibility and effectiveness of the proposed method, the data of station 1 were utilized to conduct a prediction experiment of the maximum temperature for the next five days. In addition, to investigate the applicability of the model in temperature prediction correction, this study predicted and analyzed the maximum temperature for the next day at other stations. Firstly, the model had been trained with site 1 data, and the model revision improvement rate (the ratio of the difference between the MAE before and after revision to the MAE before revision) was also introduced to the model performance except for the RMSE and MAE as the evaluation criteria. At the same time, in order to explore the necessity of real data, the following experiments were designed: LSTM and E-LSTM, A-LSTM and EA-LSTM, GRU and E-GRU. As a whole, the specific descriptions of each model were shown in Table 2.

Table 2. Model Descriptions

Model name	Model Description
LSTM	Using only forecast data
ECMWF-LSTM(E-LSTM)	Using integrated data of forecast and real-world observations
Attention-LSTM(A-LSTM)	Using only forecast data
ECMWF- Attention-LSTM(EA-LSTM)	Using integrated data of forecast and real-world observations
GRU	Using only forecast data
ECMWF-GRU(E-GRU)	Using integrated data of forecast and real-world observations

4.2 Experimental Results for Station 1

Table 3 presented the revision results of different models in details for the maximum temperature in one day at site 1 in 2018. After a comprehensive analysis of 30 groups, all types of models showed a revised effect compared to the original temperature forecast at the site.LSTM model played a revised role but took the minimum effect, only reducing the overall MAE error to 1.60 ℃ andshowing a decrease up to 0.07 ℃ (3.91%). By contrast, EA_LSTM model had the best revised effect, reducing the overall MAE error to 1.43 ℃, with a decreasing degree up to 0.24 ℃ (14.18%), becoming the unique model with an improvement rate of more than 10%, which was due to the fact that the attention mechanism assigned different weights to each element, enabling the model to focus on the important parts of the data, and the models with the addition of the attention mechanism all outperform the models without the addition of the attention mechanism in the single-variable experiments.It is obviously that the attention mechanism owns the ability to enhance the performance of the model in terms of its. According to the analysis of the data processing, the model with real data was better than the model with only ECMWF data, for real data made the data more consistent with the laws of physics.

In summary, the EA_LSTM model showed the best revision effect in temperature revision, which strongly proved the importance and effectiveness of combining real data and introducing the attention mechanism to enhance the model performance and improved the prediction accuracy when predicting the temperature.

Table 3. Evaluation metrics for different model revisions for Site 1 in 2018

Forecast Duration/Days	1		2		3		4		5	
Model Name	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
ECMWF	1.35	1.62	1.55	1.86	1.68	2.01	1.79	2.28	1.97	2.51
LSTM	1.37	1.73	1.47	1.79	1.51	1.90	1.74	2.26	1.92	2.57
E_LSTM	1.12	1.36	1.35	1.75	1.57	2.07	1.65	2.15	1.98	2.58
A_LSTM	1.12	1.47	1.37	1.76	1.68	2.21	1.72	2.20	1.99	2.68
EA_LSTM	1.05	1.37	1.18	1.62	1.41	1.92	1.65	2.25	1.87	2.46
GRU	1.29	1.50	1.63	1.92	1.47	1.92	1.69	1.69	1.93	2.48
E_GRU	1.17	1.42	1.47	1.85	1.54	1.89	1.63	2.26	1.88	2.49

In order to verify the stability of the models, this study chose site 1 as the experimental target, and repeated each experiment ten times, and plotted the root mean square error as a box-and-line plot, as shown in Figure 4.in which the root mean square errors of the E_LSTM and EA_LSTM models showed a continuous decreasing trend. The figure indicated that these two models maintainedexcellent stability and consistency in multiple repetitive experiments and could accurately predict temperature change. While compared to the E_GRU model, it presented the largest error span and showed an increasing trend. It was indicated that the model did not perform stably enough in multiple experiments, and the fluctuation of the prediction results was relatively large.

Moreover, the analysis of the box-and-line plot showed that the model with the added attention mechanism performed the smallest error span, exhibiting high stability of the model. Therefore, by introducing the attention mechanism, the model was able to better focus on the crucial information and ignore the irrelevant noise, thus improving the accuracy and stability of the prediction results.

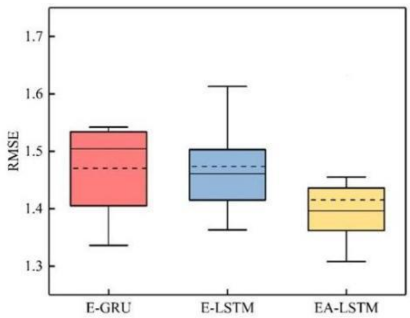

Fig. 4. Box Plot for Station 1

Learning Rate, as an important parameter in the model training process, made a significant effect on the convergence of the model prediction curve.Tofurther explore this effect, this paper carried out a systematic analysis through seven single-control-variable experiments, the result of which was shown in Figure 5. when the Learning rate lay in the range between 0.001 and 0.01, the model exhibited a stable downward trend, especially when the learning rate was 0.01, the result reached the optimal convergence state.However, when the learning rate increased to 0.25, the model performance showed abnormal fluctuations, which indicated that the model used the proposed method showed excellent stability when the learning rate was small, and will be abnormal when learning rate exceeded the limitation.

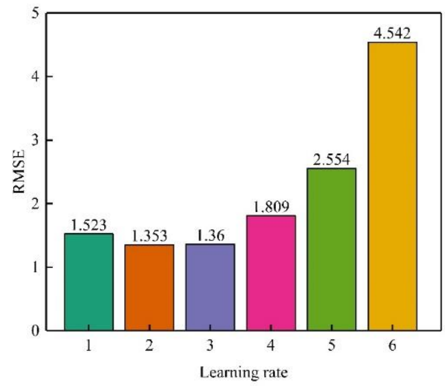

Fig. 5. RMSE of the proposed model at different learning rates

4.3 Correction of Temperature Forecasts for Other Stations

To deeply explore the significant advantages of the proposed method in temperature revision. The study specifically analyzed the maximum temperature prediction for a period of one day for stations No.2, No.3 and No.4. As shown in Fig. 6, the original station temperature predictions values were generally low at these three stations. However, after the revision of different models, the valueswere significantly increasing. Among many models, the proposed model had the highest fitness degree between the prediction curves and the observation curves, and all the curves showed a convergence trend in the early stage, which was highly consistent with the observation curves. At the end of prediction time, some curves showed a gradual divergence trend.It is worth noting that the EA_LSTM model curves and the E_LSTM model curves kept maintaining a high degree of fitting with the observed curves. Based on the detailed data in Table 4, it was further figured out that the root mean square error and mean absolute value error of the EA_LSTM model were lower than those of the other models.In addition, compared with the original temperature forecast at the station, the E_LSTM model reduced the mean absolute value error by 17.729%, while the EA_LSTM model reduced this error by 21.649%. As a whole, this significant improvement was mainly attributed to the attention mechanism in the EA_LSTM model, which effectively strengthened the influence on the critical features in the temperature prediction task while suppressing the interference of non-critical features on the model prediction results, thus substantially improving the prediction accuracy.

Table 4. Evaluation metrics for different model revisions by site in 2022

Serial Number	Site 2		Site 3		Site 4	
Model Name	MAE	RMSE	MAE	RMSE	MAE	RMSE
ECMWF	1.44	1.80	1.44	1.74	1.56	1.93
LSTM	1.22	1.64	1.24	1.64	1.27	1.69
E_LSTM	1.22	1.58	1.14	1.55	1.29	1.66
A_LSTM	1.33	1.74	1.23	1.68	1.23	1.66
EA_LSTM	1.20	1.61	1.10	1.50	1.18	1.56
GRU	1.38	1.75	1.36	1.74	1.28	1.71
E_GRU	1.37	1.73	1.30	1.64	1.27	1.63

Consequently, theproposed model not only makes up for the shortcomings of the LSTM model in some aspects, but also further highlights the effectiveness of the LSTM neural network incorporating the attention mechanism in combining the real data and the timescale (e.g., ECMWF).All in all, This model effectively surmounts the limitations of the numerical temperature prediction and proves the excellent performance and strong advantages of this scheme in temperature revision.

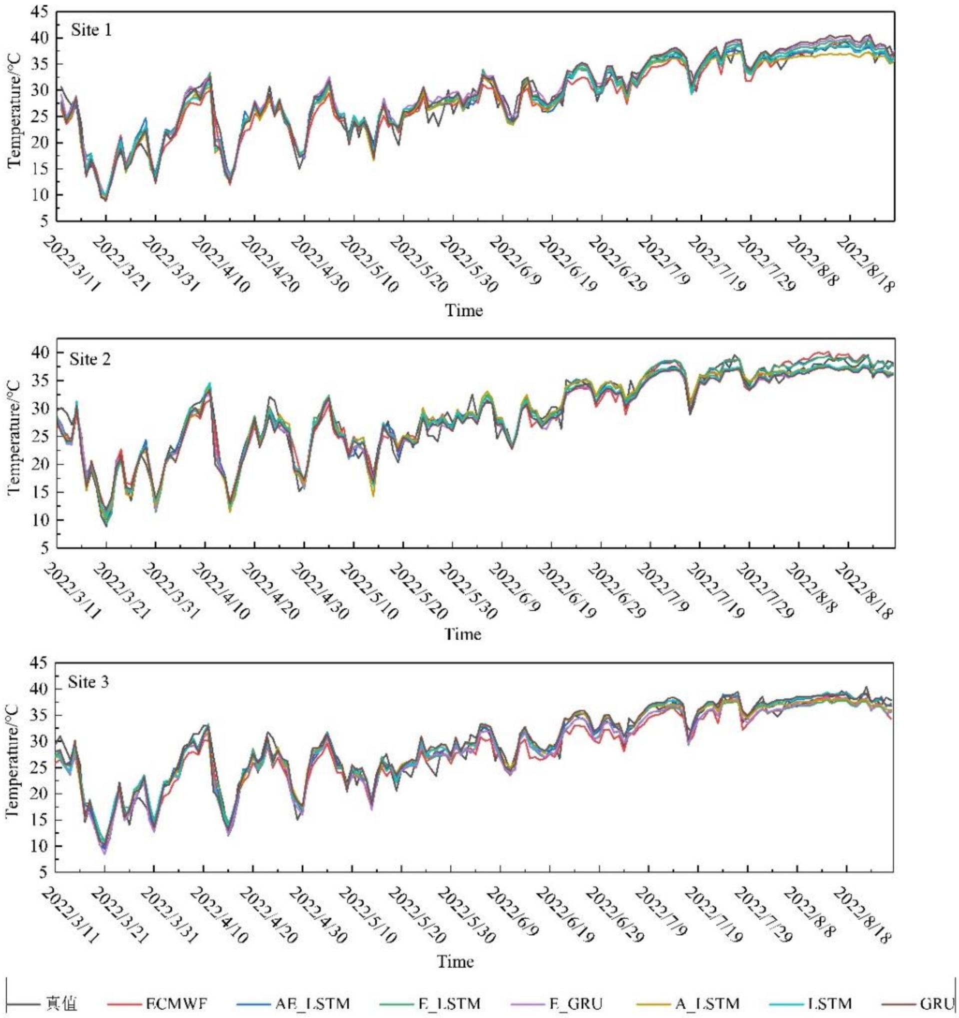

Fig. 6. Comparison of Temperature Forecasts and Observations for Stations 1, 2, and 3 on Day 1 from March 11 to August 27, 2022

V. Conclusion

In this paper, an LSTM neural network model based on the attention mechanism had been proposed, aiming at revising the maximum temperature forecasting results of the ECMWF model product and the real data. The model firstly extracted features by one-dimensional CNN, then utilized LSTM neural network for prediction and redistributed the weights by combining the attention mechanism, and finally obtains the temperature prediction curves after linear mapping of FCN. The experimental results showed that the model not only hadstrong stability and excellent performance, but also verified the effectiveness of the combination of real data and numerical forecasting model. As a whole, the conclusions were obtained as followed:

(1) Forecasting experiments at multiple sites of Nanchang City and its district area have proved that combining ECMWF data with real data can significantly improve maximum temperature forecasting accuracy.

(2) The EA_LSTM model performs best in the prediction task for its attention mechanism highlights the features that affect the prediction result, making the prediction more accurate. After ten time tests , the model error is stable and the error fluctuation can be reduced after adjusting the learning rate.

The method proposed have effectively played a revision effect on temperature, and effectively integrate real data and ECMWF data to improve the accuracy and stability of temperature prediction. Moreover, it shows a great promise in solving other meteorological data prediction problems in weather forecasting. Therefore the method can be accurately and stably applied to a wider and more complex application scenario.

References

[1]. P. Bauer, A. Thorpe, G. Brunet.(2015) The quiet revolution of numerical weather prediction. Nature, J 525,47-55，.

[2]. P. Lynch.(2008) The origins of computer weather prediction and climate modeling. Journal of Computational Physics, J 227,3431-3444，.

[3]. X. Huang, A. Ding.(2021) Aerosol as a critical factor causing forecast biases of air temperature in global numerical weather prediction models. Science Bulletin, J 66,1917-1924，.

[4]. J. Stanger, I. Finney, A. Weisheimer, T. Palmer.(2019) Optimising the use of ensemble information in numerical weather forecasts of wind power generation. Environmental Research Letters, J 14.

[5]. A. Agapiou.(2016) Remote sensing heritage in a petabyte-scale: satellite data and heritage Earth Engine© applications. International Journal of Digital Earth, J 10,85-102，.

[6]. L. Han, M. Chen, K. Chen, H. Chen, Y. Zhang, B. Lu, L. Song, R. Qin.(2021) A Deep Learning Method for Bias Correction of ECMWF 24–240 h Forecasts. Advances in Atmospheric Sciences, J 38,1444-1459，.

[7]. D.C.R. Novitasari, H. Rohayani, Suwanto, Arnita, Rico, R. Junaidi, R.D.N. Setyowati, R. Pramulya, F. Setiawan.(2020) Weather Parameters Forecasting as Variables for Rainfall Prediction using Adaptive Neuro Fuzzy Inference System (ANFIS) and Support Vector Regression (SVR). Journal of Physics: Conference Series, J 1501.

[8]. K. Zhou, J. Sun, Y. Zheng, Y. Zhang.(2022) Quantitative Precipitation Forecast Experiment Based on Basic NWP Variables Using Deep Learning. Advances in Atmospheric Sciences, J 39,1472-1486，.

[9]. C. Zhou, H. Li, C. Yu, J. Xia, P. Zhang.(2022) A station-data-based model residual machine learning method for fine-grained meteorological grid prediction. Applied Mathematics and Mechanics, J 43,155-166，.

[10]. A. Kudo.(2022) Statistical Post-Processing for Gridded Temperature Prediction Using Encoder–Decoder-Based Deep Convolutional Neural Networks. Journal of the Meteorological Society of Japan. Ser. II, J 100,219-232，.

[11]. H. Zhao, X. Huang, Y. Huang, J. Jin.(2021) A revised forecast method of ECMWF precipitation based on CNN feature extraction. IOP Conference Series: Earth and Environmental Science, J 658.

[12]. A. Sherstinsky.(2020) Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. Physica D: Nonlinear Phenomena, J 404.

[13]. G. Huan-tong, H.U. Zhong-yan, W. Tian-lei.(2022) ConvLSTM Based Temperature Forecast Modification Model for North China. Journal of Tropical Meteorology, J 28,405-412，.

[14]. J. Zhang, P. Wu, X. Xu, M. Han, B. Pan.(2022) PCS-LSTM: A hybrid deep learning model for multi-stations joint temperature prediction based on periodicity and closeness. Neurocomputing, J 501,151-161，.

[15]. J. Hou, Y. Wang, J. Zhou, Q. Tian.(2022) Prediction of hourly air temperature based on CNN–LSTM. Geomatics, Natural Hazards and Risk, J 13,1962-1986，.

[16]. M. Neshat, M.M. Nezhad, E. Abbasnejad, S. Mirjalili, L.B. Tjernberg, D. Astiaso Garcia, B. Alexander, M. Wagner.(2021) A deep learning-based evolutionary model for short-term wind speed forecasting: A case study of the Lillgrund offshore wind farm. Energy Conversion and Management, J 236.

[17]. A. Ozbek, Ş. Ünal, M. Bilgili.(2022) Daily average relative humidity forecasting with LSTM neural network and ANFIS approaches. Theoretical and Applied Climatology, J 150,697-714，.

[18]. X. Zhang, Q. Jin, T. Yu, S. Xiang, Q. Kuang, V. Prinet, C. Pan.(2022) Multi-modal spatio-temporal meteorological forecasting with deep neural network. ISPRS Journal of Photogrammetry and Remote Sensing, J 188,380-393，.

[19]. Z. Zhao, P. Zou.(2022) Attention-based Dual-channel Deep Neural Network For Aero-engine RUL Prediction Under Time-varying Operating Conditions. Journal of Physics: Conference Series, J 2386.