ISSN 2616-5775 Vol. 2, Issue 1: 120-126, DOI: 10.25236/AJCIS.010025

The Prediction and Analysis of Opioid Crisis

Zhou Xiaochen

North China Electric Power University, Beijing, China

ABSTRACT. *Using a linear fitting method to identify locations where specific opioid uses might has started in five states. Next, using the logistic model to find out the growing rules of reported opioids incidents, and getting the threshold is 268 380. Finally, this paper uses a grey model to predict the future.*

KEYWORDS: *Logistic model, Origin inference model, Linear fitting method, Grey model*

1. Background

After years of abuse either for the treatment of pain or for recreational purposes, the United States is in the midst of a national crisis related to the use of synthetic and non-synthetic opioids. What's more, this crisis is very fearful since it may cause a serious adverse effect on the U.S economy. Therefore, the government should analyze the data-intensive annual reports published by the DEA to control this situation.

2. Origin inference model

2.1 Model preparation

First, this paper processes the table data to get the number of annual reported synthetic opioid incidents in each county and their averages, and finds out which counties have greater changes. Then, this paper divides these counties into inflow and outflow counties. As figure1 shows, green represents the gathering area of outflow counties, and red represents the inflow counties.

ISSN 2616-5775 Vol. 2, Issue 1: 120-126, DOI: 10.25236/AJCIS.010025

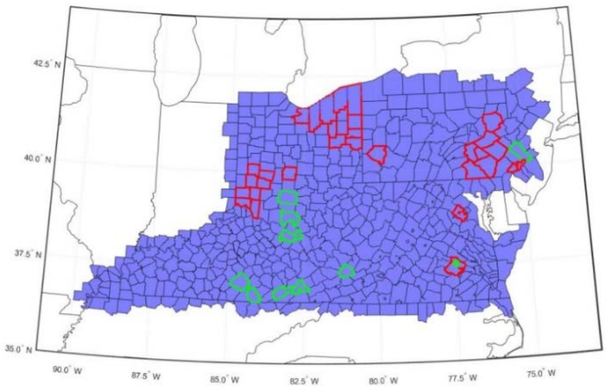

Figure 1: Gathering area

This paper uses the straight line model to infer the number of annual reports for each county in 2005 and identifies the county with the largest number of annual reports as the origin.

2.2 The model construction

In this model, it fits the total number of annual discrete cases in each county, making it best approximate the fitted model. And the criteria are as follows:

$$y = ax + b \tag{1}$$

y is the number of incidents(cases), x represents the year.

In order to achieve the best fitting effect between the model and the number, this paper has done the following work: Minimizing the maximum deviation of all points, which is $\mathrm{Max}|y_i - f(x_i)|$, i =2005, 2006, …, 2020. Minimizing the sum of these absolute deviations, which is $\sum_{i=1}^{m} |y_i - f(x_i)|$,and solving the estimation of model parameters a, b.

2.3 Model solution and analysis

By solving the model and predicting short-term past values, people can find that the number of reports in some counties in 2005 is close to zero, and that in some counties there are many reports in 2005. After analysis, it can be believed that:

1) In some counties, in or after 2005, the number of incidents is zero or very small compared with other counties. Later, it increased. It can be believed that these counties were disseminated.

ISSN 2616-5775 Vol. 2, Issue 1: 120-126, DOI: 10.25236/AJCIS.010025

2) In some counties, in or after 2005, the number of incidents is much higher than that of other counties. Later, it began to decrease or was still rising. So these counties were outward transmissions.

Therefore, the conclusion is: the origin of southern OH is Mahoning, the origin of northern OH is Scioto, the origin of southern PA is Bucks and Alleghenyall.

3. Logistic model

3.1 Model preparation

This paper takes five states as the research object and takes the total amount of drugs as the research index to make hypotheses for five states. After the data was processed, here is the total drug quantity from 2010 to 2017 as follows:

Table 1: Total drug quantity

Years	2010	2011	2012	2013	2014	2015	2016	2017
Total amount	240698	224639	233174	249419	245009	243476	253126	257636

Based on these data, this paper establishes a logistic model for drug growth.

3.2 The model construction

Introducing N_m in the logistics model to indicate the maximum number of drugs that can be tolerated by realistic conditions. Then this paper assumes that the growth rate is equal to $r(1-\frac{N_{(x)}}{N_m})$, that is, the net growth rate decreases as N_x increases. When $N_x \to N_m$, the net growth rate tends to zero.

Establishing a drug quantity prediction model based on this assumption, assuming that no major drug management measures will occur during the forecast period. Here is the following logistic model:

$$\begin{cases} \frac{dN(t)}{dt} = r\left(1-\frac{N}{N_m}\right)N \\ N(t_0) = N_0 \end{cases} \tag{2}$$

x represents the year, x_0 is the initial year, N_0 is the initial number of drugs in the initial year, and N(x) is the number of drugs in a specific year. By solving the logistic model equations above, this paper can obtain the following formula:

$$N(t) = \frac{N_m}{1+(\frac{N_m}{N_0}-1)e^{-r(t-t_0)}} \tag{3}$$

ISSN 2616-5775 Vol. 2, Issue 1: 120-126, DOI: 10.25236/AJCIS.010025

3.3 Model solution and analysis

This paper considers the data in 2012 as N_0 in the equation, and the data in 2014 and 2016 as two known $N_{(x)}$ and gets the corresponding values of r and N_m of total drug quantity in five states by MATLAB. Among this, $N_m = 268380$

Now this paper gets the logistic prediction equations of five states. Substituting r and N_m into the equation can get the equation:

$$N(t) = \frac{268380}{1+(\frac{268380}{233174}-1)e^{-r(t-2012)}} \tag{4}$$

Here is an image showing the logistic model clearly.

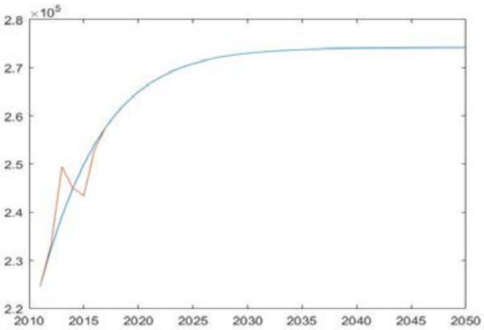

Figure2 logistic model

Through the analysis of the formula, if the threshold is set to 266 000, it can be found that the number of incidents occurred is very close to the threshold of 274260 around 2026.

4. Location prediction grey model

4.1 Model preparation

Through the analysis, it shows that most of the county's SA is small or the same, which shows that the spread of synthetic opioids and heroin is almost nonexistent in these counties.

Then this paper takes the outflow county as the starting point and the inflow county as the endpoint to get the propagation path. This paper sets the coordinate of the outflow counties on the map like $x(i), i = 1,2,3 \dots \dots n$

n is the number of the outflow county. *i* is the first outflow county. Then this paper establishes the predicted model with the help of the aggregation site of

ISSN 2616-5775 Vol. 2, Issue 1: 120-126, DOI: 10.25236/AJCIS.010025

outflow counties' coordinates.

4.2 The model construction

Step 1: $x^{(0)}$ is the original sequence. Accumulating it and getting the sequence $x^{(1)}$

Original sequence $x^{(0)}=\left(x^{(0)}(1), x^{(0)}(2), \cdots x^{(0)}(n)\right)$: ,

Accumulative sequence: $x^{(1)}(m)=\sum_{i=1}^{m} x^{(0)}(i)$, $m=1,2,\cdots n$

Step 2: calculate the mean sequence of $x^{(1)}$ ($z^{(1)}=\left(z^{(1)}(2), z^{(1)}(3), \cdots, z^{(1)}(n)\right)$,)

$z^{(1)}(m)=0.5x^{(1)}(m)+0.5x^{(1)}(m-1)$, $m=2,3,\cdots n$

$x^{0}(k)+az^{(1)}(k)=b$ is the basic form of the GM(1, 1) model. $-a$ is the development coefficient, b is gray action. b is exhumed from the background value data, which reflects the data changes.

Step 3: Determine the model expression and the time response type

If we consider the moment $m=2,3,\cdots n$ of $x^{(0)}(m)$ as a continuous variable

Let $x^{(0)}(k)=\frac{d\hat{x}^{(1)}(l)}{dl}$ $z^{(1)}(k)=x^{(1)}(l)$ So $z\frac{d\hat{x}^{(1)}(l)}{dl}+\hat{x}^{(1)}(l)=b$

then the solution of $x^{0}(k)+az^{(1)}(k)=b$ is:

$$\hat{x}^{(1)}(m+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-am}+\frac{b}{a}\ ,m=1,2,\cdots,n-1$$

From above we can know the estimated value of $x^{(0)}$ is

$$\hat{x}^{(0)}(m+1)=\hat{x}^{(1)}(m+1)-\hat{x}^{(1)}(m)=(1-e^{a})\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-am}\ ,m=1,2,\cdots,n-1$$

4.3 Model solution and analysis

According to the figure of gathering area (Figure 1), using the coordinates of the outflow counties gathering area as the data, combined with the gray location

ISSN 2616-5775 Vol. 2, Issue 1: 120-126, DOI: 10.25236/AJCIS.010025

prediction model to get the county may have a surge when the number of cases reaches the threshold after 14 years, as shown in the following figure 3.

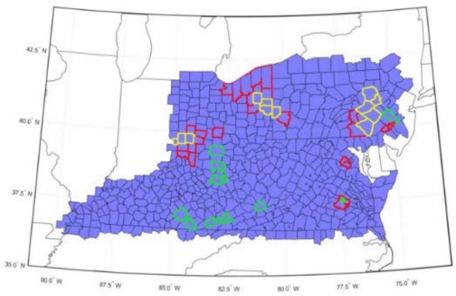

Figure 3 incidents of rapid growth county

It can be seen that as time goes by, the more cases occur, and more counties may have a surge in the use of drugs, and the overall trend is from west to east, from south to north.

4.4 Error Analysis

In the above solution, this paper can also get the predicted value by looking up the statistical method with the corresponding data. For error analysis, this paper uses it to represent the percent margin of error. The formula is as follows:

$$\bar{t} = \sum \frac{|\beta - \alpha|}{\alpha} * 100\%$$

β is the predicted value, α is the true value. According to the formula, here is the percent margin of error of actual incidents and forecasted incidents in five states. As shown in the table below:

Table 2: The percent margin of error of two kinds of incidents

	actual incidents	forecasted incidents	the percent margin of error
2010	240698	216613.3745	10.01%
2011	224639	225537.9533	0.40%
2012	233174	233174	0.00%
2013	249419	239622.5582	3.93%
2014	245009	245008.3477	0.00%
2015	243476	249465.0891	2.46%
2016	253126	253124.8894	0.00%
2017	257636	256111.3892	0.59%

ISSN 2616-5775 Vol. 2, Issue 1: 120-126, DOI: 10.25236/AJCIS.010025

By analyzing the tables, excluding the data of 2012, 2014 and 2016, and getting the average percent margin of error is 3.48%, which is acceptable.

References

[1] WANG Yong. Solution of Logistic population model[J].Journal of Harbin University of Commerce(Natural Science Edition),2006(05):58-59.
[2] KONG Xue, WANG Li, FENG Yi-hua.Review and Prospects of the Application Status of GM (1, 1) [J].Qilu University of Technology (Shandong Academy of Sciences), 2018, 32(06):49-53.
[3] Troy Quast, Melissa A. Bright,Chris Delcher. The relationship between foster care entries and high-dose opioid prescribing in California [J]. Addictive Behaviors, 2019.
[4] Diao Weikang, He Jiafen, Chen Jianzhi. The Application of Linear Regression Model in Web Development Software [J].Gansu Science and Technology, 2018, 34(18):17-18.
[5] U.S.Nationalmap. https://viewer.nationalmap.gov/basic/#productSearch.