Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2013, Article ID 597243, 8 pages
http://dx.doi.org/10.1155/2013/597243

Research Article

Risk Modelling for Passages in Approach Channel

Leszek Smolarek and Henryk Śniegocki

Gdynia Maritime University, Morska 81-87, 81-225 Gdynia, Poland

Correspondence should be addressed to Henryk Śniegocki; henryksa@wn.am.gdynia.pl

Received 22 February 2013; Revised 14 June 2013; Accepted 1 July 2013

Academic Editor: Jia-Jang Wu

Methods of multivariate statistics, stochastic processes, and simulation methods are used to identify and assess the risk measures. This paper presents the use of generalized linear models and Markov models to study risks to ships along the approach channel. These models combined with simulation testing are used to determine the time required for continuous monitoring of endangered objects or period at which the level of risk should be verified.

1. Introduction

Logistic models are arguably one of the most widely used data analysis techniques. They are the most studied discriminative models [1–3]. Logit models appear in a variety of forms in applications in biostatistics, epidemiology, economics, marketing research, and sociology. They are used to model the relationship between covariates and various types of discrete outcomes from the ubiquitous binary logit model for a two-level response to the conditional logit and multinomial (generalized) logit models concerning polytomous responses. Nested logit models allow for modeling the sequence of the decision process faced by the grouping alternatives at each stage into nests [4].

Logistic regression was applied to estimate the likelihood of mortality and severe injury in pedestrian casualties by considering the associations of such factors as demographic characteristics, injury characteristics, crash time, location, road environment, traffic control, and traffic conditions [5, 6]. Logit models were used for analyzing traffic crash severities aim at identifying and quantifying the effects of the factors which affect different crash injury severities [7, 8]. Ordered logit model were used in past studies about the great variability in conflict judgments by Air Traffic Controllers [9]. Polynomial and logistic regression and maritime transportation simulation were basis to developed an oil outflow model for collision and grounding accidents of tankers [10]. Navigation safety assessment can be carried out with the use of risk metrics analysis, multivariate statistical methods, stochastic models, and simulation methods, [11–15].

In this paper we present the possibility of using the category of ordered logit models to examine threats to vessels. Stochastic models were also proposed to allow, among other things, classification of ships with reference to the degree of risk of collision. Safe navigation requires knowledge of the whole picture of navigational, hydrometeorological situations and ship's maneuverability. An important factor affecting the safety of the vessel is the proper execution of maneuvers, [15]. Weather conditions affect safe performance of ship's maneuvers during entering the port and berthing, especially the vessels with a big surface exposed to wind. Wind direction and force are very important factors.

2. Safety of Navigation in the Approach Channels

Safety of navigation is a state of a system related to performing certain maneuvers without collisions, accidents in restricted waters. Among all the hazards we can distinguish [12]

(i) collision with another vessel or seamark in the fairway,

(ii) running aground or hitting the bottom.

Probability of activating any of these risks can be represented by the following formula:

$$\Pr = f\left(A_i, S_i, H_i, M_i, I_i, R_i\right), \tag{1}$$

(see [12]) where i is the index of hazard type, A_i are sea area parameters, S_i are vessel's parameters, H_i are hydrometeorological parameters, M_i are parameters of the performed maneuver, I_i are parameters of traffic density, and R_i are parameters of traffic control system.

It is a dependent function; the variables A_i, S_i, H_i, M_i, I_i, R_i, which reflect a number of factors, describe certain states of the system the vessel—water area—the environment.

Hydrometeorological conditions have a significant impact on the safety of navigation in restricted areas. You can, amongst other weather factors, include wind, waves, and currents.

Nature of the fluctuations in water level depends on what the cause is. They can be characterized as follows [16, 17]:

(i) short-term fluctuations in water level caused by

(a) tides,

(b) positive or negative surges,

(ii) seasonal changes in water levels caused by long-term hydrometeorological changes in the sea area (region).

Water levels lowered or raised by wind are taken into consideration by adopting characteristic water level when determining fairways. Basing on these assumptions, procedures meeting the conditions of safe navigation are established to use a given sea area. Wind affects safe navigation in restricted waters and is always described by means of wind direction and force.

Sea currents influence the safety of the ship's maneuver as they act on the part of the hull which is below water. There are three important types of currents and streams in restricted waters [12]:

(i) tidal streams,

(ii) sea currents,

(iii) wind current.

Wind sea is an important factor in determining the safety of maneuvering in restricted waters. Parameters of waves affecting the size of the safe maneuvering area for a vessel are the following:

(i) height, length, and period of the waves,

(ii) directions of waves flow in relation to the available navigable water,

(iii) distributions of direction and wave height in available navigable waters.

In order to determine the effect of variables A_i, S_i, N_i, H_i, M_i, I_i, and R_i on the safety of navigation, it is necessary to adopt quantitative indicators. They are different for different types of sea areas on which navigation takes place (or maneuver is performed). The limit values of these indicators are the basis for assessing the safety of navigation.

There are two types of evaluation criteria of safety of navigation in restricted waters:

(i) basic criteria for assessing the safety of navigation determined by a single parameter,

(ii) complex evaluation criteria of safety of navigation which take into account the effects of an accident and which are a function of a number of variables.

Navigational risk is a complex criterion of assessing safety of navigation, [13, 14, 18]. In this work a deviation from the approach channel is assumed to be the hazard which could result in hitting the bottom or in collision. The level of risk depends on the distance from the centerline of the fairway.

Let us assume that the classification of vessel's states in terms of the degree of risk will be carried out basing on the value of a random variable Y using vector of limits. Elements of limits (thresholds) $\boldsymbol{\delta}^T = [\delta_0, \delta_1, \ldots, \delta_{r-1}]$ vectors are monotonic, nondecreasing sequence $\delta_0 \le \delta_1 \le \delta_2, \ldots, \delta_{r-2} \le \delta_{r-1}$.

The different levels of risks and therefore states of the system are determined by

$$\begin{gathered} \tilde{y} \le \delta_0, \quad \text{then } z = 0, \\ \delta_{k-1} < \tilde{y} \le \delta_k, \quad \text{then } z = k \text{ for } k = 1, \ldots, r-1, \\ \tilde{y} > \delta_{r-1} \quad \text{then } z = r. \end{gathered} \tag{2}$$

3. Logit and Probit Models of Ordered States

Generalized linear models extend classical linear models. Logistic regression is an alternative to ordinary linear regression especially when we have discrete variables which describe categories in a given classification or states which may hold [19]. Sequential logit and probit models are often sequences of binary outcome models [20].

Let us assume that for a vessel at the time t value of a variable $y_i(t)$ deciding about assessing the degree of the ith threat to be defined for the moment $t + dt$ depends linearly on the vector of features $\mathbf{x}_i(t)$ [16],

$$y_i(t + dt) = \mathbf{x}_i^T(t)\,\boldsymbol{\beta} + \varepsilon_i(t), \tag{3}$$

where $\boldsymbol{\beta}$ is a vector of unknown structural parameters of the model (3), ε has a distribution independent of i, and t of the expected zero value and of a constant variance and is defined with distribution function of the logistic distribution. Write the model as

$$Y = \begin{cases} 1 & \text{if } y_i^* \ge 0 \\ 0 & \text{if } y_i^* < 0. \end{cases} \tag{4}$$

Let us divide the features vector into two "subvectors"

$$\mathbf{x}_{it}^T = \left[\mathbf{x}_{1i}(t),\ \mathbf{x}_{2i}(t)\right], \tag{5}$$

where $\mathbf{x}_{1i}(t)$ is a vector of features defining the level of threat for a vessel i at a given moment t and $\mathbf{x}_{2i}(t)$ is a vector of quantifiable characteristics.

In this situation, it describes only the value of the unobservable variable y^*, that if the random variable/component has a continuous distribution, (3), it can be written as

$$P\{y_i = 1\} = 1 - F_\varepsilon\left(-\mathbf{x}_i'\boldsymbol{\beta}\right), \tag{6}$$

where F_ε is the cumulative distribution function of the random variable.

An important issue is the way in which the various explanatory variables affect the dependent/explained variable and the extreme results in case of no unit variance of variable ε.

Let us consider a situation in which the ship enters the approach channel and we do not have any information about this channel. What is the probability that "the vessel makes an error" (navigational error)?

Option 1. Let us assume that the vessel is typical and for which the value of an individual effect (resulting from a vessel type-operating characteristics (5)) is zero.

Then

$$P\left(y_{it} = 1\right) = \Phi\left(\mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i\right), \tag{7}$$

for a standard distribution ε, or generally

$$P\left(y_{it} = 1\right) = \Phi\left(\frac{\mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i}{\delta_\varepsilon}\right). \tag{8}$$

Option 2. In normal and independent distribution ε and α their sum $v_{it} = \alpha_i + \varepsilon_{it}$ is normally distributed with parameters $(0;\ \delta_\varepsilon^2 + \delta_\alpha^2)$. This implies that we can write

$$P\left(y_{it} = 1\right) = P\left(y_{it}^* \geq 0\right) = 1 - \Phi\left(\frac{\mathbf{x}_{it}'\boldsymbol{\beta}}{\sqrt{\delta_\varepsilon^2 + \delta_\alpha^2}}\right). \tag{9}$$

Of course, the parameter values and the variances are unknown, so you can take advantage of their assessment.

Let us consider a vessel about which we know how far it deviated from the centerline of the fairway in the approach channel in the past and what corrective maneuvers were performed. What is the probability that the vessel will make navigational error in the next period of time (T)?

Knowing the vessel's history and being able to determine the current level of risk, we get information enabling to identify the individual effects. Let us write the probability in the form of conditional probability (3):

$$\begin{aligned} &P\left(y_{it} = 1 \mid y_{i1}, \ldots, y_{i,t-1}, \mathbf{x}_{it}, \boldsymbol{\beta}\right) \\ &\quad = \frac{P\left(y_{it} = 1, Y_{i1} = y_{i1}, \ldots, Y_{i,t-1} = y_{i,t-1} \mid \mathbf{x}_{it}, \boldsymbol{\beta}\right)}{P\left(Y_{i1} = y_{i1}, \ldots, Y_{i,t-1} = y_{i,t-1} \mid \mathbf{x}_{it}, \boldsymbol{\beta}\right)}. \end{aligned} \tag{10}$$

We use the probability P calculated in accordance with Options 1 or 2, respectively, and then we take a threshold value (p_p). If $P > p_p$ we forecast an error and otherwise lack of an error in the next period of time, [16].

Let us consider probit model with delayed response variable as explanatory variable [4]:

$$\begin{gathered} y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it}, \\ y_{it} = \begin{cases} 1 & \text{if } y_{it} \geq 0 \\ 0 & \text{if } y_{it} < 0. \end{cases} \end{gathered} \tag{11}$$

Since the probabilities in the subsequent periods for the same vessel are not independent, the model cannot be estimated with the presented method. However, conditional probabilities are independent with respect to the current value of the explanatory variable and to the individual effects.

To solve the initial conditions problem we can use the Heckman solution, [21], noting the model (11) in the form:

$$\begin{gathered} y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it} \quad \text{for } t \geq 2, \\ y_{i1} = \mathbf{z}_{it}'\lambda + \mu\alpha_i + \varepsilon_{i1}, \\ y_{it} = \begin{cases} 1 & \text{if } y_{it} \geq 0 \\ 0 & \text{if } y_{it} < 0. \end{cases} \end{gathered} \tag{12}$$

In this model, for the first period an additional static equation defining the value of y_{i1} was used. The explanatory variables in the equation for the first period need not be the same as in subsequent periods.

An assumption that the random variable has a logistic distribution leads to a logit model in which the probability of adopting the value 1 by the dependent variable with given values of explanatory variables is given as

$$P\left(y_i = 1\right) = P\left(\mathbf{x}_{it}'\boldsymbol{\beta} \leq \varepsilon_{it}\right) = F_L\left(\mathbf{x}_{it}'\boldsymbol{\beta}\right) = \left(1 + e^{-\mathbf{x}_{it}'\boldsymbol{\beta}}\right)^{-1}. \tag{13}$$

Elements of the vector of structural parameters of the model and vector of limit values can be estimated by maximum probability method using numerical procedures. Let us consider ordered model in which the observable dependent variable can take of n different values:

$$\begin{gathered} y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i + \varepsilon_{it}, \\ y_{it} = \begin{cases} 0 & \text{if } \delta_0 \leq y_{it} < \delta_1 \\ 1 & \text{if } \delta_1 \leq y_{it} < \delta_2 \\ \vdots & \\ n & \text{if } \delta_{n-1} \leq y_{it} < \delta_n. \end{cases} \end{gathered} \tag{14}$$

The above model is not appropriate if the variable y_{it} has values which cannot be arranged according to a particular scale.

Probability of accepting by the variable values of m can be written as

$$\begin{aligned} P\left(y_{it} = m\right) = {} & P\left(\mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i + \varepsilon_{it} < k_{m+1}\right) \\ & - P\left(\mathbf{x}_{it}'\boldsymbol{\beta} + \alpha_i + \varepsilon_{it} < k_m\right) \end{aligned} \tag{15}$$

or by means of cumulative distribution of the random variable in the form of

$$P(y_{it} = m) = F_\varepsilon\left(k_{m+1} - \mathbf{x}_{it}'\boldsymbol{\beta} - \alpha_i\right) - F_\varepsilon\left(k_m - \mathbf{x}_{it}'\boldsymbol{\beta} - \alpha_i\right), \tag{16}$$

where F_ε is the cumulative distribution function of the random variable.

In this way we obtain an ordered probit, and the estimation is based on maximizing the probability function. Estimation of the model requires the scale, taking a specific value of the variance of variable ε and finding the maximum of the probability function.

4. Examples of Models Used for Examining Risks in the Approach Channel

The methodology and logic of ordered logit model can be presented using a following example. Let there be N ships in the channel, and assume that each ship's navigator "degree of risk acceptance" affects the probability of being in state z. Sometimes the outcomes in a response variable are perceived as a sequence with stages (2), states z. The related probabilities can be written as [20]

$$\begin{aligned}
P_0 &= F\left(\sum_i \mathbf{x}_{i0}\boldsymbol{\beta}\right),\\
P_1 &= \left[1 - F\left(\sum_i \mathbf{x}_{i0}\boldsymbol{\beta}\right)\right] F\left(\sum_i \mathbf{x}_{i1}\boldsymbol{\beta}\right),\\
&\vdots\\
P_{r-1} &= \prod_{k=0}^{r-2}\left[1 - F\left(\sum_i \mathbf{x}_{ik}\boldsymbol{\beta}\right)\right] F\left(\sum_i \mathbf{x}_{ir}\boldsymbol{\beta}\right),\\
P_r &= \prod_{k=0}^{r-1}\left[1 - F\left(\sum_i \mathbf{x}_{ik}\boldsymbol{\beta}\right)\right].
\end{aligned} \tag{17}$$

In the logit model the F is logistic cumulative distribution function and subscripts ik, for $k = 0, 1, \ldots, r$ indicate the sets of $\mathbf{x}$ variables included in states $z = 0$, $z = 1, \ldots$, $z = r$, respectively. The parameters $\boldsymbol{\beta}$ can be estimated by dividing the sample into groups according to levels of risk.

4.1. Prioritizing Vessels due to the Risk That They Generate. Let us examine the case of two vessels along the approach channel, and let us ask the question which of them generates the greater threat?

Let us assume that the risk of the first ship is described by the random variable X and the other by random variable Y. Without any information, we assume that both these variables have the same probability distribution with cumulative distribution function $F(x)$ and independent (if ships are not interacting). In addition, without changing the generality, we can assume that both variables are continuous type, $P(X = Y) = 0$.

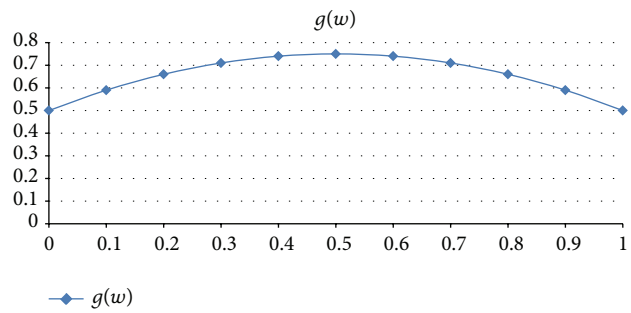

Figure 1: Graph of the function $g(w)$.

Let us calculate the probability $P(X > Y)$

$$P(X > Y) = \int_{-\infty}^{\infty} P(x > Y)\, dF(x) = \left.\frac{F^2(x)}{2}\right|_{-\infty}^{\infty} = 0.5. \tag{18}$$

We obtained the result stating that without knowledge (for any distribution $F_X(x)$) the probability of analyzed events for each vessel is the same and is 0.5.

Let us consider the following algorithm, Figure 2:

(i) let us assume a certain level of d,

(ii) assessing the ship for example, the first one, during the remaining period of time, we check if $X < d$,

(iii) if so, we predict that $X < Y$,

(iv) otherwise, we have $X > Y$.

Let us calculate probability $P(d)$ of a correct assessment using the adopted assumptions about the type of random variables and probability properties we obtain:

$$\begin{aligned}
P(d) &= P(X < d, X < Y) + P(X > d, X > Y)\\
&= \int_{-\infty}^{d} P(x < Y)\, dF(x) + \int_{d}^{\infty} P(x > Y)\, dF(x)\\
&= -F^2(d) + F(d) + 0.5.
\end{aligned} \tag{19}$$

Since $F(x)$ is the cumulative distribution then for each level d value of $F(x) \in [0, 1]$, so let us examine the function $g(w) = -w^2 + w + 0.5$, Figure 1. When calculating we have $g'(w) = -2w + 1 = 0$ for $w = 0.5$, $g''(w) = -2$.

This function has the following maximum value for $w = 0.5$. This means that an optimal level d_{opt} is the median of the distribution $F(x)$, $F(d_{\text{opt}}) = 0.5$. Then $P(d) = 0.75$.

4.2. Generalized Linear Model (GLM) of Leaving the Approach Channel. In the present example we will consider the following distance ranges $\delta_0 = 30$, $\delta_1 = 50$, $\delta_2 = 90$, $\delta_3 = 120$, $\delta_4 = 150$, and $\delta_5 = 150$. Range below 30 therefore corresponds to the level of threat $z = 0$.

The vessel's domain is defined by a rectangle of a length and width corresponding to the length and breadth of the

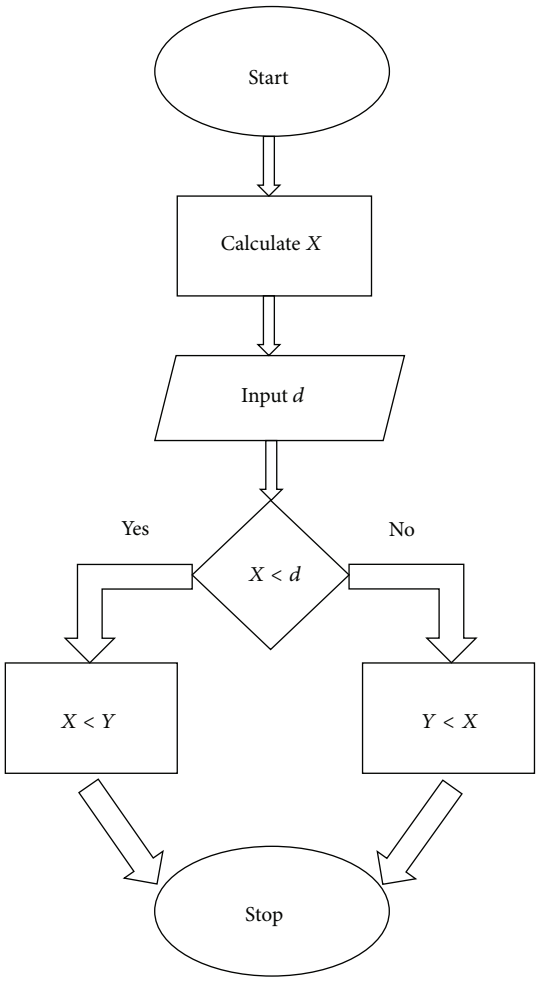

Figure 2: The block diagram for the presented algorithm.

vessel, where γ-heading, d-maximum distance of the domain point from the centerline of the approach channel, Figure 3.

The center of the domain, which is described around the ship, coincides with the geometric center of the vessel. A $d[m]$ was assumed as the maximum distance which is the maximum distance measured from the domain vertices to the fairway centerline.

Simulation studies were carried out in the laboratories of Gdynia Maritime University, using navigational-maneuvering simulator. This simulator in a very realistic way that represents sea areas and behaviour of vessels which are almost identical with real vessels. It is possible because these models use six degrees of freedom. The model of the vessel used was the model of a loaded LNG carrier of the following characteristics: $L = 315$ m, $B = 50$ m, and $T = 12$ m. The ship was on an even keel. Simulation of the ship's behaviour was carried out in the fairway and was influenced by interfering factors such as wind and swell—constant NE wind, period: 9 s., height of the swell: 1.5 m.

The value of 6 and 8 knots was taken as the nominal speed of the vessel. The wind direction adopted in simulations

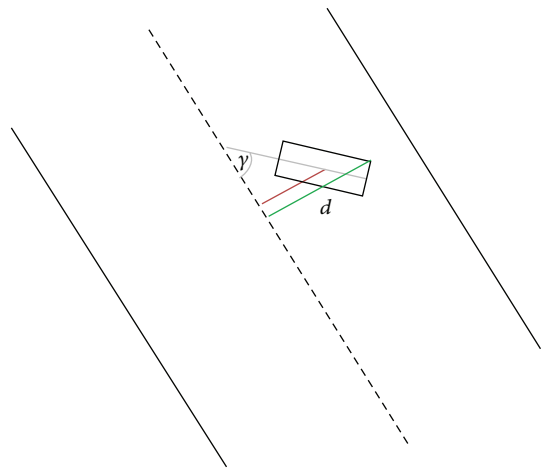

Figure 3: Determination of the maximum distanced for the ship along the approach channel, [15].

changed every 45 degrees starting from the N. The simulation scenario of the behaviour of the model under way along the approach channel assumes that the initial position of the vessel is in the centerline of the existing fairway in its northern most part and that the line of symmetry of the model and the course over ground overlaps with the centerline of the fairway. The parameters of the generated waves were determined, and the simulation results were obtained for winds (no squalls) of speed up to 15,0 m.

For the assumed parameters of the fairway the distanced is given by

$$d = \frac{\left|\tan\left(99.8\left(\pi/180\right)\right) x_{\mathrm{LONG}} - y_{\mathrm{LAT}} + 3254.384\right| \cdot 1852}{\sqrt{\tan\left(99.8\left(\pi/180\right)\right)^2 + 1}} + \sin\left(\gamma\frac{\pi}{180} + a\tan\left(\gamma\frac{\pi}{180} + \frac{50}{315}\right)\right) \cdot \frac{315}{2}, \tag{20}$$

where y_{LAT} and x_{LONG} Cartesian coordinates converted from geographic ones, [15].

In the simulation, in particular, the course of the function of the distance from the model vessel to the centerline of the fairway with given parameters of influencing factors (hydrometeorological conditions) was examined.

Thanks to the research on the simulator some information was received according to which the following variables were determined:

(i) Y_i taking value of 1 ($z = i$) if the vessel at the time t was found in i class of the distance from the centerline of the fairway and 0 otherwise, $i = 0, 1, \ldots, 5$

(ii) X_1 vessel's speed,

(iii) X_2 wind speed,

(iv) X_3 wind direction.

Table 1: Models obtained for X_1, X_2, and X_3 variables.

z	Eta	Chi-square test of compatibility P value
0	$1.9929765 - 0.1864415X_1 - 0.188414X_2 - 0.00511045X_3$	0.8604
1	$2.633185 + 0.14835385X_1 - 0.217938X_2 - 0.001503915X_3$	0.03981
2	$-6.14982 + 0.17340395X_1 + 0.236026X_2 + 0.00403941X_3$	0.1479
3	$-3.884925 - 0.2309091X_1 + 0.239662X_2 + 0.001281675X_3$	0.11903
4	$0.910562 - 0.7587995X_1 + 0.1048175X_2 + 0.001195145X_3$	0.46453
5	$-4.06442 - 0.28746895X_1 + 0.2784895X_2 + 0.0005145X_3$	0.6430

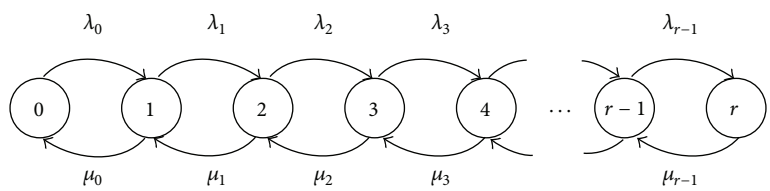

Figure 4: Graph of the system states.

For X_1, X_2, and X_3 variables logistic models were obtained

$$Y = \frac{\exp(\text{eta})}{(1 + \exp(\text{eta}))}, \tag{21}$$

where eta is given in Table 1.

Test of compatibility χ^2 determines whether logistic function adequately fits the observed data. Since the value of p is greater than or equal to 0.05, there is no reason to reject the adequacy of the fitted model at least 95.0% of confidence level.

4.3. Markov Model of Leaving the Approach Channel. Let us consider birth-death process (BD) with finite space of states $\{0, 1, 2, \ldots, r\}$. Graph of state of this process is shown in Figure 4.

Birth and death processes is a continuous-time Markov chain for which transitions may take place only between neighbouring states (i.e., i to $i+1$ or $i-1$ only).

We use the following notation:

(i) λ_i—birth rate, $i = 0, 1, 2, 3, \ldots, r-1$;

(ii) μ_i—death rate, $i = 1, 2, 3, 4, \ldots, r$.

State transition probability:

(i) $p_{01} = 1$

(ii) birth before death:

$$p_{i,i+1} = \frac{\lambda_i}{\lambda_i + \mu_i} \tag{22}$$

(iii) death before birth:

$$p_{i,i-1} = \frac{\mu_i}{\lambda_i + \mu_i}. \tag{23}$$

Consider a system in which arrivals and departures occur one at a time. Let $X(s)$ be the risk level of the system at time s. We now consider fitting a BD model to data collected over an interval $[0, t]$. Let $\lambda_i(t)$ and $\mu_i(t)$ be direct estimates of the birth rates and death rates based on sample averages over the time interval $[0, t]$. Similarly, let $\gamma_i(t)$ be estimates of the stationary distribution based on sample averages over the time interval $[0, t]$.

Moreover let $A_i(t)$ be the number of up state changes (from $z = i$ to $z = i+1$) during the interval $[0, t]$ when the system is in state i; let $D_i(t)$ be the number of down state changes (from $z = i$ to $z = i-1$) during the interval $[0, t]$ when the system is in state i; and let $T_i(t)$ be the total time during the interval $[0, t]$ in which the system is in state i. Then, [22] we have

$$\lambda_i(t) = \frac{A_i(t)}{T_i(t)}, \quad \mu_i(t) = \frac{D_i(t)}{T_i(t)}, \quad \gamma_i(t) = \frac{T_i(t)}{t}, \quad t \geq 0. \tag{24}$$

This estimation procedure need not produce an irreducible BD process, because there can be initial and final transient states. However, under the simplifying assumption of irreducibility, this estimated BD process has the unique stationary probability distribution. For example if there are some constants a_1, a_2, d_1, and d_2 such that $\lambda_i(t) > 0$ for $a_1 \leq i \leq a_2$ with $\lambda_i(t) = 0$ otherwise and $\mu_i(t) > 0$ for $d_1 \leq i \leq d_2$. From [22], this estimated BD process has the unique stationary probability distribution

$$\gamma_i^e(t) = \frac{s_i(t)}{\sum_{j=1}^{r} s_j(t)}, \quad 0 \leq i \leq r, \tag{25}$$

where $s_0(t) = 1$ and $s_j(t) = \prod_{i=0}^{j-1} \lambda_i / \prod_{i=1}^{j} \mu_i$, $0 \leq j \leq r$.

4.4. System with Blocking. Let α_i be the probability that a ship will enter the higher class $i+1$ if it is already at class i.

$$\lambda_i = \alpha_i \lambda, \quad i \geq 0; \qquad \mu_i = \mu, \quad i \geq 1. \tag{26}$$

Transition probability function is given by

$$p_{i,i+1} = \frac{\alpha_i \lambda}{\alpha_i \lambda + \mu}, \qquad p_{i,i-1} = \frac{\mu}{\alpha_i \lambda + \mu}. \tag{27}$$

Transition time T_i, the time required to go from state i to state $i+1$ has the mean given by [13]

$$\begin{gathered} E[T_0] = \frac{1}{\alpha_0 \lambda}, \qquad E[T_i] = \frac{1}{\alpha_i \lambda} + \frac{\mu}{\alpha_i \lambda} E[T_{i-1}], \\ \operatorname{Var}(T_0) = \frac{1}{\alpha_0 \lambda^2}, \\ \operatorname{Var}(T_i) = \frac{1}{\alpha_i \lambda (\alpha_i \lambda + \mu)} + \frac{\mu}{\alpha_i \lambda} \operatorname{Var}(T_{i-1}) \\ + \frac{\mu}{(\alpha_i \lambda + \mu)} (E[T_{i-1}] + E[T_i])^2. \end{gathered} \tag{28}$$

From (25) we have stationary probability distribution

$$\gamma_i^e(t) = \frac{\prod_{k=0}^{i-1} \alpha_k (\lambda/\mu)^i}{\sum_{j=1}^{r} \prod_{i=0}^{j-1} \alpha_i (\lambda/\mu)^j}, \quad 0 \leq i \leq r. \tag{29}$$

The properly selected sets of explanatory variables in the model (3) allow you to define Markov chains with different transition probability matrix. If we assume that the set of variables was limited to two variables that as a result of the estimation of the logit model, we obtain the assessment of probabilities that can be interpreted as the transition probability of a one threat category to another category. The use of the full set of explanatory variables of the form (5) gives us the transition probabilities for any ship i and a time t.

5. Conclusion

The study proves suitability of logit models used to estimate the hazards to navigation. They can be used in the development of guidelines for traffic safety management system and large ships manoeuvring in restricted waters, to ensure obtaining the assumed level of safety. The approach proposed in this work can be used in Markov models by using in estimations the transition matrix of other variables than just defining the transition between states. This will allow the use of semi-Markov models to estimate safety of navigation.

Immediate work in simulation follows better evaluation measures and improvement of duration modeling, model system, and model confidence levels. The presented approach has to be further developed by more comprehensive experimental evaluations, examples of applications, and analytical models relating selected simulation responses with model parameters.

Discrete event models allow inclusion of individual variables without creating compound states, which could improve the precision of the model. Some interesting questions are still open. For example possible questions can relate to correlation between random variables.

References

[1] M. I. Jordan, "Why the logistic function? A tutorial discussion on probabilities and neural networks," Computational Cognitive Science Report 9503, MIT Press, Boston, Mass, USA, 1995.

[2] M. I. Jordan and A. Y. Ng, "On discriminative versus generative classifiers: a comparison of logistic regression and naive Bayes," in *Proceedings of the 16th Annual Conference on Neural Information Processing Systems*, 2002.

[3] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer Series in Statistics, Springer, New York, NY, USA, 2001.

[4] J. C. Gardiner and Z. Luo, *Logit Models in Practice: B, C, E, G, M, N, O. . ., SAS Global Forum 2011 Statistics and Data Analysis*, Michigan State University, East Lansing, Mich, USA, 2011.

[5] K. K. W. Yau, "Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong," *Accident Analysis and Prevention*, vol. 36, no. 3, pp. 333–340, 2004.

[6] K. K. W. Yau, H. P. Lo, and S. H. H. Fung, "Multiple-vehicle traffic accidents in Hong Kong," *Accident Analysis and Prevention*, vol. 38, no. 6, pp. 1157–1161, 2006.

[7] R. O. Mujalli and J. de Ona, "Injury severity models for motor vehicle accidents," in *Proceedings of the Institution of Civil Engineers*, pp. 1–16, University of Granada, Granada, Spain, 2011.

[8] S. Patil, S. R. Geedipally, and D. Lord, "Analysis of crash severities using nested logit model—accounting for the under-reporting of crashes," *Accident Analysis and Prevention*, vol. 45, pp. 646–653, 2012.

[9] P. Averty, K. Guittet, and P. Lezaud, "An ordered logit model of air traffic controllers' conflict risk judgment," *Air Traffic Control Quarterly*, vol. 16, no. 2, pp. 101–125, 2008.

[10] G. van de Wiel and J. R. van Dorp, "An oil outflow model for tanker collisions and groundings," *Annals of Operations Research*, vol. 187, no. 1, pp. 279–304, 2011.

[11] S. Greenland, "Principles of multilevel modelling," *International Journal of Epidemiology*, vol. 29, no. 1, pp. 158–167, 2000.

[12] S. Gucma, "Model of vessel's manoeuvring in limited sea areas in navigational risk aspect," *Archives of Transport*, vol. 12, no. 1, 2000.

[13] A. Blokus-Roszkowska Smolarek L, "Collision risk estimation for motorways of the sea," *Reliability: Theory and Applications*, vol. 1, no. 2 (25), pp. 58–68, 2012.

[14] Z. Smalko and L. Smolarek, "Modelling a ship safety according to collision threat for ship routes crossing," *Scientific Journals Maritime University of Szczecin*, vol. 20, no. 92, pp. 120–127, 2010.

[15] L. Smolarek and H. Sniegocki, "The use of logit models to navigational risk analysis at restricted areas" (Polish), Gdynia Maritime University, 2012.

[16] K. V. Borooah, *Logit and Probit: Ordered and Multinominal Models*, Sage, Thousand Oaks, Calif, USA, 2002.

[17] J. S. Cramer, *Logit Models from Economics and Other Fields*, Cambridge University Press, Cambridge, UK, 2003.

[18] H. Sniegocki, "Errors in the presentation of the vessels course and speed for the Vts operator," *Annual of Navigation*, vol. 4, pp. 81–90, 2002.

[19] M. S. Lewis-Back, *Applied Regression*, Sage, Newbury Park, Calif, USA, 1980.

[20] T. F. Liao, *Interpreting Probability Models Logit, Probit and Other Generalized Linear Models*, Sage, Thousand Oaks, Calif, USA, 1994.

[21] J. J. Heckman, "Dummy endogenous variables in a simultaneous equation system," *Econometrica*, vol. 46, no. 4, pp. 931–959, 1978.

[22] W. Whitt, "Fitting birth-and-death queueing models to data," *Statistics and Probability Letters*, vol. 82, no. 5, pp. 998–1004, 2012.