Research on Solutions to Opioid Crisis Based on Time-Society Model

Yifei Ma, Yuanyuan Zhang, Wei Zhao

Statistic Institute, Shanxi University of Finance and Economics, Wucheng Road, Taiyuan, China

Keywords: Time-society model, Time series algorithm, Factor analysis model, MATLAB

Abstract: In this paper, using the NFLIS data and the U.S. Census socio-economic data and a variety of scientific ideas and algorithms to construct a time- society model (TS model). The dissemination patterns and characteristics of opioids and heroin are analyzed, while the future development trend is forecasted. At the same time, potential influencing factors leading to the current situation are excavated and effective strategies are proposed to fight the opioid crisis.

1. Introduction

Opioid abuse and addiction – including prescription painkillers, synthetic opioids such as heroin and fentanyl – are serious national crises that affect public health as well as social and economic stableness. The increase of medical expenses and crime rate and the decrease of productivity all increase the economic burden of the country. Based on NFLIS data and the US Census socio-economic data, this paper calculates the entropy weight and constructs practical mathematical models through a series of methods of data mining, such as time series model, to solve the opioid crisis.

2. Prediction of Opioid propagation and feature

2.1 Determine Weight by Entropy Weight Method

Entropy weight method is an objective weighting method to determine the weight according to the information contained in the index. It avoids the influence of human subjectivity on the conclusion and ensures the scientific and reliability of the evaluation results. [1].

Use the extremely poor standardization method to standardize data:

$$r_{ij} = \frac{x_{j\max} - x_{ij}}{x_{j\max} - x_{i\min}}$$

Calculate the proportion p_{ij} of the indicator value of the item i under the indicator j:

$$p_{ij} = \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}}$$

Calculate the entropy value e_{ij} of the indicator j:

$$E_j = -k\sum_{i=1}^{m} p_{ij} \cdot \ln p_{ij}, k = \frac{1}{\ln m}$$

Calculate the entropy weight of the indicator j:

$$W_j = \left(1 - E_j\right) \Big/ \sum_{j=1}^{n} \left(1 - E_j\right)$$

Use each year as a sample to count the quantity of each opioid. Through the above steps using MATLAB we calculate the weight of each opioid, and then calculate the opioid index of each county according to the weight, finally, calculate the opioid index of the five states.

  DOI: 10.25236/etmhs.2019.308

The following is a part of counties opioid index and their rankings, such as Tab.1.

Tab. 1 A Part of Counties Opioid Index and Rankings

County	Opioid Index	Ranking	County	Opioid Index	Ranking
PHILADELPHIA	508.6303	1	…		
HAMILTON	496.6654	2	LEXINGTON CITY	0.0516	454
ALLEGHENY	294.2671	3	HIGHLAND	0.0509	455
CUYAHOGA	266.9186	4	BEDFORD CITY	0.0507	456
MONTGOMERY	234.7680	5	TAYLOR	0.0365	457
FRANKLIN	171.6813	6	WIRT	0.0258	458
JEFFERSON	139.2082	7	LINCOLN	0.0250	459
LAKE	103.0030	8	NORTON CITY	0.0247	460
	…		COVINGTON CITY	51580	461

The above are the 8 counties with the largest index and the 8 counties with the smallest index. It can be clearly seen from the image, most of the counties with a large amount of opioid are concentrated in Ohio and Pennsylvania, at the same time, these counties are at the border of the state, or the coastal areas which the transportation is convenient, so that these counties which are considered are important areas for spreading opioids to other regions.

According to 2010-2017 NFLIS data of the counties, there are more than 20 synthetic opioids that have existed since 2010, and more than 30 synthetic opioids started in 2016 or 2017. Appeared, through analysis, these new synthetic opioids appeared for the first time in OH or PA. From the processed data, we found that the amount of heroin far exceeds that of other synthetic opioid and the number remained at a high level. We consider that the spread of synthetic opioids is related to the supply of heroin. Considering the situation in the five states, the relationship between states and counties is showed in the form of a bubble chart such as Fig.1.

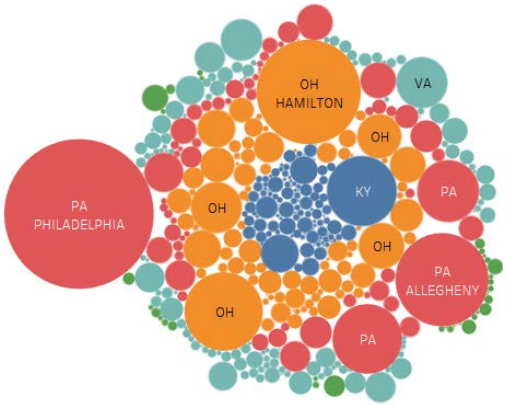

Fig. 1 The Bubble chart

2.2 Method and prediction based on Time Series Model

Use the first-order differential exponential smoothing method to establish a time series model for each state and county. The historical observations are based on a weighted average to predict the characteristics and patterns of transmission of opioids and heroin. The steps are as follows:

$$\nabla y_t = y_t - y_{t-1}$$

$$\nabla \hat{y}_{t+1} = \alpha \nabla y_t + (1-\alpha)\nabla \hat{y}_t$$

$$\hat{y}_{t+1} = \nabla \hat{y}_{t+1} + y_t$$

Time series data is a type of rising trend. It should take a larger value between 0.6 and 1. Query the literature to understand the time series, the initial range of values, and then select several values for

trial calculation. Compared different α. The prediction standard error under the value is selected as the minimum of the prediction standard error, and the value is placed at 0.78.

The exponential smoothing value plus the actual value of the current value is used as the predicted value, so that the forecast value fluctuates around the actual value. In this way, we fundamentally solve the problem that the result obtained by one exponential smoothing method is always behind the issue of the actual value in the case of a linear growth trend.

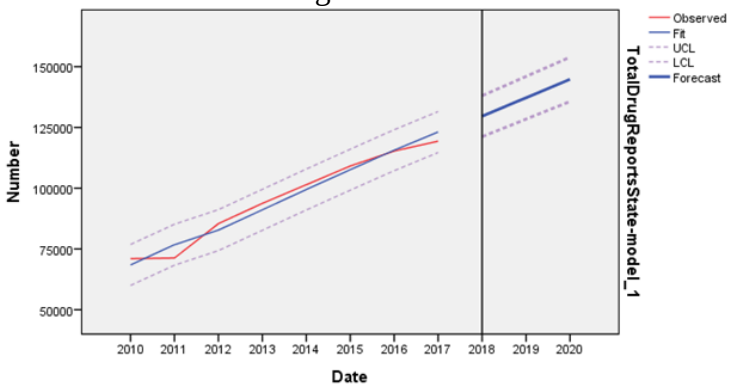

Fig. 2 Time series fitting and prediction of OH

It can be seen from the figure that the fitting effect is good. At the same time, the R^2 – adjusted of the model is 0.875 through the test, so the model prediction is accurate.

In addition, it can be observed that the time series exhibits a quadratic curve growth state and can be predicted by a second-order differential exponential smoothing model. The steps are as follows [2]:

$$\begin{cases} \nabla y_t = y_t - y_{t-1} \\ \nabla^2 y_t = \nabla y_t - \nabla y_{t-1} \\ \nabla^2 \hat{y}_{t+1} = \alpha \nabla^2 y_t + (1-\alpha)\nabla^2 \hat{y}_t \\ \hat{y}_{t+1} = \nabla^2 y_{t+1} + \nabla y_t + y_t \end{cases}$$

The time series data fluctuates, but when the long-term trend does not change much, the value is often between 0.1 and 0.4. After the quantitative change of the prediction object is made, the prediction error is calculated, considering that the prediction sensitivity and the prediction accuracy are contradictory. After comprehensive consideration, it is set at 0.24.

Heroin is an opioid which is different from other synthetic opioids and it can cause more serious crisis. So, we built a model to analyze the changes in heroin alone. For the time series prediction of heroin, the adaptive filtering method is adopted to reduce the error by repeatedly calculating the prediction error adjustment weight. Until you find the best set of weights, the error is reduced to a minimum [3].

The basic prediction formula is:

$$\hat{y}_{t+1} = w_1 y_t + w_2 y_{t-1} + \ldots w_N y_{t-N-1} = \sum_{i=1}^{N} w_i y_{t-i+1}$$

Among them, $\hat{y}_{t+1}$ is the predictive value of $t-i+1$, w_i is the weight of observation value of $t-i+1$, y_{t-i+1} is the observation value of $t-i+1$, N is the number of weights. The formula for adjusting the weight is:

$$w_i^{'} = w_i + 2k \cdot e_{i+1} y_{t-i+1}$$

$w_i^{'}$ is the adjusted weight i, k is the learning constant, e_{i+1} is the prediction error of $t+1$.

Through the adjustment process of the computer, the error is reduced to 0 and the weight is stable. The optimal weight obtained at this time is: $w_1^{'}$=0.7, $w_2^{'}$=0.3. Utilizing the best weight to predict the value of 2018 is 51530.8, the forecast for 2019 is 52215.3, and the forecast for 2020 is

52899.8.

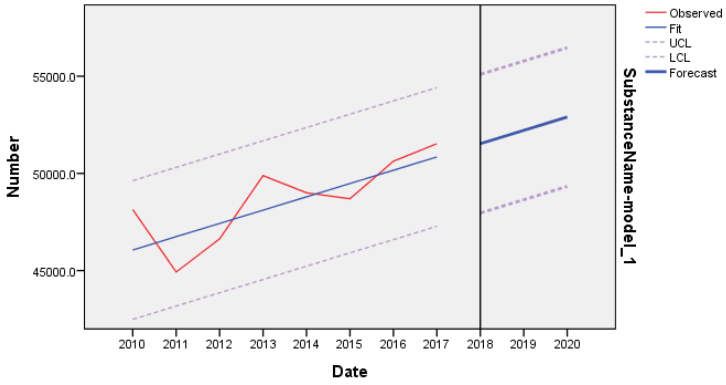

Fig. 3 Time series fitting and prediction of Heroin

At the same time, the same situation exists in other synthetic opioids, especially strong opioids. The time prediction and position prediction are as follows:

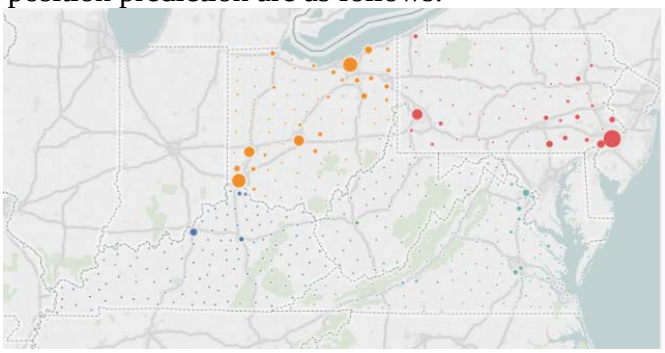

Fig. 4 Time series fitting and prediction

2.3 Factor Analysis Model

The purpose of factor analysis is to reduce the dimension and extract a few potential unobservable random variables to describe the covariance relationship between the original variables under the condition of little information loss [4]. The factor analysis can be used to mine the potential effects of the use of opioids and the growth of opioid addiction. In order to analyze whether the trend of its use or use is related to the socio-economic data of the US Census to some extent, we have factored the data from 2010 to 2016 separately. The KMO value of the data for 2016 is 0.72 for factor analysis.

The seven factors are named as human resource factor, psychological factor, gender factor, living environment factor, family size factor, ethnic factor and population expansion factor. The human resource factor includes veterans, disabled people, housewives, and educational attainment. Psychological factor includes marital status, with or without children, relationship, and other variables. The gender factor includes the age composition of men and women, the distribution of male and female populations; The environmental factor includes place of birth, place of residence, time of residence, and common language. The family size factor includes the total number of households, the number of households, and the average population which reflects the size of the family. The ethnic factor includes entry time and ancestral conditions. And the population expansion factor includes fertility and Age-appropriate male and female population. The use of opioids is at a current level which has a strong relationship with seven potential factors. We find that the areas of poor human resources, especially with low levels of education, the areas with low psychological factors, especially divorced or widowed and the areas which the number of men is high are high-risk areas where opioids and heroin are used. Calculate factor scores.

$$F = \frac{(0.36283F_1 + 0.19534F_2 + 0.15450F_3 + 0.09573F_4 + 0.05869F_5 + 0.04025F_6 + 0.02870F_7)}{0.90734}$$

The comprehensive evaluation model is constructed with the variance contribution rate of each

common factor, and the samples are sorted and compared according to the size of the F value. The ranking is shown as Tab.2.

Tab.2 the Ranking of F-score

County	Ranking	County	Ranking
PHILADELPHIA	1	…	
HAMILTON	2	FRANKLIN	6
ALLEGHENY	3	JEFFERSON	7
CUYAHOGA	4	LAKE	8
MONTGOMERY	5	DELAWARE	9
…		MAHONING	10

3. Conclusion and Strategies

The improved time series model can quantitatively analyze the expansion rate of opioids and heroin over time due to socio-economic factors, geographical location, and potential factors such as urbanization and industrialization. The important influencing factors are analyzed by factor analysis model as human resource factor, psychological factor, gender factor, environmental factor, family size factor, racial factor, and population expansion factor. Therefore, we should start from this aspect and propose feasible strategies against opioids.

◆Strengthen education and re-employment for disabled and veteran training. And rationally distribute benefits.

◆Strengthen the control of opioids and heroin, strengthen the inspection of transportation, and make people aware of their harm through vigorous publicity, focusing on rectifying more serious counties.

◆Use financial means to optimize the ratio of men and women, improve marital status, and enhance reproductive ability and increase people's sense of responsibility for future generations and education.

◆Strengthen entry and exit inspections and improve entry standards.

References

[1] Guangcai Liu, Luyuan Yang. Evaluating the Development Trend of Shared Vehicle Based on AHP and Entropy Weight Method [J]. Commercial Economic Research, 2019 (01): 190-192.

[2] Y. Shuai, Integrated parallel forecasting model based on modified fuzzy time series and SVM, Journal of Systems Engineering and Electronics, 28 (4) (2017).

[3] Jidong Yuan, Zhihai Wang. Overview of Time Series Representation and Classification Algorithms [J]. Computer Science, 2015, 42 (03): 1-7.

[4] S. Gong, Energy efficiency evaluation based on DEA integrated factor analysis in ethylene production, Chinese Journal of Chemical Engineering , 25(6)2017 793-799

[5] Yi Liu, Jiawen Peng, and Zhihao Yu. 2018. Big Data Platform Architecture under The Background of Financial Technology: In The Insurance Industry As An Example. In Proceedings of the 2018 International Conference on Big Data Engineering and Technology (BDET 2018). ACM, New York, NY, USA, 31-35