For office use only	Team Control Number	For office use only
T1 ________________	**1921733**	F1 ________________
T2 ________________		F2 ________________
T3 ________________	Problem Chosen	F3 ________________
T4 ________________	**C**	F4 ________________

2019
MCM/ICM
Summary Sheet

Summary

The US has experienced an opioid crisis over the years. In order to analyze the spread and characteristics of the reported synthetic opioid and heroin incidents (cases) in and between the five states and their counties over time, we establish Quadratic Mean Time Series Model. In order to correlate the three factors of personnel flow, family and education with drug use, we establish Grayscale Correlation Analysis Model. We also establish Grayscale Prediction Model to improve the model one to get more accurate results.

According to model 1, we analyze the trends of opium use in five states and predict them. According to the results of prediction, we find that opium use in OH State is deteriorating year by year. So we believe that the OH state will reach the threshold first in 2025. In addition, we also obtain the locations where specific opioid use might have started in each of the five states. They are HAMILTON, JAFFERSON, PHILADELPHIA, FAIRFAX and KANAWHA in turn.

We believe that the border control was not good enough between OH and PA states which led to the spread of opiates from PA to OH states through population movements in earlier years. As far as KY is concerned, we believe that education and family are the most important factors affecting opioid abuse according to the gray correlation analysis model.

Border surveillance between states should be more stringent. Schools should increase adolescents' awareness on drug hazards. And society should care for the unfortunate family members.

We make error analysis on Grayscale Prediction Model. We find that the prediction error can be controlled within 10%. We believe that there are three factors contributing to the prevalence of opioids in the United States. They are: (1)Abuse of prescription drugs (2)Expensive medical costs (3)Addiction characteristics of drugs themselves.

There is still a long way to fight with the opioid crisis. But we believe that we will eventually eliminate this disaster through our efforts.

Contents

Memo to the Administrator

To: Group of Administrators
From: Team 1921733
Date: January 28, 2019

Dear administrator,

The United States has experienced an opioid crisis over the years, a crisis that has had a significant impact on people's health and socio-economic causes, and has attracted widespread attention in recent years. In this paper, we analyze and model the data provided by DEA, NFLIS and the United States Census Bureau, and draw some important conclusions, thus giving strategies to deal with opioids, and discussing the models.

First, in order to determine the spread and characteristics of the reported synthetic opioid and heroin over time in five states and their counties, we establish a quadratic mean time prediction model. The model is based on the data from 2010 to 2017, using the law of the deviation after moving average to establish a quadratic average time series model. Based on the analysis of the results obtained from the model, we think that the opioid use trend in OH state is the most serious, followed by the PA state, but its situation has improved year by year; The situation is close in KY and VA states, probably because the policies of the two states are similar; the situation in the state of WV is the most optimistic. According to the predict results, the use of opioid in OH State has deteriorated year by year, so we believe that the threshold will be reached first in 2025. At this time, the relevant departments in US must take measures to intervene in the epidemic. In addition, we also get the origins of the five states of OH, KY, PA, VA and WV as HAMILTON, JAFFERSON, PHILADELPHIA, FAIRFAX and KANAWHA.

We believe that Ohio and Pennsylvania may lack strict border controls that has led to the spread of opioids from PA state to OH State through population movements in the early years. We suggest that the goverment should strengthen border regulation in each state to limit drug transmission to a minimum.

Based on the socio-economic data provided by the U.S. Census Bureau, we establish a grayscale association model that links the three factors of people mobility, family and education to drug use. The so-called Grayscale Association model is to measure the degree of correlation between factors according to the similarity or degree of different development situation between factors, which explains the characteristics and the degree of dynamic association of things. In the case of KY state, based on the results of the model, we believe that education and the family are the biggest factors affecting their opioid abuse, and the conclusions of other states are also given in the paper. In short, when regulation is not strict, the movement of people can lead to the spread of drugs; if family status and education level are not in place. This may lead to people not knowing the dangers of taking large amounts of drugs. This tells us that border surveillance in each state should be stricter and that education on drug hazards should be implemented.

In order to improve the model one, we establish a grayscale prediction model, which gives the prediction results of opioid use in five states, and is more accurate. Finally, we discuss the error of this model and the sensitivity of the grayscale correlation model to the parameters.

After consulting the relevant information, we conclude that the prevalence of opioid in the United States is due to the following reasons:

(1) **Abuse of prescription drugs.** Because there are so many patients with chronic pain in the United States, doctors are opening up more and more opioids for patients, which has led to a significant increase in the risk of opioid addiction.

(2) **The medical system in the United States is conducive to the development of prescription drugs.** A large proportion of the population in the United States cannot afford expensive treatments, and they have to take medication to treat them, which increases the probability of addiction to opioids.

(3)**The characteristics of the drug itself.** Because of its pain-relieving effect, when taken in large quantities, it leads to dependence on the drug, which ultimately leads to addiction.

In short, there is a long way to go in the fight against the opioid crisis, all sectors of the United States community should be actively involved. The relevant State departments should enact stricter laws to sanction the wrongdoers. We believe that after taking positive measures, the disaster will eventually pass.

Thank you for considering our model and wish you a bright future!

Sincerely,
Team 1921733

1 Introduction

1.1 Background

The opioid crisis refers to a health disaster caused by a sharp rise in the use of prescription and non-prescription opioids in the United States and Canada in the 2010 years. In addition to morphine and heroin, opioids also include hydroxyl two hydrogen codeine ketone, hydrogen ketone and fentanyl. The CDC says the deaths of 42,249 Americans in 2016 were linked to opioid overdoses[1]. There are reports that the prescription of opioid drugs has placed an economic burden on the United States by about $78.5 billion over a 2013-year period[2]. On October 26, 2017, US President Donald Trump announced that the United States entered a national public health emergency to cope with the spread of opioids. Trump said "more people die from drug overdoses than from gun killings and car crashes to death combined." He called drug overdoses now the "leading cause of accidental death" in the United States[3]. In order to more clearly illustrate the effects of death caused by opioids, please pay attention to figure1.

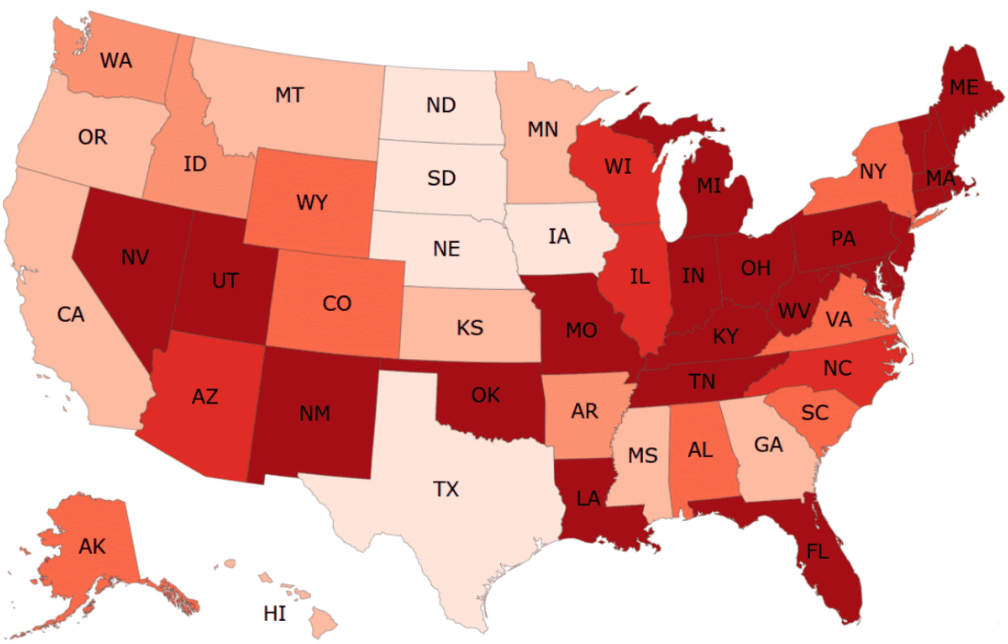

Figure 1: The redder color indicates the more serious the death of people in the area due to opioid

As can be seen from the figure1, in the United States, the problem of death from opioid is so serious that the economic situation in the United States will also be greatly affected.

We can know that the situation of the opioid crisis in the United States is very bad, and it is imperative to take measures to deal with it.

1.2 Problem Restatement

In order to study the use of opioids in the United States and propose corresponding countermeasures, we need to analyze the data provided by NFLIS. For this problem, we focus on the individual counties located in five states: Ohio, Kentucky, West Virginia, Virginia, and Pennsylvania. We need to do as follows:

Part1. Using the NFLIS data provided, we need to build a mathematical model to describe the spread and characteristics of the reported synthetic opioid and heroin incidents in and between the five states and their counties over time and then identify any possible locations where specific opioid use might have started in each of the five states.

If the patterns and characteristics we identified continue, we need to determine where and when the drug threshold will be met, and the US government needs to take the necessary interventions.

Part2. Using the U.S. Census socio-economic data provided, we need to modify the model of Part1 to determine the following issues: Who is using opioids, what contributes to the growth in opioid use and addiction, and why opioid use persists despite its known dangers. Is use or trends-in-use somehow associated with any of the U.S. Census socio-economic data provided?

Part3. Using a combination of our Part 1 and Part 2 results, identify a possible strategy for countering the opioid crisis. Use our model to test the effectiveness of this strategy; identifying any significant parameter bounds that success (or failure) is dependent upon.

1.3 Our Work

- First, we performed a statistical analysis of the total data of the five states to obtain trends in the use of opioid drugs in each state. We constructed a **Quadratic Mean Time Series Model (QMTSM)** to predict changes in the use of opioid drugs in the states in the coming years.

- Secondly, for a specific state, we classify each county to select the county with more serious disasters for analysis, and regard other counties as the whole to determine the birthplace of the state's opioid drugs use. Based on current trends, we predict when and where the state will reach the threshold.

- In order to study the relationship between opioid drugs and education, family, population mobility and other factors, we statistically analyze the socio-economic data of the US census and establish a **Grayscale Correlation Analysis Model (GCAM)**.

- To correct the **QMTSM**, we introduced the **GM(1,1) theory**[4] to make the predictions take into account the relevant factors of the US Census socioeconomic data.

- Finally, combined with the results of the model, we present the measures the government needs to address the opioid crisis. Then, the reliability of the strategy given by the model is evaluated and a sensitivity analysis is performed.

2 Assumptions and Justfication

To simplify the problem, we have the following basic assumptions, which are properly justified.

- The provided data is realistic and accurate to a certain degree. Despite the incompleteness of the data and some tolerant error in statistics, we make this assumption to guarantee one valid solution.
- The trend of opioid use change is approximately linear in a short time. In our model, we optimize the time series model. And the predicting model is an equation with upward or downward trend. So it is necessary that trend is nearly linear.
- We ignore the county, where the the opioid are lessly used. It's necessary to simplify our model by neglecting secondary factors.
- We assume that the relative importance of various indicators doesn't have drastic change overtime. If the relevant indicators change dramatically, it will be difficult to simplify the model and solve our problem.
- In the grayscale correlation analysis model, in order to simplify the model, we mainly consider the influence of education, family, and population mobility on opioid use.

3 List of Notation

Table 1: The List of Notation

Notation	Meaning
t	time node
a_t	Intercept of trend Lines
b_t	Slope of trend lines
m	Time Independent variable
y_t	Observation sequence values for t moments
$\hat{y}_{t+m}$	Predictive values for $t+m$ moments
w_i	The weight of y_{t-i+1}
$M_t^{(1)}$	One weight moving average

$M_t^{(2)}$	Secondary weight moving average
N	Number of items
γ	The proportion of drug users to the total population
$y_t \in (0, 1000]$	The drug use y_t is relatively light
$y_t \in (1000, 5000]$	The drug use y_t is generally the case
$y_t \in (5000, 10000]$	The drug use y_t is serious
$y_t \in (10000, +\infty)$	The drug use y_t is very serious
$x_i(t)$	i-th time series
$z_i(t)$	Normalization of $x_i(t)$
x_0	Reference sequence
ρ	Resolution factor
ξ_i	Correlation coefficient
β_i	Correlation (The summation of ξ_i)
σ^2	Variance
HBT	HOUSEHOLDS BY TYPE - Family households (families) -one-parent family- With own children under 18 years
RN	RELATIONSHIP - Nonrelatives
MS	MARITAL STATUS-Separated&Widowed&Divorced
GN	GRANDPARENTS - Number of grandparents responsible for own grandchildren under 18 years
PEA	Population 25 years and over-EDUCATIONAL ATTAINMENT-University degree and over
$PEAL$	Population 25 years and over-EDUCATIONAL ATTAINMENT-less than University degree
$PYAS$	Population 25 years and over-EDUCATIONAL ATTAINMENT-less than University degree
$PYAD$	RESIDENCE 1 YEAR AGO - Different house in the U.S. - Different county - Different state

4 Data Processing

- In order to get a trend graph of the use of opioid in each state over time, we extract the data provided by NFLIS, focus on TotalDrugReportsState, and remove other data. Here we use MATLAB to map figure2 in the use of opioid in each state over a period of eight years.
- For specific states, such as Ohio, we need to analyze trends in the use of opioid its counties over time. We find that opioid use in a few counties was very serious, such as HAMILTON reaching 13,780 in 2010. And more counties use less, so we classify these counties into one category and take the average. And we define several intervals to characterize their severity of the disaster, please pay attention to Table1. Here we also give pie charts

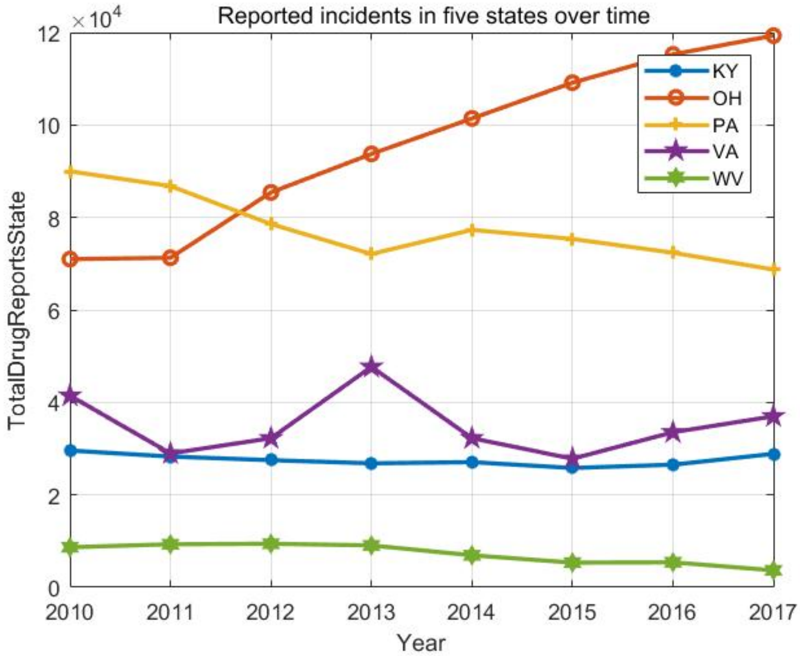

Figure 2: Reported incidents in five states over time

of the use of opioid in the counties of Ohio between 2010 and 2017, see the Appendix.

- When dealing with data from the US Census Bureau, we selected three aspects of family, education, and population mobility in socioeconomic factors. For family status, we select four aspects: HBT, RN, MS, and GN. For the level of education, the data given is the education status of people aged 25 and over. After processing the data, we divide it into PEA and PEAL. For the floating population, we analyze PYAS and PYAD.

5 Establishment of Our Model

To describe the spread and characteristics of the reported synthetic opioid and heroin events over time in five states and their counties, we need to establish a mathematical model. According to figure2, we consider that we might be able to build a time series model.

5.1 Weighted Average Moving Model

In the simple moving average formula, the effect of each period of data on averaging is equivalent. However, the amount of information contained in each issue is different, and recent data contains more information about the future. Therefore, it is not reasonable to treat each period of data as equivalent. The importance of each period of data should be considered, and the recent data should be given a larger weight. This is the basic idea of the weighted moving average method.

Let the time series be $y_1, y_2, ..., y_t, ...$; the weighted moving average formula is

$$M_t^{(1)} = \frac{w_1 y_t + w_2 y_{t-1} + ... + w_N y_{t-N+1}}{w_1 + w_2 + ... + w_N}, \qquad t \geq N$$

Where $M_t^{(1)}$ is the one weighted moving average at time t; w_i is the weight of y_{t-i+1}, it embodies the importance of the corresponding y_t in the weighted average.

Using a weighted moving average to make a prediction, the prediction formula is

$$\hat{y}_{t+1} = M_t^{(1)}$$

That is, the t th period weighted moving average is used as the predicted value of the $t+1$ th period.

We make $w_1 = w_2 = ... = w_N = 1$, through the above theoretical analysis, we can obtain data of 2018 based on data from 2010 to 2017, and the table is as follows:

TotalDrugState Year / State	2010	2011	2012	2013	2014	2015	2016	2017	**2018**
KY	29588	28285	27502	26820	27077	25811	26530	28870	**27560.4**
OH	70999	71282	85415	93747	109150	115276	115276	119349	**95830.1**
PA	89981	86793	78577	72096	77318	75351	72376	68751	**77655.4**
VA	41462	28969	32251	47694	32265	27819	33539	36994	**35124.1**
WV	8668	9310	9429	9062	6926	5345	5405	3672	**7227.1**

Table 2: 2018 data predicted by this model

According to this table, we can see that the predicted values of the five states in 2018 are $27560.4, 95830.1, 77655.4, 35124.1, 7227.1$. After qualitative analysis, the changes in these five data relative to the previous year are contrary to the trend of our curve in figure2. Therefore, we suspect that this model has defects in prediction.

5.2 Quadratic Mean Time Series Model (QATSM)

In the previous model, we think that the predicted values it processes may deviate significantly from the actual values. According to assumption2, when the time series shows a trend of linear change, the simple moving average method and the weighted moving average method can be used to predict the lag deviation. Therefore, it needs to be revised by means of second moving average. So next we will discuss the new model: **Quadratic Mean Time Series Model**, referred to as **QATSM**.

From the analysis of the previous model, we can know that the average number of moves is

$$M_t^{(1)} = \frac{y_t + y_{t-1} + ... + y_{t-N+1}}{N}$$

Performing a moving average on the basis of a moving average is a secondary moving average, and the calculation formula is

$$M_t^{(2)} = \frac{1}{N}(M_t^{(1)} + ... + M_{t-N+1}^{(1)}) = M_{t-1}^{(2)} + \frac{1}{N}(M_t^{(1)} - M_{t-N}^{(1)})$$

Let the time series y_t have a linear trend from a certain period, and think that the future period also changes according to this straight line trend, then the linear trend prediction model can be set as

$$\hat{y}_{t+m} = a_t + b_t \cdot m, \qquad m = 1, 2, ...$$

According to the above analysis, the calculation formula of a_t and b_t can be obtained as

$$\begin{cases} a_t = 2M_t^{(1)} - M_t^{(2)} \\ b_t = \frac{2}{N-1}(M_t^{(1)} - M_t^{(2)}) \end{cases}$$

where $N = 3$.

We calculated the parameters of the five states separately and the predicted values for 2018, which are shown in the table below:

	KY	OH	PA	VA	WV
$\boldsymbol{a_t}$	2.7436	12.0966	7.0016	3.2262	0.3678
$\boldsymbol{b_t}$	0.0366	0.0735	-0.2413	-0.0521	-0.1129
$\boldsymbol{TotalDrugState_{2018}}$	121701	27802	67603	31742	2486

Table 3: Linear analysis and prediction results

We make 5 fitting straight lines as shown in figure3. This fit curve is more convincing than the results from the previous model analysis, considering the data and trends of each state. According to figure3, we can see that OH is the most severely used state of opioid, and it has a trend of increasing year by year; PA state is second only to OH state, and its situation has improved, which may be due to the migration of PA state to OH state population the result of. The situation in the two states of KY and VA is quite close, which may be related to the local policies of the two states, while the state of WV is the best and stable. Figure4 shows the geographical distribution and opioid use of the five states. The darker the color, the more severe the use of opioid in the state.

As the number of drug users in the OH state is increasing year by year, the state can be considered to be the first to reach the drug identification threshold. Once this value is reached, the relevant US government should take corresponding measures to intervene.

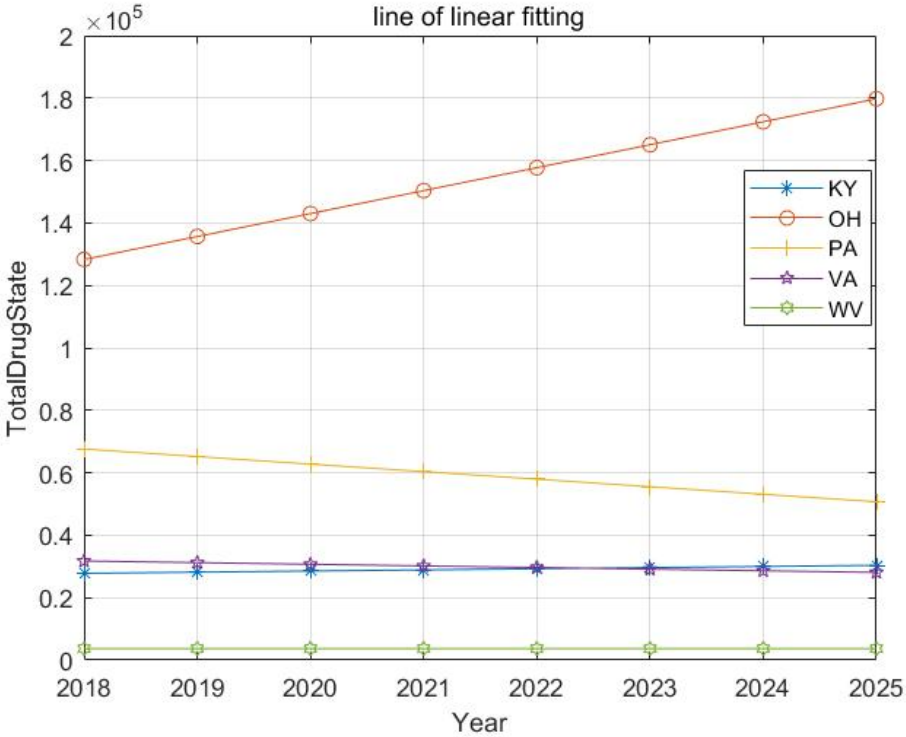

Figure 3: line of linear fitting

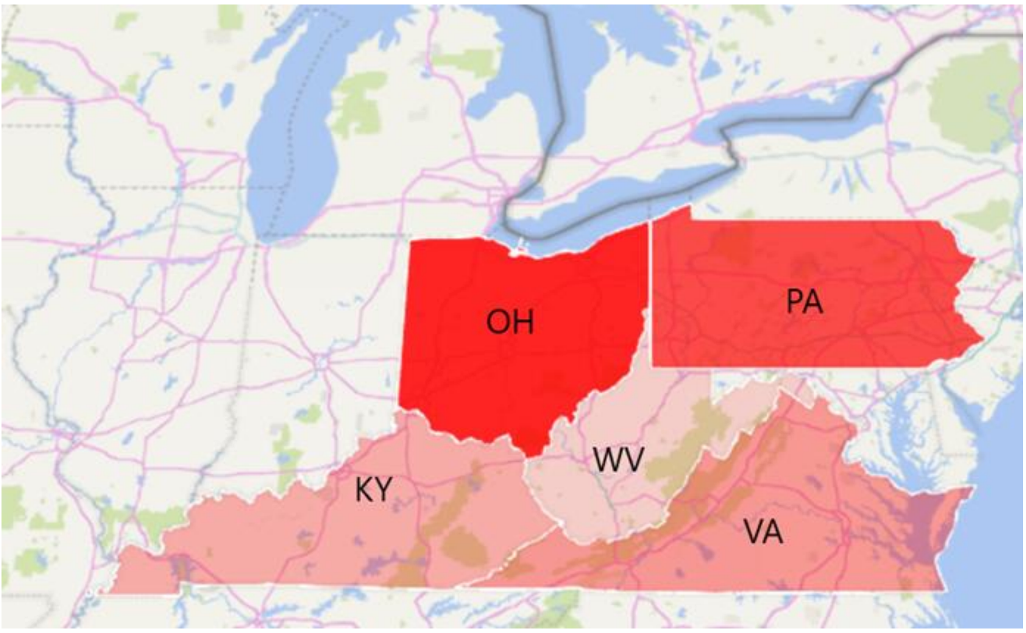

Figure 4: Geographical location and opioid use in five states

The total population of OH State in 2017 is 11.66 million[4]. We define a series of problems such as social stability when γ reaches the drug identification threshold of 1.5%. In order to simplify the analysis, we believe that the total population remains unchanged. The calculation results are given in table4.

Year	2018	2019	2020	2021	2022	2023	2024	2025
γ	1.10%	1.16%	1.23%	1.29%	1.35%	1.42%	1.48%	1.54%

Table 4: Predict situation of γ

It can be seen that OH State will reach the drug identification threshold by 2025. Let's continue to analyze the specific conditions of the counties in this state by **QATSM**.

According to assumption3 and the discussion of intervals in the symbol table, we focus on four counties: CUYAHOGA, HAMILTON, MONTGOMERY, and FRANKLIN. Based on the statistical results (see appendix), we draw the following opioid use maps for each county between 2010 and 2017:

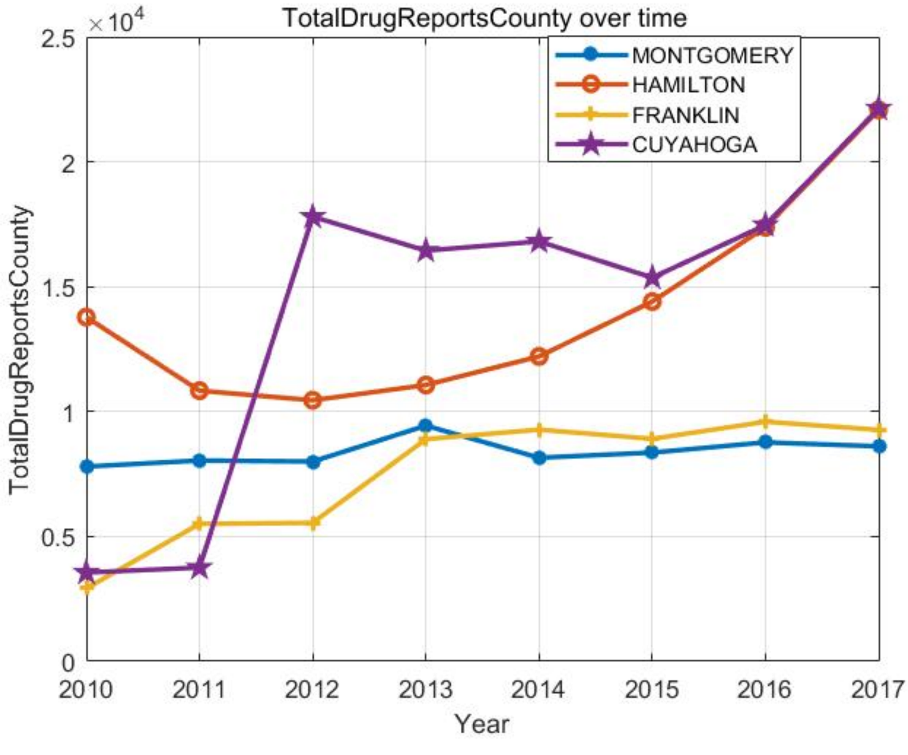

Figure 5: TotalDrugReportsCounty over time

From figure5 we can see that the two counties of MONTGOMERY and FRANKLIN are very similar, and the law of change over time is about the same. HAMILTON's data increased year by year after a small decline in the first year, and the growth rate increased; CUYAHOGA's number in the second year increased sharply, and then it fluctuated within a certain range, we guess that in the second year, there are a large number The opioid invades this state.

Using this model, we can plot the predicted lines of these counties in figure6.

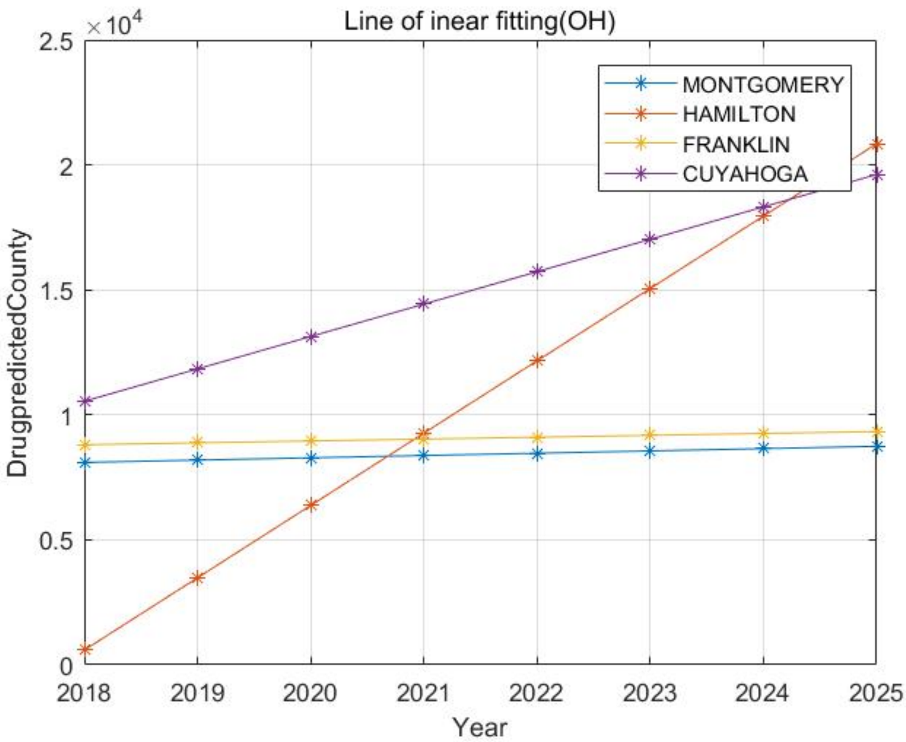

Figure 6: Line of linear fitting(OH)

According to figure5, we believe that the origin is HAMILTON, and the value of CUYAHOGA increases sharply due to propagation; based on figure6, HAMILTON will reach the threshold in 2025.

Using the same method, we also analyzed the situation of the other four states. The statistic line chart and predict curve are in the appendix. After analysis, the origins of the four states KY, PA, VA, and WV are JEFFERSON, PHILADELPHIA, FAIRFAX, and KANAWHA.

5.3 Grayscale Correlation Analysis Model (GCAM)

In order to solve the problem in part2, we introduce the **Grayscale Correlation Analysis Model**, referred to as **GCAM**. Let's introduce it.

5.3.1 Data Transformation

We find that the data in the statistics table cannot directly perform correlation operations. So we introduce data changes to eliminate the dimension of the data for data processing. Suppose there is a sequence

$$x = \{x(1), x(2), ..., x(n)\}$$

We define the map

$$f(x(t)) = \frac{x(t)}{x(1)} = z(t), \qquad t = 1, 2, ..., n$$

is the data transformation from sequence x to sequence z.

5.3.2 Correlation Analysis

We define the reference sequence

$$x_0(t) = \{x_0(1), x_0(2), ..., x_0(n)\}$$

Suppose there are m comparison sequences

$$x_i = \{x_i(1), x_i(2), ..., x_i(n)\}, \qquad i = 1, 2, ..., m$$

we say

$$\xi_i(t_0) = \frac{\min_s \max_t |x_0(t) - x_s(t)| + \rho \max_s \max_t |x_0(t) - x_s(t)|}{|x_0(t_0) - x_i(t_0)| + \rho \max_s \max_t |x_0(t) - x_s(t)|}$$

is the correlation coefficient of the comparison sequence x_i to the reference sequence x_0 at time t_0, where $\rho\in$[0, 1] is the resolution coefficient.

In general, the larger the resolution coefficient ρ, the greater the resolution. We make ρ be 0.8 in this paper. $\xi_i(t_0)$ is an indicator describing the degree of association between the comparison series and the reference sequence at t_0. Since there is an associated number at each moment, the information appears to be too scattered and inconvenient to compare. For this reason, we make

$$\beta_i = \frac{1}{N} \sum_{t=1}^{N} \xi_i(t)$$

be the degree of association of the sequence x_i with the reference sequence x_0. The degree of association is to concentrate the correlation coefficients at each moment into an average value, that is, to concentrate the information that is too scattered. Using the concept of relevance, we can factor the various issues.

In the calculation formula of $\xi_i(t_0)$, we can't use $|x_0(t) - x_i(t)|$ to distinguish whether the factor correlation makes positive correlation or negative correlation. In order to solve this problem, we have

$$\sigma = \sum_{t=1}^{N} kx_i(t) - \sum_{t=1}^{N} x_i(t) \sum_{t=1}^{N} \frac{t}{n}, \qquad i = 1, 2, ..., n$$

so:

(1)When $sign(\sigma_i) = sign(\sigma_j)$, then x_i and x_j are positively correlated;

(2)When $sign(\sigma_i) = -sign(\sigma_j)$, then x_i and x_j are negatively correlated.

As far as KY, we use the $\boldsymbol{TotalDrugState}$ between 2010 and 2016 as a reference sequence, and compare the data of HBT, RN, MS, GN, PEA, PEAL, PYAS and PYAD as normalization sequences, and then perform correlation analysis. We calculate the correlation coefficient matrix of the comparison sequence and the associated sequence

$$\boldsymbol{\xi} = \begin{bmatrix} 1 & 0.82 & 0.72 & 0.67 & 0.69 & 0.63 & 0.68 \\ 1 & 0.80 & 0.67 & 0.58 & 0.53 & 0.46 & 0.44 \\ 1 & 0.82 & 0.73 & 0.66 & 0.65 & 0.58 & 0.61 \\ 1 & 0.76 & 0.62 & 0.57 & 0.61 & 0.54 & 0.54 \\ 1 & 0.76 & 0.65 & 0.53 & 0.55 & 0.49 & 0.49 \\ 1 & 0.87 & 0.83 & 0.88 & 0.85 & 0.75 & 0.83 \\ 1 & 0.99 & 0.92 & 0.82 & 0.0.72 & 0.59 & 0.58 \end{bmatrix}$$

We can also get the value of σ for each sequence, as shown in the Table5, where σ_0 is the value obtained from the reference sequence, $\sigma_1 ... \sigma_8$ is the value

σ_0	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8
-14547	9975	194863	217938	21601	940854	-384546	72380	-43334

Table 5: $\boldsymbol{\sigma_i}$(KY)

obtained from the comparison sequence. We conclude that PEAL and PYAD are positively correlated with drug use in the state, and the other six factors are negatively correlated or weakly correlated.

We also average each row of the above matrix to get the relevance of each factor, as listed below:

Factor	HBT	RN	MS	GN	PEA	PEAL	PYAS	PYAD
$\boldsymbol{\beta_i}$	-0.77	-0.64	-0.72	-0.66	-0.64	0.86	-0.80	0.89

Table 6: $\boldsymbol{\beta_i}$(KY)

The closer $\boldsymbol{\beta_i}$ is to 1 the greater the impact this factor has on the state's opioid use. As can be seen from the above table, for KY, PYAD has the greatest impact, and RN and PEA have the least impact. Specifically, population movements can have a significant impact on the surge in opioid users, and when opioid users go out, the increased volume of activity and the increase in exposure will greatly increase the probability of the use of opioids by those around them. At the same time, as the level of individual education increases, awareness of the dangers of opioids increases and it is not easy to use opioids. As a result, the utilization rate of opioid use may reduce as the level of education increases.

We also list the σ_i and β_i values of other states in Table7, 8.

	σ_0	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8
OH	224575	-1043	374553	267355	18023	1998321	-1024323	87545	75424
PA	-76958	-56205	447317	210020	30579	2078509	-693230	-17638	-83191
VA	-26055	-763	389655	250122	19683	1908801	-1009413	82820	77662
WV	-20222	4769	61375	4589	3904	152637	-340966	2378	-35284

Table 7: $\boldsymbol{\sigma_i}$(OH,PA,VA,WV)

Factor	HBT	RN	MS	GN	PEA	PEAL	PYAS	PYAD
$\boldsymbol{\beta_i}$(OH)	-0.7	0.74	0.71	0.72	0.73	-0.68	0.69	0.71
$\boldsymbol{\beta_i}$(PA)	0.73	-0.63	-0.70	-0.67	-0.67	0.73	0.78	0.80
$\boldsymbol{\beta_i}$(VA)	0.67	-0.63	-0.66	-0.65	-0.64	0.69	-0.66	0.65
$\boldsymbol{\beta_i}$(WV)	-0.77	-0.75	-0.78	-0.76	-0.77	0.78	-0.76	0.79

Table 8: $\boldsymbol{\beta_i}$(OH,PA,VA,WV)

5.4 Grayscale Prediction Model

To improve the **QMTSM**, we introduce the **GM(1,1) theory** to make the predictions take into account the relevant factors of the US Census socioeconomic data.

5.4.1 Introduction of GM (1,1)

GM(1,1) uses the original discrete data column to generate a more regular new discrete data column that weakens the randomness by one accumulation, and then builds the differential equation model to obtain the solution at the discrete point. An approximate estimate of the raw data, thereby predicting the subsequent development of the original data.

Let's model KY first.

(1) Make an accumulation of raw data $x^{(0)}$, this is

$$x^{(1)} = \{29688, 57873, 85375, 112195, 139272, 165083, 191613, 220483\}$$

(2) Construct a data matrix $\boldsymbol{B}$ and a data vector $\boldsymbol{Y}$

$$\boldsymbol{B} = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(3)) & 1 \\ \dots & \dots \\ -\frac{1}{2}(x^{(7)}(1) + x^{(1)}(8)) & 1 \end{bmatrix}, \boldsymbol{Y} = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(8) \end{bmatrix}$$

(3) Calculate $\hat{u}$

$$\hat{u} = (a,b)^T = (\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{Y} = \begin{bmatrix} 0.0016 \\ 2.75 \times 10^4 \end{bmatrix}$$

(4) Modeling

$$\frac{dx^{(1)}}{dt} + 0.0016x^{(1)} = 2.75 \times 10^4$$

In this way, we sequentially derive the models of OH, PA, VA, and WV:

OH:

$$\frac{dx^{(1)}}{dt} - 0.0768x^{(1)} = 6.96 \times 10^4$$

PA:

$$\frac{dx^{(1)}}{dt} + 0.0301x^{(1)} = 8.69 \times 10^4$$

VA:

$$\frac{dx^{(1)}}{dt} - 0.0069x^{(1)} = 3.31 \times 10^4$$

WV:

$$\frac{dx^{(1)}}{dt} + 0.1389x^{(1)} = 1.22 \times 10^4$$

5.4.2 Improved Predict Results

Solutions of the model:

KY:

$$x^{(1)}(k+1) = -45803745e^{-0.0016k} + 4583333$$

OH:

$$x^{(1)}(k+1) = 835251e^{0.0768k} - 906250$$

PA:

$$x^{(1)}(k+1) = -2797062e^{-0.0301k} + 2887043$$

VA:

$$x^{(1)}(k+1) = 4755639e^{0.0069k} - 4797101$$

WV:

$$x^{(1)}(k+1) = -79164e^{-0.1389k} + 87833$$

To illustrate more vividly, we plot the above equation in figure7.

It can be seen that this result is more consistent with the data from 2010 to 2017, and gives the predicted value for the next three years.

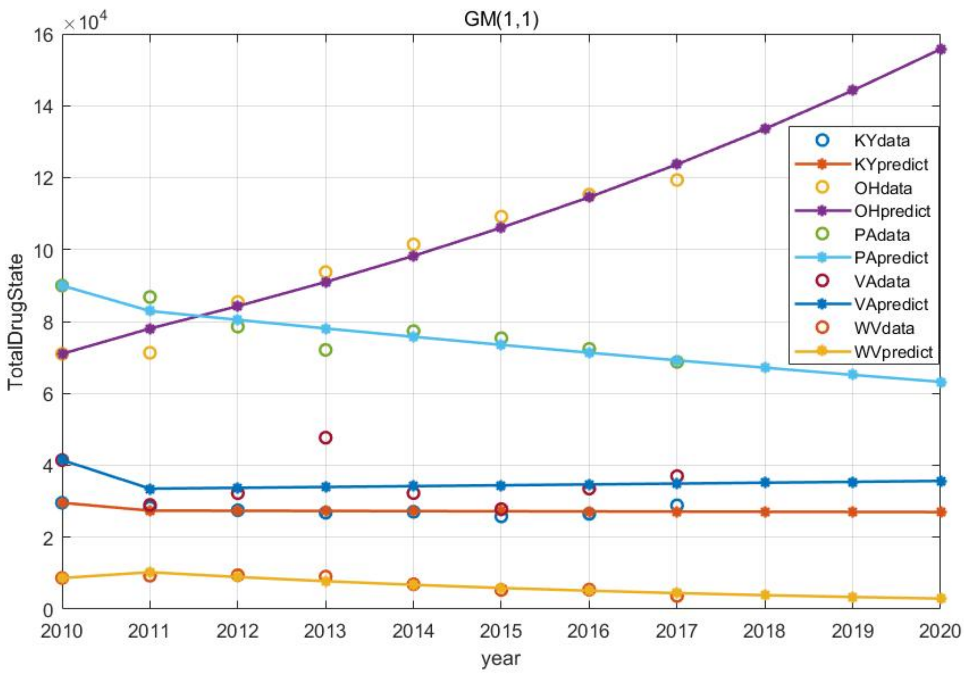

Figure 7: Grayscale prediction

6 Conclusion

- (1) OH is the most severely used state of opioid and is increasing year by year; PA is second only to OH, and its condition is improving slightly. Combined with **GCAM**, we believe this is due to the migration of the drug population from PA to OH. The situation in the two states of KY and VA is very close, while the situation in WV is the most optimistic.

  (2) For the OH state, the two counties of MONTGOMERY and FRANKLIN are very similar, and its growth rate is gradually increasing. HAMILTON's data increased year by year after a small decline in the first year, and the growth rate is getting bigger; CUYAHOGA deteriorated sharply in the second year and then fluctuated within a certain range. We believe that a large amount of opioid has invaded the state.

  (3) For the OH state, we believe that the source of the drug is HAMILTON County, and due to the spread, the situation of CUYAHOGA has deteriorated drastically. And in 2025, we believe that HAMILTON of OH State will reach the threshold. Before this time, the US government must take relevant measures to intervene.

  (4) The drug origins of the four states KY, PA, VA and WV are JEFFERSON, PHILADELPHIA, FAIRFAX and KANAWHA.

- According to the conclusion of **GCAM**, we believe that the reasons why opioid is so popular in the United States are as follows:

(1) **Abuse of prescription drugs.** Because there are so many patients with chronic pain in the United States. Doctors presciribe paticients more and more opioids, which leads to a significant increase in the risk of opioid addiction.

(2) **The medical system in the United States is conducive to the development of prescription drugs.** A large proportion of the population in the United States cannot afford expensive treatments. So they have to take some medicine such as opioid, which increases the probability of addiction to opioids.

(3)**The characteristics of the drug itself.** Because of its pain-relieving effect, it's easy to be taken in large quantities. And it leads to dependence on the drug, which ultimately leads to addiction[5].

7 Model Analysis

7.1 Error Analysis

Considering that the **Grayscale Prediction Model** is an improvement of **QMTSM**, we describe its error by analyzing the relative residual value of the model.

	2010	2011	2012	2013	2014	2015	2016	2017
KY	0.0585	0.0612	0.0629	0.0645	0.0639	0.0671	0.0652	0.0600
OH	0.0612	0.0610	0.0509	0.0464	0.0429	0.0398	0.0377	0.0364
PA	0.0050	0.0052	0.0057	0.0063	0.0058	0.0060	0.0062	0.0066
VA	0.0498	0.0713	0.0641	0.0433	0.0640	0.0743	0.0616	0.0558
WV	0.0915	0.0852	0.0841	0.0875	0.1145	0.1483	0.1467	0.2159

Table 9: Relative residual value

7.2 Sensitivity Analysis

We analyze the sensitivity of the parameters of **GCAM**. By adjusting the value of the correlation coefficient ρ, the original value is 0.8, and then we make it be 0.2, 0.4, 0.6, 0.8. We find that when ρ=0.8, the variance of the model is smaller and the model is more stable. Results are shown in figure8 and table10.

ρ	0.2	0.4	0.6	0.8
σ^2	0.0156	0.0146	0.0119	0.0096

Table 10: The relationship between ρ and σ^2

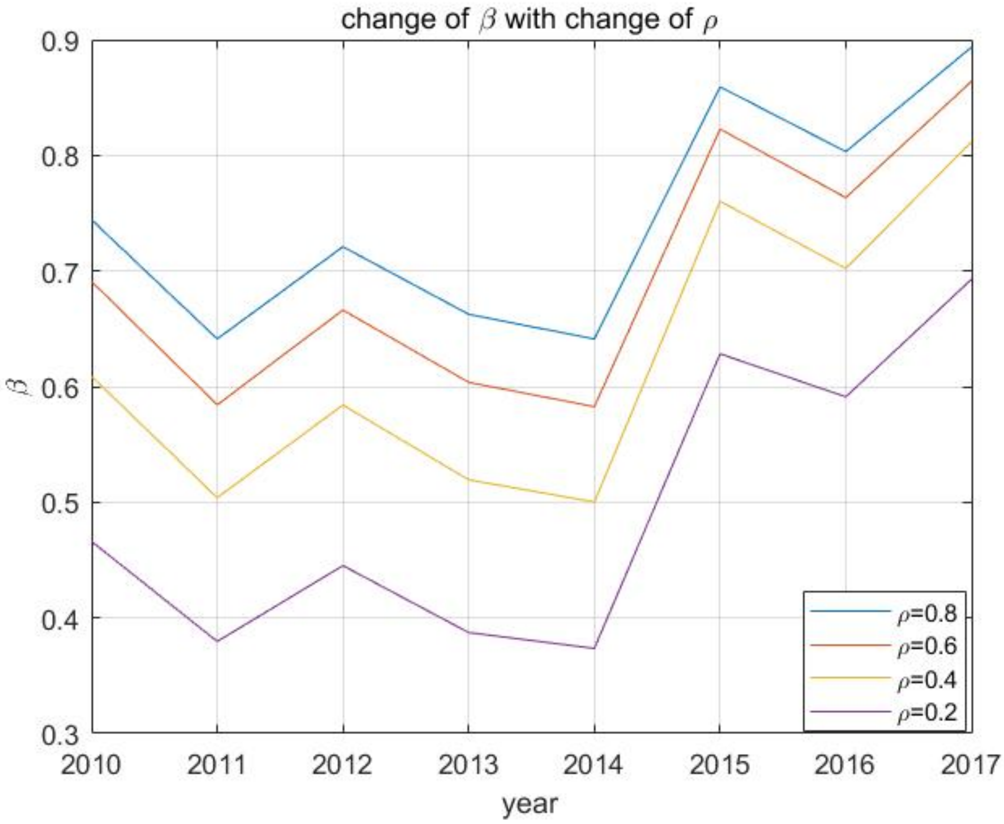

Figure 8: Sensitivity Analysis

8 Evaluations of Models

8.1 Strengths

- **QMTSM** is more accurate in predicting the time series of monotonous rise or fall trends, such as the change of the use of opioid in OH.
- **GCAM** is a multi-factor statistical analysis method, which is based on the sample data of various factors to describe the strength, size and order of the relationship between factors, if the sample data reflects the two factors change the situation is basically the same, then the correlation between them is large; conversely, the correlation degree is small. The advantage of this method is that the idea is clear, which can reduce the loss caused by information asymmetry to a great extent, and the data requirement is low and the workload is less.

  GCAM is analyzed according to the development trend. Therefore, there is no excessive requirement for the sample size, nor the typical distribution law is required, and the calculation amount is relatively small, and the results are in good agreement with the qualitative analysis results. Therefore, **GCAM** is a relatively simple and reliable method of analysis in system analysis.

- **Gray Prediction Model** requires less data and the predictions are accurate. The sample distribution does not need to be regular, the calculation is simple, and the test is convenient.

  Gray Prediction Model can solve the problem of less historical data, sequence integrity and low reliability when processing less eigenvalue data, and the sample space without data is large enough, which can generate the irregular raw data and obtain the law strong generation sequence.

8.2 Weaknesses

- **QMTSM** does not apply to time series where the change is not monotonic, and can only be applied to recent predictions, not for long-term predictions.
- The main disadvantage of **GCAM** is that it requires the current determination of the optimal value of each index, subjectivity is too strong, and the optimal value of some indicators is difficult to determine. Due to the large number of correlation coefficients, the information is too scattered and it is not convenient to compare.
- In the process of predicting with GM(1,1), **Gray Prediction Model** uses the result of accumulating the original data as a known condition to solve the prediction formula. This lacks a strict theoretical basis, so the final result is not necessarily the most Good prediction formula.

  After error analysis, we find that not all relative residuals are less than 10%, so this model is not perfect in terms of prediction.

9 What we hope

Based on the above conclusions and relevant information[7][8], we now give strategies against opioids.

- **Improving opioid prescribing**. The best ways to prevent opioid overdose deaths are to improve opioid prescribing, reduce exposure to both prescription and illicit opioids, prevent misuse, and to treat opioid use disorder.
- **Improving Data Quality and Tracking Trends**. Timely, high-quality data are critical to help public health officials effectively respond to the opioid overdose epidemic. Data help us understand the extent of the problem, focus resources where they are needed most, and evaluate the success of prevention and response efforts.
- **Building State, Local, and Tribal Capacity.** States, local communities, and tribes play an important role in preventing opioid overdoses and related

harms. They run prescription drug monitoring programs, regulate controlled substances, license healthcare providers, respond to drug overdose outbreaks, and run large public insurance programs such as Medicaid and workers compensation.

- **Improving the level of education and drumbeating the harm of opioids.** Schools and society should actively assume responsibility by opening relevant education subjects, and increasing the publicity of the harm of opioids.

- **Supporting Healthcare Providers and Health Systems**. Improving the way opioids are prescribed can ensure patients have access to safer, more effective pain treatment while reducing the number of people who misuse, abuse, or overdose from these drugs. Providers and the health systems in which they work are critical when it comes to promoting safer and more effective opioid prescribing.

- **Increasing border supervision.** If border surveillance is not strict enough, opioids may spread from one state to another, and the consequences are unimaginable. Therefore, border supervision must be strengthened.

- **Caring for the people with a unfortunate family.** Some family unfortunate people, in order to make themselves happy, are likely to use opioids, so society must give them care so that they can regain the confidence of life.

- **Partnering with Public Safety.** The opioid overdose epidemic has worsened with a rise in the use of illicit opioids. This fast-moving epidemic does not distinguish between age, sex, or location, and increases in deaths across states indicate the need for better coordination. Many different responders come together to prevent opioid overdoses and deaths, including health departments, law enforcement, and community-based organizations. Improving communication and collaboration between public health and public safety can help identify changes in illicit drug supply and coordinate a more timely and effective response.

References

[1] BBC News. Opioid crisis linked to two-year drop in us life expectancy. https://www.bbc.com/news/world-us-canada-42452733. Accessed December 22, 2017.

[2] Medical Xpress. Costs of us prescription opioid epidemic estimated at $78.5 billion. https://medicalxpress.com/news/2016-09-prescription-opioid-epidemic-billion.html. Accessed September 14, 2016.

[3] Steve Herman. Trump declares public health emergency to fight opioid crisis. https://www.voanews.com/a/trump-declare-public-health-emergency-opioid-crisis/4087269.html. Accessed October 27, 2017.

[4] J. Deng. Modeling of the GM Model of Grey System.Essential Topics on Grey System theory and Application. *China Ocean Oress*, (40-53), 1988.

[5] Wikipedia. Ohio. https://en.wikipedia.org/wiki/Ohio. Accessed December 14, 2018.

[6] Wikipedia. Opioid_epidemic. https://en.wikipedia.org/wiki/Opioid_epidemic. Accessed January 30, 2019.

[7] CDC. Confronting opioids. https://www.cdc.gov/features/confronting-opioids/index.html. Accessed April 16, 2018.

[8] PharmD Beth Han, Wilson M. Compton and Christopher M.jones. Non-medical Prescription Opioid Use and Use Disorders Among Adults Aged 18 Through 64 Years in the United States, 2003-2013. *JAMA*, (40-53), 2015.

Appendices

This is the pie charts of the use of opioid between 2010 and 2017 in the county of Ohio.

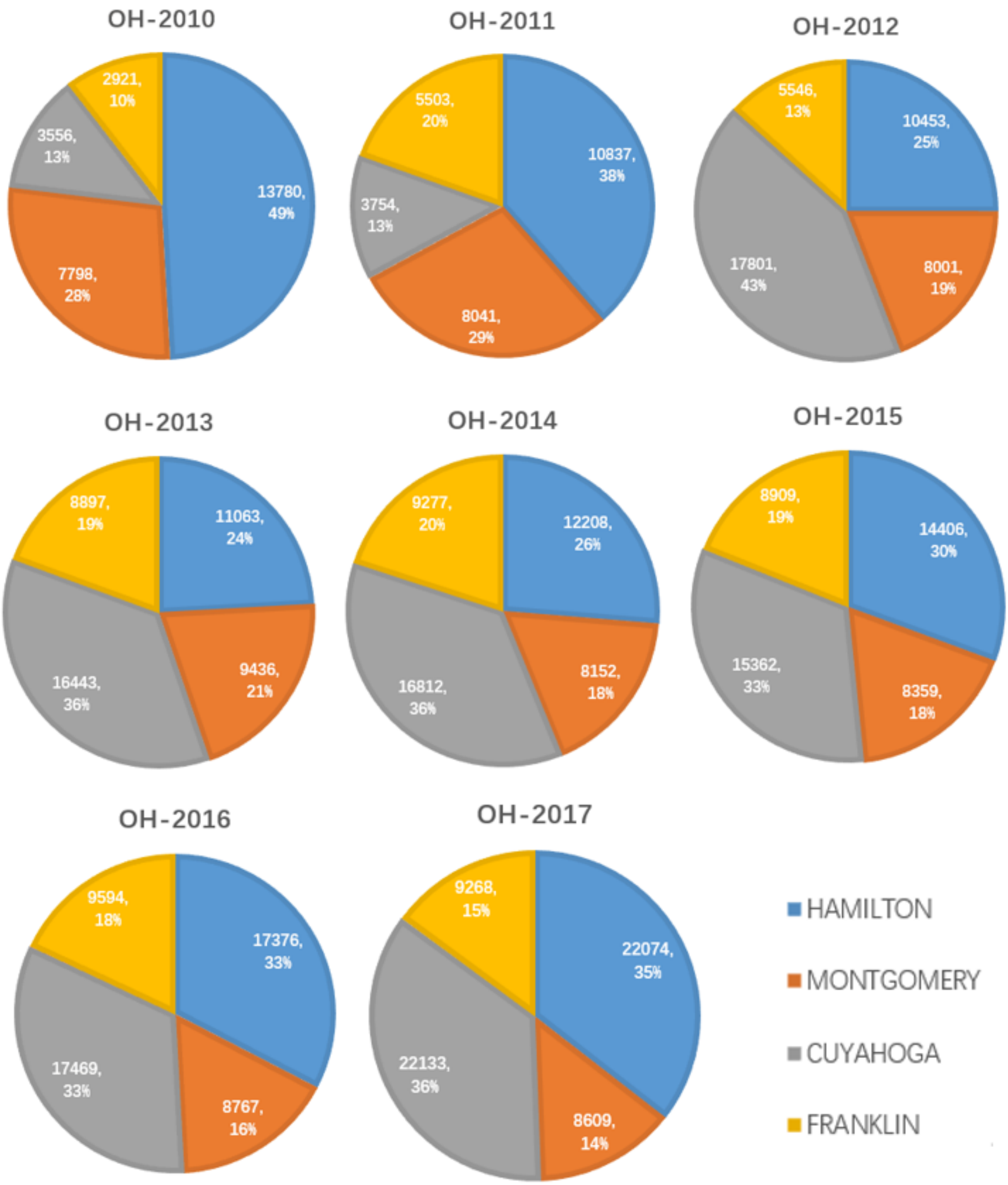

Figure 9: Pie charts of the use of opioid in the counties(OH)

Here's a show of historical statistics for each county of the four states in KY,PA,VA,WV.

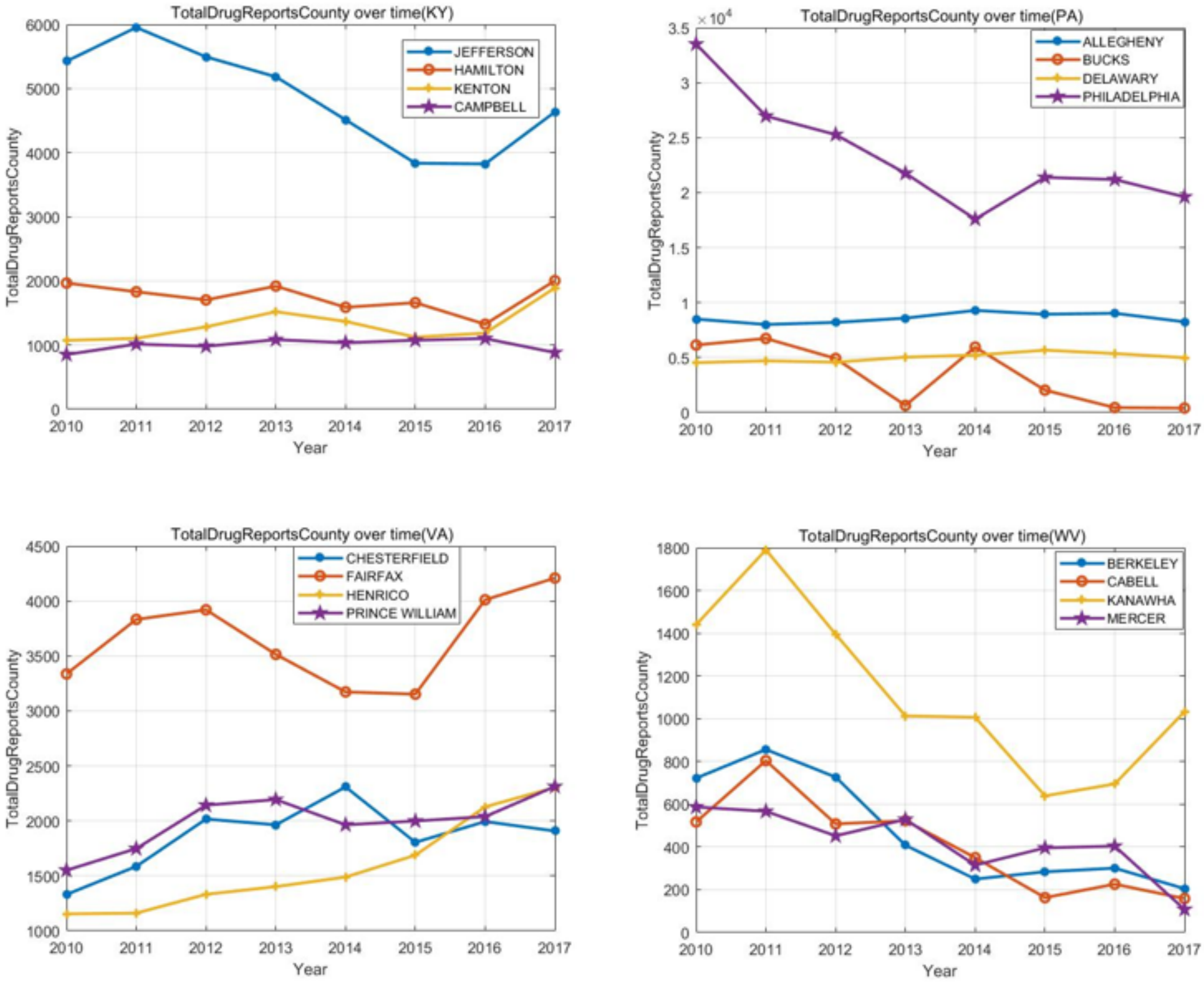

Figure 10: TotalDrugReportsCounty over time(KY,PA,VA,WV)

Here are the predict graphs for these four states.

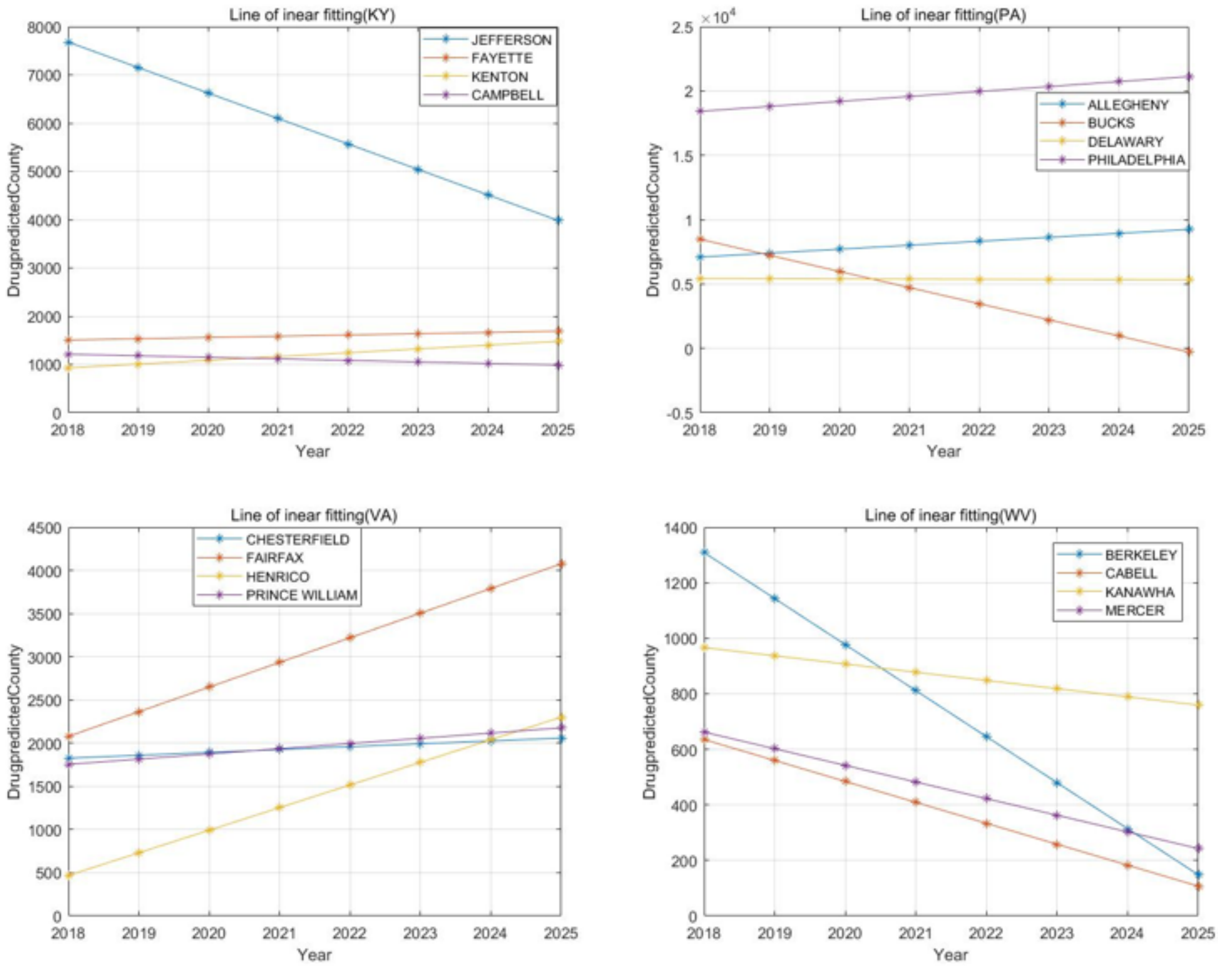

Figure 11: Line of linear fitting(KY,PA,VA,WV)