ISSN 2706-655X Vol.3, Issue 1: 1-7, DOI: 10.25236/IJFET.2021.030101

Research on Forecast of Airport Visibility Based on Multivariate Statistics

Weibing Feng, Licheng Jiang* and Danya Liang

College of Science, Xi'an University of Science and Technology, Xi'an 710000, China
**Corresponding author e-mail: 879137768@qq.com*

Abstract: *Studying the correlation between various meteorological elements and visibility is of great significance to the establishment of visibility prediction methods. This paper first uses a quadratic linear polynomial model to fit the relationship between visibility and seven meteorological factors, and uses Matlab's stepwise regression function to solve the problem. Finally, the relationship expression between visibility and various meteorological factors is obtained. And the expression can be better The forecast dominates the trend of visibility.*

Keywords: *Visibility, Meteorological Factors, Polynomial Fitting*

1. Introduction

Visibility prediction is a matter of great concern to highway management departments and airlines [1-3]. In fact, the formation and dissipation of heavy fog has its own laws, which are usually related to meteorological factors near the ground. The video data contains a wealth of information [4-6], especially covering the changing process of the fog. Making full use of this information can not only improve the accuracy of visibility estimation, but also predict the dissipation of heavy fog [7].

This article selects hourly observation data from a domestic airport from December 15, 2019 to March 13, 2020, and establishes a mathematical model to describe the relationship between visibility and ground meteorological observations. Based on this, first calculate the two based on the principle of multivariate statistical analysis. Pearson correlation coefficient of a random vector X and Y, and then use a quadratic linear polynomial model to fit the relationship between visibility and 7 indicators, and use Matlab's stepwise regression function to solve it, and finally obtain the relationship expression between visibility and various meteorological elements, Specifically analyze the weights of the influencing factors of visibility, so as to provide help for the forecast of airport visibility.

2. Analysis of Factors Affecting Visibility

2.1 Scatter diagram of visibility and the distribution of various meteorological elements at that time

By observing the attached data, we found that due to the large amount of meteorological observation information and various forms, here we only consider the data related to the formation and dissipation of fog [8]. First, we organize the useful data together, remove some useless fields, and draw a scatter plot of visibility and meteorological elements (as shown in Figures 1 and 2 below).

It can be seen from the above figure that the greater the visibility, the larger the distribution value range of various meteorological elements, and the smaller the visibility, the smaller the corresponding meteorological element value range. The black box area is below 1000m (inclusive) the distribution of elements. When in low visibility, most of the corresponding element value intervals have a clear range, as shown in Figure 2, the wind speed on December 15, 2019 is less than or equal to 4m/s. Therefore, it shows that the change of visibility is not the result of a single factor, but the result of multiple factors.

ISSN 2706-655X Vol.3, Issue 1: 1-7, DOI: 10.25236/IJFET.2021.030101

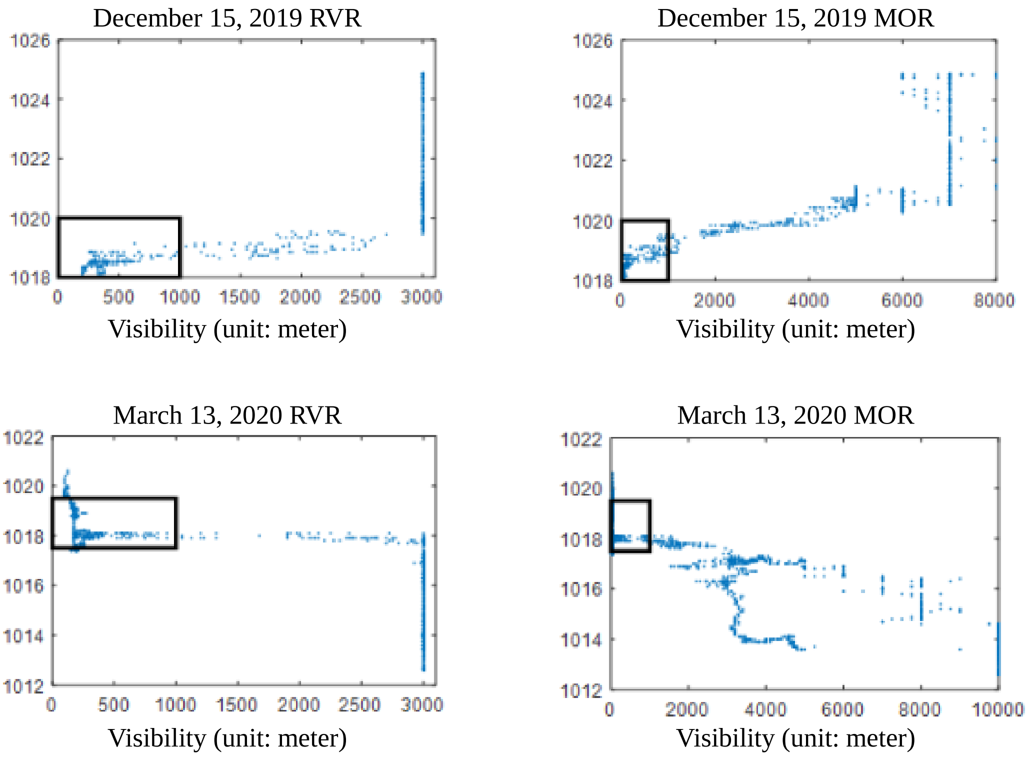

Figure. 1 Scattered distribution of visibility and air pressure

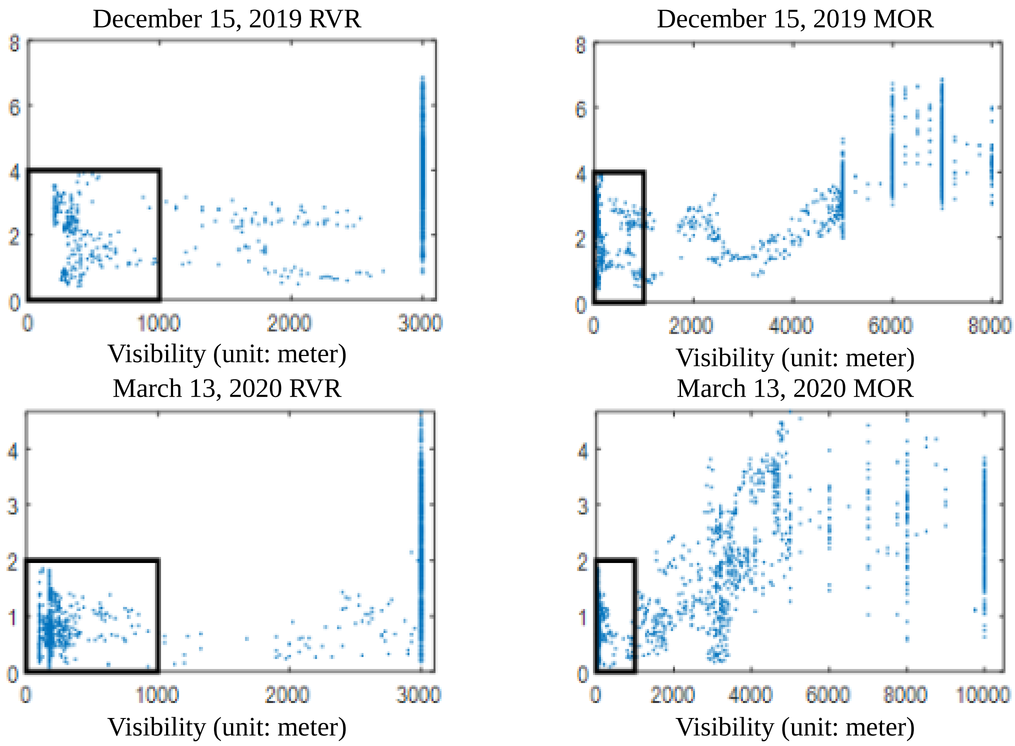

Figure. 2 Scattered distribution of visibility and wind speed

2.2 Time Distribution and Law of Visibility and Its Factors

2.2.1 Time distribution and law of average RVR value in one minute

Figure 3 shows the evolution of the one-minute average RVR value over two days. From the figure, we can see that the RVR value of the two days remained unchanged first, then began to fluctuate and decline, and finally fell to about 200. The difference between the two days is that the fluctuations began

ISSN 2706-655X Vol.3, Issue 1: 1-7, DOI: 10.25236/IJFET.2021.030101

at around 3 o'clock on March 13th, and the fluctuations started around 7 o'clock on December 15th. The different fluctuation times of the two are caused by the difference in temperature, humidity and wind speed in different seasons.

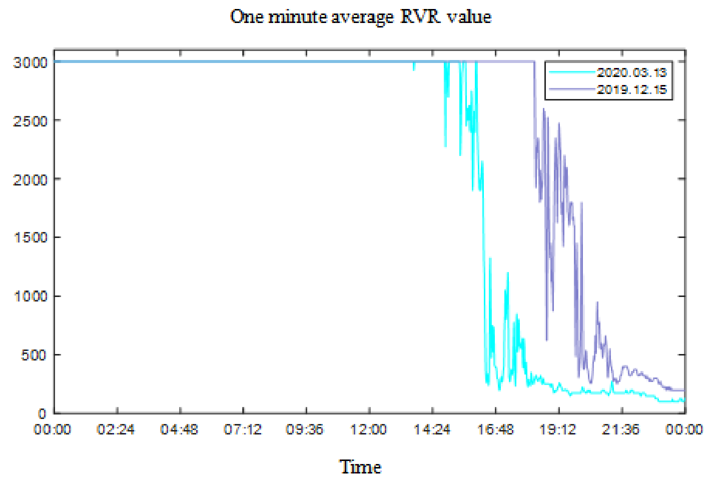

Figure. 3 Evolution trajectory of RVR value in two days

2.2.2 Time distribution and law of one-minute average MOR value

It can be seen from Figure 4 that the MOR value on December 15 showed a fluctuating state around visibility of 7000 around 0-8 o'clock, then began to decline slowly at 8-14 o'clock, and began to show a downward trend after 14:00. In the end, the visibility stayed at around 100. The MOR value on March 13 showed an upward trend at 0-5 o'clock, then showed a steady state around 5-9 o'clock, and began to show a downward trend after 9 o'clock. In the end, it stayed at around 100.

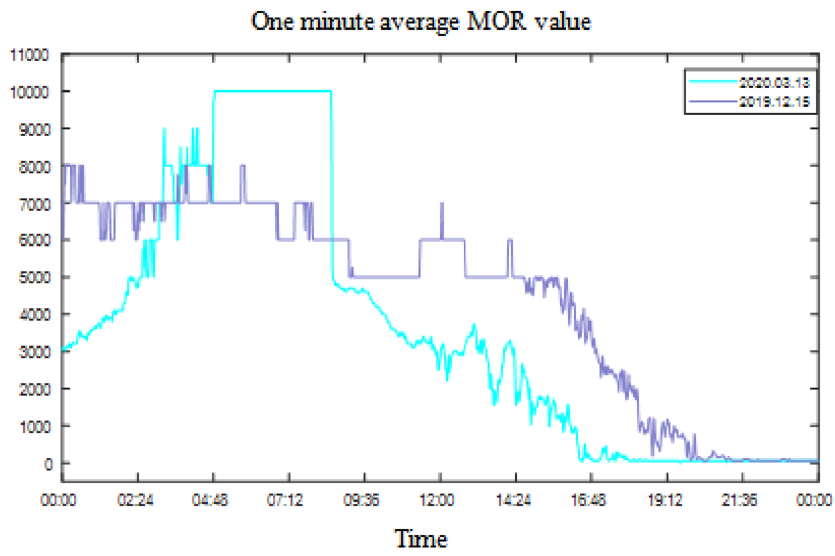

Figure. 4 Evolution trajectory of MOR value in two days

2.2.3 Time distribution and law of two-minute average wind speed

Figure 5 shows the change of wind speed throughout the day on March 13, 2020 and December 15, 2019. From the figure, we can see that the wind speeds of the two days fluctuate, and the wind speed is relatively high at 0-9. After hours the wind speed began to weaken.

ISSN 2706-655X Vol.3, Issue 1: 1-7, DOI: 10.25236/IJFET.2021.030101

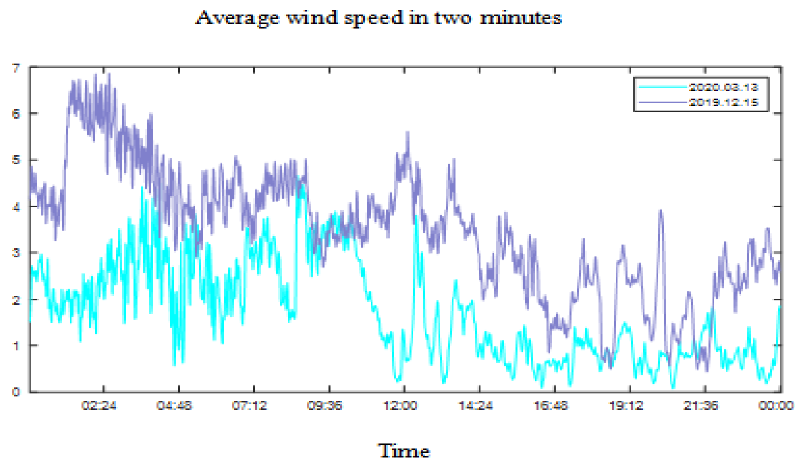

Figure. 5 Wind speed evolution trajectory in two days

2.2.4 Correct the time distribution and law of sea level pressure

Figure 6 shows the corrected sea level pressure changes throughout the day on March 13, 2020 and December 15, 2019. It can be seen from the figure that the pressure on December 15 has been showing a downward trend, and the pressure on March 13 The trend of falling first and then rising.

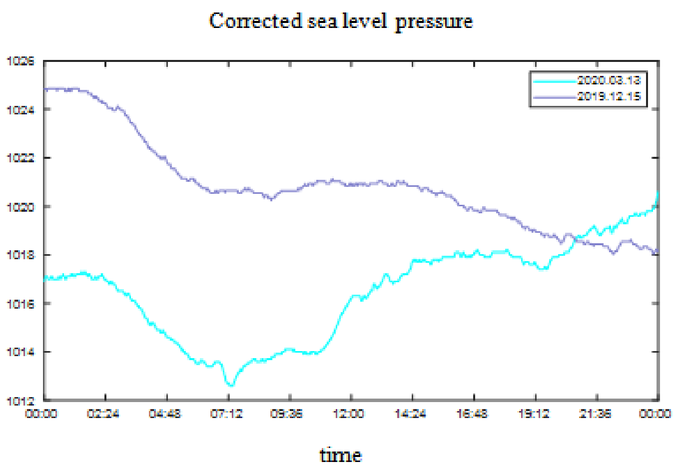

Figure. 6 Modified sea level pressure evolution trajectory in two days

3. Relationship Expression between Visibility and Related Factors

According to the principle of multivariate statistical analysis [9], the Pearson correlation coefficient of two random vectors X and Y can be obtained by the following formula:

$$r_{xy} = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\sum X^2 - \frac{\left(\sum X\right)^2}{N}\right)\left(\sum Y^2 - \frac{\left(\sum Y\right)^2}{N}\right)}} \tag{1}$$

Among them N is the number of samples, and all sums are from 1 to N; the value of the correlation coefficient r is between –1 and +1, that is, $-1 \le r \le +1$ can generally be divided into three levels[10]:

A. $|r|<0.4$ is low-degree linear correlation;

B. $0.4 \le |r| < 0.7$ is a significant correlation;

C. $0.7 \le |r| < 1$ is highly linear correlation.

ISSN 2706-655X Vol.3, Issue 1: 1-7, DOI: 10.25236/IJFET.2021.030101

3.1 Analysis of related factors based on data on December 15, 2019

First use the data on December 15, 2019, first perform outlier detection on the data, and delete the processing of outliers and missing values to obtain a data set for statistical analysis. In addition, because some of the data given in the table is within one minute After measuring the data four times, we averaged this part of the data for calculation, as shown in Table 1 below (due to space limitations, only the first two and last two observations are schematically listed).

Table. 1 Test data on January 15, 2019

RVR	MOR	WS2	CW	QNH	QFE0	PAI	TEM	RH	DEW
3000	7000	3.79	1.85	1024	1023	1023	8.6	90	7.06
3000	7000	3.77	1.91	1024	1023	1023	8.6	90	7.06
…	…	..	..	…	…	..	…	…	…
200	50	2.70	1.41	1018	1016	1016	10.6	100	10.6
200	50	2.60	1.37	1018	1016	1016	10.6	100	10.6

After the above data is preprocessed to remove incomplete data, use Matlab's corrcoef function (set the significance level) to obtain the correlation matrix of temperature, humidity, wind speed, air pressure and visibility in the data as follows:

$$R_1 = \begin{bmatrix} 1.0000 & 0.8469 & 0.5299 & 0.4903 & 0.6490 & 0.6491 & 0.6491 & 0.6214 & -0.6410 & 0.0557 \\ 0.8469 & 1.0000 & 0.7702 & 0.7353 & 0.8139 & 0.8135 & 0.8135 & 0.8114 & -0.8689 & -0.0674 \\ 0.5299 & 0.7702 & 1.0000 & 0.9045 & 0.7510 & 0.7506 & 0.7506 & 0.7072 & -0.7453 & -0.0225 \\ 0.4903 & 0.7353 & 0.9045 & 1.0000 & 0.7402 & 0.7398 & 0.7398 & 0.6880 & -0.7190 & -0.0276 \\ 0.6490 & 0.8139 & 0.7510 & 0.7402 & 1.0000 & 1.0000 & 1.0000 & 0.4601 & -0.6020 & -0.3048 \\ 0.6491 & 0.8135 & 0.7506 & 0.7398 & 1.0000 & 1.0000 & 1.0000 & 0.4593 & -0.6011 & -0.3041 \\ 0.6491 & 0.8135 & 0.7506 & 0.7398 & 1.0000 & 1.0000 & 1.0000 & 0.4593 & -0.6011 & -0.3041 \\ 0.6214 & 0.8114 & 0.7072 & 0.6880 & 0.4610 & 0.4593 & 0.4593 & 1.0000 & -0.9482 & 0.2005 \\ -0.6410 & -0.8689 & -0.7453 & -0.7190 & -0.6020 & -0.6011 & -0.6011 & -0.9482 & 1.0000 & 0.1143 \\ 0.0557 & -0.0674 & -0.0225 & -0.0276 & -0.3048 & -0.3041 & -0.3041 & 0.2005 & 0.1143 & 1.0000 \end{bmatrix}$$

At the same time, the corresponding saliency matrix is obtained as follows:

$$R_2 = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0345 \\ 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0105 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.3925 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.2948 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0345 & 0.0105 & 0.3925 & 0.2948 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

According to the correlation matrix, we can obtain the correlation coefficients between the one-minute average RVR value and the one-minute average MOR value and various influencing factors, as shown in Table 2 below:

Table. 2 RVR correlation coefficient on December 15, 2019

Symbol	regression variable	correlation coefficient	Symbol	regression variable	correlation coefficient

ISSN 2706-655X Vol.3, Issue 1: 1-7, DOI: 10.25236/IJFET.2021.030101

X_1	Average wind speed in two minutes	0.5299	X_5	Local pressure	0.6491
X_2	Vertical wind speed in two minutes	0.4903	X_6	temperature	0.6214
X_3	Corrected sea level pressure	0.6490	X_7	Relative humidity	-0.6410
X_4	Air pressure at the highest point in the landing	0.6491	X_8	Dew point temperature	0.0557

It can be seen from Table 2 that, except for the dew point temperature, the other factors and the one-minute average RVR are all between 0.4 and 0.7, which are significant correlations. It can be seen from the first row of the significance test matrix that, except for the last element, the other elements are all $0.000 < \alpha$. According to statistical principles, the linear correlation between visibility and other indicators is significant at the significance level α = 0.01.

3.2 RVR visibility model on December 15, 2019

According to the above correlation matrix, we can know that some of the variables are highly correlated. In order to avoid the impact, we select factors from them for analysis based on the correlation. According to the 4 regression variables selected above, the calculated regression equation is as follows:

$$y = -193895000 + 32.5463x_1^2 - 186.36x_2^2 + 0.768x_4^2 + 380189x_2 - 12050.6x_4 - 13.0962x_1x_3 + 0.212x_2x_4$$

From the above formula, the relevant statistics are: coefficient of determination is 0.911694, statistics F=2112.06, significance probability p=0, model variance is 300.553, indicating that the regression equation is significant, especially the coefficient of determination has been closed to 1, indicating that the regression fitting is very close to reality. The residual diagram of stepwise regression is shown in Figure 7 below.

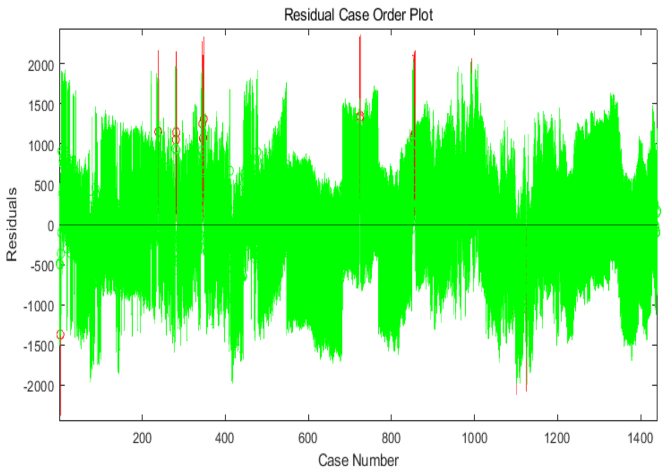

Figure. 7 The residual analysis of the MOR regression equation

4. Conclusion

Grasping the temporal and spatial characteristics of visibility at this stage, mining the influencing factors of visibility changes, and accurately predicting them is of great significance for ensuring traffic safety and promoting social and economic development. This paper fits the relationship between

ISSN 2706-655X Vol.3, Issue 1: 1-7, DOI: 10.25236/IJFET.2021.030101

visibility and seven meteorological factors, and obtains the expression of the relationship between visibility and each meteorological element, and this expression can better predict the trend of visibility, and can provide a solid theory for the prediction of airport visibility in accordance with.

References

[1] Gao Song. Application of Deep Learning in Airport Visibility Prediction. Computer Products and Circulation, 2020(04):260.（in Chinese）

[2] Wang Nan, Zhu Lei, Zhou Jianjun, Liu Qian, Yang Hongru. Interpretation and application of EC fine grid products in the low visibility prediction of Urumqi Airport. Desert and Oasis Meteorology, 2020, 14(02): 81-89 .（in Chinese）

[3] Zhu Guoliang. Airport visibility prediction model based on MLP neural network. Science and Technology Innovation and Application, 2018(18):1-4.（in Chinese）

[4] Natalia Fedorova, Vladimir Levit, Aliton Oliveira da Silva, Deydila Michele Bonfim dos Santos. Low Visibility Formation and Forecasting on the Northern Coast of Brazil. Pure and Applied Geophysics,2013,170(4).

[5] Lei Zhu,Guodong Zhu,Lei Han,Nan Wang. The Application of Deep Learning in Airport Visibility Forecast. Atmospheric and Climate Sciences,2017,7(3).

[6] Hongru Yang, Nan Wang, Yulin Qin, Dawei Fan. Interpretation of EC Fine Grid Numerical Prediction Products in the Forecast of Frozen Fog in Urumqi Airport. Journal of Geoscience and Environment Protection,2018,6(9).

[7] Wang Nan, Zhu Lei, Zhou Jianjun, Liu Qian, Yang Hongru. Interpretation and application of EC fine grid products in the low visibility prediction of Urumqi Airport. Desert and Oasis Meteorology, 2020, 14(02): 81-89 .（in Chinese）

[8] S. Goswami,S. Chaudhuri,D. Das,I. Sarkar,D. Basu. Adaptive neuro-fuzzy inference system to estimate the predictability of visibility during fog over Delhi, India. Meteorological Applications,2020, 27(2).

[9] Feng Hanzhong, Chen Yongyi, Cheng Yongqin, Luo Kesheng. Research on low-visibility weather forecasting methods at Shuangliu Airport. Journal of Applied Meteorology, 2006(01): 94-99.（in Chinese）

[10] J. Güldner. A model-based approach to adjust microwave observations for operational applications: results of a campaign at Munich Airport in winter 2011/2012. Atmospheric Measurement Techniques,2013,6(10).