Application of Kalman Filter Method in the Forecast of Temperature in Nanchang

Feifei WU, Xiaoyou LONG, Chuanshi TANG, Landi ZHONG*

Nanchang Meteorological Bureau, Nanchang 330038, China

Abstract A temperature forecasting model was created firstly based on the Kalman filter method, and then used to predict the highest and lowest temperature in Nanchang station from October 27 to November 1, 2017. Finally, according to the empirical forecasting method, guidance forecasts were established for the northern, central, and southern parts of Nanchang City. After inspection, it was found that the temperature prediction model established based on the Kalman filter method in Nanchang station had good prediction performance, and especially in the 24-hour forecast, it had advantages over the European Center. The accuracy of low temperature forecast was better than that of high temperature forecast.

Key words Kalman filter method; Temperature forecast; Nanchang City

DOI 10.19547/j.issn2152 – 3940.2024.04.009

The change of temperature is closely related to people' life, and temperature forecast can guide citizens to wear clothes and travel. With the improvement of living standards, citizens' demand for refined temperature forecast is increasing day by day. At present, temperature forecast mainly relies on numerical model results, and no in-depth research has been conducted on how to correct numerical model results in local areas. The research on the refined forecasting method fails to meet the current social requirements.

The temperature in different places of a region is often affected by terrain, underlying surface, environment and other factors. The numerical forecast data based on the average of regional ideal state has certain differences in the forecast of such daily extreme value. In daily work, forecasters need to revise numerical forecast products according to experience, but the correction effect is often greatly reduced because the revision method is not systematic and quantitative.

Nanchang is located in the Poyang Lake area, and borders Meiling in the north. There is heat island effect in the urban area. Environmental factors make temperature forecast more complicated. Especially, the weather forecast of Nanchang as a provincial capital has received more extensive attention, so it is urgent to carry out relevant analysis and research.

In this paper, based on the monitoring data of temperature over the years and the summarized forecast correction experience of forecasters, a systematic quantitative temperature forecast correction method which can be used in daily forecast work was studied to provide technical support for future forecast. It is of great significance to study the local correction scheme of temperature for further improving the refined forecast of villages and towns.

Received: July 13, 2024 Accepted: August 20, 2024

* Corresponding author.

1 Data and methods

1.1 Data sources The data used in this study are from Jiangxi Provincial Information Center, including ECGQ numerical forecast data of Nanchang City during 2016 – 2017, temperature data of national stations in Nanchang City from 2012 to 2016, and data of regional automatic stations in Nanchang during 2012 – 2016.

1.2 Research method At present, MOS forecasting method is widely used in forecasting business, and it is a forecasting method based on numerical forecasting model products. With the development of numerical prediction, numerical models are constantly updated and improved, and there will be some prediction errors in the process of using MOS prediction method. Therefore, it is needed to explore a prediction model that can adapt to the change of a numerical model. In recent years, Kalman filter method has been used more and more in the establishment of prediction models, and has achieved satisfactory results. In this paper, Kalman filter method was combined with local forecasting experience to establish a model suitable for the local maximum and minimum temperature forecast methods in Nanchang.

2 Kalman filter principle

Kalman filter is a modern dynamic system analysis technique. It takes the system state space model as the analysis object, and uses modern stochastic estimation theory to give the recursive estimate of unbiased minimum variance of the system state according to the system model affected by noise and the system observations containing noise interference. By using the basic idea of filtering, the prediction equation can be modified by the feedback information of the prediction error at the transfer time, and the best estimate of the required physical parameters can be obtained by processing a series of actual measurement data with errors. The advantage of using Kalman filter method is that it does not need too much historical data value of numerical forecast products, and can constantly revise the calculation equation, so it has a wide application prospect.

It is assumed that the filtered objects constitute a certain form of stochastic dynamic system, and based on this stochastic dynamic system, recursive filtering is carried out. The composition of stochastic dynamic system and the method of recursive filtering are introduced as follows.

2.1 Composition of stochastic dynamic system The stochastic dynamic system consists of two sets of equations: state equation (1) and prediction equation (2):

$$\beta_t = \beta_{t-1} + \varepsilon_{t-1} \quad (1)$$

$$Y_t = X_t\beta_t + \varepsilon_t (t \geqslant 1) \quad (2)$$

State equation (1) describes the changing rule of the state of a stochastic dynamic system. In the equation, β_t is the coefficient of the regression equation; ε_{t-1} is the dynamic noise; Y_t is the predicted value at time t, and can be expressed as a n-dimensional variable:

$$Y_t = [y_1, y_2, \cdots, y_n]$$

X_t is the predictor at time t, and can be expressed as a matrix of $n \times m$ dimensions:

$$X_t = \begin{bmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2m} \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ x_{n1} & x_{n2} & \cdots & \cdots & x_{nm} \end{bmatrix}$$

Based on the random vector of state at time $t-1$, the desired random vector of state at time t can be estimated by equation (1). In the process of change from the state at time $t-1$ to the state at time t, the change of the random vector of the state is affected by random perturbations (*i.e.* dynamic noise).

Prediction equation (2) describes the relationship between state variable β_t and prediction variables Y_t and X_t. At the same time, this relationship is affected by measurement noise v_t, and its mean is 0, $v_t \approx N(0, V)$, where V is variance.

It is supposed that sums are uncorrelated white noise, and their mean is 0. Therefore, no matter how the number of calculations is increased, there is no need to store a lot of historical data and data, and it is only needed to add, subtract, multiply and invert the matrix, which is the advantage of Kalman filter method.

2.2 Kalman filter recurrence method The generalized least square method is used, and the Kalman filter recurrence formula is as follows:

$$Y'_t = X_t\beta_{t-1}' \quad (3)$$

$$R_t = C_{t-1} + W \quad (4)$$

$$\delta_t = X_t R_t X_t^T + V \quad (5)$$

$$A_t = R_t X_t^T \delta^{-1} \quad (6)$$

$$\beta'_t = \beta'_{t-1} + A_t(Y_t - Y'_t) \quad (7)$$

$$C_t = R_t - A_t\delta_t A_t^T \quad (8)$$

Formula (3) is the prediction equation, where Y'_t is the forecast value; X_t is the forecast factor; β_{t-1}' is the regression number. In formula (4), R_t is the error variance matrix of β value at time T, and W is the variance matrix of error setting value ε_{t-1}. In formula (5), δ_t is the prediction error variance matrix, and X_t^T is the transposed matrix of X_t; V is the variance matrix of measurement noise. In formula (6), A_t is the gain matrix, and δ^{-1} is the inverse matrix of δ_t. In formula (7), Y_t is the actual observed value of the forecast quantity. In formula (8), C_t is the error variance matrix of β_t. By repeated calculation from formula (3) to (8), the coefficients in the prediction equation can be corrected while making the prediction.

β_t, C_t, W and V are the four important parameters during the calculation, and the initial values and must be determined to achieve the recursion process by repeatedly calculation.

2.3 Establishment of prediction equations As the temperature of Nanchang station represents that of surrounding areas, it is decided to establish the prediction equation of temperature in Nanchang station. The forecast of temperature in other stations can be revised by empirical means based on the results of Nanchang station. The 2M temperature, 1 000 hPa temperature, 925 hPa temperature and 850 hPa temperature predicted by ECGQ numerical model at 20:00, as well as the daily minimum and maximum temperature in the past 24 h are selected as the forecasting factors, and the prediction equations of the minimum and maximum temperature in Nanchang were established by using multiple regression method. Next, the minimum temperature in Nanchang station is taken as an example.

Based on the sample data in 2016, the above regression equation was established, and the regression coefficient was estimated: $\beta_1 = [6.310\,7, 0.090\,0, -0.381\,6, 1.221\,4, 0.108\,2, 0.534\,9]$.

Because it is obtained from the sample data through strict mathematical calculation, so it is considered to be accurate, and it is the zero matrix:

$$C_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

There is no particularly effective method to determine W and V. The research shows that the initial values of these two parameters are not sensitive to the correction of the regression coefficient, so the empirical method is used here to determine them subjectively. According to the assumption of white noise, the initial values of W and V can be determined as follows:

$$W = \begin{bmatrix} 50 & 0 & 0 & 0 & 0 \\ 0 & 50 & 0 & 0 & 0 \\ 0 & 0 & 50 & 0 & 0 \\ 0 & 0 & 0 & 50 & 0 \\ 0 & 0 & 0 & 0 & 50 \end{bmatrix} \quad V = \begin{bmatrix} 10 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 & 10 \end{bmatrix}$$

After the four parameter values of the recurrence start are determined, the prediction equation of the minimum temperature can be obtained according to the recurrence formulas (3) - (8):

$$T_{min} = 6.310\,7 + 0.09T_2 - 0.381\,6T_{1\,000} + 1.221\,4T_{925} + 0.108\,2T_{850} + 0.534\,9T_{min-1}$$

Similarly, the prediction equation of the maximum temperature can be obtained as follows:

$$T_{max} = 5.4339 + 0.1266T_2 - 0.0644T_{1000} + 1.1504T_{925} + 0.2501\ T_{850} + 0.3041T_{max-1}$$

3 Subjective revision of district and county temperature based on Nanchang station

Due to the differences in latitude and longitude, geographical location and surrounding environment of the five national stations in Nanchang City, there are some differences in the monitored temperature of each station. There is a certain difference in temperature under different weather conditions, so it is particularly complicated to predict temperature. Hence, the temperature in the five national stations in Nanchang City is predicted based on the forecast of temperature in Nanchang station. By judging different weather phenomena, weather processes, and sky conditions, the maximum and minimum temperatures are distinguished and empirically corrected.

3.1 Representative analysis of the national stations in Nanchang The annual values of each element from 2012 to 2016 are compared and analyzed, respectively, and the average temperature of all stations in the city (including the national stations in Nanchang) and national stations in Nanchang are shown in Table 1.

Table 1 Average temperature of all stations in the city and national stations in Nanchang

Average	Average temperature//℃	Maximum temperature//℃	Minimum temperature//℃	Average wind speed//m/s	Extreme wind speed//m/s	Relative humidity//%
The whole city	18.87	38.70	-2.55	1.94	17.00	70.99
National stations	18.87	37.77	-1.93	1.75	14.43	74.08

Seen from the three-year average temperature, the average temperature of the national stations was basically the same as that of all stations in the city. From the maximum temperature, the maximum temperature of the national stations was nearly 1 ℃ lower than that of the whole city. The minimum temperature of the national stations was slightly higher than that of the whole city, and the difference was less than 1 ℃.

The average wind speed of the national stations was about 0.19 m/s smaller than that of the whole city, which was related to the location of some regional stations on the banks of rivers and lakes and on the roofs of buildings. The extreme wind speed of the whole city was obviously larger than that of the national stations.

Seen from the average of relative humidity (only five regions stations have stable data of relative humidity), the average relative humidity of the national stations was higher than that of regional stations, and the difference was up to about 4%.

3.2 Comparative analysis of elements of each station in the city From the above analysis, it can be seen that there were some differences in the elements between urban and suburban areas of Nanchang City, especially the average temperature, the maximum temperature and the minimum temperature.

It can be seen from the average temperature in all stations of the city (Fig. 1) that the average temperature in urban areas was higher than that in mountainous suburbs. However, the increase did not exceed 1.5 ℃, and the reduction did not exceed 2 ℃. With the continuous expansion of Nanchang urban area in recent years, the slightly low-temperature areas in the suburbs were only some regional stations in Wanli District.

Similarly, the maximum temperature and minimum temperature in all stations of the city were analyzed. It is found that the average of the maximum temperature in the national stations of Nanchang was 0.93 ℃ lower than the average of the city, and the absolute values of anomaly in most stations were within 1 ℃. The minimum temperature in all stations of the city was 0-1 ℃ higher than the average of the whole city. The minimum temperature in suburban stations was significantly lower, 1-4 ℃ lower than the average of the whole city.

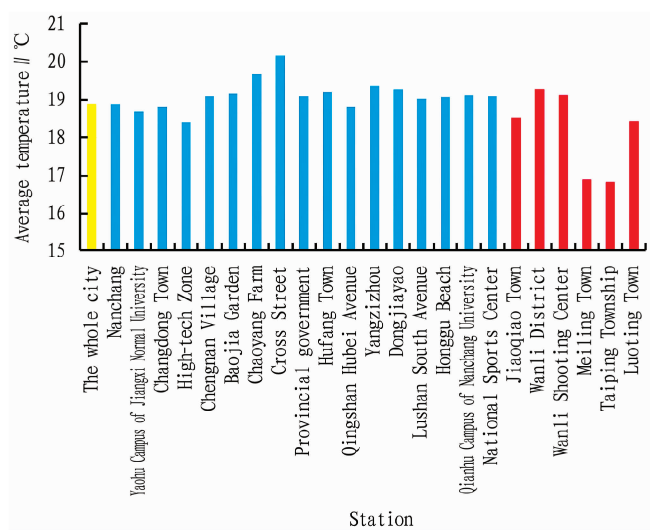

Fig. 1 Average temperature in each station from 2012 to 2016

4 Trial run and test of the method

The new temperature forecast models were used to forecast the temperature from October 27 to November 1, 2017, and the forecast results were tested.

As shown in Fig. 2, when the temperature within 24 h was predicted, the score of the maximum temperature predicted by Kalman filter method was 91, and the score of European Center was 73. The score of the minimum temperature predicted by Kalman filter method was 65, and the score of European Center was 74. As the time is 48 h, the maximum temperature was scored 56 by Kalman filter method and 61 by European Center. The minimum temperature was scored 70 by Kalman filter method or European Center. When the time increased to 72 h, the maximum temperature was scored 65 by Kalman filter method and 50 by European Center. The minimum temperature was scored 62 by Kalman filter method and only 23 by European Center. It is concluded that Kalman filter had a certain advantage over European Center in prediction of the maximum temperature, especially in 24 h. In the forecast of the minimum temperature, Kalman filter method also had certain advantages. In the first 48 h, Kalman filter method was comparable to European Center, and became better after 72 h.

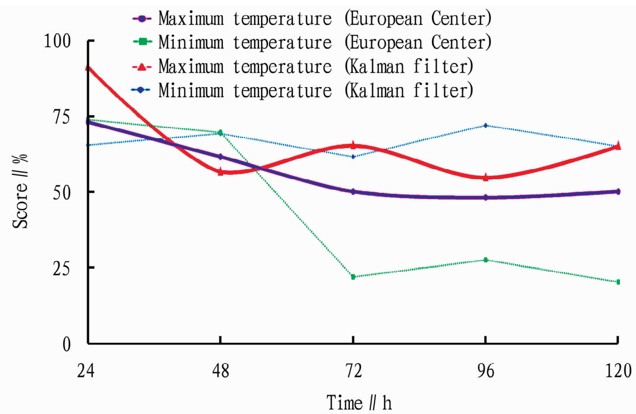

Fig. 2 Score of the prediction of the maximum and minimum temperature from October 27 to November 1, 2017 based on Kalman filter method and European Center

The prediction of Kalman filter method was relatively good, which is closely related to the selection of factors in the equation. In general, Kalman filter method achieved good results in this process. However, the current factors make the system unstable, and more research on the selection of factors is needed in the future, which is worth further discussion in the future work.

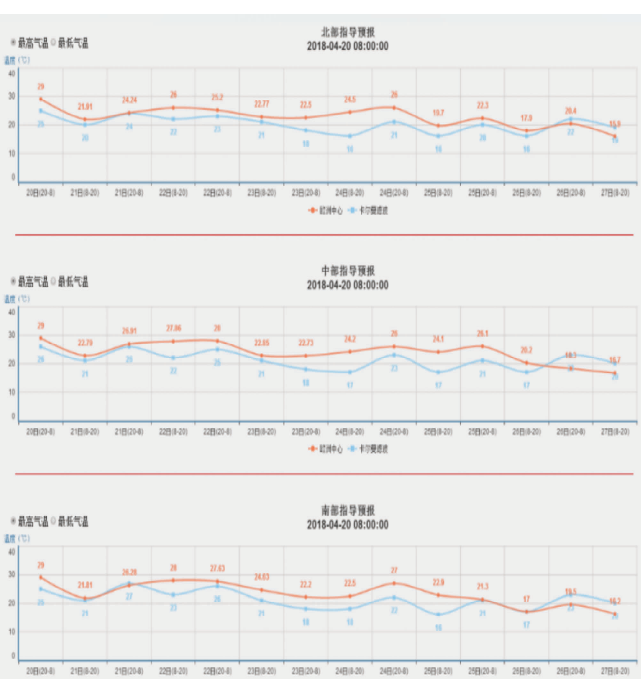

Fig. 3 Example of temperature forecast interface

After the forecast equation was established by Kalman filter, it was corrected by subjective forecast method. According to the predicted weather conditions, the north guide forecast, the middle guide forecast and the south guide forecast were established. As shown in Fig. 3, the new temperature forecast software has a simple interface that allows all values to be seen at a glance and allows comparisons with the prediction of European Center.

5 Conclusions

The temperature forecast model based on Kalman filter method was used to forecast the maximum and minimum temperature of Nanchang station from October 27 to November 1, 2017. Finally, according to the empirical forecasting method, the north guide forecast, the middle guide forecast and the south guide forecast were established. Some conclusions were drawn as follows.

(1) The temperature of Nanchang was higher in the center (near the urban area) than the south and north, which is related to the urban heat island effect to a certain extent. The elements of Nanchang station were representative in the whole city, and the temperature anomaly was about 1 ℃ compared with the average of the whole city.

(2) The temperature forecast model of Nanchang station was established based on Kalman filter method, and was corrected it by the empirical method to obtain the temperature forecast method of the whole city. The test show that it had certain forecasting ability, and especially the forecast within 24 h had advantages over European Center. The accuracy of low temperature forecast was better than that of high temperature forecast.

(3) Kalman filter method requires less historical data and is simple, but the prediction results are closely related to the selection of factors in the equation, so it is worth further research in the selection of factors.

References

[1] QIU XX, WANG DY, ZHU HF. A refined forecast method of township minimum and maximum air temperature[J]. Journal of Meteorology and Environment, 2013, 29(3): 92 –96.

[2] HE CF. Comparative study of ground and 500 hPa multi-model integrated forecast based on Kalman filter[D]. Nanjing: Nanjing University of Information Science and Technology, 2016.

[3] SHAN JS, YUAN ZG, ZHOU JX. Application of Kalman filter method to medium range temperature forecast[J]. Journal of Chengdu University of Information Technology, 2001, 16(1): 12 –16.

[4] SHENG Y, YUAN ZP, CHEN YQ, *et al*. Villages and towns weather forecast operational system in Liaoning Province[J]. Journal of Meteorology and Environment, 2008, 24(3): 13 –17.

[5] ZHANG B, WEI JS, PEI HY. Verification and evaluation of rainfall and temperature forecasting for T213L31 global numerical model in Jiangsu area in 2006[J]. Journal of the Meteorological Sciences, 2008, 28(4): 468 –472.

[6] WU JQ, GUO PW. Refined temperature prediction based on statistical downscaling technique[J]. China Science and Technology Information, 2009(12): 44 –45.

[7] CHEN FJ, JIAO MY, CHEN J. A new scheme of calibration of ensemble forecast products based on bayesian processor of output and its study results for temperature prediction[J]. Meteorological Monthly, 2011, 37(1): 14 –20.