Journal of Statistical Modelling: Theory and Applications
Vol. 3, No. 2, 2022, pp. 39-50
Yazd University Press 2022

JSMTA

Research Paper

Predicting intensity function of nonhomogeneous Poisson process

Masoud Yarmohammadi*[1], Ada Afshar[1], Rahim Mahmoudvand[2], Parviz Nasiri[1]
[1]Department of Statistics, University of Payame Noor, 19395-4697, Tehran, Iran
[2]Department of Statistics, Bu-Ali Sina University, Hamedan, Iran

Received: April 16, 2023/ Revised: June 17, 2023/ Accepted: June 18, 2023

Abstract: The nonhomogeneous Poisson process is commonly utilized to model the occurrence of events over time. The identification of nonhomogeneous Poisson process relies on the intensity function, which can be difficult to determine. A straightforward approach is to set the intensity function to a constant value, resulting in a homogeneous Poisson process. However, it is crucial to assess the homogeneity of the intensity function through an appropriate test beforehand. Failure to confirm homogeneity leads to an infinite-dimensional problem that cannot be comprehensively resolved. In this study, we analyzed data on the number of passengers using the Tehran metro. Our homogeneity test showed a nonhomogeneous arrival rate of passengers, prompting us to explore different functions to estimate the intensity function. We considered four functions and used a piecewise function to determine the best intensity function. Our findings showed significant differences between the two models, highlighting the effectiveness of the piecewise function model in predicting the number of metro passengers.

Keywords: Hypothesis testing; Intensity function; Nonhomogeneous Poisson process; Poisson process.
Mathematics Subject Classification (2010): 93D05, 60G55, 11J20, 62F03.

1 Introduction

Counting processes are a type of random process that takes on values which are always correct, non-decreasing, and non-negative values. One of the most commonly used counting processes is the Poisson process, which models the occurrence of random

*Corresponding author: masyar@pnu.ac.ir

events over time. In a homogeneous Poisson process with a constant rate parameter λ , the time intervals between events are independent and follow an exponential distribution. However, this assumption of a constant rate may not hold true in real-world scenarios where there are events that occur at varying rates over time. To address this limitation, a nonhomogeneous Poisson process (NHPP) has been developed as a more general version of the homogeneous Poisson process. NHPP is particularly useful for analyzing events whose occurrence rate changes over time. It allows for a variable rate parameter that can capture the changing nature of the process.

The NHPP is commonly used in a wide variety of events. Several examples include seismographic event modeling (Hong and Guo, 1995; Sumiati et al., 2019), repairable damages (Lindqvist, 2006), daily measurements of ozone gas (Vicini et al., 2012), ambient noise emissions (Guarnaccia and Quartieri, 2014), evaluating the buying behavior of customers (Letham et al., 2016), software reliability and improvement (Wang et al., 2016; Shinde and Kumar, 2017), and the number of accidents (Grabski, 2017, 2018) can be listed.

Several methods have been proposed for estimating the parameters of the intensity function in univariate NHPPs. These include parametric techniques (Zhao and Xie, 1996), Bayesian methods (Guida et al., 1989) and non-parametric approaches (Law and Kelton, 1991).As the field of NHPPs has expanded, additional studies have been conducted to develop and refine these methods. For example, Fathi and Khoshkar (2014) devised an algorithm for simulating NHPPs with a multivariate intensity function and evaluated its performance. Cifuentes-Amado and Cepeda-Cuervo (2015) proposed a new class of NHPPs that consider seasonal factors and improve the estimation of patient admissions.

When modeling a NHPP, it is necessary to estimate both the unknown parameters and the functional form of the intensity function. The functional form of the intensity function is usually determined based on the initial description of the data under investigation, which may take various forms, such as periodic, ascending, or descending. For data related to repairable systems, the Poisson process with a power intensity function is often applied as a common approach (Karbasian and Ibrahim, 2010). In recognizing the functional form of intensity function, polynomial functions ($\lambda(t) = a_0 + a_1 t + a_2 t^2 + ... + a_n t^n ; \lambda(t) \geq 0$) and periodic functions ($\lambda(t) = [A \sin(\omega t + \phi) + c] ; \lambda(t) \geq 0$) are among the well-known and common families considered for intensity functions in this process. Moreover, the NHPP for modeling the number of passengers in public transportation systems has been investigated. Moreira-Matias et al. (2013) modeled the demand for taxi passengers using a NHPP. To model the number of passengers entering the Valparaiso subway line in Chile, Allende-Bustamante et al. (2016) used two types of counting processes, including the Hawkes-Phan process and the NHPP, with the intensity function viewed as a sum of normal functions. The results indicated that estimating the intensity function in the NHPP was more straightforward and required fewer calculations than the Hawkes-Phan process; however, the Hawkes-Phan process provided more accurate estimates of points with high variability (including passenger peak hours). Menon and Lee (2017) employed the NHPP to model the passenger flow in short time intervals whose intensity function was determined by a single-layer neural network and, in a case study, with the data of bus passengers entering an area. A large city in Australia was used to investigate the model.

Apart from studies that have used NHPP as a modeling tool, several other studies have been conducted to predict the number of passengers in rail transportation systems. These studies can be classified into two categories: parametric and non-parametric methods. Parametric methods such as autoregressive integrated moving average (ARIMA) models, multiple regression, exponential smoothing methods, gray models, and Markov models are commonly used to predict the number of passengers. Non-parametric methods such as neural network and nearest k-neighbor are also employed for this purpose. One such study by Sun and Mou (2018) proposed a composite model for predicting the volume of passenger traffic in high-speed rail transportation systems. The model combined partial least squares regression (PLS) and gray models, and was evaluated based on the sum of squared errors.

In addition to analyzing rail passenger flow, Milenkovic et al. (2018) used the seasonal ARIMA (SARIMA) model to predict the number of passengers in the Serbian rail network using data from January 2004 to June 2014, and the results demonstrate the model's excellent performance. In the section on non-parametric methods, Wang et al. (2019) predicted the number of passengers in the Qingdao rail transportation system using the generalized regression neural network model. Borucka and Guzanek (2022) have predicted the number of passengers in the rail transport system using data from Polish railway lines and exponential smoothing and ARIMA models. According to the findings of their study, the ARIMA model is superior at estimating the trend and seasonal component.

As stated in the background, different methods, such as time series analysis methods, can be used to predict the number of passengers. However, these methods may not be suitable for analyzing the data under consideration, which involves counting processes that do not occur at a constant rate over time. Therefore, the approach taken in this paper is to use the NHPP method to predict the number of passengers entering the Tehran metro.

The structure of this paper is as follows: Section 2 provides an overview of the fundamental theories of the NHPP, as well as a statistical hypothesis test to determine whether the number of passengers is homogeneous or inhomogeneous. We also outline a method for estimating the intensity function of the NHPP. In Section 3, we apply the homogeneity test to the number of passengers and apply the univariate NHPP to estimate that number on the Tajrish-Kahrizak metro line in Tehran using two different methods and compare their accuracy in terms of RMSE and MAPE criteria. Finally, Section 4 concludes the paper and summarizes the key findings and implications of the study.

2 Poisson process

2.1 Counting process

The counting process $\{N(t), t \geq 0\}$ is a random process with the following conditions:
1. $N(t) \geq 0$;
2. $N(t)$ is integer valued;
3. If $s < t$ then $N(s) \leq N(t)$ and

4. For $s < t$, then $N(t) - N(s)$ equals the number of occurrences in the interval $(s, t]$. (Ross, 2014)

If $t_1 > t_0$, in the counting process $\{N(t), t \geq 0\}$, then, $N(t_1) - N(t_0)$ is the increments of the process for in time Δt, and if $t_n > ... > t_1 > t_0$, the random variables $N(t_1) - N(t_0), N(t_2) - N(t_1), ..., N(t_n) - N(t_{n-1})$ are independent of one another, the process has $N(t)$ independent increments.

The following subsection discusses the homogeneous Poisson process for counting random events over time.

2.2 Homogeneous Poisson process

A counting process $\{N(t), t \geq 0\}$ is called a homogeneous Poisson process with the rate λ ($\lambda > 0$) if the following conditions apply:
1. $N(0) = 0$;
2. $N(t)$ has independent increments;
3. $P(N(t+h) - N(t) = 1) = \lambda h + o(h)$ and
4. $P(N(t+h) - N(t) \geq 2) = o(h)$, where $h > 0$ and $o(h)$ is a small amount applied in the condition $\lim_{h \to 0} \frac{o(h)}{h} = 0$. (Ross, 2014)

Under the conditions stated above, the distribution of events in the time interval t follows the Poisson distribution with the rate λt. In other words, if $N(t)$ is a Poisson process with the rate λ ($\lambda > 0$), then for every $t, s > 0$, $N(t+s) - N(s)$ is a Poisson random variable with a mean λt.

The assumption underlying the homogeneous Poisson process is a constant rate of occurrence. This assumption may not hold for the majority of natural phenomena. Therefore, the homogeneous Poisson process cannot be utilized. One can model the intensity function as a random variable and estimate it using Bayesian methods to solve this issue. Another approach is considering the intensity function as a random process and employing multiple random processes. However, both of these methods require adding a complex probabilistic structure. A more straightforward solution is to use the NHPP.

2.3 Non-Homogeneous Poisson process

The process $\{N(t), t \geq 0\}$ is a non-homogeneous Poisson process (NHPP) with intensity function $\lambda(t)$ if the following four conditions apply:
1. $N(0) = 0$;
2. $N(t)$ has independent increments;
3. $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$ and
4. $P(N(t+h) - N(t) \geq 2) = o(h)$, where $h > 0$ and $o(h)$ is a small amount that satisfied in condition $\lim_{h \to 0} \frac{o(h)}{h} = 0$.

The mean function of the NHPP $\{N(t), t \geq 0\}$ with the intensity function $\lambda(t)$ is defined as Equation 1:

$$m(t) = \int_0^t \lambda(x) \, \mathrm{d}x. \tag{1}$$

In addition, according to Equation 1, the estimate of the average $N(t)$ is equal to:

$$\hat{m}(t) = E(N(t)) = \int_0^t \hat{\lambda}(x)dx, \tag{2}$$

where $E(.)$ is the expected value.

Given the conditions outline above, it can be shown that for each $t, s > 0$, $N(t+s) - N(s)$ has the Poisson random variable with mean given by Equation 3 (Ross, 2014):

$$m(t+s) - m(s) = \int_s^{t+s} \lambda(x)\ \mathrm{d}x. \tag{3}$$

As a result, $\{N(t), t \geq 0\}$ generally has a nonhomogeneous Poisson distribution with the distribution of Equation 4:

$$P(N(t) = k) = \frac{(\int_0^t \lambda(x)\ \mathrm{d}x)^k}{k!} e^{-\int_0^t \lambda(x)\ \mathrm{d}x} \quad k = 0, 1, 2, \ldots. \tag{4}$$

Before addressing the estimation of the intensity function, we should formally examine the assumption of homogeneity so that if the data are accepted as inhomogeneous, the NHPP can be applied to the examined data. Consequently, a test to examine the homogeneity of observations is presented in the subsequent section.

2.4 Poisson process homogeneity test

Brown and Zhao (2002) proposed a new homogeneity test based on Anscombe's variance stability transformation (1948), which outperforms the likelihood ratio, conditional chi-square, and Niemen-Scott tests. In this paper, we employ the test developed by Brown and Zhao (2002) to examine the homogeneity of the metro passenger arrival rate. If $N_1(t), N_2(t), ..., N_n(t)$ are the independent random variables and are non-negative and integers, the following test can be utilized:

$$H_0 : N_i \sim Poiss(\lambda_i), \qquad \lambda_1 = \lambda_2 = \cdots = \lambda_n,$$

Versus the assumption

$$H_1 : N_i \sim Poiss(\lambda_i), \sum_{i=1}^{n} (\lambda_i - \bar{\lambda})^2 > 0.$$

Anscombe (1948) showed that if $N \sim Poiss(\lambda)$, then:

$$Var_\lambda(\sqrt{N + \frac{3}{8}}) = \frac{1}{4} + o(\frac{1}{\lambda}).$$

Using this information and by defining $Y_i = \sqrt{N_i + \frac{3}{8}}$ and considering the test statistic as follows:

$$T = 4 \sum (Y_i - \bar{Y})^2 \sim \chi_{(n-1)}(4 \sum (\nu(\lambda_i) - \bar{\nu}_n)^2). \tag{5}$$

H_0 is tested against H_1; wherein

$$\bar{\nu}_n = \frac{1}{n}\sum \nu(\lambda_i), \nu(\lambda_i) = E_{\lambda_i}(Y_i) = E_{\lambda_i}(\sqrt{N_i + \frac{3}{8}}).$$

and so H_0 is rejected if $T > \chi^2_{n-1;1-\alpha}$.

2.5 Estimation of the intensity function

The intensity function, plays a crucial role in characterizing the probabilistic behavior of the NHPP. Maximum likelihood (ML) is a well-established technique for estimating the parameters of the model. To estimate the intensity function, we must consider the nature, limitations, periodicity, and non-periodicity of the data. Commonly used functions for intensity function estimation include trigonometric, polynomial, exponential, trigonometric exponential, and spline functions. However, when dealing with non-negative counting data, the exponential function is often preferred due to its ability to ensure non-negativity. In this paper, we evaluate both single and piecewise functions, which are represented by Equations 6 and 7, respectively:

$$\lambda(t) = \exp(\beta_0 + \sum_{i=1}^{p} \beta_i t^i), \tag{6}$$

$$\lambda(t) = \begin{cases} \exp(\beta_{01}t^2 + \beta_{11}t), & t = \text{spring}, \\ \exp(\beta_{02}t^2 + \beta_{12}t), & t = \text{summer}, \\ \exp(\beta_{03}t^2 + \beta_{13}t), & t = \text{fall}, \\ \exp(\beta_{04}t^2 + \beta_{14}t), & t = \text{winter}. \end{cases} \tag{7}$$

To estimate the intensity function, we utilized the maximum likelihood (ML) method. We considered various forms of the intensity function, including trigonometric, polynomial, exponential, trigonometric exponential, and spline functions, but the details of these functions are not discussed in this paper. However, we found that applying these functions produced either constant or negative estimates, which is not in line with the characteristics of the data we studied.

3 Application

The impact of passenger's volume on subway planning, costs, optimization, and service improvement highlights the need for modeling and forecasting. The number of passengers entering the Tehran metro is registered hourly in an online system for each station between 5 am and midnight. This study considered the data set from March 20th, 2016, to March 19th, 2020, before the COVID-19 pandemic in Iran.

Figure 1 shows the monthly number of passengers who entered the Tehran metro between 7 and 8 am. The intensity of passenger arrivals exhibits periodic behavior. For instance, the number of passengers decreases during the Nowruz holiday in April. In contrast, the number of passengers increases in October and November at the start of the academic year.

In addition, Figure 2 illustrates the intensity of passenger arrivals from March 2016 to March 2020. As shown in the Figure 2, the exponential functions given in Equations 6 and 7 appear to be well suited for the data at hand.

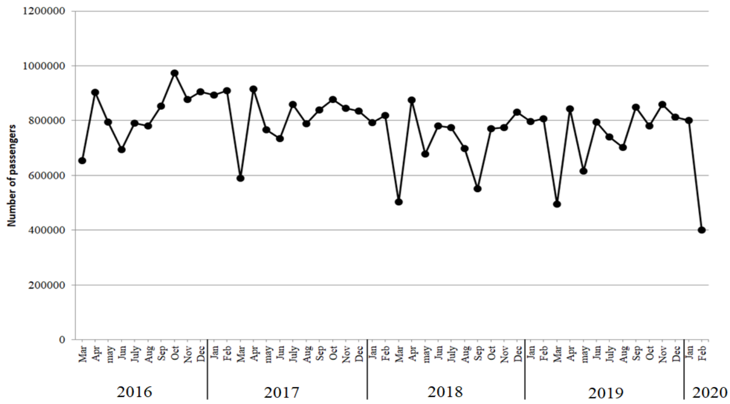

Figure 1: The number of passengers entered Tehran metro between 7 and 8am from March 2016 to March 2020.

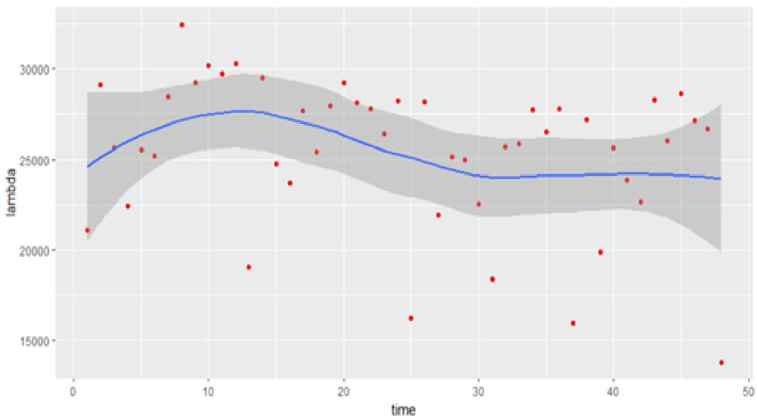

Figure 2: The intensity of the arrival of passengers from March 2016 to March 2020.

This section begins by testing the homogeneity of passenger arrival rates. Next, we apply the NHPP to model and predict the number of passengers in the Tehran metro, using the monthly passenger data from the Tajrish-Kahrizak line between March 2016 and March 2020 as a case study. We employ two methods for this purpose and compare their outcomes.

3.1 Comparison criteria

For a practical comparison of the two fitted models, the data is divided into two parts:

The first part consists of the data used for data modeling from March 2016 to March 2020. The second one consists of the data from March 2019 to March 2020 used to validate and compare the models.

Using the root mean square error (RMSE) and the mean absolute percentage error

(MAPE), a more suitable model is then suggested. The following relationships illustrate these criteria.

• The Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(N(t_i) - \hat{N}(t_i)\right)^2}. \tag{8}$$

• The Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\left(N(t_i) - \hat{N}(t_i)\right)\right|}{N(t_i)} \times 100. \tag{9}$$

In Equations 8 and 9, $N(t_i)$ is the total number of passengers in time t_i, $\hat{N}(t_i)$ is its estimate, and also n is the number of observations.

3.2 Homogeneity test of passenger arrival rate

To test the homogeneity of passenger arrival rate, we consider the monthly total number of passengers from March 2016 to March 2020 and calculate the test statistic of Equation 5 by converting the data into $Y_i = \sqrt{N_i + \frac{3}{8}}$:

$$T = 4\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = 131990.5.$$

Considering $\alpha = 0.05$ that

$$T = 131990.5 > \chi^2_{35;0.95} = 22.465.$$

Therefore, at a significance level of 0.05, the homogeneity hypothesis of the total number of metro passengers is rejected, and the NHPP can be used to model the data in Subsections 3.3 and 3.4.

3.3 The first model

The first model considered the total number of passengers from March 2016 to March 2019, and the intensity function was estimated using Equation 6. The best model was obtained as follows:

$$\hat{\lambda}(t) = \exp(-0.026t^2 + 1.12t).$$

In Table 1, RMSE value is 1186486.48, which is a large value. Additionally, MAPE equals 176%, representing the deviation from the total number of incoming passengers.

Using Equations 2 and 4, the mean function and the distribution of the NHPP using the estimated model are obtained as follows:

$$\hat{m}(t) = \int_0^t \hat{\lambda}(t)dx = \int_0^t \exp(-0.024t^2 + 1.09t)dt.$$

Table 1: Comparative Criteria of the First Model

criteria	*Amount*
$RMSE$	1186486.48
$MAPE$	176.89

$$\begin{aligned} P(\hat{N}(t)=k) &= \frac{\hat{m}(t)}{k!}e^{-\hat{m}(t)} \\ &= \frac{1}{k!}(\int_0^t \exp(-0.024t^2+1.09t)dt)^k e^{\int_0^t \exp(-0.024t^2+1.09t)dt}. \end{aligned}$$

Considering that the time t from March 2016 to March 2019 is numbered from 1 respectively, for instance, to predict the average number of passengers entering the Tehran metro line 1 between 7 and 8 am on March 21st, 2021, to April 20th, 2021 for $t = 61$ and using Equation 2, yields:

$$\hat{m}(61) = E(\hat{N}(61)) = \int_0^{61} \exp(-0.024t^2 + 1.09t)dt = 2712163.$$

This estimate is far from its actual value $(281, 463)$.

3.4 The second model

In the second model, the intensity function is estimated by dividing the number of passengers from March 2016 to March 2019 into spring, summer, fall, and winter seasons using Equation 7, the best estimate of the intensity function as a piecewise function for each season. The following is obtained:

$$\hat{\lambda}(t_i) = \begin{cases} \exp(-0.3725t^2 + 4.2339t), & t = \text{spring}, \\ \exp(-0.035t^2 + 1.208t), & t = \text{summer}, \\ \exp(-0.0133t^2 + 0.741t), & t = \text{fall}, \\ \exp(-0.0064t^2 + 0.515t), & t = \text{winter}. \end{cases}$$

Considering that each season consists of three months and that the test data range from March 2016 to March 2019, the time t contains the values for each season.

Table 2: Comparative criteria of the second model

criteria	*spring*	*summer*	*fall*	*winter*
$RMSE$	244878.62	422532.08	451098.22	218625.27
$MAPE$	36.59	50.65	60.02	21.50

In Table 2, the spring model's RMSE and MAPE are less than other seasons.

If the objective is to predict the mean number of passengers entering Tehran metro line 1 between 7 and 8 am between March 21st, 2021, and April 20th, 2021, for example, $t = 16$ and then:

$$\hat{m}(16) = E(\hat{N}(16)) = \int_0^t \hat{\lambda}(x)dx$$

$$= \int_0^{16} \exp(-0.3725t^2 + 4.2339t)dt = 487446.2 \cong 487446.$$

This prediction is significantly more accurate than the forecast obtained from the first model compared to the actual outcome (281463).

Comparing Tables 1 and 2 reveals that the second model has smaller RMSE and MAPE values. In addition, the calculation of the percentage improvement in the mean estimate of the number of passengers from March 21st to April 20th, 2021, reveals an improvement of 82% over the first model. By considering the intensity function in a multi-rule manner and by separating the seasons of the year, the estimation of the intensity function has become more precise. The actual value and prediction of the number of passengers in the first and second models for the data from March 2019 to March 2020, which form the basis of the validity check, are plotted in Figure 3 to facilitate a comparison of the outcomes.

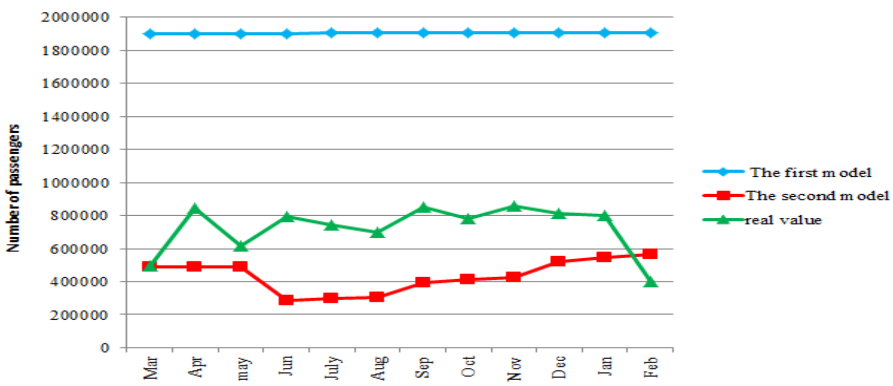

Figure 3: Comparison of the mean prediction of the number of incoming passengers in the two models presented with the actual value

4 Discussion and conclusions

In this paper, we propose a NHPP-based model for estimating the number of passengers entering the Tehran metro. The monthly number of metro passengers can be modeled as a NHPP because it fluctuates and is independent of the previous month. We demonstrated that the data set is not homogenous using Anscombe's homogeneity test. In a case study, the intensity function was estimated using two models based on the number of passengers entering line 1 of the Tehran metro between 7 and 8 am from March 2016 to March 2019. The results demonstrate that the piecewise function intensity function is more accurate and has the lowest RMSE and MAPE.

Acknowledgement

The authors are grateful to the Editor in Chief and referees for the helpful comments and suggestions to improve this paper.

References

Allende-Bustamante, J., Yuz, I. and Rodo, J. (2016). Application of point processes estimation to a metro system. In *2016 Australian Control Conference (AuCC)*, 232–237.

Anscombe, F.J. (1948). The transformation of Poisson, binomial and negative-binomial data. *Biometrika*, **35**:246–254.

Borucka, A. and Guzanek, P. (2022). Predicting the seasonality of passengers in railway transport based on time series for proper railway development. *Transport Problems*, **17**(1):51–61.

Brown, L.D. and Zhao, L.H. (2002). A test for the Poisson distribution. *Indian Journal of Statistics,* **64**(3):611–625.

Cifuentes-Amado, M.V. and Cepeda-Cuervo, E. (2015). Non-homogeneous Poisson process to model seasonal events: Application to the health diseases. *International Journal of Statistics in Medical Research,* **4**(4):337–346.

Fathi-vajargah, B. and Khoshkar-foshtomi, H. (2014). Simulating non-homogeneous Poisson point process based on multi-criteria intensity function and comparison with its simple form. *Journal of Mathematics and Computer Science,* **9**:133–138.

Guarnaccia, C. and Quartieri, J. (2014). Modeling environmental noise exceedances using non-homogeneous Poisson processes. *Journal of the Acoustical Society of America,* **136**(4):1631–1639.

Guida, M., Calabria, R. and Pulcini, G. (1989). Bayes inference for a non-homogeneous Poisson process with power intensity law (reliability). *IEEE Transactions on Reliability,* **38**(5):603–609.

Grabski, F. (2017). Non-homogenous Poisson process application to modeling accidents number at Baltic sea waters and ports. *Journal of Polish Safety and Reliability Association,* **8**(1):39–47.

Grabski, F. (2018). Non-homogeneous stochastic processes connected to Poisson process. *Scientific Journal of Polish Naval Academy,* **2**(213):5–15.

Hong, L.L. and Guo, S.W. (1995). Nonstationary Poisson model for earthquake occurrences. *Bulletin of the Seismological Society of America,* **85**(3):814–824.

Karbasian, M. and Ibrahim, Z. (2010). Estimation of parameters of the power-law-non-homogenous poisson process in the case of exact failures data. *International Journal of Industrial Engineering and Production Research,* **21**(2):105–110.

Law, A.M. and Kelton, W.D. (1991). *Simulation Modeling and Analysis.* New York: McGraw-Hill.

Letham, B., Letham, L.M. and Rudin, C. (2016). Bayesian inference of arrival rate and substitution behavior from sales transaction data with stockouts. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 1695–1704.

Lindqvist, B.H. (2006). On the statistical modeling and analysis of repairable systems. *Statistical Science,* **21**(4):532–551.

Menon, A.K. and Lee, Y. (2017). Predicting short-term public transport demand via inhomogeneous Poisson processes. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* , 2207–2210.

Milenkovic, M., Svadlenka, L., Melichar, V., Bojovic, N. and Avramovic, Z. (2018). SARIMA modelling approach for railway passenger flow forecasting. *Transport,* **33**(5):1113–1120.

Moreira-Matias, L., Gama, J., Ferreira, M., Mendes-Moreira, J. and Damas, L. (2013). Predicting taxi–passenger demand using streaming data. *IEEE Transactions on Intelligent Transportation Systems,* **14**(3):1393–1402.

Ross, S.M. (2014). *Introduction to Probability Models.* Eleventh edition, USA: Academic Press.

Shinde, V. and Kumar, J. (2017). Enhance non-homogeneous Poisson process models incorporating testing effort with coverage function. *Journal of Statistics and Management Systems,* **20**(3):297–308.

Sumiati, I., Rahmani, U., Supian, S. and Subiyanto, S. (2019). Application of the nonhomogeneous poisson process for counting earthquakes. *World Scientific News,* **127**(3):163–176.

Sun, L. and Mou. H.B. (2018). A high-speed railway traffic forecast based on IOWA combined model. *Railway Transport And Economy,* **40**(9):74–79.

Vicini, L., Hotta, L.K. and Achcar, J.A. (2012). Non-homogeneous Poisson processes applied to count data: a bayesian approach considering different prior distributions. *Journal of Environmental Protection,* **3**(10):1336–1345.

Wang, J., Wu, Z., Shu, Y. and Zhang, Z. (2016). An optimized method for software reliability model based on non-homogeneous Poisson process. *Applied Mathematical Modelling,* **40**:6324–6339.

Wang, X.F., Zhu, Y.Q. and Pan, F.Q. (2019). Prediction of railway passenger volume based on GRNN neural network. *Journal of Luoyang Institute of Science and Technology,* **29**(1):40–50.

Zhao, M. and Xie, M. (1996). On maximum likelihood estimation for a general nonhomogeneous Poisson process. *Scandinavian Journal of Statistics,* **23**(4):597–607.