2019
MCM/ICM
Summary Sheet
(Your team's summary should be included as the first page of your electronic submission.)
Type a summary of your results on this page. Do not include the name of your school, advisor, or team members on this page.

Perhaps most people think drugs crisis is further from them. But now we might think different after we view the data provided by the DEA (the U.S. Drug Enforcement Administration) and the NFLIS (National Forensic Laboratory Information System).

In the first part, we give an introduction about drugs and the opioid addiction crisis now the U.S. is experiencing, especially in the five states, i.e. Kentucky, Ohio, Pennsylvania, Virginia and West Virginia.

In the second part, we focus on the model efficiently describing the characters of spread of drug addiction. Noticing there are similarities between epidemiology and drug addiction, we build a spread model based on the SIS model. To deal with the main difficulty, the inverse problem about differential equation parameters, we apply the trust region method, so that the relative error of parameters is less than 5%. To be more efficient and easier to compute, we develop a spread matrix computed by GA algorithm to evaluate the performance of the trend to spread drug addiction between every two places.

In the third part, we turn to find out which socio-economic factors might affect the spread of drug addiction remarkably. From data provided, we consider three factors that may play an important role on it. We come out a modified drug spread model, test its fitting effect, and give a reasonable social analysis on the relation between drugs addiction and these factors.

Finally, we combine the results from both two models so as to produce a regulation strategy to avoid the expand of the opioid crisis. The result of our model indicates the importance of controlling the flow of opioid addiction from other states, and also proposes a state to improve its recovery rate to decrease the speed of increasing reports number.

Also, we discuss the possible improvement direction of our model and the existing strengths of it.

Contents

MEMORANDUM

To: the Chief Administrator, DEA/NFLIS Database

From: Team#1904783

Subject: A Brief Summary and Strategy for the Opioid Crisis

Date: Monday, January 28, 2019

Our team has proposed a modified epidemic model to analyze the characteristics of the opioid crisis that five states are facing, i.e. Kentucky (KY), Ohio (OH), Pennsylvania (PA), West Virginia (WV) and Virginia (VA). In this model, we take the elements both in the state and out of the state into consideration. Our model is based on the SIS model, people in which model are assumed to be possible to get infective again after the recovery.

More specifically, we develop a spread matrix to evaluate the performance of the trend to spread opioid use between every two places. When computing the value of this matrix, we set up a reasonable assumption that the recovery rate will change in linear distribution over time. Therefore, there are a spread matrix and coefficient matrix of linear distribution to be identified in our training.

After the training, we get the robust results and successfully fit the data of five states, which suggest the important information about the flow of opioid use in and between these five states. The rate of contacting into State KY is significantly higher than any other state, followed by that into State OH. This can explain the large number of drug reports in these two states, as well as predict the possible location and time that opioid crisis might have occurred in these five states.

We also develop our modified model to fit the influence that socio-economy factors may have on the trends-in-use. We identify three main important factors, i.e. educational degree, marital condition and disability as correction factors to further adjust our model. Afterwards, our model successfully improves the fitting effect of the predictions.

Combining with both two results from our model, we successfully produce a regulation strategy to avoid the expand of the opioid crisis. The result of our model indicates the importance of controlling the flow of opioid use from other states, and also proposes a state to improve its recovery rate to decrease the speed of increasing reports number.

We have a strong belief that our model can effectively solve the opioid crisis no matter in what places and provide an effective and feasible way to best solve this problem.

1 Introduction

1.1 Problem Background

The United States is experiencing an epidemic of drug overdose deaths. According to Centers for Disease Control (CDC), drug overdoses have dramatically increased over the last two decades, with deaths more than tripling between 1999 and 2016[1]. In 2017, more than 72,000 Americans died from a drug overdose. At 197 people each day, this is more than the number of lives lost in car accidents or gun-related homicides, which is declared by U.S. DEPARTMENT OF AGRICULTURE[2]. An overwhelming majority of these overdose deaths involved opioids.

Opioids are a type of narcotic pain medication. And morphine, which is well known, is the most widely used opioid analgesic. In addition, opioids include fentanyl, codeine, methadone, etc. In addition to analgesia, opioids can also be used for cough, diarrhea, and sedation. In 2000, WHO proposed: "In all analgesic treatments, opioid analgesics are essential for the treatment of cancer pain. Especially for patients with moderate and severe cancer pain, opioid analgesics have an irreplaceable position."

However, they may cause serious side effects if used incorrectly. Using opioids frequently may cause both mental and physical dependence; once people are addicted, it would be difficult to treat. For people who have an opioid addiction, their problem often starts with a prescription. They compulsively seek out the pain medications, which usually leads to negative consequences in their personal lives or workplace. They might take someone else's pills or buy them off the street, which is especially dangerous since those drugs are often laced with lethal amounts of fentanyl. In addition, recent studies have found that opioids may also cause other health problems, such as an increased risk of hepatitis C and HIV infection with the use of injectable opioids for private use.

In the United States, opioid abuse has been around for a long time. As early as the 1980s, some scholars believed that opioids were not severely addictive, while most opioid manufacturers in U.S. also aggressively promoted their products, promising that opioids were not addictive. This has led to a surge in opioids. Furthermore, the huge opioid market in the United States may also be the motivation for many pharmaceutical companies to take risks. From the perspective of medical insurance expenses, addiction treatment, criminal justice investigation, etc., the United States suffers economic losses of up to 80 billion US dollars due to the abuse of opioids every year.

The United States has set up a drug control bureau, the U.S. Drug Enforcement Administration (DEA), to set corresponding production quotas for con-

---

[1]See from Centers for Disease Control website, (https://www.cdc.gov/features/confronting-opioids/index.html), accessed 25 January 2019.

[2]See from U.S. Department of Agriculture website, (https://www.usda.gov/topics/opioids), accessed 25 January 2019.

trolled substances such as opioids to meet the practical requirements of medical treatment, scientific research, export, and storage, etc. In August 2017, the United States officially listed opioids as "national emergency".

Figure **1** shows the number and age-adjusted rates of drug overdose by state in 2016 in the U.S. From the perspective of the problem requested, particularly, the five targeted states surrounded by blue contour lie in the district which suffered the drug overdose most severely, among which, comparatively speaking, Virginia was damaged the most slightly, from 13.6 to 16.0.

1.2 Literature Review

Researches about the opioid use are abundant, varying from the mathematical model establishment to the evaluation of social consequence, and also from the association with other diseases to the strengths for treatment. Paper [1] validates the feasibility of the epidemiologic method to the spread of heroin use, with particular emphasis on the analogy between heroin use and communicable disease. This inspires us to transfer the epidemic model for this problem.

Rather than use the epicdemic model, the authors [3] designed Hierarchical Bayesian Poisson space–time analysis to relate the prescription opioid (PO) poisoning spread with spatial contagion and other social properties. In addition, the distribution of age-related and work-place-related sources of medical need for POs in rural area can partly explain the growth of PO poisoning.
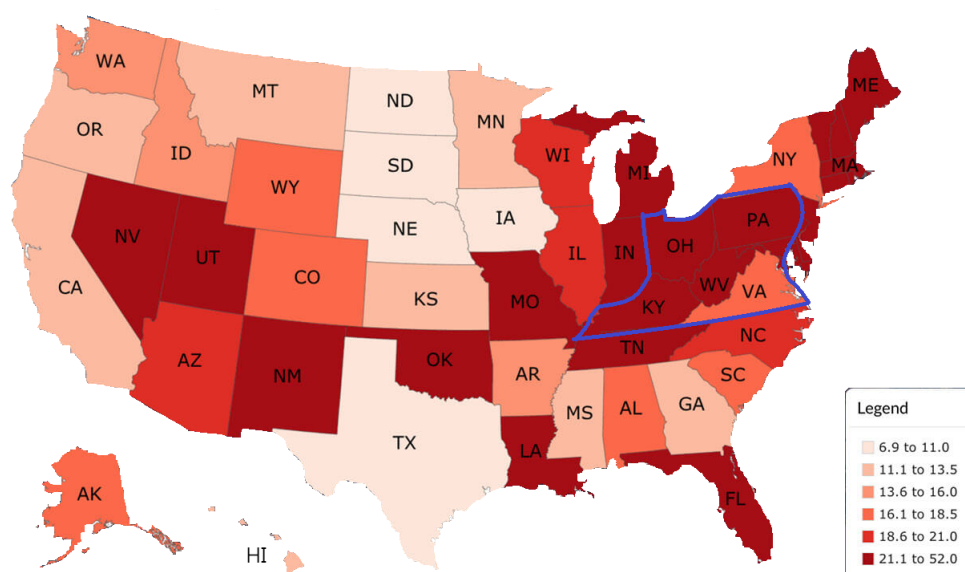

Fig. 1. Number and age-adjusted rates of drug overdose by state, US 2016.

From the pespective of consequence caused by it, authors of paper [2] investigated the association of maximum prescribed daily opioid dose and dosing schedule with risk of opioid overdose death. They suggested that among patients receiving opioid prescriptions for pain, higher opioid doses were more associated with increased risk of opioid overdose death.

Considering the kinds of epidemic models, the author [4] surveyed previous models for the spread of infectious diseases in populations with age structure, heterogeneity, and spatial structure. Similar results with new expressions for the basic reproduction number R_0 are obtained in his survey.

1.3   Our work

In this paper, we make a quantitative and qualitative analysis on the reported synthetic opioid and heroin incidents between these five states and their counties over time and identify some strategies for countering the opioid crisis.

We establish an SIS model to stimulate the spread of the drug incidents between and inside five states. This model derives from the epidemic model, which fits the spread characteristics of the opioid use.

Furthermore, we develop the solution to the inverse problems about differential equation parameters of the SIS model with the Genetic Algorithm **GA**, for the difficulty solving enormous dimension of the parameter matrix. This model enhances the robustness and the convergence of solution, yet at the cost of the accuracy and computation capacity.

As for the first problem, we need to propose a mathematical model to fit the drug reports of five states, and to identify any possible places that might have started the opioid trends now. To solve these problems , we predict the trend of drug reports (take fentanyl as an example) by our model and divide the state into several blocks to narrow up the target places.

As for the second problem, we need to identify any possible factors from given socio-economy data that are associated with the opioids overdosing, as well as the specific degree of this relation. So we improve our model, considering these factors in it. And then, our model fits the real data better than before, with train error varying from 0.077 to 0.045.

As for the third problem, we need to produce a strategy that can solve the opioids crisis and identify its effectiveness by our model as well as the optimal boundaries of parameters. We analyze the parameter characteristics of our model, and then find the important impact of the flow of opioid use from other places on a state. We propose to pay more emphsis on it. Finally, we analyze the sensitivity of our model.

2 Preparation of the Models

2.1 Assumptions

According to the paper [6], we set up the following assumptions for the SIS model.

A. The population N is constant in a fixed place, including states and counties. People are divided into two groups: Susceptible group and Infective one.

B. The infectious ratio in and between two places (i, j) is considered as, take $j \to i$ as an example,

$$\frac{a_{ij} I_j}{N_j}$$

where a_{ij} is the weight that Place j affects Place i, I_j is the infective number in Place j.

C. Define the average effective contact number of infective members in Place i as λ_i,

$$\lambda_i = \sum_{j=1}^{n} \frac{a_{ij} I_j}{N_j}$$

where n is the number of places that have impact on Place i.

D. Any state (or county) we know only affects those states (or counties) that are no more than 1 state (or county) away from it, no matter how small they are.

E. The number of reports for both each possible factor and each sub-division of the factor is linearly distributed over time. This hypothesis is obtained by observing the characteristics of the data set.

F. For the sake of simplicity, and considering the macro regulation of the government policies, we assume that the recovery rate in each place varies linearly over time.

G. We assume that it's reasonable to narrow down the range of parameters based on prior knowledge or experience, such as each state's self-impact is greater than that of other states on itself. This is because, compared with the number of parameters of the model, the size of given dataset samples is too small, which would result in an unsatisfactory fit in the absence of specific conditions.

H. The number of drug reports in a state will not be affected by the economic conditions, marital status and other social factors in other states, but only affected by those factors in the state itself. This is an naturally acceptable hypothesis.

2.2　Notations

The primary notations used in this paper are listed in **Table 1**.

Table 1: Notations

Symbol	Definition
a_{ij}	Effective contact ratio per patients in Place j with people in Place i
A	$\{a_{ij}\}_{n \times n}$
I_i	Drug identification (Infective) number in Place i, $i = 1, 2, 3, \cdots, n$
I	$\begin{pmatrix} I_1 & I_2 & \cdots & I_n \end{pmatrix}^T$
n_c	Number of given counties
n_s	Number of given states
N_i	Total population in the Place i
I/N	Infective ratio in places at the same level: $\begin{pmatrix} \frac{I_1}{N_1} & \frac{I_2}{N_2} & \cdots & \frac{I_n}{N_n} \end{pmatrix}^T$
S_i	Susceptible number in Place i
S	$\begin{pmatrix} S_1 & S_2 & \cdots & S_n \end{pmatrix}^T$
γ_0^i	The constant coefficient of recovery rate of Place i
γ_1^i	The linear coefficient of recovery rate of Place i
γ_t^i	The recovery rate of Place i in the year t, $\gamma_t^i = \gamma_0^i + \gamma_1^i t$
Γ	$\begin{pmatrix} \gamma_0^1 & \gamma_0^2 & \cdots & \gamma_0^n \\ \gamma_1^1 & \gamma_1^2 & \cdots & \gamma_1^n \end{pmatrix}^T$
λ_i	Average effective contact number of infective members in Place i
σ_i	Average effective contact number of patients in Place i in infective period
KY	Kentucky
OH	Ohio
PA	Pennsylvania
WV	West Virginia
VA	Virginia

Notes: If the parameter is superscripted with c, it means the place is at the level of counties; if it is superscripted with s, it means the place is at the level of states. Otherwise, it is considered to be at a comprehensive level.

3　The Models

3.1　Model 1: SIS Model

3.1.1　Detail about Model 1

Due to the spread of the synthetic opioids and heroine incidents among states, we take the factor of impacts from other locations into consideration when setting

the equation of the SIS model. Paper [4] has indicated the feasibility of applying epidemic model to the spread of heroine utility; therefore, we determine to apply SIS model which fits the practical situation of drug spread.

In specific, we can get the following differential equation array:

$$
\begin{aligned}
\frac{\mathrm{d}S_i}{\mathrm{d}t} &= -\sum_{j=1}^{n} \frac{a_{ij}I_j}{N_j} S_i + \gamma_i I_i \\
\frac{\mathrm{d}I_i}{\mathrm{d}t} &= -\frac{\mathrm{d}S_i}{\mathrm{d}t}
\end{aligned}
\qquad (i = 1, 2, \cdots, n)
\tag{1}
$$

We first experiment the traditional approach to solving the differential equations. That is, we apply the trust region method [5] for the inverse problem about differential equation parameters. According to the algorithm introduced in paper [5], we transform the problem into a minimization problem, the objective function of which is as following:

$$
f\left(A, \gamma_1, \gamma_2, \cdots, \gamma_n\right) = \frac{1}{ns} \sum_{i=1}^{n} \sum_{t=0}^{s} \left[\frac{\widehat{I}_{i,t}\left(A, \gamma_1, \gamma_2, \cdots, \gamma_n\right) - I_{i,t}}{I_{i,t}}\right]^2
\tag{2}
$$

where $A, \gamma_1, \gamma_2, \cdots, \gamma_m$ are parameters, $\widehat{I}_{i,t}\left(A, \gamma_1, \gamma_2, \cdots, \gamma_m\right)$ is the corresponding predicted value of I_i in the year t ($t = 0, 1, \cdots, s$) after $A, \gamma_1, \gamma_2, \cdots, \gamma_m$ are obtained, and $I_{i,t}$ is the practical value I_i in the year t ($t = 0, 1, \cdots, s$).

$$
\begin{aligned}
\min \ &\varphi_i\left(a_{i1}, a_{i2}, \cdots, a_{in}, \gamma_i\right) \\
=&f_i\left(\left(a_{i1}^{(k)}, a_{i2}^{(k)}, \cdots, a_{in}^{(k)}, \gamma_i^{(k)}\right)\right) + \\
&\nabla f_i\left(\left(a_{i1}^{(k)}, a_{i2}^{(k)}, \cdots, a_{in}^{(k)}, \gamma_i^{(k)}\right)\right)^T \left(a_{i1}, a_{i2}, \cdots, a_{in}, \gamma_i\right) + \\
&\frac{1}{2}\left(a_{i1}, a_{i2}, \cdots, a_{in}, \gamma_i\right)^T \nabla^2 f_i\left(\left(a_{i1}^{(k)}, a_{i2}^{(k)}, \cdots, a_{in}^{(k)}, \gamma_i^{(k)}\right)\right)\left(a_{i1}, a_{i2}, \cdots, a_{in}, \gamma_i\right)
\end{aligned}
\tag{3}
$$

where $\|a_{i1}, a_{i2}, \cdots, a_{in}, \gamma_i\| \leqslant r_k, i = 1, 2, \cdots, n$.

To program this algorithm, we compute the $Hessian$ matrix of points in the trust region. And then substitute the $Hessian$ matrix, along with the gradient matrix, into equation (3) to find out the point corresponding to the minimum value, which should be the optimal solution currently. Thirdly, re-divide the trust region according to different conditions of the optimal solution, and iterate above steps until the gradient of the point is lower than the threshold so that we obtain the final optimal solution.

However, the weakness of this algorithm is the high complexity for computation. As is known, the $Hessian$ matrix is at the dimension of $n^2 \times n^2$ for n parameters, which will cause a heavy burden to the computation with the increase of n. What's more, both the gradient matrix and the $Hessian$ matrix are time-consuming to conduct, thus being impossible to for us to solve the inverse problems at the level of counties particularly.

Therefore, we decide to develop the approach to solving this model in an attempt to decrease the complexity when facing the enormous data to be computed. Specifically, we apply the GA model to search the solution to parameters.

For the part 1 problem, we choose the following function as the objective function.

$$\min \ f\left(A^s, \gamma_1^s, \gamma_2^s, \cdots, \gamma_{n_s}^s\right) = \frac{1}{n_s s} \sum_{i=1}^{n_s} \sum_{t=0}^{s} \left| \frac{\widehat{I}_{i,t}^s \left(A^s, \gamma_1^s, \gamma_2^s, \cdots, \gamma_{n_s}^s\right) - I_{i,t}^s}{I_{i,t}^s} \right| \quad (4)$$

This is one at the level of states, which, with the help of the GA model, can offer us the convergent solution to parameters. The result of the parameter matrix is given below.

$$A = \begin{bmatrix} 0.50713517 & 0.02263958 & 0.07114156 & 0.40748377 & 0.24293688 \\ 0.00092936 & 0.64209954 & 0.16669959 & 0.17898226 & 0.14229662 \\ 0.00250955 & 0.24609878 & 0.50347999 & 0.02277512 & 0.20417946 \\ 0.01678149 & 0.0914572 & 0.01279008 & 0.68407223 & 0.06001597 \\ 0.00268578 & 0 & 0.01794575 & 0.00094326 & 0.78687606 \end{bmatrix} \quad (5)$$

$$\Gamma = \begin{bmatrix} 0.86816492 & 0.99901323 & 0.9356767 & 0.96664902 & 0.6449881 \\ 0.02645234 & -0.05330094 & 0.07823487 & 0.0726793 & 0.06921307 \end{bmatrix} \quad (6)$$

where index i refers: $1 \to KY, 2 \to OH, 3 \to PA, 4 \to VA, 5 \to WV$. We transfer the spread characteristics into the map (See Fig. 2) so that we can have a clear view on it.

We implement the modeled parameters into the equation 1 or **??**, getting the predicted situation of specific drug reports in five states respectively. The comparation between the real reports and predicted reports is showed in the Fig. 3. We can find that our model fits the real situation well. The training error is 0.077 for our model.

Analysis From the perspective of the epidemic model, namely SIS model, we can get the relation between the σ_i and λ_i, γ_i. That is $\sigma_i = \lambda_i / \gamma_i$. Then from the article [6], we can learn the different change trend in the different extent of σ. To be specific, the trend goes like the Fig. 4. We can get the conclusion that the trend in five states can match the part of the curves in Fig. 4 respectively. For example, the predicted line of OH corresponds to the situation that $\frac{I(0)}{N} < 1 - \frac{1}{\sigma}$ and $\sigma > 1$, while that of WV is more likely to match the situation that $\sigma \leqslant 1$. This suggest that our model can fit the synthetic opioids crisis well in these five states.

Next, we propose a further study on a specific opioid: fentanyl. Fentanyl has been an issue that can not be ignored in recent years. The number of fentanyl drug reports from 2010 to 2017 is projected on a county-scale map as shown in Fig. 5.

The figure indicates that state OH is suffering from a severe fentanyl crisis. Therefore it is necessary to model the fentanyl spread and characteristics of OH,
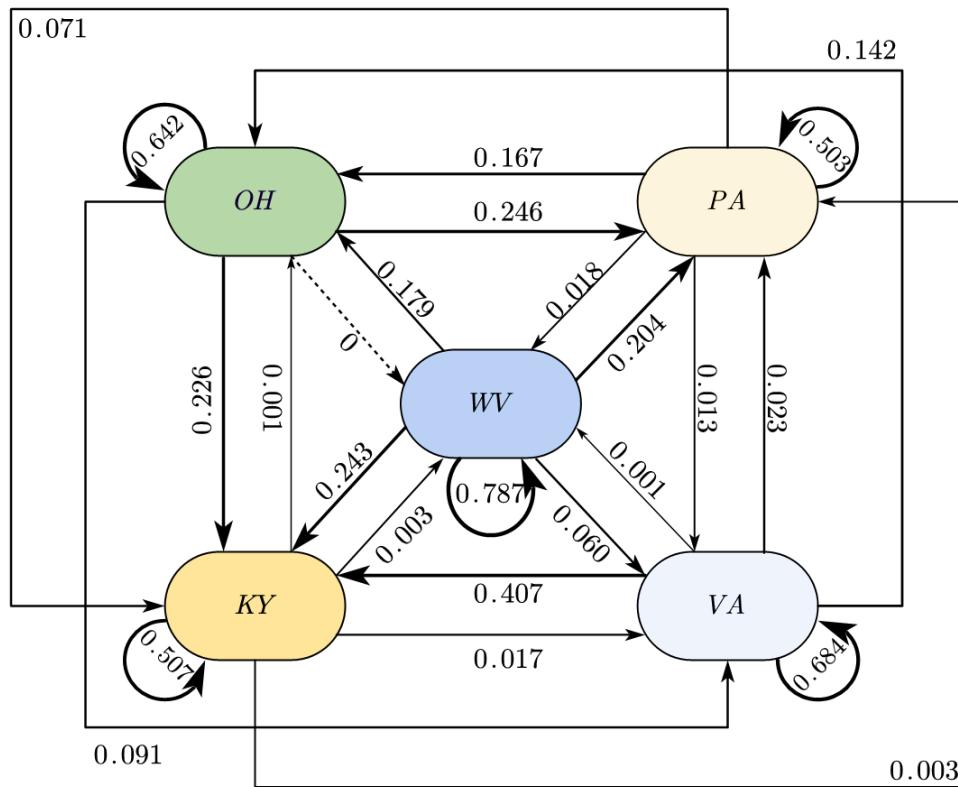
Fig. 2. Map of five states and Weight of Spread Lines.

and then find the proper drug identification threshold levels to contain the crisis. However, fitting all the counties of OH separately into the model is not feasible as a huge amount of computational resources will be needed. Therefore, we choose to divide the whole OH state into six large blocks and put them into the model for prediction. The way of dividing is shown in Fig. 6.

After obtaining the model parameters between the six blocks, we predict their trend in the next few years. According to result showed in the Fig. 7 with blocks 3 and 4 rising at the fastest rate, indicating that the two regions are likely to explode in the next few years.

If the patterns and characteristics continue, i.e. the γ_0, γ_1, A parameters of the model remain unchanged, the number of fentanyl drug reports in the 3rd and 4th block is expected to have a sharp increase at around 2019 (Fig. 6). The U.S. government should draw enough attention to prevent the incident from happening.

To identify the location and the time that might have started in these five states, we hold the view that where there is a growing trend of specific drug reports in the next few years, there might have been the start of this drug reports. And regarding to the threshold of the occurrence, we think that the higher the recovery rate in a state is, the more possible the drug reports cases will be to decrease in the next few years. As a result, we consider the threshold of the recovery rate, namely γ_i in each state, as the threshold of the occurrence, which means searching the maximum value of γ_i that makes the reports number increase in the
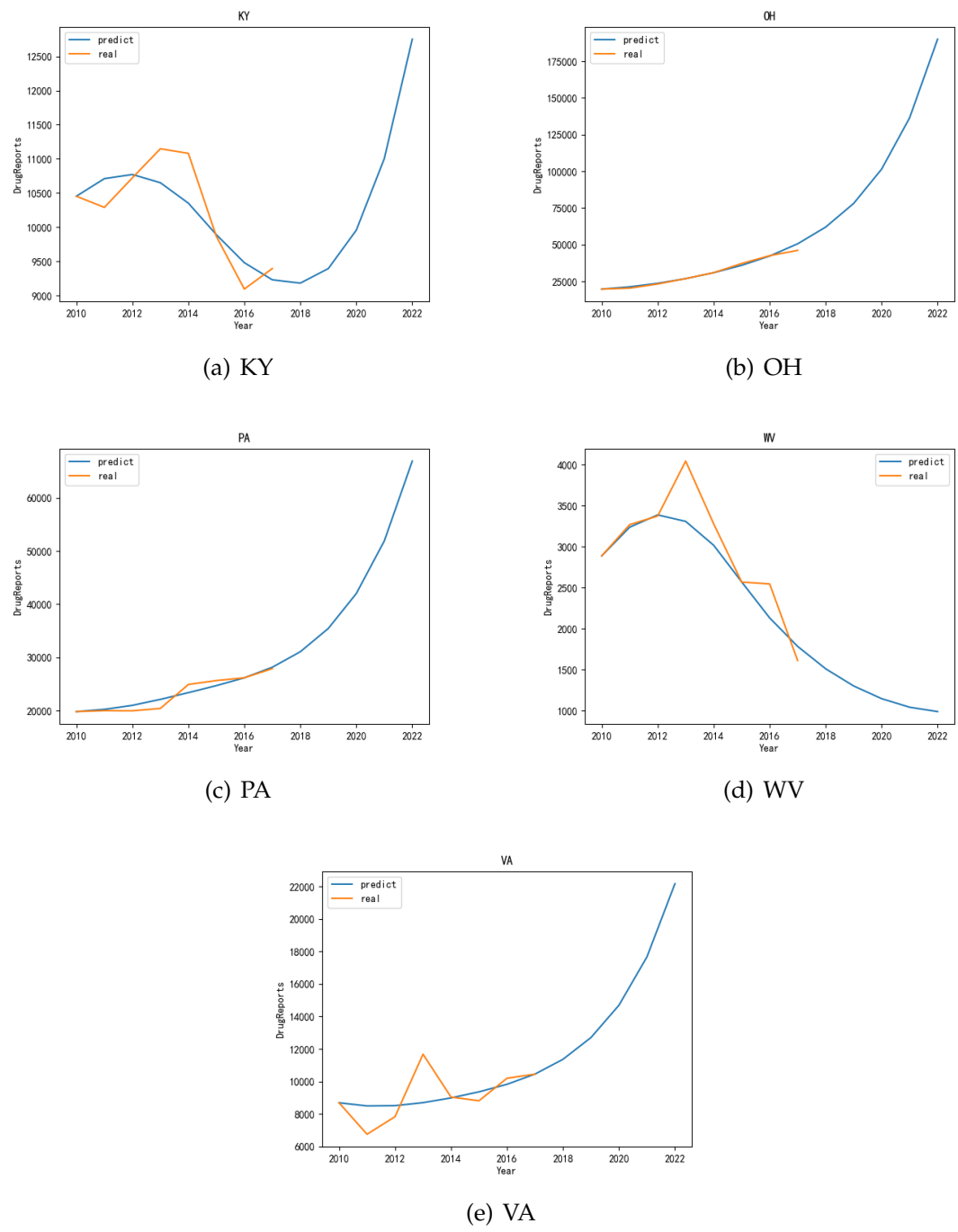
(a) KY

(b) OH

(c) PA

(d) WV

(e) VA

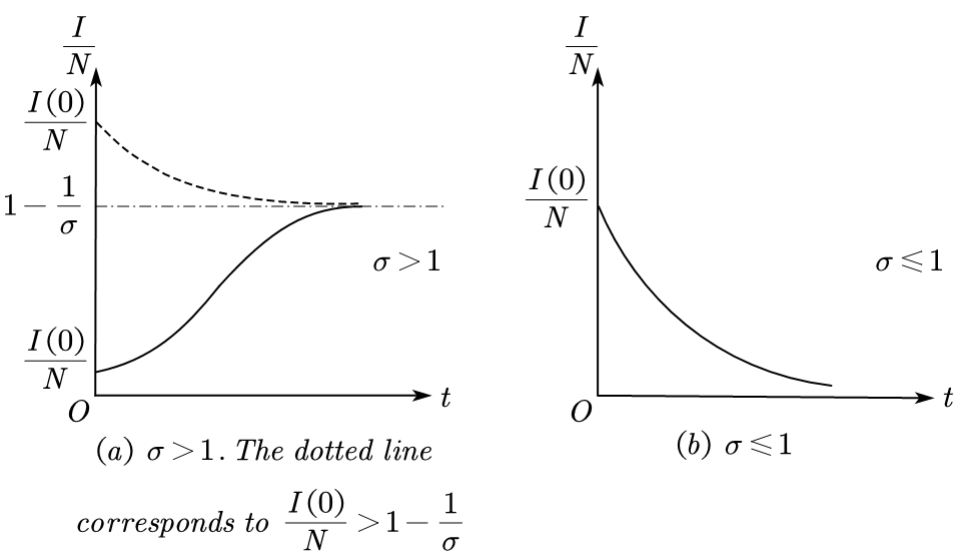**Fig. 3.** The drug reports in five states.

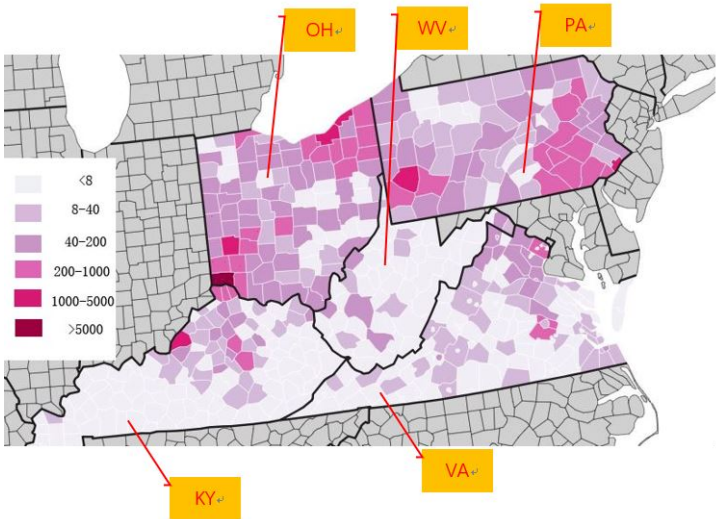**Fig. 4.** $\frac{I}{N}$-t figures in SIS.

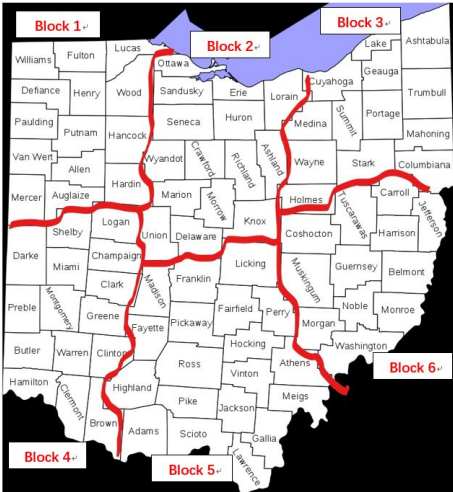**Fig. 5.** Fentyl reports number of all the counties in five states.

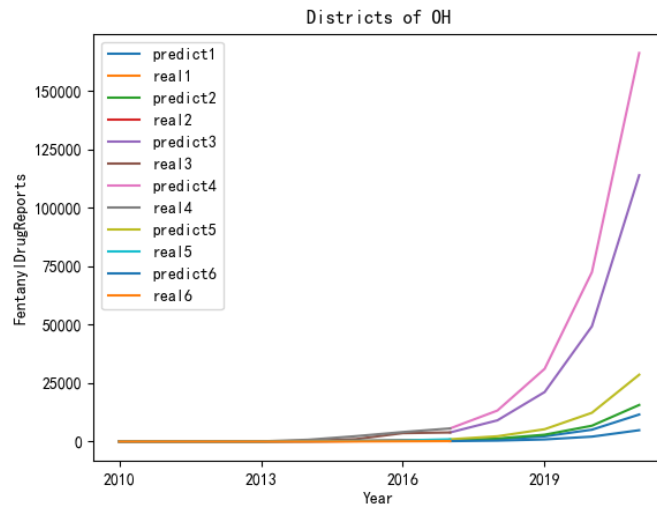**Fig. 6.** The way of dividing State OH into six parts.

Fig. 7. The trend of fentanyl reports of six parts.

next few years.

Fig. 8 represents 12 predictions with different drug identification threshhold levels taken after 2017. The threshhold level is represented by the γ_0, γ_1 parameters of the model. Hard thresshholds can be decribed by large value of γ_0, γ_1 while soft thresshholds can be described by small ones. When the drug identification threshhold is softer than that of $(\gamma_0, \gamma_1) = (0.9, 0)$ or $(0.8, 0.05)$ or $(0.7, 0.1)$, the incident is likely to occur at around 2021, in the 3rd and 4th block of State OH.
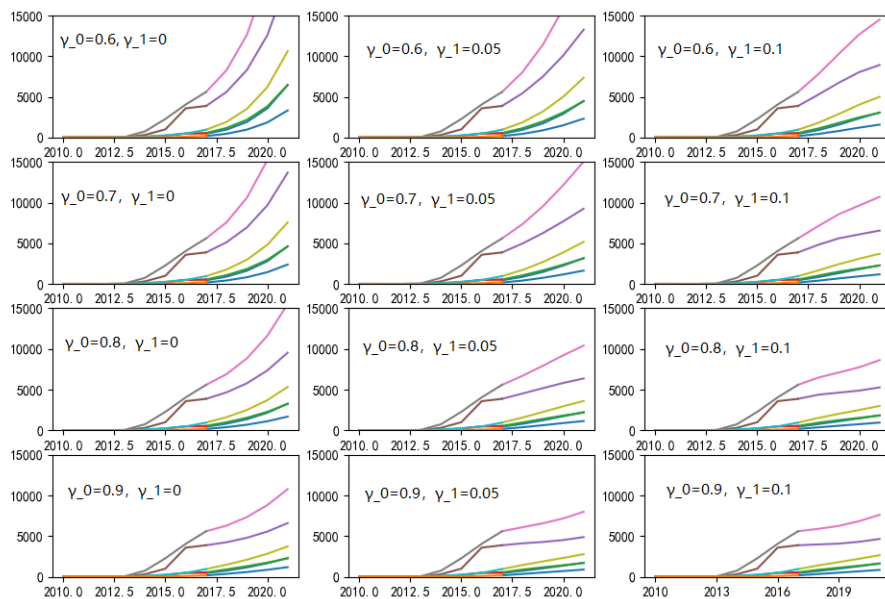

Fig. 8. The trend of fentanyl reports of six parts.

3.2   Model 2: Developed SIS Model

3.2.1   Detail about Model 2

Concerning the good number of factors that we do not take into consider but do have contribution to the drug report trend, we choose to improve our model. That is to add the weight of those possible factors into the model. By training the parameter matrix, we can finally research the correlation between the factors and the drug reports in each states. To be specific, we can get the following differential function array:

$$
\begin{aligned}
\frac{\mathrm{d}S_i}{\mathrm{d}t} &= -\left(\sum_{j=1}^{n}\frac{a_{ij}I_j}{N_j}\right)\exp\left(\frac{1}{\sum_{\iota=1}^{L}K_\iota}\sum_{\iota=1}^{L}\sum_{\kappa=1}^{K_\iota}\frac{\widehat{\beta_0^{\kappa\iota}}+\widehat{\beta_1^{\kappa\iota}}t}{\widehat{\beta_0^{\iota}}+\widehat{\beta_1^{\iota}}t}\rho_{i\kappa\iota}\right)S_i+\left(\gamma_0^i+\gamma_1^i t\right)I_i \\
\frac{\mathrm{d}I_i}{\mathrm{d}t} &= -\frac{\mathrm{d}S_i}{\mathrm{d}t}
\end{aligned}
\tag{7}
$$

where the parameter:

- L is the number of the factors selected from given data;

- K_ι is the number of the subdivision of factor ι;

- $\widehat{\beta_0^{\kappa\iota}}$ is the constant coefficient of the linear regression of the kth sub-factor of factor ι.

- $\widehat{\beta_1^{\kappa\iota}}$ is the linear coefficient of the linear regression of the kth sub-factor of factor ι.

- $\frac{\widehat{\beta_0^{\kappa\iota}}+\widehat{\beta_1^{\kappa\iota}}t}{\widehat{\beta_0^{\iota}}+\widehat{\beta_1^{\iota}}t}$ is the ratio of the fitted sub-division portion of factor ι.

- ρ is the weight of the correlation between the factor and the local drug reports number.

In this model, we establish a statistic in the form of

$$
\exp\left(\frac{1}{\sum_{\iota=1}^{L}K_\iota}\sum_{\iota=1}^{L}\sum_{\kappa=1}^{K_\iota}\frac{\widehat{\beta_0^{\kappa\iota}}+\widehat{\beta_1^{\kappa\iota}}t}{\widehat{\beta_0^{\iota}}+\widehat{\beta_1^{\iota}}t}\rho_{i\kappa\iota}\right)
$$

to evaluate the performance of factors in improving the model. Firstly, utilizing the exponential function is in an attempt to magnify those significant drug reports, so that it can perform well in evaluating the degree of significance. On the other hand, we sum up the weight value of all possible factors (here we select three elements- education degree, disability, marital condition) for the fact that this method fits the dataset better than other models we have tried and have a good performance in improving the training error of our model.

Our objective function is the same as the equation 2. In this train procedure, we substitute the weight parameter A and the coefficient matrix of recovery rate Γ obtained from the model 1 into the equation 7. The parameter to be optimized is $\rho_{i\kappa\iota}$.

By training the processed data, then we get the result of the parameter matrix. It is showed in the Table. 9. And the train error reduces to 0.045, which performs better than model 1 (0.077). And the predicted curves of five states are showed in the Fig. 10

3.2.2 Analysis

From the table 9, we can get several conclusions:

$\iota = 1$ **(Marital Condition):**   The poor marital condition has more correlation with the trends-in-use of opioids. In particular, those who live separated or widowed are more likely to use opioids in their life. This is mainly due to their seeking for things that can relieve them from their frustration in marriage.

$\iota = 2$ **(Educational Degree):**   The high educational degree can effectively decrease the risk of using opioids, which can be obtained by its significantly negative value.

$\iota = 3$ **(Disable or not):**   Except from the value of WV, we can find that disability is also an important factor that has an impact on the trends-in-use of opioids. Some prescription drugs should be responsible for this factor because many of them like antalgic contain synthetic opioids that can cause patients to be tolerant with opioids if they take them over a long time. On the other hand, there is no significant relation between a health body and a low use of opioids.

Finally, combining the results from both model 1 and model 2, we can get the conclusion that the combination of the spread of use from other states and the socio-economy condition in the state when modelling the opioid incidents can represent the main characteristic of the progress of the opioid cases in and between these five states. As for the external factors, people across the borders, along with the potentially large requirements for prescription drugs from other

1	ι=1					ι=2							ι=3	
i	i=1	i=2	i=3	i=4	i=5	i=1	i=2	i=3	i=4	i=5	i=6	i=7	i=1	i=2
KA	26.639	-7.739	20.402	48.708	22.855	2.118	-0.928	0.206	6.299	-2.436	-48.168	-2.709	15.164	-8.955
OH	7.946	-21.229	43.324	15.009	0.420	2.674	-1.779	2.125	17.558	-17.470	-0.010	-37.695	18.757	4.918
PA	42.000	-10.011	41.445	15.780	17.462	10.342	-6.463	10.660	-18.151	-33.586	-8.287	-27.127	1.271	-7.761
VA	26.313	-2.406	10.282	32.274	35.017	0.368	-19.050	-5.202	-8.614	-26.119	-9.208	-30.335	11.725	-1.397
WV	26.136	-9.127	44.816	48.700	2.276	-0.249	15.508	9.530	-17.284	-49.652	-38.632	-34.088	-7.490	1.962

Fig. 9. The matrix of $\rho_{i\kappa\iota}$, $\iota = 1$ corresponds to marital conition, $\iota = 2$ corresponds to educational degree,$\iota = 3$ corresponds to disable or not. As for the meanings of index i in each situation, see Table 12.

(a) KY

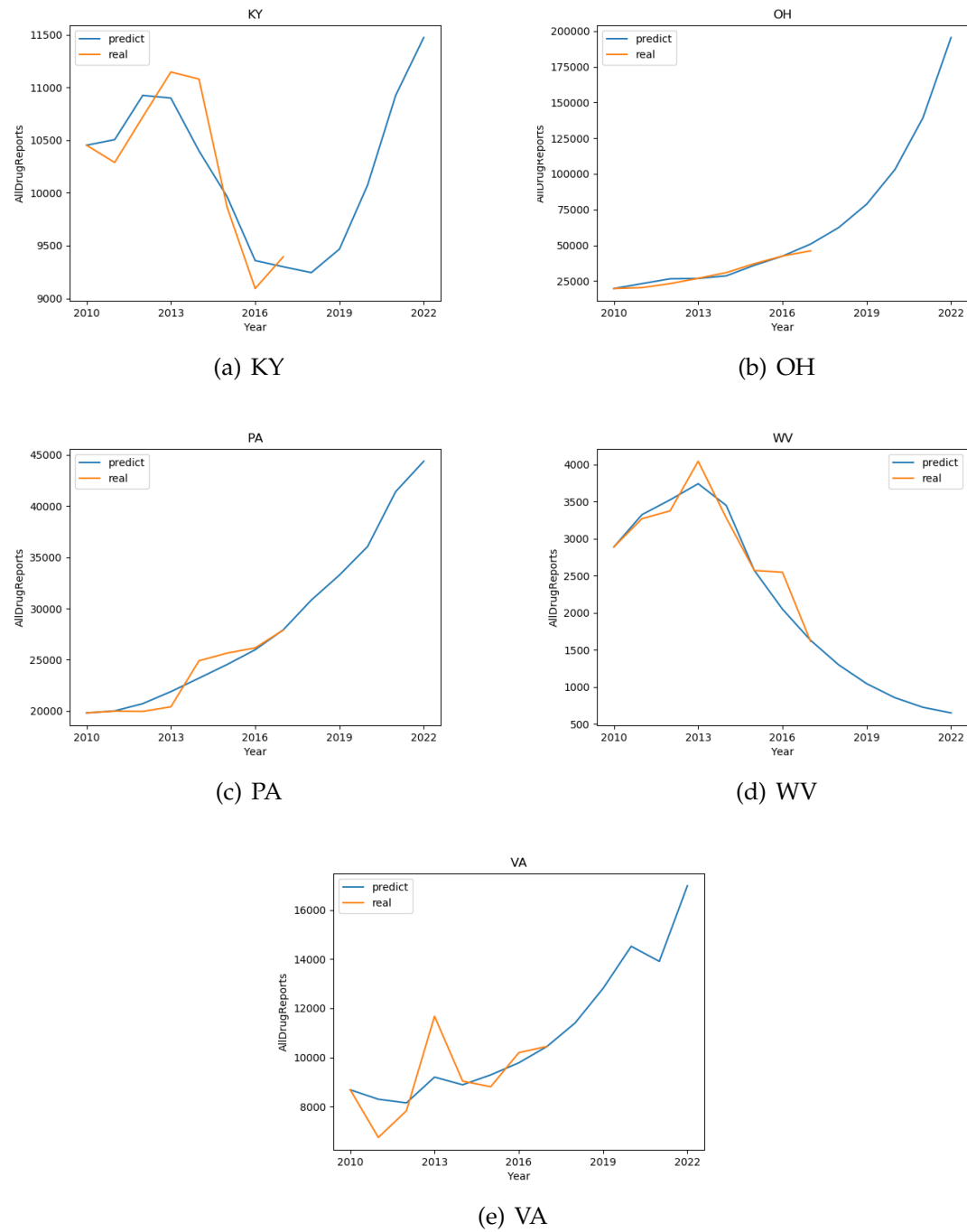(b) OH

(c) PA

(d) WV

(e) VA

Fig. 10. The drug reports with modified model in five states.

places, are more associated with the spread of the opioids use, while as for the internal factors, besides the high possibility to access opioids users for the dense flow, many macro regulations are key to the future of such trends-in-use; for example, the development of local education to acknowledge the damage of such crisis, the access level to health care, the happiness level of local people and so on. Such statistical results indicate different problems that five states are facing according to our model results.

3.2.3   Strategy for Crisis

According to the results we predict, if we should devise a strategy that countering the opioid crisis, we would suggest that the spread of the opioids uses should be the main concern for these five states, because when we consider the internal factors, we find that the recovery rate of each state is commonly high, which means a coming decrease if the governments control the flow of opioid across the borders. In fact, the gamma coefficients of other states have increased with time, and only the gamma coefficient of OH state has become smaller and smaller with time. This may be because fentanyl has been popular in OH in recent years, and OH has not had time to do good countermeasures.

Therefore, it is imperative that OH State take effective measures against the fentanyl crisis and increase its own gamma coefficient to above 0.75. Other states should ensure that their gamma coefficient does not decrease. On the other hand, the government should restrict the spread of drugs between states, reducing the non-diagonal elements of the A matrix to less than 80% of the original value.The predicted results after the strategy implementation are shown in Fig. 11.

The results suggest the effectiveness of our strategy. The increase rate of opioid report in State OH plunges to nearly zero in the coming years after the improvement of its recovery rate. And for other states, after we reduce the portion of opioid use from other states, for the high recovery rate, their drug reports plunge into a micro number comparative to their past data.

4   Conclusions and Discussion

Conclusions: This paper manages to develop a fully data-based model to produce a regulation strategy for countering the opioid crisis. Our model exhibits a great potential in drawing the following conclusions:

- We formulate a performance matrix for drug spread in and between five given states with the GA algorithm and improve the classic SIS model by considering the spread correlation of infection with other places and the dynamic characteristics of recovery rate.

- We identify three main performance contributing variables that generate a
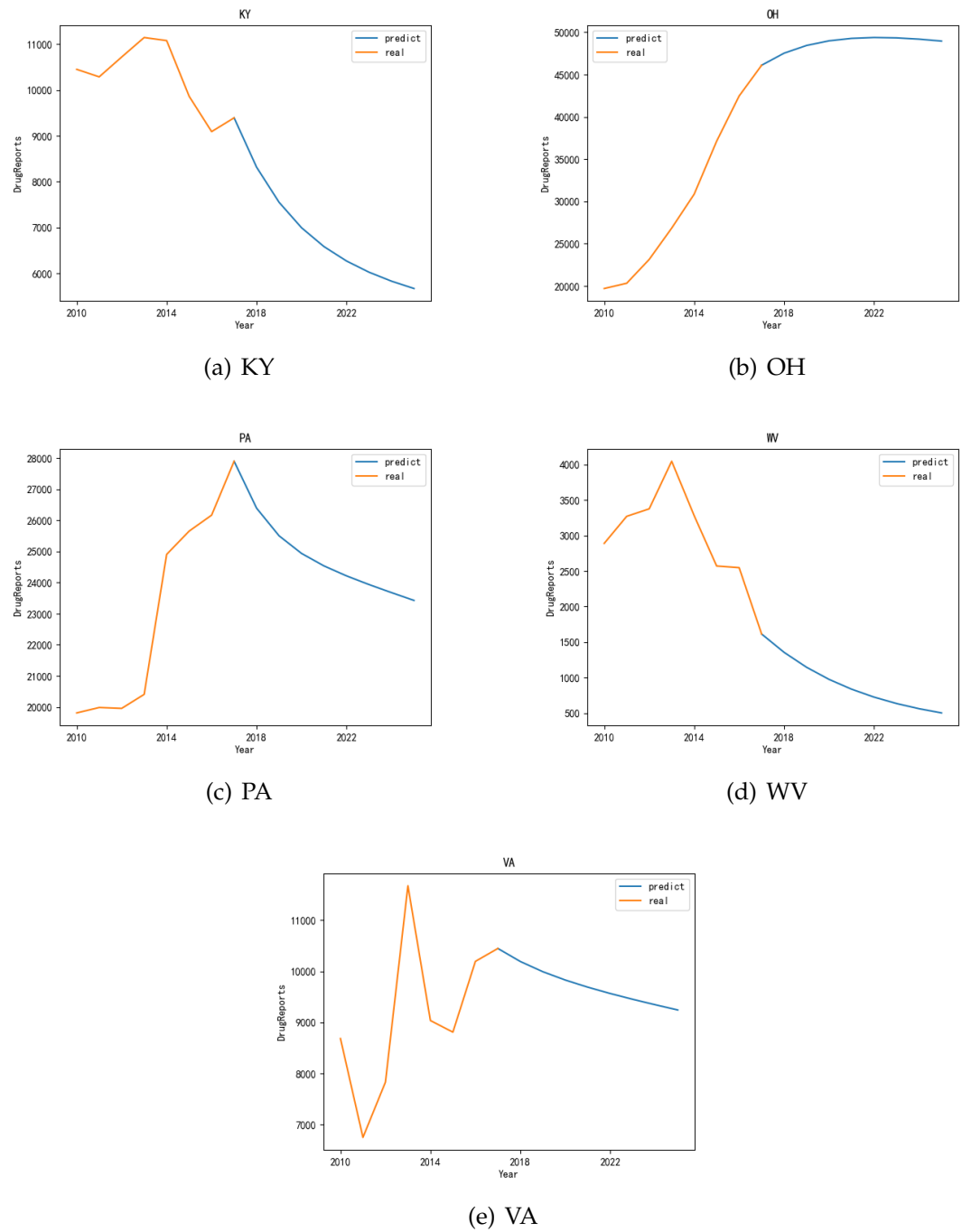
(a) KY

(b) OH

(c) PA

(d) WV

(e) VA

Fig. 11. The drug reports in five states after implementing the strategy.

significant impact on the performance of the trends-in-use of opioids with the estimation of matrix ρ.

- We manage to devise a more practical epidemic model that predicts the system under the consideration of both inland and outland characteristics.

Limitation and discussion: Though our model successfully identifies the characteristic of the opioid crisis in five specific states, it can be improved from the following several respects:

- As for the model 1, the performance of its prediction is affected by the number of total drug reports. This is due to the limitation from the objective function; for example, if we apply this function for the prediction of counties, then its train error will be certainly greater than that of states, since many places just have less than 10 reports, which as a result, makes the relative error, namely the value of the objective function, easy to become larger than 100%, and even more. So, if time allows, we attempt to develop the model 1 to decrease the sensitivity of reports base.

- As for the model 2, local economy level should always be an important element to consider for a social issue; however, since the data about economy development of five states are absent in the given dataset, we have no way utilizing them objectively. So, if possible, local economy data will improve the results more and indicating more information about this crisis.

References

[1] Greene, Mark., An Epidemiologic Assessment of Heroin Use. *American journal of public health*, 1975(64): 1-10.

[2] Bohnert, Amy S. B. *et al.*, Association Between Opioid Prescribing Patterns and Opioid Overdose-Related Deaths, *JAMA*, 2011(305):1315-1321.

[3] Magdalena, Cerdá *et al.*, Prescription opioid poisoning across urban and rural areas: identifying vulnerable groups and geographic areas, *ADDICTION*, 2017(112): 103-112.

[4] Herbert W. Hethcote, The Mathematics of Infectious Diseases, *Society for Industrial and Applied Mathematics*, 2000(42): 599-653.

[5] MIN Tao, CHENG Yao, GU Mingli, et al. Parameter estimation of nonlinear dynamic system and sensitivity. *Computer Engineering and Applications*,2013, 49(4): 47-49.

[6] Jiang Qiyuan, Xie Jingxing, Jun Ye. *Mathematical Models*. Beijing. Higher Education Press, 2011.

Appendices

Appendix A   Results

l=2

	High School Graduate (k=3)		Some College, No Degree (k=4)		associate's degree (k=5)		Bachelor's Degree (k=6)		Graduate or Professional (k=7)		BASE	
	1418.28571	984710.714	9162.21429	568457.5	186392.536	8419.64286	74562.393	7262.96429	234005.964	18831.9643	2861397.96	
	-13696.679	2734246.04	11216.2143	1562524.93	16478.4286	22653.5	17823.3214	671330.179	32121	18690.9	7654147.43	
	-7558.4643	10056.7857	1389509.93	14849.5714	621044.714	30141.8571	1361177.43	49474.25	8561600.25	1582739		
	6532.42857	1348458.71	15657.9643	11632.3929	344713.821	25935.6071	1037135.46	26660.25	5211200.36	810755.3		
	-1488.1429	530793.714	1866	229893.857	1986.46429	2951.39286	136509.393	2209.46429	84958.3214	77.3214286	1288176.18	1114

l=1

L / k	Never Married (k=1)		Now Married (k=2)		Sparated (k=3)		Widowed (k=4)		Divorced (k=5)		BASE
KA	15845.5	913622.79	-2649.25	1790680.9	944.53571	-692.35714	239125.79	7531.3214	429761.18	20079.75	
OH	42135.25	2740183.8	-23263.071	4640474.5	169380.43	-3108.8214	622610.61	12335.036	1074116.6	27805.107	
PA	48335.5	3265170.5	-15754.286	5147312.9	335.28571	-6569.5	767678.93	13734.929	932670.36	40081.929	
VA	40936.536	19212219.4	18639.75	3272987.3	256.85714	2558.2143	361242.64	14097.786	608427.79	76489.143	
WV	4437.2143	392527.07	-7381.5	800563.21	-224.57143	-1154.8214	125376.61	3207.3214	18446.89	-1116.3571	

l=3

	Yes (k=1)		No (k=2)		BASE	
	8955.2	691420.1	8671.4	3531794.2	17626.6	4232214.3
	18690.9	1458529.7	-5059.5	9874463.5	13631.4	11332993.2
	22655.5	1582739	-4258.2	10892964.9	18397.3	1247503.9
	18037.1	810755.3	58689.4	68549902.7	76726.5	7665748
	1114	344854.5	-8381.8	1497347.1	-7267.8	1842201.6

	BASE	less than 9th grade (k=1)		9th to 12th grade, no (k=2)	
	3447753	-7363.3214	234371.39	-7227.6429	307878.93
	9246766	-4627.4643	263131.68	-17726.321	696102.25
	10340907	-5668.8214	338637.75	-18831.964	729653.61
	63295522.7	-4202.3571	288253.07	-9960.7857	436215.79
	15288893.2	-4065.1429	87823.143	-3382.7143	143192.43

(a)   (b)

Fig. 12. The coefficient of linear regression for all sub-factors. The first column of each sub-factor is the linear coefficient, while the second one is the constant coefficient.