

One-dimensional Poisson marked point process model and Its Random Characteristic Analysis in Haze Weather

Yonghua Zhu, Shunzi Li

North China Electric Power University, 2 Beinong Road, Huilongguan Town,
Changping District, Beijing, China

zhuyonghua@ncepu.edu.cn, 18010039283@163.com

Abstract: In this paper, a One-dimensional Poisson marked point process model with haze intensity as the marked point is established. The parameters of the model are estimated based on the measured data of haze. Finally, the probability of haze occurrence with different intensities and the important digital characteristics of haze occurrence above severe are calculated.

1. Introduction

Frequent haze weather in China has attracted public attention. In many areas, haze weather has been used as a warning and forecast for severe weather. Fog is an aerosol system consisting of a large number of tiny water droplets or ice crystals suspended in the air near the ground. Haze refers to the deterioration of visibility caused by aerosol pollution in the near-surface atmosphere. In the study of hazy weather, the main research areas at present are the causes of haze, the components of haze, the hazards of haze, and the statistical description of haze. In fact, the haze phenomenon should essentially be a random process that changes over time. In this paper, the Poisson point process model will be used to describe the random characteristics of the haze and the probability of haze occurrence. Moreover, the numerical characteristics of haze will be calculated.

2. Model establishment

At present, when describing haze phenomenon, haze phenomenon refers to the atmospheric phenomenon when the air index (AQI) exceeds the intensity h_0 (AQI = 100). The severity of haze is described by the magnitude of intensity, so the intensity of haze (expressed with maximum intensity H) is the most important characteristic parameter of haze. This paper will establish a Poisson point process with haze intensity as a marked point to describe haze phenomena. We assume the AQI index H from $H > h_0$ to $H < h_0$ as a haze event, and One-dimensional Poisson marked point process model with the marked point H is as follows:

First, we make the following assumptions:

(1) The number of haze events with the marked point $H > h_0$ in the time range of $(0, t]$ is a homogeneous Poisson process with the intensity λ , denoted by $\{N(t), t \geq 0\}$, and we obtain that:

$$P\{N(t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \quad (k = 0, 1, 2, \dots) \quad (1)$$



(2) The point event with the marked point H occurred in the time range of $(0, t]$. Among them, H is the continuous random variable valued in the range of $[a, +\infty)$ ($a > 0$), and its density function is $f(h)$. Assume that $f(h)$ is continuous in the range of $[a, +\infty)$, and $h_0 \geq a$, $p_0 = P\{H > h_0\} > 0$.

Let $B = \{H > h_0\}$ where M is a positive integer. h_1, h_2, \dots, h_M satisfy the conditions as below:

$$\begin{aligned} h_0 < h_1 < h_2 < \dots < h_{M-1} < h_M = +\infty \\ B_i = \{h_{i-1} < H \leq h_i\}, \quad p_i = P\{B_i\}, \quad (i=1, 2, \dots, M) \end{aligned} \quad (2)$$

where B_1, B_2, \dots, B_M are the division of B . From the decomposition of Poisson point process and superposition principle, we can know that, in the range of $(0, t]$, the number of occurrence of event B can be described by the Poisson process $\{N(t), t \geq 0\}$ with the intensity λ , and the number of occurrence of event B_i is the Poisson process $\{M_i(t), t \geq 0\}$ ($i=1, 2, \dots, M$) with the intensity $\lambda_i = \lambda p_i$. The Poisson process $\{N(t), t \geq 0\}$ can be decomposed into independent Poisson processes $\{M_i(t), t \geq 0\}$ ($i=1, 2, \dots, M$) with the intensity λ_i , $\lambda = \sum_{i=1}^M \lambda_i$. On the contrary, the sum of M independent Poisson process $\{M_i(t), t \geq 0\}$ ($i=1, 2, \dots, M$) with the intensity λ_i is the Poisson process $\{N(t), t \geq 0\}$.

In the observation interval of $(0, t]$, the Poisson point process $\{N(t), t \geq 0\}$ is the homogeneous Poisson process with the intensity λ whose inter-point spacing $\{T_n\}$ is independent random variable sequence of exponential distribution (parameter is λ). This means that the interval of haze occurrence obeys the exponential distribution with a parameter of λ . In the same way, in the observation interval of $(0, t]$, the number of haze occurrence with the intensity λ_i is the Poisson process $\{M_i(t), t \geq 0\}$ ($i=1, 2, \dots, M$) with the intensity λ_i . The inter-point spacing $\{T_n^i\}$ between two adjacent haze occurrences with the intensity λ_i is a sequence of independent random variables, and the distribution of $\{T_n^i\}$ is also an exponential distribution with a parameter of λ_i ($i=1, 2, \dots, M$).

3. Model calculations

The data analyzed in this paper comes from the Beijing Environmental Protection Monitoring Center. The time range of the original data is from January 1, 2014 to December 31, 2016. The data is the daily air condition of an observation point in the center of Beijing, and the current data is classified according to the season. In the time range of $(0, t]$, we record the Air Quality Index as the marked point H of One-dimensional Poisson marked point process model during one of haze events. First of all, the marked point H is defined as: $h_1: 100 < \text{AQI} \leq 150$, mild pollution; $h_2: 150 < \text{AQI} \leq 200$, moderate pollution; $h_3: 200 < \text{AQI} \leq 300$, heavy pollution; $h_4: \text{AQI} > 300$, severe pollution. According to the definition, the number of haze events with a marked point h_i is a Poisson process $\{M_i(t), t \geq 0\}$ ($i=1, 2, 3, 4$) with the intensity λ_i , and the number of haze events with marked point

$H > 100$ is a Poisson process $\{N(t), t \geq 0\}$ with the intensity $\lambda = \sum_{i=1}^4 \lambda_i$.

In the period from January 1, 2014 to December 31, 2016 about 1096 days for three years, the haze phenomena about totally 172 times of all types occurred. From the parameter estimation formula:

$$\hat{\lambda} = N(T)/T, \quad \hat{\lambda}_i = N_i(T)/T, \quad (i=1, 2, 3, 4) \quad (3)$$

where $T = 1096$, $N(T)$, $N_i(T)$ is the number of haze occurrence with the intensity λ and λ_i .

The estimated value of Poisson process intensity λ_i about Beijing annual and seasonal haze occurrences of various types Poisson process intensity

	λ_1	λ_2	λ_3	λ_4
Whole year	0.0437	0.0538	0.0374	0.0219
Spring	0.0507	0.0688	0.0362	0.0109
Summer	0.0580	0.0688	0.0217	0.0109
Autumn	0.0403	0.0256	0.0477	0.0293
Winter	0.0185	0.0406	0.0296	0.0406

According to the definition of Poisson process, the formula for the probability of haze occurrence with the intensity λ_i is:

$$p_i = \sum_{k=1}^{\infty} P\{M_i(t) = k\} = \sum_{k=1}^{\infty} \frac{(\lambda_i t)^k}{k!} e^{-\lambda_i t} = 1 - P\{M_i(t) = 0\} = 1 - e^{-\lambda_i t} \quad (4)$$

As a result, it is possible to calculate the probability of haze occurrence with monthly intensities that are respectively like λ_i ($i = 1, 2, 3, 4$) as follows:

The probability of haze occurrence with various types monthly in Beijing ($t=30$ days)

	h_1	h_2	h_3	h_4
Whole year	0.7304	0.8009	0.6744	0.4816
Spring	0.7815	0.8730	0.6624	0.2789
Summer	0.8245	0.8730	0.4785	0.2789
Autumn	0.7015	0.5361	0.7609	0.5846
Winter	0.4259	0.7042	0.5885	0.7042

For haze pollution problems, the harm of haze above severe must be much greater than that of light and moderate haze. Severe haze can lead to flight cancellations, traffic accidents and acute respiratory diseases. Therefore, the analysis of severe haze above has important practical significance. According to the superposition of homogeneous Poisson process, the intensity of haze above heavy is $\lambda_{zhong} = \lambda_3 + \lambda_4$. The estimated haze intensity for severe haze over the whole year and in each season of spring, summer, autumn and winter is calculated. The result is as follows:

The probability of severe haze occurrence monthly in Beijing ($t=30$ days)

Whole year	Spring	Summer	Autumn	Winter
0.8312	0.7566	0.6239	0.9077	0.8727

In the observation interval of $(0, t]$, the Poisson point process $\{N(t), t \geq 0\}$ is the homogeneous Poisson process with the intensity λ whose inter-point spacing $\{T_n\}$ is independent random variable sequence of exponential distribution (parameter is λ). The number of haze occurrence $\{M_i(t), t \geq 0\}$ ($i = 1, 2, \dots, M$) with the intensity λ_i is the Poisson process with the intensity λ_i , and the inter-point spacing $\{T_n^i\}$ between two adjacent haze occurrences with the intensity of λ_i is a sequence of independent random variables. The distribution of $\{T_n^i\}$ is also an exponential distribution with a parameter of λ_i ($i = 1, 2, 3, 4$), and $E(T_n) = 1/\lambda$, $\lambda = \sum_{i=1}^4 \lambda_i$.

We can find the average number of consecutive days for which no haze occurs throughout the year and seasons. The result is as follows:

The average number of consecutive days that haze does not occur in Beijing (Unit: days)

Whole year	Spring	Summer	Autumn	Winter
6.37	6	6.63	7	7.78

We can also calculate the total days that haze above severe occurred in the observation interval of $(0, t]$. We denote that the number of severe haze occurrences in the range of $(0, t]$ is a Poisson process

$\{M(t), t \geq 0\}$ with the intensity of λ_{Zhong} . And we can obtain $T = \sum_{k=1}^{M(t)} Z_k$ where Z_1, Z_2, \dots, Z_k is an independent, discrete random variable whose distribution law is:

Z	1	2	...	k
P	p_1	p_2	...	p_k

where $p_1 + p_2 + \dots + p_k = 1$, and Z_i indicates the number of days that a severe haze lasts on the i -th time, $i = 1, 2, \dots, k$.

$$E(T) = E[M(t)]E(Z) = \lambda t \sum_{i=1}^k i \cdot p_i \quad (5)$$

As a result, the average days of haze occurrence above severe in one year and in spring, summer, autumn and winter are as follows:

The continuous average days of haze occurrence above severe in Beijing (Unit: days)				
Whole year	Spring	Summer	Autumn	Winter
94.8	21.8	17.8	27.7	27.5

4. Conclusion

For the first time, through using the haze intensity as a marked point, this paper applies the random point process theory to the research of practical problems of haze and defines the Poisson point process model of the single-parameter value of haze. The model was used to study the haze data in Beijing, the probability of occurrence of various types of haze and the important numerical characteristics of severe haze occurrence were obtained. Through studying the model, we found that the Poisson point process model of the haze phenomenon established in this paper is a very effective model to study the random characteristics of haze weather. Many of the random features of haze we want to know can be calculated through model analysis. If there are more haze sample data, multivariate point Poisson process models can also be established, which can be used to refine the different types, analyze and compare the results of different parameter estimation methods. As a result, the model will be more accurate, the analysis will be more complete, and the conclusion will be more realistic.

References

- [1] Deng Yonglu, Liang Zhishun. Random point process and its application [M]. Beijing: Science Press, 1992: 1-319
- [2] Zhu Yonghua. A Poisson marked point process model and its properties in flood risk analysis [J]. Engineering Journal of Wuhan University, 2000, 33(3): 33-35
- [3] Zhu Yonghua, Dong Yajuan. A two-dimensional marked point Poisson process model in flood risk analysis [J]. Acta Mathematica Scientia, 2011, 21(4): 527-532
- [4] Sun Liang. Hazard and Treatment of Haze Weather [J]. Environmental Science and Management, 2012, 37(10): 71-75
- [5] Wu Dui. 2012. Haze weather research in China in the last decade: A review [J]. Acta Scientiae Circumstantiae, 32 (02): 257-269.
- [6] Dong X L, Liu D M, Yuan Y S, et al. Pollution characteristics and influencing factors of atmospheric particulates in Beijing during the summer of 2005. Chinese Journal of Environmental Engineering, 2007, 1(9): 100-104
- [7] Zhang Y, Li G, et al. The spatial characteristics of ambient particulate matter and daily mortality in the urban area of Beijing, China. Science of the Total Environment, 2012, 435: 14-20.