Advances in Social Science, Education and Humanities Research (ASSEHR), volume 199
2018 International Conference on Sports, Arts, Education and Management Engineering (SAEME 2018)

Selection of Ship Traffic Accidents Forecasting Models in Downstream Waterway of Yangtze River

Shukui Zhang[a], Keyin Miao[b], Jing Li[c]

Navigational Department Jiangsu Maritime Institute, Nanjing 211170, Jiangsu, China.

[a]zhangshkfy@163.com, [b]2091391569@qq.com, [c]158142476@qq.com

Keywords: Waterway safety, accidents forecasting, Poisson models, analysis.

Abstract. In order to explore the key factors causing vessel traffic accidents, historical accident data and vessel traffic data at 9 waterways of Yangtze River in Jiangsu are employed, and the accidents are verified to follow lognormal distribution. Five Poisson-lognormal accident forecasting models are established taking the vessel traffic accidents from 2009 to 2014 as dependent variables and vessel traffic, waterway length, and design consistency (including waterway depth, waterway width, curvature changing rate, vessel speed difference and vessel speed changing rate) as explanatory variables. Among the five models, the one with vessel speed changing rate as design consistency measure has the best effect.

1. Introduction

There are many traffic accidents in the downstream of Yangtze River r very year, and its security situation is severely. According to statistics from Jiangsu Maritime Bureau, there were 48 common and above-grade water traffic accidents in Jiangsu section of Yangtze River in 2014, among them, 34 occurred in major waterways. The data shows that the safety of important waterways in the downstream of Yangtze River is in urgent need of improvement.

In terms of accident prediction, support vector machines [1], grey system theory [2], Markov [3], fractal theory are used to predict number of future traffic accidents based on relevant historical data [4]. Although various methods have good prediction results, there are lacks of further analysis on the causes of accidents and thus no effective solutions have been proposed. Recently, Poisson lognormal model has been increasingly used in traffic accident prediction and have achieved very good results [5].

2. Traffic Accident Prediction Model

2.1 Poisson Model.

Suppose x_k traffic accidents occurred on the k-the waterway, and x_k follows Poisson distribution with parameter λ_k, then its probability distribution function can be expressed as:

$$P(X_k = x_k) = P(x_k) = \frac{\lambda_k^{x_k} e^{-\lambda_k}}{x_k\,!} \tag{1}$$

Here λ_k is average value of X_k, and called Relationship function. Traffic accident prediction factors are composed of waterway characteristics variables, the form of traffic accident prediction function is determined by traffic accident prediction factors and relation function. If accident prediction factors are linear, then

$$\lambda_k = esp(\theta_0 + \theta_1 y_{k1} + \theta_2 y_{k2} + \cdots + \theta_k y_{ki}),\ k = 1, 2, \cdots, n \tag{2}$$

Here $\theta_0, \theta_1, \cdots, \theta_k$ are coefficient, $y_{k1}, y_{k2}, \cdots, y_{ki}$ are waterway characteristics variables.

If waterway characteristics variables composed of traffic accident prediction factors are logarithmic, then Poisson model is

$$\lambda_k = \exp(\theta_0 y_{k1}^{\theta_1} y_{k2}^{\theta_2} \cdots y_{ki}^{\theta_i}), \mathrm{k} = 1, 2, \cdots, n \tag{3}$$

2.2 Poisson Lognormal Model.

Divide traffic accidents that occur in n waterway into m categories, define vector $X_k = \{x_{k1}, x_{k2}, \cdots, x_{km}\}$, here x_{km} indicates the number of type m traffic accidents that occurred at location k, assume that X_k is an independent Poisson distribution with a given value λ_{km}, then there is function

$$f(x_{km} \mid \lambda_{km}) = \lambda_{km}^{x_{km}} \frac{e^{-x_{km}}}{x_{km}!}, k = 1, 2, \cdots, n; m = 1, 2, \cdots, M \tag{4}$$

Now suppose $Ln\lambda_{km} = Ln\beta_{km} + \varepsilon_{km}$, and there is

$$Ln\beta_{km} = \theta_{m0} + \theta_{m1}y_{k1} + \theta_{m2}y_{k2} + \cdots + \theta_{mJ}y_{kJ} \tag{5}$$

Here $Y_{kJ} = \{y_{k1}, y_{k2}, \cdots, y_{kJ}\}$ is waterway characteristics variables, E_{km} is multiple random errors and obeys normal distribution ($E_{km} : N_m(0, \omega)$.

$$E_{km} = \begin{Bmatrix} \varepsilon_{k1} \\ \varepsilon_{k2} \\ \vdots \\ \varepsilon_{km} \end{Bmatrix}, \mathrm{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1M} \\ u_{21} & u_{22} & \cdots & u_{2M} \\ \vdots & \vdots & \cdots & \vdots \\ u_{M1} & u_{M2} & \cdots & u_{MM} \end{bmatrix}$$

Let Y is variable matrix and μ_m is regression coefficient vector, note $\mu = \{\mu_1, \mu_2, \cdots, \mu_M\}$, given (Y, μ, U), then the probability density function of M-dimensional Poisson lognormal distribution is

$$f(\lambda_k \mid \mathrm{Y}, \mu, \mathrm{U}) = \frac{e^{-0.5(\lambda_k^* - \beta_k^*)^T} U^{-1}(\lambda_k^* - \beta_k^*)}{(2\pi)^{M/2}(\prod_{m=1}^{M} \lambda_{km}) \left|U\right|^{1/2}} \tag{6}$$

Here $\lambda_k = \{\lambda_{k1}, \lambda_{k2}, \cdots, \lambda_{ki}\}$, $\lambda_k^* = \{\mathrm{Ln}\,\lambda_{k1}, \mathrm{Ln}\,\lambda_{k2}, \cdots, \mathrm{Ln}\,\lambda_{km}\}^T$, $\beta_k = \{\beta_{k1}, \beta_{k2}, \cdots, \beta_{ki}\}$, $\beta_K^* = \{\mathrm{Ln}\,\beta_{k1}, \mathrm{Ln}\,\beta_{k2}, \cdots, \mathrm{Ln}\,\beta_{km}\}^T$, when U is a diagonal matrix, the Poisson logarithmic distribution is a single variable lognormal distribution.

3. Traffic Accident Data of Important Waterways of Yangtze River

The downstream of Yangtze River is the most important channel in China, its relevant data of traffic flow, waterway length, watercourse curvature change rate and numbers of traffic accidents of them are listed in table 1, here data of traffic accidents refers to numbers of annual average accidents that occurred in 9 important waterways in 2011-2016, including the sum of accidents in near waterway entrance, exit and main body.

4. Traffic Accident Poisson Lognormal Prediction Model

4.1 Determination of Model and Model Parameters.

In the article traffic conditions, natural conditions, waterway design and traffic environment of waterway are focused on to establish model in the paper. Considering that traffic accidents have characteristics of random low probability features and their distributions obey lognormal distribution, Poisson lognormal model is used to predict the number of traffic accidents in important waterways of Yangtze River in this paper.

Aiming to the relationship between traffic accidents and influencing factors, historical accident data of downstream of Yangtze River in recent years are used to analyze the correlation with SPSS method in the paper, calculation results show that the value of correlation coefficient $|r|$ between traffic volume, waterway length, waterway depth, waterway width, watercourse curvature change rate, speed indicators and the number of traffic accidents are between 0.5 and 0.8, it shows that the indicators have significant relationship with traffic accidents, among them, traffic volume describes traffic environment conditions of waterway, and waterway length reflects the size of ship's voyage, waterway depth, waterway width and speed indicators are related to waterway design and can be used to characterize good or bad of waterway design consistency. Therefore, according to formula (4) and considering ship traffic volume T, waterway length L, and waterway design consistency, Poisson lognormal model is

$$LnA = \theta_0 + \theta_1 LnT + \theta_2 LnL + \theta_3 D_{ci} \tag{7}$$

Here D_{ci} is waterway design consistency indicator.

The consistency indicators are designed as follows: D_R , W_R , C_R , ΔV and ΔV_R , and relevant parameters in equation (6) are calculated and determined using data in Table 1,, the results are shown in table 2, as known from Table 2, the coefficients of LnT, LnL, C_R,W,W_R are all positive, it shows that with the gradual increase of changes of traffic volume, waterway length, curvature rate and ship speed, the number of traffic accidents will gradually increase. The coefficients of D_R is negative, it shows that as the depth of waterway increases, the number of traffic accidents will gradually decrease. The coefficient of W_R is also positive, it is inconsistent with the actual situation, because the wider the waterway is, the less likely it is that vessel traffic accident will occur, the reason may be that once boats encounter a wider water surface, they will not follow the traffic rules and will easily induce traffic accidents.

Based on 5 consistency indicators and relevant coefficients in Table 3, the following five accident prediction models are constructed:

Table 1. Data of 9 waterways in Yangtze River

variable	Minimum	Maximum	average value	Standard deviation
accidents logarithm $\ln A$	2.29	4.58	4.49	0.35
Traffic logarithm $\ln T$	13.44	13.57	13.48	0.07
waterway length $\ln L$	2.46	3.69	3.26	0.42
curvature change rate C_R (./km)	0	1.62	0.94	0.71
Waterway depth D_R (m)	10	40	18.4	0.69
Waterway width W_R (m)	642	3100	1250	72.4
Visibility(d/a)	89	132	100	28
ship speed difference ΔV	0	8	4.65	3.26
ship speed change rate ΔV_R	0.091	0.182	0.124	0.042

Table 2. Parameters of accident prediction models

consistency indicator / parameters	D_R	W_R	C_R	ΔV	ΔV_R
θ_0	-0.954	-0.997	-1.638	-3.105	-2.316
θ_1	0.113	0.109	0.228	0.241	0.203
θ_2	0.247	0.239	0.293	0.388	0.436
θ_3	-7.242×10-4	6.452×10-4	5.672×10-3	6.634×10-2	2.416

Model 1(D_R)

$$LnA = -0.954 + 0.113 LnT + 0.247 LnL - 7.242 \times 10^{-4} D_R \tag{8}$$

Model 2(W_R)

$$LnA = -0.997 + 0.109LnT + 0.239LnL + 6.452 \times 10^{-4} W_R \quad (9)$$

Model 3(C_R)

$$LnA = -1.638 + 0.228LnT + 0.293LnL + 5.672 \times 10^{-3} C_R \quad (10)$$

Model 4(ΔV)

$$LnA = -3.105 + 0.241LnT + 0.388LnL + 6.654 \times 10^{-2} \Delta V \quad (11)$$

Model 5(ΔV_R)

$$LnA = -2.316 + 0.203LnT + 0.436LnL + 2.435\Delta V_R \quad (12)$$

4.2 Model Selection.

Five traffic accident prediction models are evaluated respectively by using R^2 increment, significance test, F-test and $D-W$ test, and the values of various test are used as good or bad standards for judging models, the test values are shown in Table 3, as can be seen from table 3, five prediction models can better predict annual average number of traffic accidents, however, there are also differences in five models, mainly as follows: (1) If ΔV and ΔV_R are used as consistency indicators, R^2 and R^2 increments are large, the fitting of prediction model is better than that using D_R, W_R and C_R as consistency indicators, both of them , ΔV_R is the best;(2) At the 95% confidence level, D_R, W_R, ΔV and ΔV_R have significant explanation for accident, as can be seen in F test, the value of model 3 is 2.308, and the value of other models is about 7, and in significance test, the value of model 3 is 0.121, and the others are less than 0.05;(3) $D-W$ test results show that all 5 models have strong autocorrelation and their values are between 0~4. By comparing synthetically parameters of the five prediction models in table 5 it can be seen that model 5 is the best choice.

Table 3. Data of four test methods

model	Model 1	Model 2	Model 3	Model 4	Model 5
R^2 increment	0.683(0.063)	0.637(0.012)	0.648(0.019)	0.885(0.277)	0.902(0.296)
F test	7.022	6.947	2.308	6.863	6.974
significant	0.003	0.004	0.121	0.003	0.002
$D-W$ test	1.827	1.862	1.951	1.808	1.722

5. Conclusion

Based on data of ship traffic accidents from 2011 to 2016 in the downstream of Yangtze River, Poisson Lognormal forecasting model for traffic accidents is constructed by using waterway depth, curvature change rate, ship speed difference and ship speed change rate as consistency indicators in the design of waterways, and it is concluded that the model 5 is the best by comparison. It should be pointed out that the results of this study have certain reference values for the design and management of downstream waterway of Yangtze River. For other waterways, a certain amount of data is required for verification.

Acknowledgments

Jiangsu Province University Natural Science Fund Project (17KJD570001)

References

[1]. LI Jun. The research of maritime accidents based on the support vector machines [J].Wuhan: Wuhan University of Technology, 2008. (In Chinese).

[2]. Kumar C N, Parade M, Jain S S. Poisson family regression techniques for prediction of crash counts using Bayesian inference [J]. Procedia-Social and Behavioral Sciences, 2013, 104(2):982-991

[3]. Wang Z, Lu J J. Zhang Z. Exploring impacts of factors contributing to injury severity at freeway exit ramps [C]// Transportation and Development Institute Congress. Chicago, USA, 2011:1266-1285.

[4]. ONDER E,KUZU S. Forecasting air traffic volumes using smoothing techniques[J].Journal of China Aeronautics and Space Technologies,2014,7(1):65-70.

[5]. Isla R. Maximum likelihood fitting of the Poisson log normal distribution [J]. Envion Ecol Stat, 2008, 15(2):143-156.