ISSN 2616-5872 Vol.3, Issue 1: 20-24, DOI: 10.25236/AJEE.2021.030104

Analyze and extract public data on the correct spread of Asian Hornet

Xinze Li, Yunxiao Li, Hongbing Li

College of Information Science and Engineering, Shandong Agricultural University, Taian, Shandong 271000, China

Abstract: *Since September 2019, the Asian Hornet has had a serious impact on the ecological balance of Washington State. First select and integrate the data, sort, filter, and match the data in the table to obtain valuable information. We adopt the strategy of spatial area analysis, based on historical data, and use the quadratic exponential smoothing method to fit the ARIMA parameters, and finally get a robust time series forecasting model. Furthermore, according to the predictive model, the possible quantity in the next year is simulated, which provides a theoretical basis for controlling the spread of this pest.*

Keywords: *ARIMA, Vespa mandarinia propagation, Data visualization*

1. Introduction

In the eyes of the Chinese, bees often symbolize hard work, practicality, self-discipline, dedication and team spirit, but they seem to be unfriendly to recent Americans. Giant Asian hornets have recently made their first landfall in North America and are spreading rapidly across the U.S., where a bite from the killer bees can put them at risk of death. In response, several media and research organizations have warned that the invasion of the Asian giant bumblebee could decimate the native bee population of the United States and pose a deadly threat to humans.

At more than five centimeters long, the giant Asian bumblebee is five times the size of the average bee. Now, the killer bees have spread so rapidly across the United States that they have been found north of the Canadian border. When they fly through a colony of ordinary bees, they kill all the adults and spend the next few hours feasting on the young bees [1]. This would cause a precipitous decline in native bee populations in the United States, severely affecting the ecological balance and possibly even hurting agricultural production.

Although their venom is toxic, wasps "generally do not attack people or pets" unless threatened. Washington State Department of Agriculture entomologist Sven Eric Spiechger says killing bee populations could jeopardize the future of American agriculture [2]. Pollination is an important part of agriculture and agricultural systems. So if this becomes entrenched and spreads, it could be catastrophic. We were asked to predict future transmission of this pest, to classify and prioritize existing sample reports, to develop prevention strategies, and to discuss the status of the Asian bumble bee in Washington State.

2. Model building

2.1 Data

Since 2021MCM_ProblemC_DataSet. XLSX has a large amount of data, it is not intuitive, so we can visualize some data directly for display.

2.2 Data cleaning

Data were grouped by month, and the number of Asian bumblebees in detection data with Lab Status of Positive ID was calculated. For missing monthly values in the data, we try to skip it and just look for valid averages. For the complete miss group from May to October, the values are recorded as the miss group.

ISSN 2616-5872 Vol.3, Issue 1: 20-24, DOI: 10.25236/AJEE.2021.030104

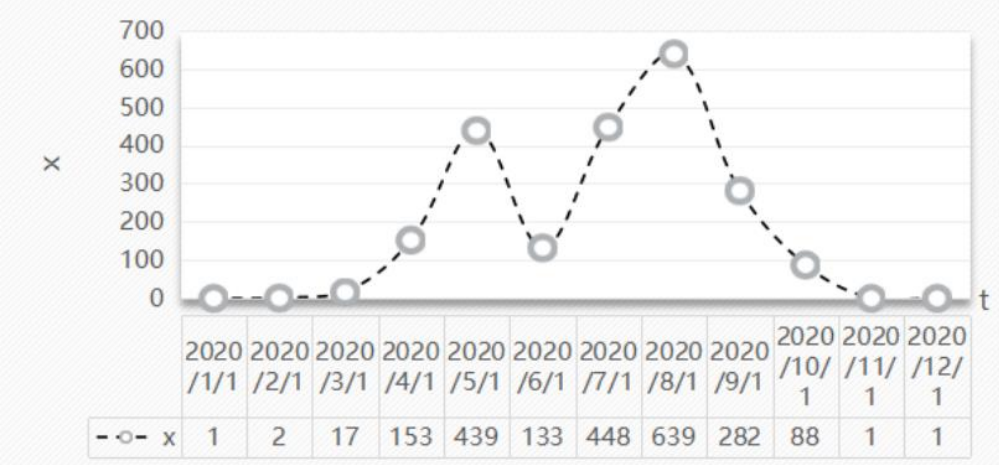

Figure 1: The wasp population varies month by month

Wherein, in the above image, T represents the month, and X represents the number of Vespa Mandarinia in each month. From the figure, we can see the trend of the number of Vespa Mandarinia over time.

2.3 Conversion between longitude and latitude and distance

Spherical coordinates are used to represent points on the data set. In order to obtain the true distance relationship on the map, we can approximately think of the observed map area as a planar quadrilateral. After solving the quadrilateral length by using the geodesic equation (on the GRS80 sphere) [3], the corresponding relationship between spherical coordinates and plane coordinates is obtained by fitting the projection transformation. In this way, the Euclidean distance between the points approximates the geodesic distance on the sphere.

It is mentioned in the question that a new queen nest in the range of 30 kilometers, now has detected 14 confirmed found bumblebees, the latitude and longitude of these 14 places as the center, draw the queen bee range. After calculation, the distance increases by 111.12km for every 1 °increase in latitude, and 72.90km for every 1 °increase in longitude. It can be seen from this that the colony propagation area is an ellipse, and the coverage area is shown in yellow.

Since both unprocessed and unverified can not be confirmed due to the lack of data, we put them together when we collect the data and draw each monitoring point with scattered points. Dots within the yellow region are more likely to have Asian bumblebees.

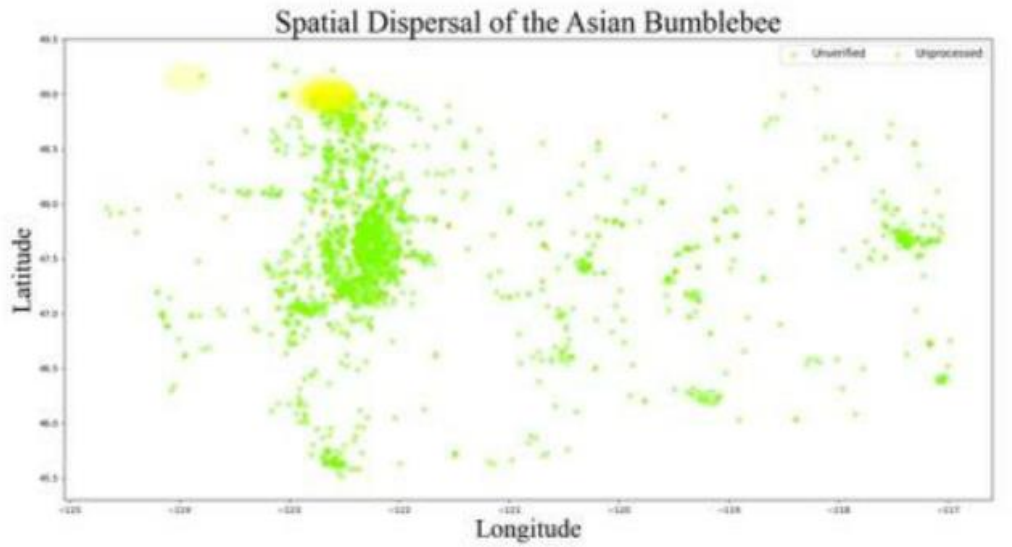

Figure 2: Dispersion map of the Asian bumblebee

ISSN 2616-5872 Vol.3, Issue 1: 20-24, DOI: 10.25236/AJEE.2021.030104

2.4 Establishment and solution of bumblebee migration prediction model

The migration of bumblebees is determined by a number of factors, namely environmental factors, natural predators, food chains and climate factors. Therefore, for such a complex dynamic system, the multi-time series vector autoregressive method is used to solve the problem. Formally, vector autoregressive algorithm can simultaneously consider the spatiotemporal correlation of each variable, and mining data information to the maximum extent without introducing exogenous factors. Therefore, the prediction based on the autoregressive integrated moving average model (ARIMA) is a good approximation.

2.5 Temperature prediction

According to the temperature data of different years, the ARIMA(p, d, 0) model is used to model the temperature series data. For the ith temperature series, the general situation of the model is as follows:

$$\Delta^{(\mathrm{d})}u_{i,t} = \sum_{j=1}^{p} a_{i,j}\Delta^{(d)}u_{i,t-1} + \varepsilon_{i,t} \tag{1}$$

Where, Δ(d) represents the order difference $\varepsilon \sim N(0,\varepsilon_{i,t}^2)$ operator. Is the residual value of the model. Therefore, the first-order difference prediction value of the ith sequence after the next t can be obtained:

$$\Delta^{(\mathrm{d})}u_{i,t} = \sum_{j=1}^{p} a_{i,j}\Delta^{(d)}u_{i,t-1} + \varepsilon_{i,t} \tag{2}$$

In addition, according to the assumptions and practical experience of the ARIMR model, the predicted values can meet the normal distribution, that is, $u_{i,t} \mid u_{i,t-1} \sim N^{-}(\mu_{i,t},\delta_i^2)$, so different predictions can be obtained. These temperature-related results can lead to changes in colony migration. As a result, colonies can spread out in different areas over several years, which can have a big impact on the American population.

3. Quantity forecast

The life cycle of wasps is similar to that of many other wasps. The fertilized queen emerges in the spring to start a new colony. In the fall, the new queen leaves the nest and spends the winter in the soil, waiting for spring to arrive. When we understand the changes in wasp populations, we can better understand their migration.

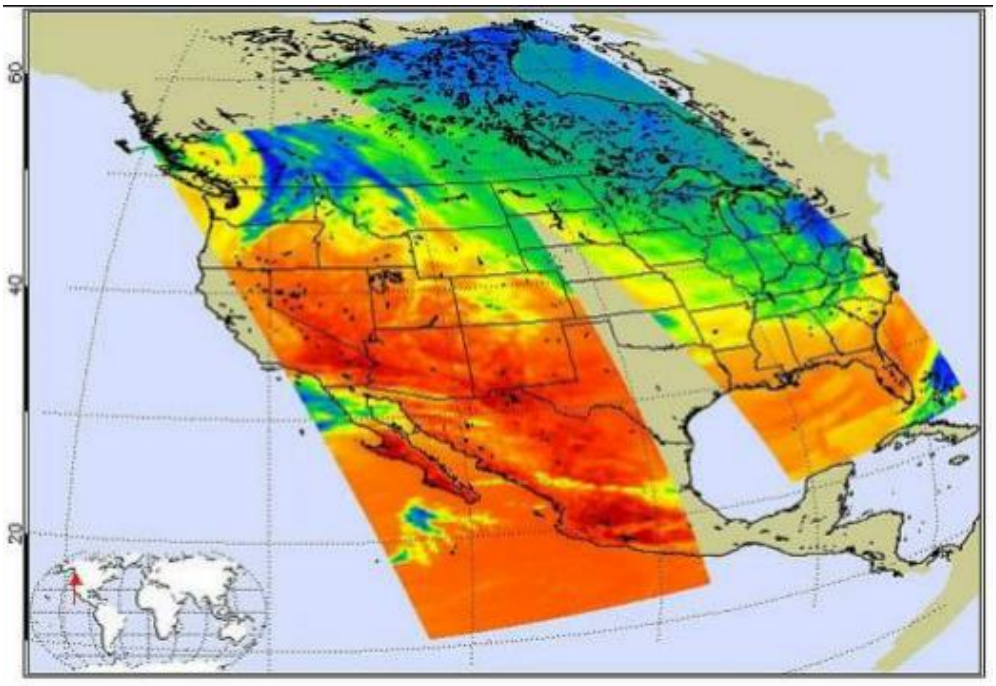

Figure 3: Washington state temperature map

ISSN 2616-5872 Vol.3, Issue 1: 20-24, DOI: 10.25236/AJEE.2021.030104

We use the exponential smoothing method to overcome the shortcomings of the moving average method. However, when the change of time series shows a linear trend, there is still obvious lag deviation when using the first exponential smoothing method to forecast. This, too, must be amended. The correction method is the same as the trend moving average method, that is, the quadratic exponential smoothing is done again, and the linear trend model is established by using the law of lag deviation. That's the quadratic exponential smoothing. Its calculation formula is as follows:

$$S_t^{(1)} = ay_t + (1-a)S_{t-1}^{(1)} \tag{3}$$

$$S_t^{(2)} = aS_t^{(1)} + (1-a)S_{t-1}^{(2)} \tag{4}$$

4. Colony migration

Expression of migration speed of honeybee:

$$(dx, dy) = \varphi dt \tag{5}$$

The update formula for position (x,y) is:

$$(x_t, y_t) = (x_{t-1}, y_{t-1}) + (dx, dy) \tag{6}$$

$$\nabla v(x, y, t) = \left(\frac{\partial v(x, y, t)}{\partial x}, \frac{\partial v(x, y, t)}{\partial y}\right) \tag{7}$$

5. Result analysis

According to the calculation results of Model, the average monthly distribution number of Asian bumblebees after 1 year can be predicted as follows:

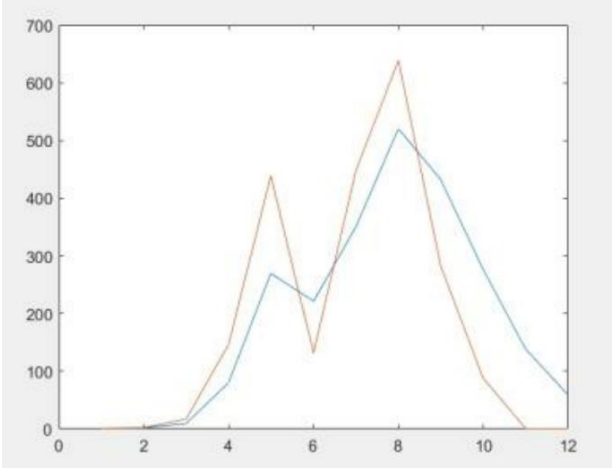

Figure 4: The number of Asian bumblebees in one year and the predicted average value of the next year

6. Conclusion

Among them, the red line represents the number of Asian bumblebees in 20201, and the blue line represents the fitting value of the average forecast of future opinions. It can be seen that the overall trend of the two lines is consistent, which proves that the quadratic exponential smoothing method is suitable for bumblebee population prediction.

Advantages of ARIMA model: The model is very simple and only requires endogenous variables without the help of other exogenous variables. Data with short forecast time period has higher accuracy. Wide range of application. Non-stationary time series can be transformed into stationary time series.

ISSN 2616-5872 Vol.3, Issue 1: 20-24, DOI: 10.25236/AJEE.2021.030104

References

[1] An Yonggui, Han Lei, Huang Xiaoyu, Gao Yang. The first launch of the prediction network of saturated vapor pressure difference in Northwest China based on exponential smoothing and ARIMA model [J]. Arid Zone Research, 2021-02-05: 31760236 34460220.

[2] K. Jarrett, K. Kavukcuoglu, M. Ranzato, and Y. LeCun, "What is the best multistage architecture for object recognition? BT - Computer Vision, 2009 IEEE 12th International Conference on," Comput. Vision, 2009 ..., pp. 2146–2153, 2009.

[3] Jiang Qiyuan, Xie Jinxing. —A successful practice of higher education reform — An exploration and practice of mathematical modeling teaching and competition activities [J]. China Institute of Higher Education, 2011-12-20.:1004-3667.