Advances in Computer Science Research, volume 91
International Conference on Modeling, Simulation and Big Data Analysis (MSBDA 2019)

A New Analysis and Prediction Model Based on American Opioid Crisis

Zongna Deng, Deyi Li*, Yuanyuan Wang, Quan Lv and Ning Chen
Wuhan University of Science and Technology, Wuhan, China
*Corresponding author

Keywords: Opioid crisis, Linear interpolation, Time-varying prediction model, Factor analysis, Coping strategy.

Abstract. Based on the existing data, the diffusion model and time-varying prediction model of opioids are established by linear interpolation method, the transmission characteristics of synthetic opioids and non-synthetic opioids are analyzed, and the specific time and place of each state to start using specific opioids in different states is determined. Then, in order to judge the correlation between the trend of opioid use and the economic data of the United States census, 10 evaluation indexes were selected and empirical analysis was carried out by factor analysis. It was found that the factors of elderly, neonatal number, disability and population size were used to analyze opioid drugs in this region. Use has had a significant impact. In addition, the most widely used opioid groups have the characteristics of lack of health insurance, unemployment, low income and low level of education. Finally, according to these conclusions, this paper presents an effective strategy to combat opioid crisis, and tests the effectiveness of the model. It is found that the fitting degree of the model is about 78%, and the fitting effect of the model is good.

Introduction

The annual use of opioids has continued since they entered the United States market on a large scale in the late 1990s increase. Not only does the abuse of opioids cause hundreds of deaths every day in the United States, the growth rate of drug abusers in the United States is increasing year by year. Increase the fiscal expenditure of the whole country. The United States government declared a public health emergency. How to scientifically control drug abuse, solve drug addiction and reduce the impact of opioid crisis has become one of the most serious problems facing the United States.[1][2][3][4]

Assumptions

(1)Assume that the county location data are correct.
(2)Assume that the non-synthetic drugs studied in this paper are only morphine, codeine, heroin.
(3)Assume that the non-synthetic opioid drugs caused by heroin events.

Establishment and Solution of Diffusion Model

In order to visually analyze the spread and characteristics of opioid heroin incidents, we did the following: first, separate the data from the NFLIS by state. Then, each state was then separated by synthetic opiates and non-synthetic opiates. Finally, the diffusion model and time-varying prediction model of opioids were established.

From the Figure 1, we can see that the heroin consumption of one county in KY is the highest, which is much higher than that in other counties of KY; secondly, the heroin consumption is higher in the three counties codenamed as 38, 67,118, and the consumption in other counties is less and morestable. As can be seen from the above Figure 2, heroin consumption increased significantly in KY counties from 2010 to 2014, and decreased significantly in KY counties from 2014 to 2016, but the decline was small and then leveled off.

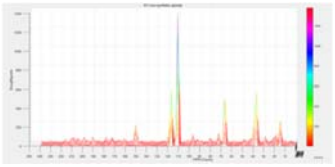

Figure 1. The heroin consumption in KY(county)

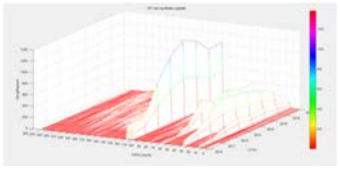

Figure 2. The heroin consumption in KY(time)

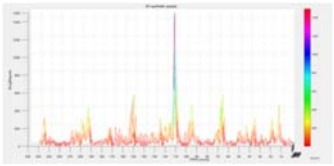

Figure 3. The synthetic opioids consumption in KY(county)

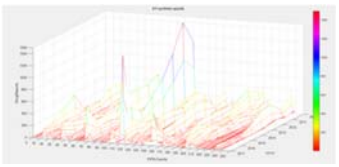

Figure 4. The synthetic opioids consumption in KY(time)

From the above Figure 3, we can see that the amount of synthetic opioids of one county in KY state is the highest, which is much higher than that in other counties of KY state; secondly, the amount of synthetic opioids drugs in several counties, the amount of synthetic opioids in other counties is less, but the curve is more irregular. From the above Figure 4, we can see that the amount of synthetic opioids in the counties of KY increased significantly from 2010 to 2011, and the amount of synthetic opioids in counties of KY decreased significantly in the period of 2011-2015. In 2016, the amount of synthetic opioids in the counties of the state of KY decreased significantly. The amount of synthetic opioids in the counties of KY was also slightly increased.

From the above analysis, we can see that the transmission and characteristics of synthetic opioids and heroin are not the same, the increase and decrease of heroin consumption shows a steady curve, while the curve of synthetic opioid drugs shows a large fluctuation. This shows that there is a great deal to do with the policies of the states and counties at that time.

It can be seen from Figure 5 that the heroin consumption of one county in PA County is the highest, which is much higher than that in other counties of PA State; secondly, the heroin consumption is higher in three counties of PA State, and the amount of heroin in other counties is less and more stable. It can be seen from Figure 6: heroin consumption in PA counties was more stable from 2010 to 2014, while heroin consumption in counties in PA state increased significantly, but by a relatively small margin during the period from 2014 to 2016.

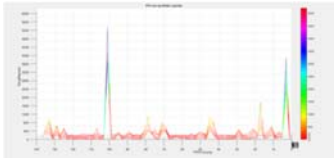

Figure 5. The heroin consumption in PA(county)

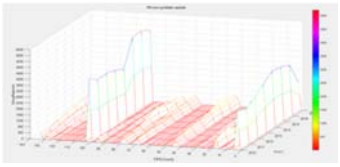

Figure 6. The heroin consumption in PA(time)

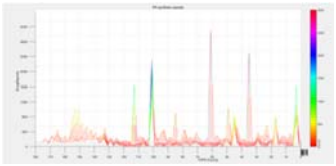

Figure 7. The synthetic opioids consumption in PA(county)

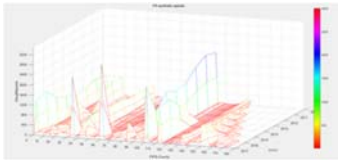

Figure 8. The synthetic opioids consumption in PA(time)

From the above Figure 7, we can see that the amount of synthetic opioids in the one county of PA is the highest, which is much higher than that in other counties of PA; secondly, the amount of synthetic opiates in several counties is higher than that in other counties, and the dosage in other counties is less. It can be seen from Figure 8 that the amount of synthetic opiates used in counties of PA state was stable from 2010 to 2016, and the amount of synthetic opioids in some counties of PA state increased significantly during 2016-2017, and the amount of synthetic opioids in some counties of PA state increased by a large margin during the period of 2016-2017.

It can be seen from Figure 9 that the heroin consumption of one county in VA is higher than that in other counties of VA, and the consumption of heroin is slightly higher in three counties codenamed as 150,180, while the consumption in other counties is less and more stable. It can be seen from Figure

10 that heroin consumption increased steadily in all counties of VA from 2010 to 2015, and decreased significantly in all counties of VA, in some counties, and in some counties in the period of 2015-2017.

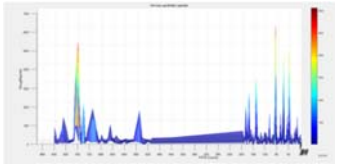

Figure 9. The heroin consumption in VA(county)

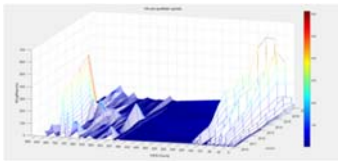

Figure 10. The heroin consumption in VA(time)

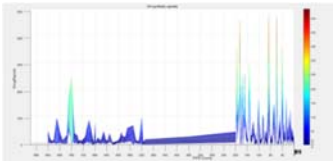

Figure 11. The synthetic opioids consumption in VA(county)

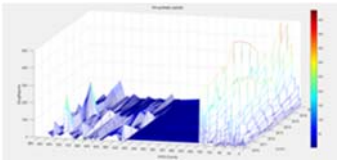

Figure 12. The synthetic opioids consumption in VA(time)

It can be seen from above Figure 11 that the amount of synthetic opioids in the one county of VA is higher than that in other counties of VA, and the amount of synthetic opiates in several counties is slightly higher. Other counties use less, the whole state showed an irregular curve. It can be seen from the above Figure 12 that the amount of synthetic opiates used in counties of VA was stable from 2010 to 2014, and the amount of synthetic opioids in some counties of VA decreased significantly during 2014-2016, and increased in 2016-2017.

From the above Figure 13, we can see that the heroin consumption in one county of WV is the highest, which is much higher than that in other counties of WV; secondly, the heroin consumption in other counties is slightly higher than that in other counties, and the consumption in other counties is relatively stable. It can be seen from Figure 14: between 2010 and 2016, heroin consumption increased more steadily in only 112 counties of WV state, while heroin consumption in other counties was less stable than that in other counties.

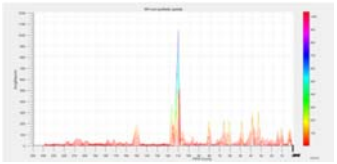

Figure 13. The heroin consumption in WV(county)

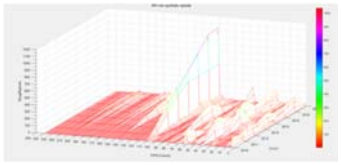

Figure 14. The heroin consumption in WV(time)

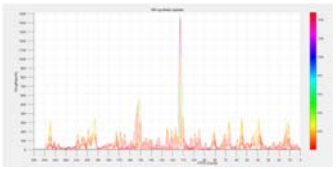

Figure 15. The synthetic opioids consumption in WV(county)

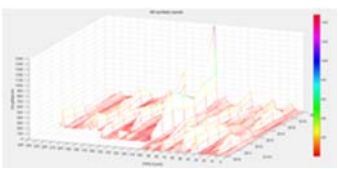

Figure 16. The synthetic opioids consumption in WV(time)

It can be seen from Figure 15 that the amount of synthetic opioid of one county in WV state is higher than that in other counties of WV state, and the amount of synthetic opioid drugs is slightly higher in

the counties codenamed as WV state as 112 county, while in other counties of WV state, the amount of synthetic opioids drugs is slightly higher. Other counties use less, the whole state showed an irregular curve. It can be seen from Figure 16 that the amount of synthetic opiates in the 102 counties of WV state increased significantly during 2016-2017, while the amount of synthetic opioids in other counties was stable.

Factor Analysis Model

Factor Analysis Method

Factor analysis was first proposed by Thurstone in 1931. Its concept originated from the statistical analysis of intelligence tests by Kard Pearson and Charles Spearmen et al in the early 20th century.[5] In recent years, with the development of computer and the popularity of statistical analysis software, factor analysis has been used in various fields. The basic principle of factor analysis is to synthesize a number of variables with certain correlation into a few factors, and to study how a group of intricate measured indexes are controlled by a few internal independent factors. Factor analysis belongs to multivariate division a commonly used statistical method to deal with dimensionality reduction in analysis. Suppose that there are p standardized variables $X_1, X_2, ..., X_p$, then the general model of factor analysis is as follows:

$$\begin{cases} X_1 = a_{11}F_2 + a_{12}F_2 + \cdots + a_{1k}F_m + \varepsilon_1, \\ X_2 = a_{21}F_2 + a_{22}F_2 + \cdots + a_{2k}F_m + \varepsilon_2, \\ \cdots \\ X_p = a_{p1}F_2 + a_{p2}F_2 + \cdots + a_{pk}F_m + \varepsilon_p. \end{cases}$$

Among them, $F_1, F_2, ..., F_m$ is a common factor, ε_1 , ε_2 ,.., ε_p is a special factor and they are all unobservable random variables. The above matrix can be expressed as $X = AF + \varepsilon$, where $F = (F_1, F_2, ..., F_m)^T$ is a common factor vector, $\varepsilon = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_p)^T$ is a special factor variable, A is called factor load matrix. In addition, it is usually assumed that $E(F) = E(\varepsilon) = 0$. $\text{cov}(F, F) = I_m$, ε_1 , ε_2 ,.., ε_p and F are independent of each other. The sum of squares of line i elements in the factor load matrix $h_i^2 = \sum_{j=1}^{m} a_{ij}^2$ Called the i th change. The common degree or common variance of quantity, which reflects the influence of common factor on variable X_i , can be regarded as the contribution of common factor to the variance of X_i , and reflects the explanatory ability of factor j to explain the total variance of original variable. Is a measure of the importance of the common factor F_1 .

General steps of Factor Analysis:

1. The index system is established according to the original data of the sample.
2. Standardized processing of sample index data and factor analysis adaptability test.
3. The correlation coefficient matrix is calculated and the eigenvalue of the correlation matrix is obtained. The number of common factors is determined according to the contribution rate of cumulative variance.
4. The factor load matrix A is obtained, and the matrix B is obtained by using the maximum variance orthogonal rotation method, and the common factor scores are calculated.
5. According to the size of factor score and variance contribution rate, the comprehensive score is calculated. Finally, the comprehensive score is sorted and analyzed according to the comprehensive score.

Factor Analysis in Each Region

And try to find out the reason why opioid drug use is still dangerous from many influencing factors. Based on the analysis of the population indexes of 463 states and counties collected by the United

States Census Bureau from 2012 to 2016, according to the principles of representativeness, comprehensiveness, maneuverability and comparability of the index evaluation system, 10 key evaluation indexes are obtained. According to the basic principle of factor analysis, R statistical software is used to carry out the real evaluation. Evidence research. Tables 1 and 2 are the names and original data of 10 evaluation indexes selected in this paper. Analysis on the level of Economic Development in different regions.

Table 1. Original variable name

Variable name
X1-Households with one or more people under 18 years
X2-Households with one or more people 65 years and over
X3-Population in households
X4-Population in households - Child
X6-Females 15 years and over
X7-Number of women 15 to 50 years old who had a birth in the past 12 months
X8-Number of grandparents living with own grandchildren under 18 years
X9-Percent high school graduate or higher
X10-Disability status of the civilian population

The variables selected in table 1 are the basic indicators of the census, reflecting the basic demographic situation of the United States in a number of ways:

1. Reflects the total number of families in all regions, including families with children and families with older persons;

2. It reflects the population structure of families in various regions, and counts the population of men, women and minors over 15 years of age.

3. It reflects the new number of families in various regions, the number of children left behind and the number of disabled people;

4. Reflects the education level of the family population and the economic situation to a certain extent in various regions.

Table 2. Original data from the United States Census Bureau for 463 states and counties, 2012-2016

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
Adair County, Kentucky	1953	2211	17887	5308	7800	8149	219	306	76	18870
Allen County, Kentucky	2575	2259	20208	6303	8058	8313	221	413	76.4	20236
Anderson County, Kentucky	2855	2222	21762	6153	8436	9068	314	554	88.6	21759
Ballard County, Kentucky	971	1157	8097	2206	3313	3428	83	208	86.2	8091
Barren County, Kentucky	5306	4737	42539	12528	16793	18046	482	923	81.7	42548
					……					
Wayne County, West Virginia	4680	5463	41060	12296	16386	17560	341	957	79.4	41093
Webster County, West Virginia	1154	1351	8763	1883	3608	3716	85	479	73.9	8763
Wetzel County, West Virginia	1435	2333	15880	4502	6477	6855	191	378	83.1	15899
Wirt County, West Virginia	545	777	5826	1543	2446	2347	71	37	83.3	5826
Wood County, West Virginia	10278	11546	85367	22845	34039	37140	1175	1718	89.6	85759

Table 3. Correlation matrix

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
X_1	1.000	0.438	0.860	0.000	0.447	0.518	0.036	0.797	0.647	0.385
X_2	0.438	1.000	0.141	0.652	0.933	0.946	0.632	0.598	0.498	0.919
X_3	0.860	0.141	1.000	-0.303	0.102	0.230	-0.311	0.408	0.281	0.043
X_4	0.000	0.652	-0.303	1.000	0.714	0.584	0.784	0.360	0.297	0.688
X_5	0.447	0.933	0.101	0.714	1.000	0.926	0.753	0.652	0.517	0.946
X_6	0.518	0.946	0.230	0.584	0.926	1.000	0.628	0.627	0.458	0.912
X_7	0.036	0.632	-0.311	0.784	0.753	0.628	1.000	0.373	0.284	0.766
X_8	0.797	0.598	0.408	0.360	0.652	0.627	0.373	1.000	0.900	0.611
X_9	0.647	0.498	0.281	0.297	0.517	0.485	0.284	0.900	1.000	0.500
X_{10}	0.385	0.919	0.043	0.688	0.946	0.912	0.766	0.611	0.500	1.000

From the correlation coefficient matrix of Table 3, we can see that the variables have strong correlation directly.

The test value of KMO is 0.776, which is suitable for factor analysis, and the Sig value of Bartlettt test is 0.000, which indicates that the data come from the general distribution of the data, which is suitable for further analysis.

Factor Extraction

Table 4. Variance contribution matrix

	Factor1	Factor2	Factor3
Common variance	4.914	2.215	2.091
Variance contribution rate	49.14	22.15	20.91
Cumulative variance contribution rate	49.14	71.29	92.20

As can be seen from Table 4, the cumulative variance contribution rate of the three factors has reached more than 90%, and the three common factors are not related to each other, thus avoiding the multiple collinearity of the original index. The three common factors contain the information which can be expressed by the 10 index data of the original data.

The result of factor load estimation after the rotation of variance maximization is shown in Table 5.

Table 5 shows that the factor load coefficient after rotation has been polarized and has obvious practical significance. The load value of common factor F_1on X_2,X_4,X_5,X_6,X_7,$X_{10}$is very large, which can be defined as income factor. The load value of common factor F_2 onX_8,$X_9$is very large, which can be defined as opening factor.The load value of common factor F_3 on $X_1, X_3$is very large, which can be defined as capital input factor.

Table 5. Factor load matrix

Variable	F_1	F_2	F_3
X_1	0.220	0.539	0.793
X_2	0.915	0.218	0.193
X_3	-0.052	0.177	0.951
X_4	0.770	0.214	-0.399
X_5	0.931	0.287	0.115
X_6	0.909	0.189	0.295
X_7	0.808	0.191	-0.370
X_8	0.403	0.841	0.311
X_9	0.247	0.933	0.151
X_{10}	0.930	0.262	0.068

Factor Score

The total score of each region can be obtained from the above factor score, where the comprehensive score formula is the sum of the product of each factor score and the contribution rate of each factor variance, and then divided by the sum of the difference contribution rate. So

$$F = \frac{0.4914F_1 + 0.2215F_2 + 0.2091F_3}{0.922}$$

Add the factor scores to this formula to get the combined scores and rankings of the states and counties, as shown in Table 6.

From the factor scores of regional population economy in Table 6, it can be seen that the population and the level of economic development in each region of the United States are unbalanced and multi-level. In combination with the above analysis of the use of opioid drugs in five states, we can see that:

1. New York, Los Angeles, Chicago, Houston, Phoenix, Philadelphia, and other regions in the United States have a positive overall score, indicating that the population and economic strength of these regions are above the national average, with negative scores. Its comprehensive strength lies below the national average and the population is concentrated in economically developed areas.

2. The peak value of opioid drug use first appeared in the densely populated and economically developed areas, the population is the most basic influencing factor, which will directly determine the total amount of opioid use in state and county. Through further comparison, it can be found that the total amount of opioid drugs used in the counties with close population but different economic level is close, and there is no obvious regularity difference. However, the economic level had a significant effect on the proportion of non-synthetic opiates and synthetic opioids, and the proportion of synthetic opiates was higher in the economically developed areas, and the reverse was observed in the less developed regions.

Table 6. Factor score matrix

State and county	F_1	F_2	F_3	Total Points F	Ranking
Adair County, Kentucky	3.150	0.155	-1.150	1.455	19
Allen County, Kentucky	1.769	-0.623	-0.091	0.772	109
Anderson County, Kentucky	-0.639	-0.438	0.838	-0.256	157
Ballard County, Kentucky	-0.431	-0.313	-0.524	-0.424	319
Barren County, Kentucky	0.196	-0.705	0.178	-0.025	243
……					
Wayne County, West Virginia	-0.554	-0.087	0.049	-0.305	186
WebsterCounty,West Virginia	-0.734	-0.174	-0.893	-0.636	307
Wetzel County, West Virginia	-0.412	--0.437	-0.847	-0.517	296
Wirt County, West Virginia	-0.244	-0.460	-0.875	-0.439	179
Wood County, West Virginia	-0.521	-0.270	-0.597	-0.478	271

3. Among the basic data indicators of the census, in addition to the factors affecting the total population, the number of older persons, the number of newborns and the number of people with disabilities have a significant impact on the use of opioid drugs in the region, The quantity of these three population indicators and their proportion in the total population have a significant effect on the increase of opioid use.

4. The number of males and females in census data and their proportions have no significant effect on the use of opioid drugs; at the same time, the overall education level in the United States is relatively high. Compared with the less developed regions, the population education level in the economically developed regions is relatively high, and the demand for opioids in this region is usually greater. However, there is no direct relationship between educational level and opioid dosage.

5. Based on the above comparative analysis of the fluctuation of opioid drug use over different regions and census data, the following inferences can be drawn:

(1)The amount of opioid drugs in different regions fluctuated with the migration of population, but it was relatively stable in recent years.

(2)The relative decline in the proportion of non-synthetic opiates in all regions after 2014 is expected to continue in the next few years.

(3)Population and economic census data have an important impact on the improvement of the first part of the model. In order to make a more effective analysis of the use of opiate drugs in various

regions, it is necessary to provide population and economic data in addition to the original data in the first part. In addition to the total amount used, the per capita consumption of opioid drugs in a certain area is also the necessary object of investigation.

Strategies

Strategies for Coping with Opioid Crisis

In order to control drug abuse scientifically, solve drug addiction and reduce the impact of opioid crisis, and solve the opioid crisis reasonably and effectively, we give the following five suggestions:

First, publicity and education, professional training and public health education. Educate related personnel such as prescription doctors and parents. Require doctors to be trained in rational prescription of opioids before obtaining registration of DEA controlled drug prescriptions. With the help of mass media, the importance and necessity of the management of analgesic drugs are widely publicized to improve the patients' cognition of pain diseases and the effects and risks of analgesic drugs. At the same time, we should strengthen the management of narcotic and psychotropic drugs and the management of drug harm, especially the warning of recreational use of narcotic and psychotropic substances.

Second, improve relevant laws and regulations, strengthen coordination and management. Narcotics and psychotropic substances are not ordinary commodities. The health administration department should manage the anesthetics and psychotropic drugs more carefully in order to meet the need of clinical treatment. Need according to international, domestic relevant laws and regulations to carry on strict control.

Third, big data tracking and monitoring. To increase the demand for post-market data on the impact of long-term opioid use on drug companies and provide better evidence of the serious risks of misuse, abuse and long-term use of opioids.

Fourth, carry out related research to provide comprehensive and scientific treatment for patients with addiction. To study more effective ways to treat opioid overdose, to use opioid prescription drugs more safely, and to improve the treatment of pain.

Fifth, adopt a comprehensive and balanced strategy to deal with addiction-related problems. On the one hand, we should recognize that opioid drugs are indispensable to human health and well-being, and on the other hand, we should pay attention to their iatrogenic addiction and related hazards.

Sensitivity Analysis

In the first part, the linear interpolation method is used to fit the scattered point data, and the fluctuation of opioid drugs in the states and counties is analyzed intuitively, but because of the large variation of individual scattered point data, there is a certain error.

In the second part, the factor analysis method is used to deal with the data with huge cardinality, without the choice of the data, which improves the fitting degree of the model to a great extent and has a good readability. By testing the validity of the model, it can be found that the fitting degree of the model is about 78% and the effect of model fitting is good.

The strategy proposed in the third part is in accordance with the evolution of the model to a certain extent reduces the amount of opioid drugs used to ensure the effectiveness of the strategy.

References

[1] Li Jun. The crisis of opioid abuse in the United States leads to some thoughts on the regulation of related drugs J. Journal of adverse Drug reactions, Vol. 20, No. 6, December 2018.

[2] Hao Wei. Be prepared ahead of time: the American opioid crisis brings enlightenment to our country J. Chinese Journal of Psychiatry, Vol. 51, No. 5, October 2018.

[3] Sun Li. Opioid abuse and its regulatory measures in the United States J. Chinese Journal of Drug abuse Regulation and Prevention, Vol. 24, No. 4, 2018.

[4] Feng Yan. Study on the environmental sensitivity and influencing factors of opioid dependence patients J. Chinese Journal of Epidemiology, vol. 29, No. 5, May 2008.

[5] Manbin Lin. Research on the level of Regional Economic Development based on Factor Analysis J. Economic Mathematics, Vol. 35, No. 4, December 2018.