|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | For office use only | | | T1 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | T2 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | T3 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | T4 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | |  | | --- | | Team Control Number  **1911245** | |  | | Problem Chosen  **C** | | |  |  | | --- | --- | | For office use only | | | F1 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | F2 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | F3 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | F4 | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | |

**2018 Mathematical Contest in Modeling (MCM) Summary Sheet**

**Title**（此处应写论文标题）

**Summary**

美赛论文的摘要的英文一般用Summary，摘要最好在本页完成。

**标题**为16号Times New Roman字体加粗并居中

**摘要**为14号Times New Roman字体加粗并居中

**关键词**为12号Times New Roman字体加粗并居左

**行间距一般为1倍行距，为控制在一页可适当调整。**

首页不需要页眉和页码。

“00000”修改为自己的控制编号（Team Control Number），“A”改为自己的选题题号（A/B/C/D/E/F）

**Key words: Differential equation model;** **R-type cluster analysis;** **Related analysis;** **Grey-RBF neural network model**

**Contents**

[1. Problem statement 4](#_Toc471658476)

[2. Problem analysis 5](#_Toc471658479)

[3. Basic assumption 6](#_Toc471658486)

[4. Glossary & Symbols 6](#_Toc471658487)

[5. Models 7](#_Toc471658490)

[5.1 Analysis and Solving of Question One 7](#_Toc471658491)

[5.1.1 Model Preparation 7](#_Toc471658492)

[5.1.2 Model Establishment 8](#_Toc471658493)

[5.1.3 Model solving 8](#_Toc471658494)

[5.1.4 Analysis of the Result 9](#_Toc471658495)

[5.2 Analysis and Solving of Question Two 9](#_Toc471658496)

[5.2.1 Model Preparation 9](#_Toc471658497)

[5.2.2 Model Establishment and Solving 10](#_Toc471658498)

[5.2.3 Analysis of the Result 11](#_Toc471658500)

[5.3 Analysis and Solving of Question Three 11](#_Toc471658501)

[5.3.1 Model Preparation 11](#_Toc471658502)

[5.3.2 Model Establishment 12](#_Toc471658503)

[5.3.3 Model solving 13](#_Toc471658504)

[5.3.4 Analysis of the Result 13](#_Toc471658505)

[6. Evaluation and Promotion of Model 16](#_Toc471658520)

[6.1 Strength and Weakness 16](#_Toc471658521)

[6.1.1 Strength 16](#_Toc471658522)

[6.1.2 Weakness: 17](#_Toc471658523)

[6.2 Promotion 17](#_Toc471658524)

[7. Conclusions 17](#_Toc471658525)

[7.1 Conclusions of the problem 17](#_Toc471658526)

[8. References 18](#_Toc471658528)

# 1. Problem statement

National crises about the use of synthetic and non-synthetic opioids are coming to the United States, so federal organizations such as the US Centers for Disease Control (CDC) are working to save lives and prevent the negative impact of the epidemic on health and the national economy. As part of the DEA Transfer Control Office, the DEA/National Forensic Laboratory Information System (NFLIS) has published a data-intensive annual report. "The database within NFLIS includes data from crime laboratories that process approximately 88% of state and local drug cases in the country each year.

This issue focuses on individual counties located in five US states: Ohio, Kentucky, West Virginia, Virginia, and Pennsylvania. Several data sets of drug identification count data sets for anesthesia analgesics (synthetic opioids) and heroin from five states were presented. Drug identification occurs when a law enforcement agency submits evidence to a crime laboratory as part of a criminal investigation and the forensic scientist in the laboratory tests the evidence. Typically, when law enforcement agencies submit these samples, they provide location data (counties) and their incident reports. When the evidence is submitted to the crime laboratory and the location data is not available, the crime laboratory uses the location of the city/county/state investigation law enforcement organization that submitted the case.

* Question 1

Using the provided NFLIS data, a mathematical model was created to describe the spread and characteristics of synthetic opioids and heroin events (cases) between the five states and their counties. Also identify any possible locations where specific opioids may have begun to be used in five states. The problems that the US government will have when envisioning patterns and features that continue to exist. And predict that when these conditions occur at these drug recognition threshold levels, the model will need to be able to predict when and where they will occur in the future?

* Question 2

Use the provided socio-economic data creation model to address the following questions: Explain the development of opioids, what contributed to the use of opioids and the growth of addiction, and why the use of opioids still exists. Is the use or use trend related to any US Census socioeconomic data provided? If so, please modify the mathematical model in question one.

* Question 3

Combine the results of Questions 1 and 2 to identify possible strategies for combating the opioid crisis. Use this model to test the effectiveness of the strategy; identify any important parameter boundaries on which success (or failure) depends.

# 2. Problem analysis

## 2.1 Analysis of Specific Issues

The topic asks for analysis of opioid drug case data in five US states, models the drug-distribution process, and combines socio-economic factors in the state, corrects the model, analyzes the characteristics of the opioid-drug-distributing process, and provides effective policy providers. Strategy to deal with "Opioid Crisis". In response to the problem, the states are taken as nodes and combined with socio-economic data. The specific analysis ideas are as follows:

2.1.1 Analysis of Problem 1

The infectious disease model was extended to the drug-distribution analysis model to establish a differential equation model for opioid-like drug propagation, and the characteristics of the drug-distribution process and related thresholds were analyzed based on the model parameters.

2.1.2 Analysis of Problem 2

To analyze the impact of socio-economic factors in the state on the incidence of opioid cases in the state, cluster analysis was used to classify a large number of socio-economic characteristics and extract representative characteristics, and analyze the socio-economic factors affecting the development process of “Opioid Crisis”. The first model was revised.

2.1.3 Analysis of Problem 3

The inter-state data correlation prediction was performed on the same time section using the gray RBF neural network combination model, that is, the current state case rate was obtained by the case rate prediction of other states to test the model one and the second model. For the first three models, the strategy of confronting “Opioid Crisis” is proposed, and the influencing factors and corresponding thresholds that determine the success or failure of the strategy are analyzed.

# 3. Basic assumption

* Assume that the total number of people in the area examined during the Aka drug transmission period remains the same, regardless of birth rate and mortality, nor birth rate and mortality.
* The data provided by the five states from 2010 to 2017 can better reflect the trend of the total number of Aka drugs in each state in the short term, that is, there is no extreme value.
* After R-type clustering of all socio-economic data, the indicators selected from each category can accurately represent and reflect each category.
* For some missing data, the filling is fair and reasonable, and does not affect the overall.
* The data found is true and reliable.

# 4. Glossary & Symbols

|  |  |
| --- | --- |
| Symbol | Definition |
|  | Daily infection rate |
| Μ | Daily cure rate |
| r | Correlation coefficient |
|  | I-th socioeconomic indicator |
| SSE  N | Sum of squared residuals  Total number of patients in Ohia |

# 5. Models

## 5.1 Analysis and Solving of Question One

5.1.1 Model Preparation

Assume that the total number of people in the area examined during the period of Aka drug transmission remains unchanged, regardless of birth rate and mortality, or population migration. The population is divided into two categories: no taking Aka and taking Aka. , hereinafter referred to as healthy people and patients. And at the time t, the proportion of these two groups of people in the total number N is s(t) and i(t). The average number of people who can effectively contact each patient every day and make healthy people take Aka drugs is constant λ. λ is called the daily contact infection rate, which means that when the patient is in effective contact with the healthy person, the healthy person is infected and becomes a patient. The proportion of patients who are cured every day is a constant μ, called the daily cure rate, which can be expressed as the proportion of successful people who are detoxified every day. In particular, patients can still take Aka drugs repeatedly after the patient is cured. It is obvious that 1/μ is the average infection period during the drug's transmission period. In this way, we can establish a differential equation model for the spread of Aka drugs similar to the model of infectious disease transmission.

5.1.2 Model Establishment

**Step1:** Establish differential equations based on the characteristics of Aka drugs

（1）

**Step2:** Further simplification

（2）

}

**Step3:** Know the results

= （3）

**Step4:** Defining σ=λ / μ, noting the meaning of λ and 1/μ, we can see that σ is the ratio of the daily infection rate to the daily cure rate during the whole propagation period, called the number of contacts. With σ, the model can be rewritten as:

) ] （4）

**Step5:** Separate variables and further simplif

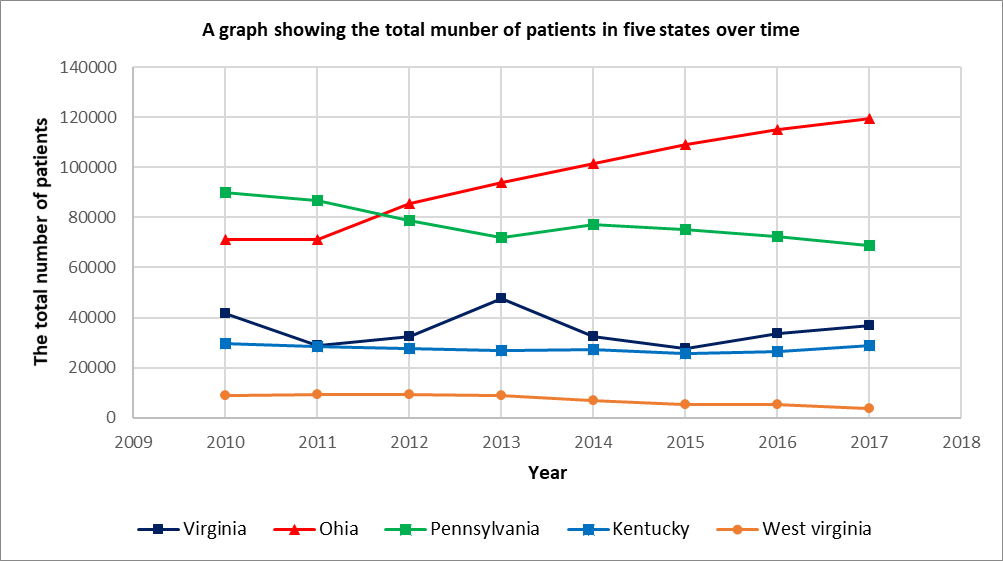
= （5）

**Step6:** Find the points on both sides and solve them.

**（c is the parameter）** （6）

5.1.3 Model soving

（1）Figure 1 shows the relationship between the total number of patients in five continents over time, and Figure 2 shows the geographical relationship between the five continents. It can be seen from the figure that the total number of patients in Ohia has been increasing, and the total number of patients in Pennsylvania is decreasing. The trend of the total number of patients in the other three states over time is not very obvious.

fig1.The total number of patients in five states over time fig2 .The geography of the five continents （2）The analysis shows that the rule of Aka drug propagation process is the same between different states. In order to make the description law more significant, the formula (6) is solved by using Ohia data as an example, and solved by Matlab software. , so that the sum of squared residuals is the smallest, that is to say.

Obtain：

，SSE=0.000014.

The function of the Aka drug infection rate in Ohia over time is

（7）

5.1.4 Analysis of the Result

（1）From σ=1.259, λ >μ, the daily contact transmission rate of the opioid-like drugs is greater than the daily cure rate, which is consistent with the increasing trend of the opioid drug infection rate in Ohia Prefecture from 2010 to 2017, and the residual square sum SSE=0.000014 is extremely small. , indicating that the model has a good fit

（2）The analysis shows that the contact number σ=1 is a threshold. When σ>1, the increase and decrease of i(t) depends on the size of i\_0, but its limit value increases with the increase of σ; When σ ≤ 1, the proportion of patients i(t) is getting smaller and smaller, and eventually tends to 0. This is because the number of healthy people who become patients during the transmission period does not exceed the original number of patients.

（3）In this way, the US government needs to pay special attention to the critical value σ. When the number of patients increases with σ>1, the social security situation may be worsened, taking into account the ratio of the number of Aca drugs taken to the total number of Aka drugs. Stable in the short term, and the cure rate, that is, the compulsory detoxification rate, is basically unchanged under certain government policies (such as social security). All of us can estimate the total infection rate λ according to the usage rate of specific Aka drugs. Compared with the cure rate μ, for example, when λ>μ, the number of patients will increase, and the government needs to develop measures such as strengthening social security. Taking Ohia as an example, it can be seen from Table 2 that the number of patients in Heroin accounts for The total number of people is 30.5%, so that the total number of patients and the infection rate can be estimated accurately based on the number of patients in Heroin, and then analyzed with the critical value, so that the problem is well solved.

**Table1.The percentage of all drug sick in 2010 Ohia**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Name** | **Heroin** |  | **Methadone** | **Morphine** | **…** | **Oxycodone** | **Buprenorphine** |
| **Percentage** | **0.305** |  | **0.0246** | **0.0209** | **…** | **0.122** | **0.377** |

## 5.2 Analysis and Solving of Question Two

5.2.1 Model Preparation

According to the analysis of the socio-economic data provided by the US population, there are too many analytical indicators. There are 149 effective analysis index categories after pre-processing, and there is a strong correlation between each index. This systematically analyzes and studies the society. The relationship between economic data and the spread of Aka drugs has caused great difficulties. Therefore, we first conduct R-type cluster analysis on 149 social and economic indicators, and classify highly relevant economic indicators into several categories to identify the main factors affecting the spread of opioids. In this way, we can select representative indicators from various analysis indicators, and then analyze the selected social and economic indicators and the spread of Aka drugs. Among them, ,,…,respectively represent 149 people's census social and economic indicators

5.2.2 Model Establishment And Solving

(1) Establishment and Solution of R-Type Cluster Analysis

**Step1:** The correlation coefficient is used to measure the correlation of the variables (j=1,2,...,m), then the sample correlation coefficients of the two variables and can be used as their similarity measure,

= （8）

**Step2:** The shortest distance method is used to cluster the variables, and the distance between the two types of variables is defined as,

R（）=min||

among them：or =1-, at this time，R（）correlation with the similarity measure between the two most similar variables in the two categories.

**Step3:** Correlation analysis of Ohia's 149 people's census social economic data indicators is conducted, and a correlation coefficient table between various factors is obtained.

**Table1.** **Table of correlation of 149 analytical indicators**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | … |  |  |  |
|  | 1 | 0.9945 | 0.9888 | … | 0.8368 | 0.8117 | 0.6925 |
|  | 0.9945 | 1 | 0.9967 | … | 0.8297 | 0.8282 | 0.6598 |
|  | 0.9888 | 0.9967 | 1 | … | 0.8087 | 0.8179 | 0.6654 |
| … | … | … | … | … | … | … | … |
|  | 0.8368 | 0.8297 | 0.8087 | … | 1 | 0.7261 | 0.6413 |
|  | 0.8117 | 0.8282 | 0.8179 | … | 0.7261 | 1 | 0.3962 |
|  | 0.6925 | 0.6598 | 0.6654 | … | 0.6413 | 0.3962 | 1 |

**Step4:** The 149 variables were systematically clustered by the maximum coefficient method.The classification results are shown in Fig. 3.

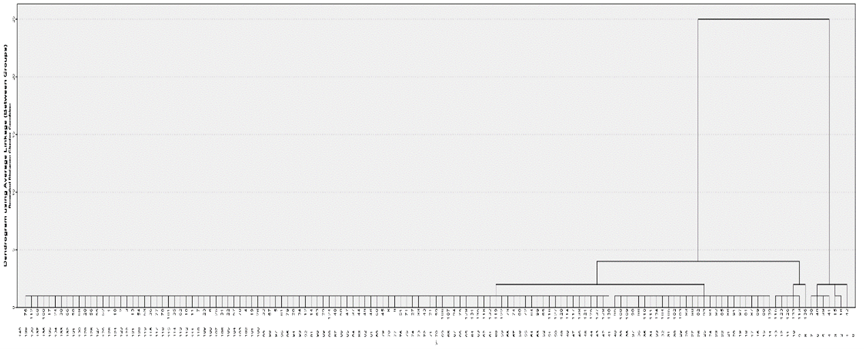


fig3. Clustering figure

Use ,,…={,,…}.

When the distance value ƒ=2, it is divided into five categories, namely

={},={}

={},

={

,}

={} ，among this =-（+++）.

The analysis shows that the first category reflects fertility information, the second category mainly reflects Average household size and Average family size information, the third category mainly reflects the information about the birthplace and the mother tongue, and the fourth category mainly reflects the year of entry. The fifth category mainly reflects information about the Year of entry and the level of education. So we select 5 analytical indicators from 149 indicators, which are ,,,,.

（2）Related analysis

Now we have a correlation analysis between the relationship between the analysis indicators ( ,,) and the number of Aka drug transmissions (y).

Table2. Ohia's five evaluation indicators and the total number of patients

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|  | 221.3 | 218.81 | 221.37 | 222.57 | 221.32 | 220.85 | 220.63 |
|  | 130349 | 129855 | 153278 | 125586 | 145998 | 139586 | 1435739 |
| , | 153782 | 150359 | 130349 | 126298 | 123940 | 122551 | 121848 |
|  | 78425 | 79963 | 84566 | 81626 | 92519 | 86580 | 90673 |
|  | 915914 | 939454 | 949903 | 968771 | 979239 | 975437 | 912807 |
| y | 70999 | 71282 | 85415 | 93747 | 101423 | 109150 | 115276 |

**Step1:** Calculate the average of five analytical indicators and the number of Aka drug transmissions

,,,,,

**Step2:** Calculate the Person correlation coefficient using the formula

(9) Correlation coefficient table

Table3. Correlation coefficient table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  | y |
|  | 1 |  |  |  |  |  |
|  | -0.13401 | 1 |  |  |  |  |
|  | -0.43901 | -0.36292 | 1 |  |  |  |
|  | 0.100381 | 0.48748 | -0.8024 | 1 |  |  |
|  | 0.328978 | -0.57424 | -0.51013 | 0.298602 | 1 |  |
| y | 0.262012 | 0.579961 | -0.92325 | 0.732652 | 0.284207 | 1 |

5.2.3 Analysis of the Result

1）It can be seen from the correlation coefficient table = -0.92325, ||>0.8,so there is a high correlation and a negative correlation between and y. According to the analysis, the number of patients in Ohia increases as people's education level decreases.

2) Further analysis of the reasons, the total number of patients changes with the degree of education of the population, probably because people will consciously reject Aka drugs as their cognitive level increases. At this time, the infection rate λ will also decrease. In this case, the assumption that the infection rate λ is constant in the first differential equation model is incorrect. Therefore, we need to correct the model one.

3) Correction of model one.

① From the previous analysis, the infection rate λ is not a constant, so we can make λ=h(t), which is expressed as a function of the infection rate as a function of time. At this time, the formula (1) should be modified to

（10）

②We can analogize to the Logistic model and make a bold assumption about h(t), let λ=a+bt(a, b be a parameter). At this point, the differential equation should be,

（11）

Simplified

=[1-]- (12)

This is the improved differential equation model.

## 5.3 Analysis and Solving of Question Three

5.3.1 Model Preparation

Based on the above two time series models, the grey RBF neural network combined interstate space prediction model is established. The number of cases in a certain state is predicted by the number of cases in other states at the same time section, and the time series model is cross-tested. The gray system mines valuable internal laws by generating and developing some known information. Set the system characteristic data sequence:

,…

Is the original sequence, where ,…represents the n variables of the decision attribute Observed value

（1－AGO）：,…

Where = ， pairs to generate sequence Establish differential equations:

The parameter column c of the equation is , and the approximate value can be obtained by the least squares method.

=

Where B, is

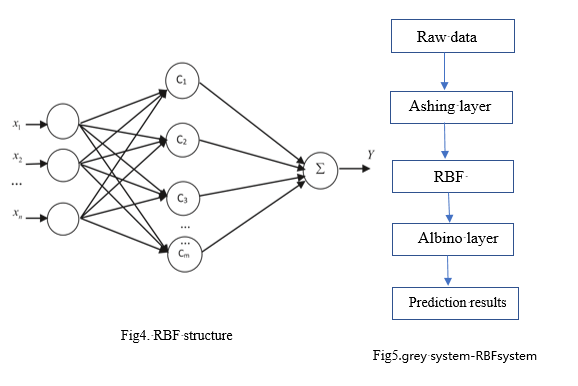
,

The solution to this equation is:)=

Where ) is the predicted value of the accumulated value, and it is reduced to the original data by one reduction:

) =)-) =

The RBF neural network is a three-layer forward neural network with a radial basis function as a hidden layer activation (nuclear) function, including input layer, hidden layer and output layer. It has strong approximation, classification and learning speed, and topology. The structure is shown below



Where =(,) is the input vector, =(,) is the hidden layer, passing the Gaussian function Implementing a nonlinear mapping of the input layer to the hidden layer, where is the center of the i-th basis function, is the norm of the vector , ie the distance between x and , is the first i perceptual variables, Y is the output vector, is the linear mapping of the hidden layer c, Y = () is the jth The weight of the node to the output layer.

5.3.2 Model Establishment

* + **Step1:** Ash the input data to eliminate the randomness of the data.
  + **Step2:** Input RBF neural network
  + **Step3:** Add a whitening layer after the output layer to restore the output data.

5.3.3 Train Results

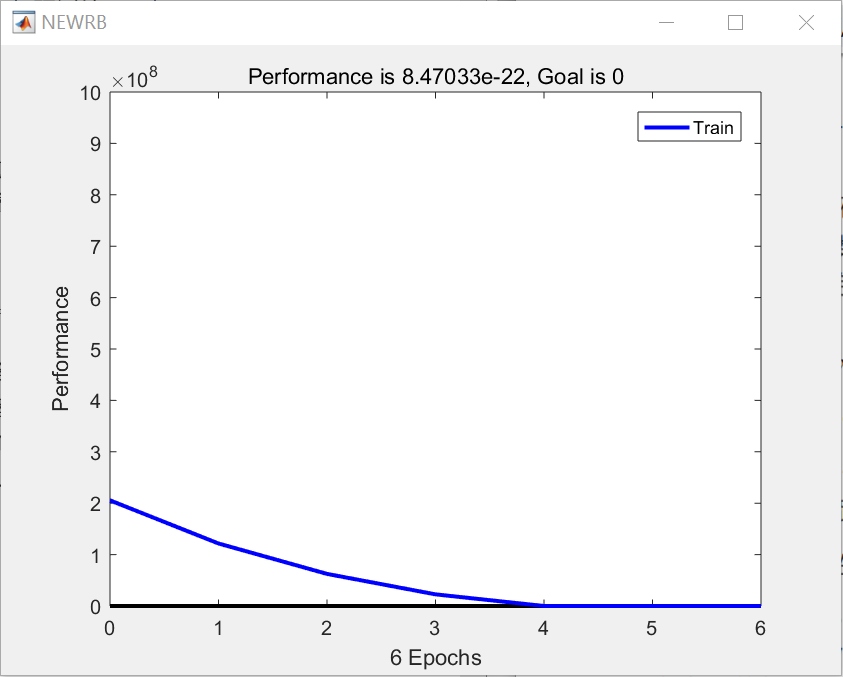


Fig6. Training result

5.3.4 Analysis of the Result

Combining the above three models, we propose the following four strategies to combat the opioid crisis, and illustrate the effectiveness of the strategy in combination with the model:

1. Increase the management and supervision of opioids through state legislation or national legislation, and reduce the chances of people getting in contact with or obtaining opioid addictive drugs. Combined with the first model, it was verified that the case incidence decreased as λ decreased.

2. Increase employment opportunities for low-income people and reduce the gap between the rich and the poor.

3. Communities and schools promote publicity on the harm of opioid abuse.

4. Open office-run schools to increase the proportion of educated residents in poor communities. Combined with the second model, when the proportion of education and education of residents rises, the rate of case-discovery within the state has a downward trend.

We conducted correlation analysis on the characteristic parameters corresponding to several strategies, and extracted the “law on the management and supervision of opioids” with the strongest correlation with the case incidence rate, corresponding to the σ parameter in model one.

Assuming that the propagation rate λ is consistent, when σ>λ, the strategy achieves the expected effect of inhibiting drug transmission, and when σ<λ, the strategy cannot achieve the desired effect.

# 6. Evaluation and Promotion of Model

## 6.1 Strength and Weakness

6.1.1 Strength

* The differential equation model and the RBF neural network model can be used to accurately predict the total number of patients and the infection rate.
* According to the differential equation model, the characteristics of Aka drug transmission can be well explained, and according to the critical value found, the spread of Aka drugs can be effectively prevented.
* We can find the socioeconomic internal factors that promote or inhibit the spread of Aka drugs based on the results of R-type cluster analysis and correlation analysis.

6.1.2 Weakness:

* The data in the paper is limited, only the data from 2010 to 2017, so the model built can not completely avoid the contingency of the data.
* The improved differential equation model was established because the data was not sufficient enough to solve the equation.

# 7. Conclusions

## 7.1 Conclusions of the problem

* The synthetic opioid propagation process satisfies the differential equation model with λ (daily contact rate) and μ (daily cure rate) as parameters. When λ (daily contact rate) is greater than μ (day cure rate), “opioid crisis” is exacerbated.
* Among the many socio-economic factors in the state, the degree of education and the place of origin have a greater correlation with the incidence of opioid-based cases. The higher the level of education, the lower the incidence of cases.
* There is a certain correlation between the incidence of opioids in neighboring states, indicating that interstate transportation is an important factor in promoting drug circulation, and proposes a strategy to deal with “opioid crisis” in terms of education level and social gap between the rich and the poor.

# I X. References

[1] Wang Shixue Liu Junli A Global Analysis of a Model of Random Heroin Drug Transmission[J]. Journal of Beihua University: Natural Science Edition, 2018, 19(5): 575-581.

[2] Wang Shixue, Liu Junli, The Global Stability of a Heroin Drug Propagation Model with Psychological Effects[J]. Journal of Basic Science in Textile Universities, 2018, 31(3): 329-334.

[3] The deep social motivation of Hu Jinye's drug production and dissemination [J]. Century Bridge, 2007 (2): 38-38.

[4] Weng Naiqun, The Social and Cultural Motivation of HIV Transmission[J].Sociological Research,2003(5):84-94.

[5] Fang Bin. Modeling and research of heroin drug transmission system [D]. Aerospace Science and Industry Group Second Research Institute; China Aerospace Second Research Institute, 2014.

[6] Yang Wei. The characteristics of new drugs spreading in China and the risks in the spread of AIDS[J]. Modern Preventive Medicine, 2010, 37(13): 2494-2495.

[7] Li Cong, Tai Hong, Wang Xueren, et al. Research on the spread of drugs and AIDS epidemic [Z]. Yunnan University, Yunnan First People's Hospital, Kunming Municipal Public Security Bureau, Kunming Municipal Disease Control Center, Kunming Maternal and Child Health Care Institute.2005.

[8] Ou Lingcheng, Li Cong, Tai Hong, et al. Research on the spread of drugs and AIDS epidemic[J].Journal of Medical Research,2006,35(1):35-36. DOI:10.3969/j.issn.1673-548X .2006.01.017.

[9] Fang Yan. Air quality prediction based on gray RBF neural network [J]. China Science and Technology Information, 2018, (22): 100-102. DOI: 10.3969/j.issn.1001-8972.2018.22.038.

[10] Jiang Qiyuan, Xie Jinxing, Ye Jun: Mathematical Model (4th Edition), Higher Education Press, 2011

[11] Mary: MATLAB Mathematical Experiment and Modeling, Tsinghua University Press, 2010

[12] Si Shoukui, Sun Zhaoliang. Mathematical modeling algorithms and applications [M]. Version 2. Beijing: National Defense Industry Press, 2016.

[13] US Centers for Disease Control and Prevention https://www.cdc.gov/drugoverdose/

[14]Get geographic data https://geology.com

[15] National Bureau of Statistics http://www.fedstats.gov/

[16]US Census Bureau http://2010.census.gov/2010census/language/chinese-simplified.php