Study on Probability Model of Freezing Damage of Walnut’s florescence in Central and Southern Hebei Province

SHI Hairui 1, SUN Donglei 1, XU Shuang 1, HAO Jufei 1, YANG Lina 1

1 Xingtai Meteorological Bureau, Hebei Xingtai 054000 Wisdom meteorological innovation team

Abstract: Based on the surface meteorological data of 11 national meteorological stations in the main walnut producing areas of central and southern Hebei province from 1960 to 2017, three distribution functions, including normal distribution, Gumbel distribution and Weibull distribution functions are used to fit the minimum temperature of walnut flowering period, and a probability model of low temperature freezing damage is proposed. The results show that the commonly used probability model is not generally applicable to the freezing damage of walnut’s florescence; different sites of freezing damage of walnut blossom obeyed different probabilistic distribution, Weibull distribution model is more fitting to Gaoc heng and Shexian than that of normal distribution fitting model, and the normal distribution model is much suitable for 9 sites such as Lianyuan than Weibull distribution model; comparing the stations in mountainous and plain areas, the mean value varies with the latitudes. The mean value of stations in the eastern foot of Taihang Mountains decreases with the increase of latitude. However, the mean value of stations in the Inner Hill is lower than that of other stations at the same altitude and latitude due to the influence of geographical topography and microclimate. Therefore, the probability of freezing damage is affected by both mean and standard deviation. The extreme weather is more likely to occur in Laiyuan, Inner Hill, Shexian and Gaoc heng. The probability of low temperature occurrence is generally consistent with the altitude and latitude and it is also greatly affected by the geographical topography and microclimate.

1. Introduction

Walnut is one of the world famous "four dried fruit", which has the rich nutritional value and medicinal value. Walnut has been widely used in our plant, at present there are 7 state walnut demonstration base in Hebei province, to promote local farmers' income, and increase the effective supply of ecological products, which is also of great significance to optimize the dietary structure. The market demand of walnut is very big, it is a very important economic forest tree species in hebei province. The florescence of walnut is between the middle and late March to early and middle April every year in Hebei province, when the weather gets warmer and the temperature rises. The walnut trees is very sensitive to the ambient temperature, spring is the transition season in the north of China air circulation from the meridional circulation to the zonal circulation, the temperature changes a lot, there often comes intermittent cold air from the northwest, and standoffs with the southern moist air, which forms continuous cold weather, leading to the freezing damage of walnuts in the bud, leaf, and anthesis stage, which affects the growth and development of walnut, resulting in a decrease of walnut production and quality, thus bring economic loss to farmers, and even affect the development of walnut industry.

There are already some researches on freezing damage of walnut and the characteristic of weather change, however, there is no study on the probability distribution models of central and southern Hebei province. In this paper, by using the probability distribution model of the lowest temperature and
analyzing the measured data the probability function model of walnut freezing damage in flowering period is proposed, which can be a useful tool for the calculation and analysis of the probability of walnut freezing damage, and provide a reference for the weather insurance rates and disaster prevention investment direction, moreover, it can be a tool to understand incidence of the south "pollen" walnut region characteristics, finally to establish a low temperature forecast and provide reliable theory basis for risk assessment of the scientific planning.

2. Data
This study selects the walnut flowering phenology observation data from the plant phenology by Guo Yanbo et al., shown in table 1, to analyze the meteorological conditions of the flowering time over the years. Considering the walnut flowering may appear freezing damage with the index of -2°C, 0°C, and 2°C, the ground meteorological data is based on the main walnut producing areas in the middle and southern Hebei province between 1960 and 2017, recorded in the country weather stations, Excluding missing data, a total of 11 sites are used for analysis. The data are strictly screened, for example, Shahe site is put into use from of 1963, so at this site, the ground meteorological data at this weather station, the daily minimum temperature data is extracted from March 20 to April 20 to form the research data sequence.

| Puberty         | Bud come unsown period | At the beginning of flowering | Full bloom | The final flowering |
|-----------------|------------------------|-------------------------------|------------|-------------------|
| The average date | 3/20                   | 3/28                          | 4/13       | 4/14              | 4/21              |
| The earliest date| 3/15                   | 3/23                          | 4/5        | 4/7               | 4/20              |
| The latest date  | 3/28                   | 4/2                           | 4/21       | 4/19              | 4/23              |

3. Methods
3.1. The probability models and parameter estimation
Normal distribution model is commonly used in temperature distribution analysis, considering the complex climate\(^1\) environment in the middle and southern Hebei Province, especially in spring, temperature changes with air spring activity is frequent, large variations in temperature causing extreme temperature values often follow some skewness distribution. Combined with the observation of the data histogram, three distribution functions are used to fit the data, including normal distribution, Gumbel distribution and Weibull distribution.

3.2. Normal distribution
Normal distribution is commonly used in meteorological\(^2\)\(^-\)\(^3\) statistics. The probability of different temperature values at a certain place appear generally is consistent with normal distribution, where the parameters of the normal distribution for the mean (μ) means the mean value of the daily lowest temperature in walnut flowering period, and standard deviation (σ) means the change of temperature respectively, the distribution function is:

\[
F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} \exp \left( -\frac{(t - \mu)^2}{2\sigma^2} \right) dt, \quad (-\infty < x < \infty) \ldots \ldots \quad (1)
\]

The probability density function is:

\[
f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right), \quad (-\infty < x < \infty) \ldots \ldots \quad (2)
\]

3.2.1. Gumbel distribution
Gumbel model of extreme values distribution, is the type of minimal extremum I with the parameters\(^4\) of the position (μ) and scale parameters (σ). The distribution function of Gumbel model is:

\[
F(x) = 1 - \exp \left( -\exp \left( \frac{x - \mu}{\sigma} \right) \right), \quad \ldots \ldots \quad (3)
\]
The probability density function is:
\[ f(x) = \frac{1}{\sigma} \exp \left( \frac{x - \mu}{\sigma} \right) \exp \left( - \exp \left( \frac{x - \mu}{\sigma} \right) \right), \quad \ldots \tag{4} \]

3.2.2. Weibull distribution

With the increase of the shape parameter, the shape of the distribution curve changes in the Weibull distribution model, which is flexible to adapt to different forms of frequency distribution, it can simulate the effect of both the skewness and approximate normal distribution, but the required data are negative, in order to be able to fit\cite{5,6} the low temperature sequences. In this paper, the parameters are extended, the location parameter \( c \) is added, the other parameters of the Weibull distribution model include the scale parameter \( a \) and the shape parameters \( b \), for its distribution function:

\[ F(x) = \begin{cases} 1 - \exp \left( -\frac{x - c}{a} \right) & x > 0, \quad \ldots \tag{5} \\ 0 & x \leq 0 \end{cases} \]

\[ F(x) = \begin{cases} 1 - \exp \left( -\left(\frac{x - c}{a}\right)^b \right) & x > c, \quad \ldots \tag{6} \\ 0 & x \leq c \end{cases} \]

The probability density function is:

\[ f(x) = \begin{cases} \frac{b}{a} \left(\frac{x - c}{a}\right)^{b-1} \exp \left( -\left(\frac{x - c}{a}\right)^b \right) & x > c, \quad \ldots \tag{7} \\ 0 & x \leq c \end{cases} \]

3.3. Parameters estimation and the fitting effect

Based on the maximum likelihood estimation function to estimate the parameters, when the sample size is large, the maximum likelihood method has high precision, the fitting parameters and reappearing period values are close to true values.

Related studies have shown that, under the conditions of large sample size, Anderson – Darling test should be used. In this paper, the Anderson - Darling test is made to test the goodness of fitting, as a result, the result is relatively accurate with high credibility. Assume \( F_n \) as the empirical distribution function, \( F_0 \) as the distribution function of the samples under the zero assumptions, the test statistics is shown in Eq. (8):

\[ A^2 = n \int_{-\infty}^{\infty} \left[ F_n(x) - F(x) \right]^2 \frac{dF(x)}{F(x)[1 - F(x)]}, \quad \ldots \tag{8} \]

Through mathematical transformation, Eq. (8) can be simplified to:

\[ A^2 = -n - \frac{1}{n} \sum_{i=1}^{n} [(2i - 1) \ln z_i + (2n + 1 - 2i) \ln (1 - z_i)], \quad \ldots \tag{9} \]

Where \( A^2 \) is the test statistics, which means the empirical distribution of measured data with a particular theory of distribution (normal distribution, Weibull distribution, and the Gumbel distribution). The larger \( A^2 \) is, the smaller the probability is with the theory of distribution. As for the specific differences, we can judge from the value of prob, which is the corresponding concomitant probability of \( H_0 \), if prob is less than 0.05, then this set of data does not obey the distribution model.

In the analysis of data fitting, based on the fitting average relative errors (RER) to judge the fitting effect, the average relative error is smaller, which illustrates the distribution function fitting is better, otherwise, the greater the average relative error is, the poorer fitting effect is. Where \( F_i \) is the various probability quantile according to the distribution theory, \( G_i \) is various probability quantile according to empirical distribution calculation, nine quantiles in this paper which are 10\%, 20\%, 30\%, 40\%, 50\%, 60\%, 70\%, 80\%, 90\%.

\[ \text{RER} = \frac{1}{N_P} \sum_{i=1}^{N_P} |F_i - G_i|, \quad \ldots \tag{10} \]
4. Results analysis
According to the above methods, fitting and testing are carried out based on the low temperature data sequence of 11 sites in the walnut region of Middle and Southern Hebei province. The fitting results of these three probability models are shown in Figure 1 and table 2, for the Gumbel distribution curve, the data is more on the left than on the right, days actually below the average temperature is more than the Gumbel distribution model, the fitting curve shows obvious negative biased, the prob value is less than 0.05, which rejects the hypothesis of the Gumbel distribution.

![Fig. 1 Part of the site fitting results](image)

| Probability model | Variable | LY | FP | JX | PS | GC | ZH | LC | NQ | SH | SX | CX |
|-------------------|----------|----|----|----|----|----|----|----|----|----|----|----|
| Normal distribution | value     | 0.408 | 1.162 | 1.327 | 1.459 | 2.004 | 1.233 | 0.911 | 1.049 | 0.798 | 1.828 | 2.042 |
|                    | prob      | 0.841 | 0.282 | 0.224 | 0.186 | 0.091 | 0.255 | 0.407 | 0.332 | 0.482 | 0.114 | 0.087 |
| Gumbel distribution | value     | 21.09 | 15.729 | 19.676 | 19.442 | 22.248 | 20.877 | 18.851 | 20.054 | 16.029 | 16.928 | 18.591 |
|                    | prob      | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| The Weibull distribution | value     | 2.731 | 1.349 | 1.418 | 3.839 | 1.29 | 1.432 | 1.368 | 1.139 | 1.194 | 1.628 | 2.526 |
|                    | prob      | 0.038 | 0.217 | 0.197 | 0.01 | 0.236 | 0.194 | 0.211 | 0.292 | 0.27 | 0.149 | 0.048 |
The value of prob of the Weibull distribution test in Laiyuan, Pingshan, Cixian Three sites is less than 0.05, rejected the Weibull distribution hypothesis. The values of prob of Weibull distribution test in 8 sites including Fuping were greater than 0.05, which obey the Weibull distribution.

All the data in 11 sites are normally distributed, based on the average relative error rate (RER) of Weibull distribution and normal distribution shown in table 3, the average relative error rate of normal distribution is between 1.58% and 3.43%, the average relative error of Weibull distribution is between 2.06% to 2.70%, which means the fitting effect is good.

The average relative error rate of normal distribution in Gaocheng and Shexian are 3.43% and 2.37% respectively, which is greater than the mean relative error of the Weibull distribution. While in Laiyuan and other nine sites, the normal distribution fitting is better than the Weibull distribution fitting.

| Probability model | LY  | FP  | JX  | PS  | GC  | ZH  | LC  | NQ  | SH  | SX  | CX  |
|-------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Normal distribution | 2.66% | 2.44% | 2.16% | 2.34% | 3.43% | 1.98% | 1.84% | 2.28% | 1.58% | 2.37% | 3.28% |
| The Weibull distribution | 6.61% | 2.70% | 2.27% | 3.11% | 2.20% | 2.06% | 2.16% | 2.62% | 2.22% | 2.25% | 2.96% |

The probability model parameters of each site is presented in table 4. The shape factors of Weibull distribution probability model for Gaocheng and Shexian are 3.09 and 3.1. Compared with the normal distribution, there is biased distribution and the peak is low, the number of days that is less than average low temperature is greater than the number of days above the average low temperature, which means the freezing damage probability in low temperature is larger.

According to the geographical position of the site locations, the mountains of Laiyuan, Fuping, Jingxing, Pingshan, and Shexian are defined as group 1, the taihang mountain foothill Zanhuang, Lincheng, Neiqiu, Shahe are defined into group 2, plain areas such as Gaocheng and Cixian are defined as group 3.

The mean values in group 1 increases with latitudes. Influenced by altitudes, the mean values in Laiyuan and Shexian, are less than others sites, which is consistent with the principle of the temperature change with latitude and altitude. Group 2 is affected by the surrounding environment, such as mountain blocks, the mean temperature increased with the latitudes. Neiqiu is below other sites, which is greatly influenced by the geographical topography and microclimate. Group 3 are greatly influenced by latitudes, the mean temperature increases with latitudes.

From the standard deviation of all stations, it can be seen that the temperature changes, greater values meant the temperature fluctuation is larger. In Laiyuan, Fuping, Jingxing, Shexian, Zanhuang and Neiqiu, the temperature fluctuation is bigger, which influences both the mean value and the standard deviation. If the mean value is smaller and standard deviation value is greater, it is more possible to have extreme weather in this site, which will bring losses to the production of walnut.

| Probability Model Parameters | LY  | FP  | JX  | PS  | SX  | ZH  | LC  | NQ  | SH  | GC  | CX  |
|-----------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| The meanμ                  | -1.257 | 5.823 | 6.416 | 6.278 | 5.35 | 6.594 | 6.342 | 5.118 | 5.914 | 5.626 | 5.86 |
| The standard deviation     | 4.746 | 4.195 | 4.2 | 4.063 | 4.215 | 4.194 | 4.111 | 4.352 | 4.165 | 4.068 | 4.108 |
| Scale factor               | 15.248 | 13.894 | 13.85 | 13.236 | 13.61 | 13.862 | 13.895 | 14.444 | 15.131 | 13.145 | 11.576 |
| Shape factor               | 3.052 | 3.195 | 3.166 | 3.022 | 3.1 | 3.164 | 3.247 | 3.186 | 3.524 | 3.09 | 2.649 |
| Coefficient of position    | -15 | -6.7 | -6.1 | -5.7 | -6.9 | -5.9 | -6.2 | -7.9 | -7.8 | -6.2 | -4.5 |

Based on the above analysis, we can determine the probability distribution model of each site. The frost damage is different with the conditions of many factors such as age, variety and so on. Generally, when the exhibition walnut leaf temperature drop to -2℃~4℃, or in sprout, flowering and young fruit
period, when the temperature drop to 1 °C ~ 2 °C, there will be production loss, besides, the days of temperature below 0 °C also reflects the freezing injury risks. We can calculate the probability of the lowest critical temperatures for each site, shown in table 5. There are low probabilities for each site, such as the temperature T is: T≤-2°C, -2°C<T≤0°C, 0°C<T≤2°C, and the probability is relatively close, which is consistent with flowering temperature rise process, but Laiyuan situation is completely on the contrary. The temperature of Shexian is higher than other sites, which may because of the site location. The probability of freezing damage in each site with different levels are between 13-74%, the worst is Laiyuan, Gaocheng, Neiqiu and Shexian, the probability of low temperature freezing injury is more than 20%. Fuping, Shahe and Cixian take the second place, the probability of low temperature freezing injury is about 18%, the rest of the sites are between 13% and 15%. Besides the influence of latitude and altitude, small stations terrain stingy is another big impact of low temperature.

Table 5 Distribution Types and Occurrence Probabilities of Minimum Temperature

| Distribution type | LY    | FP    | JX    | PS    | SX    | ZH    | LC    | NQ    | SH    | GC    | CX    |
|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| T≤-2°C           | 43.78%| 3.11% | 2.26% | 2.08% | 2.90% | 2.02% | 2.12% | 5.09% | 2.87% | 4.13% | 2.79% |
| -2°C<T≤0°C       | 16.66%| 5.15% | 4.08% | 4.03% | 6.44% | 3.77% | 4.03% | 6.88% | 4.91% | 7.34% | 4.90% |
| 0°C<T≤2°C        | 14.93%| 9.85% | 8.32% | 8.50% | 11.42%| 7.87% | 8.40% | 11.71%| 9.59% | 12.05%| 9.68% |

Figure 2 Low temperature probabilities at each site

5. Conclusion and discussion

Combined with walnut flowering phenology period data of the middle and southern Hebei province, data fitting is completed based on Normal distribution, Gumbel distribution and Weibull distribution. The results indicate:

1. The commonly used probability distribution model is not generally applicable to the walnut flowering freezing damage, the selected minimum temperature sequences between March 20 and April 20 all refused the Gumbel distribution, but are according with Normal distribution.

2. The freezing damage of walnut florescence follow different probability distributions. In the site of Gaocheng and Shexian, the Weibull distribution is more suitable than Normal distribution fitting, while in nine sites including Laiyuan the Normal distribution fitting is better than the Weibull distribution.
(3) Both in Mountain sites and plain sites, but the average low temperature increases with the latitude and altitude, the mean value in Taihang mountain site decreases with the increase of latitude. Affected by the geographical topography and microclimate, the average temperature and standard deviation in Neiqiu is lower than other sites of the same latitude and altitude. There are also low mean values but high standard deviation, such as the sites of Laiyuan, Neiqiu, Shexian and Gaocheng, which is relatively more prone to extreme weather.

(4) The probability of low temperature is generally consistent with the effects of altitude and latitude, at the same time, it is also affected by the geographical topography and microclimate. In actual production, both walnut region station data and microclimate station data should be provided for reference to avoid the deviation by elevation changes, which can improve the accuracy of the data for the cooling process more than the critical temperature, the warning level should be adjusted accordingly to the last time.

(5) Under the background of global warming, the minimum temperature of walnut florescence is increasing, on the one hand, because of the temperature rise, walnut phenophase is becoming shorter, it is possible to cause production loss by sudden chill freezing damage; On the other hand, with the temperature rises, the parameters of the probability distribution or type will be changed, the change needs to be further researched on.

(6) By statistical method to establish the probability model of freezing damage in walnut production areas, the low temperature freezing damage can predicted, which can provide scientific reference for walnut flowering cold weather index and provide reliable theory basis for the low temperature forecast by early warning model and make the risk assessment of the scientific planning.

References

[1] FAN Xiaochun, WANG Weitai, YANG Xiaohua, WU Yingjuan. Responses of apple tree’s phonology in east and west sides of Liu panshan Mountain to climate change[J]. Chinese Journal of Ecology, 2010, 29(1): 50-54.

[2] QU Zhenjiang, LIU Ruifang, GUO Zhaoxia, WANG Jinghong, LIU Lu, CHAI Qian. Study of risk assessment and prediction of apple blooming freezing injury in Shaanxi Province[J]. Journal of Natural Disasters, 2013, 22(1): 219-225.

[3] LI Meirong, LI Xingmin, BAI Qinfeng, ZHANG Weimin. Probability distribution and recurrence interval prediction of the extreme value of air temperature in extreme meteorological disasters of apples[J]. Agricultural Research in the Arid Areas, 2012, 30(3): 257-261.

[4] Singh V P. Entropy-based parameter estimation in hydrology[M]. Boston/London: Kluwer Academic Publishers, 1998.

[5] Anderson T W, Darling D A. A Test of Goodness of Fit [J]. Journal of The American Statistical Association, 1954,(49).

[6] Stephens M A. EDF Statistics for Goodness of Fit and Some Comparisons[J]. Journal of The American Statistical Association, 1974,(69).