DEVELOPMENT OF A MODEL FOR SIMULATION OF VESSEL TRAFFIC STREAMS

Agnieszka Puszcz
Lucjan Gucma
Maciej Gucma
Maritime University of Szczecin
Maritime Traffic Engineering Institute
Wały Chrobrego 1-2 70-500 Szczecin, Poland
a.puszcz@am.szczecin.pl, lucek@am.szczecin.pl,
m.gucma@am.szczecin.pl

ABSTRACT

A proper understanding of traffic stream behavior is necessary for risk analysis an efficient design of traffic facilities. The paper presents methods and models used for analysis of ships traffic in Southern Baltic on chosen areas. Ships' traffic has been analyzed by means of statistical methods with use of historical AIS data obtained from HELCOM. The paper presents probabilistic models of ships' traffic spatial distribution and its parameters. The vessel path and speed, as well as the influence of vessel size, type, and distance to danger is determined. The results could be used for safety and risk analysis models in given area and for creation of general models of ships' traffic flows.

1 INTRODUCTION

Ships' traffic is the most important factor of ships safety. The growing number of vessels and its size is related to possibility of accident determining fires, explosions and losses of life and in the environment. These negative effects raise awareness of government, industry and society on the importance of risk assessment. To assess this risk we need to know how ships traffic behaves. Traffic streams model is a input for risk assessment models. A good understanding of traffic behavior is also necessary for an efficient design of traffic facilities. In order to design such facilities better needs to be understood the stream behavior under various traffic conditions. The precise knowledge of ships' traffic phenomenon and its processes is significant for navigational safety analysis. AIS (Automatic Identification System) gives great opportunity for traffic monitoring but also for discovering basic processes which rules the traffic of ships in confined areas. Results of statistical analysis present in this paper conducted to find ships' positions spatial distribution and its parameters of ships' traffic in analyzed area.

One of the main problems posed by marine traffic engineering is to determine the optimum parameters of new constructed or modernized parts of the waterways. Depending on the type of waterway parameters that can be obtained are for example, lane width or diameter of the turning circle. These parameters are usually determined in one of two methods: a cheaper but less accurate analytical method or the more expensive and more accurate simulation method.

This paper presents how utilize historical AIS data to explore the existing maneuver pattern in the area, estimate parameters describing traffic flow and used for the construction of a new method of determining a ship maneuvering area. Paper presents first steps leading to create a general model of traffic flow- statistical analysis of ships traffic flow on selected areas of Southern Baltic.

2 STATISTICAL MODEL OF SHIPS TRAFFIC STREAMS

2.1 Data

The ship traffic data originates from Automatic Identification System (AIS) data supplied by HELCOM. AIS is an automatic system to exchange information between ships and between ships and land-based stations. A ship equipped with AIS continuously transmits information regarding its name, location, destination, speed and course. The AIS data which form the basis of the analysis cover the period from the 1st of January 2008 to the 31st of December 2008. The annual number of movements on each route is computed by analysing the number of ship crossings of report lines perpendicular to each route.

To determine the precise location of routes and the annual/monthly number of movements the AIS-data has to be processed further. This is done by examining the ship crossings of key report lines introduced across each relevant route. The location of the report lines was chosen based on an inspection of a ship traffic density plot. In Figure 1 a ship traffic density plot of the area is shown. For each report line detailed information about each ship and the specific crossing were obtained.

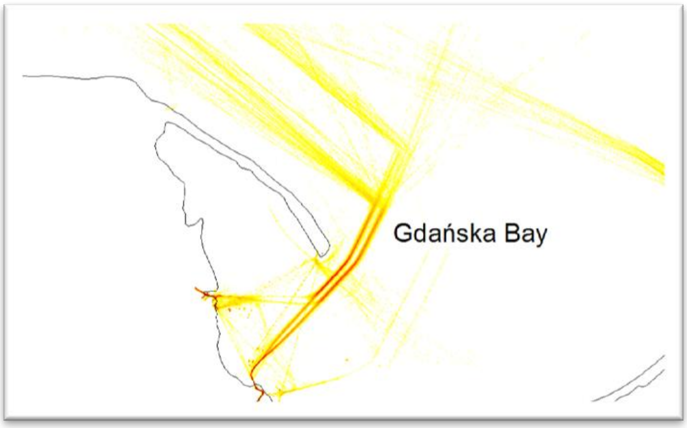

Figure 1: Ship Traffic density plot on Gdańska Bay

2.2 Vessels type and size

For each route the ship type distribution is obtained from analysing the ship types crossing each the relevant report lines. In the AIS data the ships are registered with two-digit code representing the ship type [5]. For the present study the following ship type division have been applied

- Passenger ships. Ship type code 60 to 69
- Cargo ship. Ship type code 70 to 79
- Tanker ship. Ship type code 80 to 89
- Fishing vessels. Ship type code 30 to 39
- Other. All other codes – also unknown ship types.

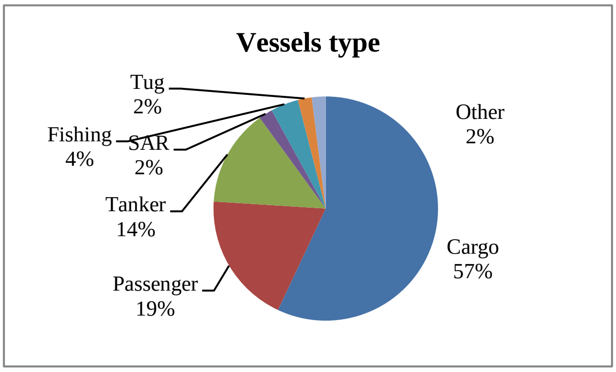

Figure 2: Ship type distribution. January 2008

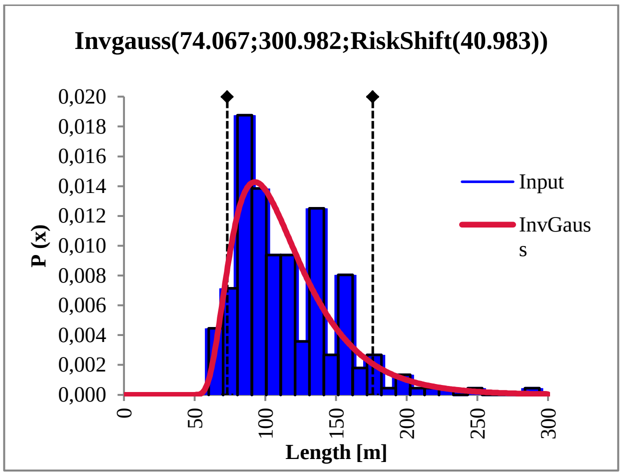

Figure 3: Cargo vessel size distribution (by length). January 2008

2.3 Spatial distribution of ships traffic

In ship traffic modeling it is common practice to model transverse ship traffic distribution by a normal distribution. This is based on the assumption that most ships try to follow the official route as close as possible and are thus normally distributed across the route. These assumptions, however, do not fully describe the behavior of the traffic. Transverse distribution of ship traffic depends largely on the type of route (bend, straight) and its character (Traffic Separation Scheme, narrow channel etc).

The theory of traffic flow of ships involved in the phenomena described the movement of many vessels through the traffic lane in the some chosen period of time. One of the main parameters describing the traffic flow is the distribution, describing the ship's hull position in relative to the axis of the track. The information about the position of the vessel's center of gravity, the shape of the waterline and the course are used to define the distribution. A simple approach to describe traffic streams is their characterization by means of a single, specific resolution. The most common and used distributions are as follows:

- normal distribution with PDF (probability density function):

$$d_l(y) = \frac{1}{\sigma_l \sqrt{2\pi}} e^{-\frac{(y-m_l)^2}{2\sigma_l^2}} \tag{1}$$

- logarithmic distribution:

$$d_l(y) = \frac{1}{y \cdot \sigma_l \sqrt{2\pi}} e^{-\frac{(\ln y - m_l)^2}{2\sigma_l^2}} \tag{2}$$

- gamma distribution:

$$d_l(y) = \frac{a^p}{\Gamma(p)} \cdot y^{p-1} \cdot e^{-a \cdot y} \tag{3}$$

- logistic distribution:

$$d_l(y) = \frac{z}{\beta(1+z)^2}, \text{ where } z = e^{\frac{v-\alpha}{\beta}} \tag{4}$$

where:

y – distance to the axis, m – average of ships distance to the waterway axis, σ – standard deviation of ships distance to the waterway axis, $\Gamma(p)$ – gamma function, $\beta(p, q)$ – beta function, a, p, q, β, α – other parameters of distributions.

The width of traffic flow is fundamental importance in its assessment. In order to describe these traffic streams is necessary to determine the characteristics of the distributions of the width of the traffic lane.

2.4 Time distribution of ships traffic

The aim of the study is to determine a statistical model that can describe the flow ships and define the parameters obtained models depending on time of day or year. The results obtained can be used to their practical application to the simulation models for traffic flow parameters, and may also facilitate the prediction of the expected values of the stream traffic on the studied region in the forthcoming years.

The basic input parameter of traffic streams for modeling vessels behavior is also the intensity. Vessel traffic along the fairway is a process that is influenced by many factors change over time and the length of the track. They make the movement becomes a random process, and his description of the probabilistic models are used. The flow of ships on the seaway can be represented on the timeline, where the transition moments of ships through the center are random events. Random stream of vessels can be analyzed by examining: distribution of number of vessels passing through this point in time ΔT; the distribution of time intervals between the vessels; the distribution of local speed vessels.

The number of ships passing given point of the waterway in case when vessels have freedom of selecting speed and manoeuvre can be described as Poisson distributed stochastic process where probability of appearance of X=n ships in Δt time is:

$$P(X=n)=\frac{(\lambda\Delta t)^n}{n!}e^{-(\lambda\Delta t)} \text{ for } \Delta t\geq 0 \tag{5}$$

where:
λ-traffic intensity [ships/s]
Probability, that in Δt time no ship will appear is:

$$P(X=0)=e^{-(\lambda\Delta t)} \quad \text{for} \quad \Delta t\geq 0 \tag{6}$$

3 ANALYZED AREA

3.1 General information

Gdansk is situated at the mouth of Wistula River, in the N of Poland, on the southern Baltic coast. There are two main Traffic Separation Schemes to Gdansk and Gdynia harbors. The daily average of ships' passages going to Gdansk is around 25 including ferries, passenger vessels and fishing boats. The number of ships passing through the area is expected to grow significantly over the years to come due to the general increase of trade, and the growing economy of the eastern Baltic countries.

Due to comprehensive researches paper presents results only for the GDANREP on the 2 gates S1,N1 “established” in TSS which is presented in figure 4.

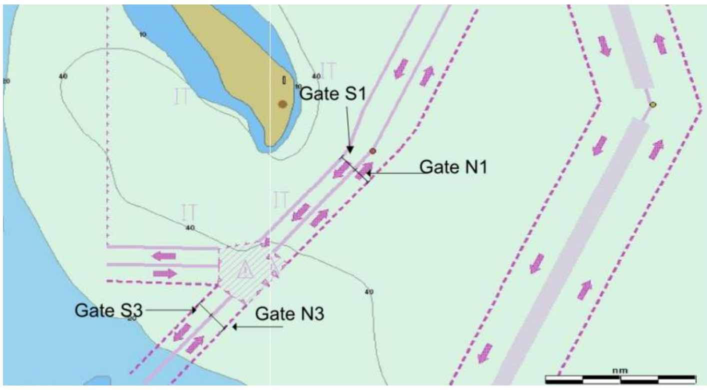

Figure 4: Map of Gdańsk Bay with the gates marked; Map source: AISLive

Located at the mouth of the river, the Port of Gdansk, as well as the adjacent area is ice-free and tide-free all year round. Despite the severe winter weather the navigation conditions are favorable, which contributes to the Port of Gdansk being one of the most approachable port in the Baltic Sea Region.

4 RESEARCH METHOD

Determine of the exact vessels path and their corresponding speed allows describing traffic stream. The results should be formulated in such a way, that they are general applicable and can be used as input for a maritime model. To reliably describe a vessel's path and speed in a model the following should be determined:

- Describe the spatial distribution of vessels in a certain section;
- Describe the lateral vessel speed distribution of vessels in a certain section;
- Describe the vessel speed distribution on a certain location;
- Take into account that the 3 distributions mentioned above depend on:
  - Vessel type;
  - Vessel size;
  - Vessel heading / destination;
  - Type of waterway segment (straight / bend)
  - Width of the (for that specific vessel type and size) navigable waterway;
  - Wind speed and wind direction;
  - Current speed and current direction;
  - Visibility;
  - Other external influences;
  - Interaction with other vessels.
- Describe the mutual dependence between two spatial successive distributions (to connect the different sections, in order to assemble an individual vessel path and correct speed development)

The main factors affecting the movement of vessels in relation to the axis of the traffic lane are the size of the vessels, meteorological conditions (waves, wind), the experience of the officer. An additional parameter is the one characterizing the intensity of shipping traffic. It depends largely on the economic situation in the market and current season.

The research consisted mainly of matching the distribution of traffic in relation to the axis of the traffic lane on GDANREP. The procedure for determining the type of distribution of selected parameters is as follows:

- Information on traffic received from the AIS (for the period 01.2008–02.2009) is used.

- Data is divided on the season, day or night time and type of traffic (inbound or outbound).
- Grouped samples studied separately as separate random variables.
- Defined a random variable as the location of the vessel in relation to the axis of the track.

The data obtained was transformed by means of the middle line method used for the bends and straight sections of the fairway [3]. This method (Fig. 5) uses as a reference the center of the track segments approximated by the length of the section *(i)*. Sections have the shape of part-circular or rectangular. Based on data about the course, waterline and geometric center of waterline are calculated coordinates of extreme points of the vessel (right and left), then their distances to the track axis (the axis of reference). Corresponding distributions have been matched by means of the distance tables [3].

The analyses have been made on a selected set of data from approach to Gdansk. The midpoint of the navigation channel or traffic lane is used as origin. Thus, mean values are the average distance from the midpoint of the navigation channel. Gates midpoint was situated in the middle of the approach channel, between buoys no 7 and no 8 in a way that the buoys 7-8 are in one line with the gate assigned. Others gates have been determined in the same way.

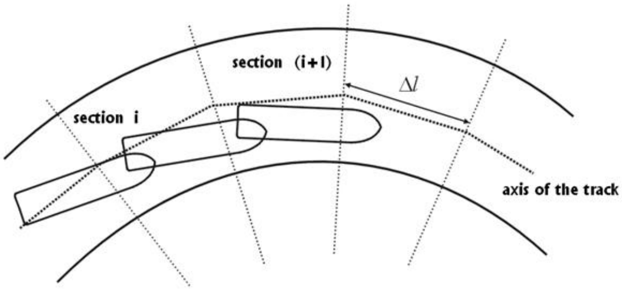

Figure 5: Method for the measure of the track, the track center line approximation to the polygon and the track division into sections [4]

4.1 Results

4.1.1 Average vessel path and speed

Figure 6-7 shows an example of a spatial distribution, derived from the empirical data at a certain gates. On the X-axis a value of zero correspondents to the middle of the axis. A positive value for X means that the vessel sails more to the starboard side.

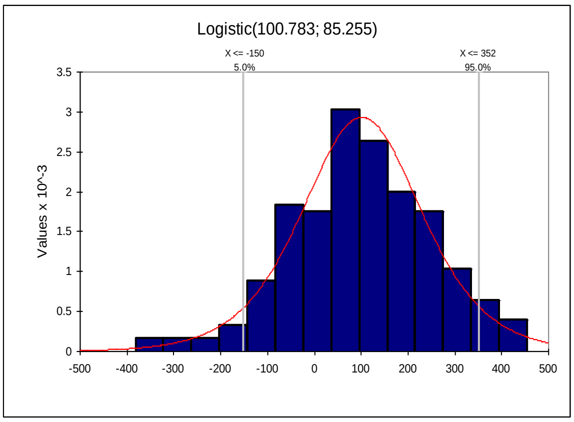

Figure 6: Distribution over the waterway at cross section S1 for vessels length L>100 m, July 2008

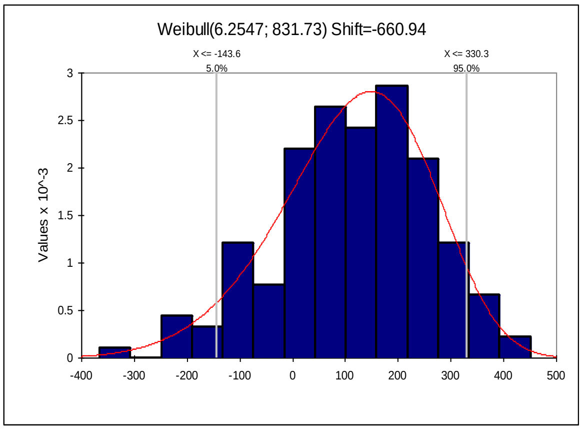

Figure 7: Distribution over the waterway at cross section S1 for vessels length L<100 m, July 2008

Their goodness of fit is first determined by performing a Chi-square test ($\chi 2$). This test determines the degree of agreement between the empirical distribution and the theoretical distribution. The hypothesis is that there is no significant difference between those distributions. The confidence level (answering the question what is significant) is set on 95%. Also Kolmogorov- Smirinov and Anderson- Darling tests have been perform. K-S statistic, the A-D statistic does not require binning. But unlike the K-S statistic, which focuses in the middle of the distribution, the A-D statistic highlights differences between the tails of the fitted distribution and input data. In tab.1 tests results are shown.

Table 1: Tests results for chosen gates in GDANREP

Test	February		July	
	S1			
	L>100	L<100	L>100	L<100
Chi sq.	6.8400	9.0110	12.430	5.1700
A-D	0.3430	0.4820	0.2228	0.1020
K-S	0.0550	0.0661	0.03075	0.0273
	N1			
	L>100 in	L<100 out	L>100 in	L<100 out
Chi sq.	5.6620	4.7500	7.5310	11.27
A-D	0.3715	0.1926	0.2157	0.3907
K-S	0.0492	0.0488	0.0255	0.04778

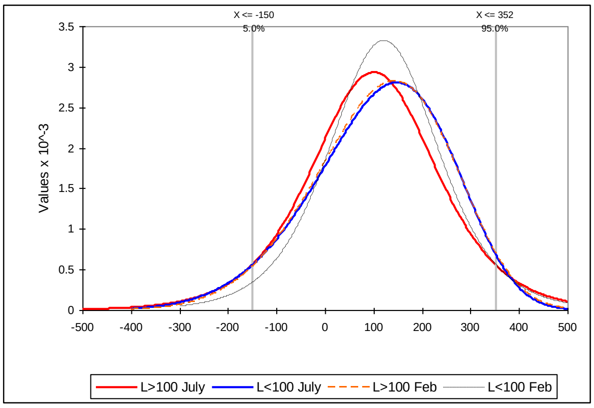

Figure 8: Distribution over the waterway at gates S1, July 2008

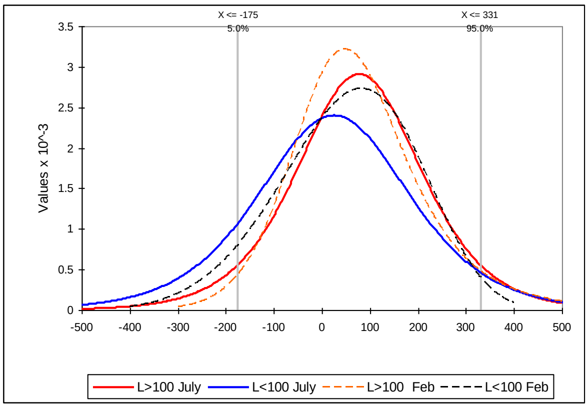

Figure 9: Distribution over the waterway at gates N1, July 2008

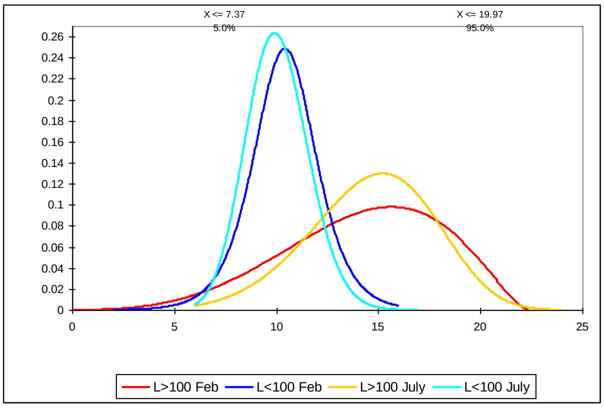

Figure 10: Vessel speed distribution on location S1, for all size vessels

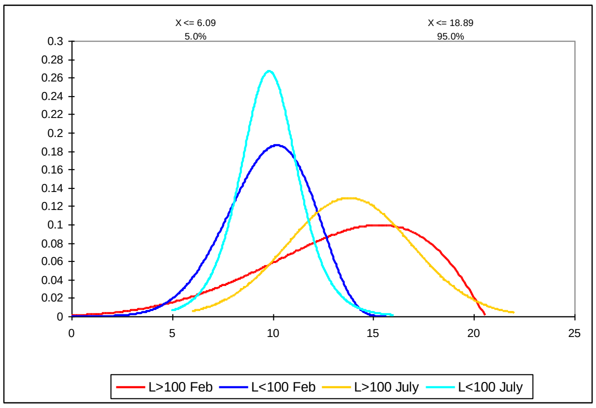

Figure11: Vessel speed distribution on location N1, for all size vessels

Can be seen on the X-axis (0 is on the port side of the incoming vessels) that inbound vessels sail on the right side of waterway for winter and summer months but for outbound vessels middle of their tracks is near the middle of the channel. Analyzing mean and standard deviation of the distributions was obtained that the largest vessel (L>100 meters) sail more to the middle of the channel. The smallest size class (L<10 meters) sails more to the outer limits.

The spatial deviation over the waterway is very well approximated by a normal distribution. This is supported by calculating values for the skewness and excess kurtosis of the empirical datasets. A χ2-test showed as well that the assumed logistic distributions are a good approximation. In 50% of obtained spatial distribution, the best describing distribution is loglogistic- and logistic-40% in other 10% of cases normal and lognormal distribution have had the best results.

Logistic distribution best matches the data around the mean value. Normal distribution is more flattened compared to the logistic distribution. There is also a difference in the description of the extreme values. Despite the good fit to the normal distribution for the value closest to the average, normal distribution adapts insufficiently to the extreme values.

Although the vessels speed entering and leaving the GDANREP is specified in the port regulations the graphs make clear that the vessel speed has a wide distribution. Speed mean value for inbound vessels is higher than for outbound which does not depend strongly on the seasons.

4.2 Intensity of ships traffic

The study focused mainly on verification of the hypothesis that the traffic flow model on the GDANREP fairway have a Poisson distribution. The random variable defined as "the number of vessels appearing on the fairway within 4 hours."- The data was transformed so as to comply with the above-defined of random variable. Data were grouped according to season, type of traffic (inbound and outbound) and time of day, so that they formed a separate random variables, which were studied separately. This was due to the variation of the intensity of vessel traffic.

The verification was carried out using χ2 test. Statistical tests were applied to the data according to:

- Season (winter and summer months)
- Time of day (day, night)
- Traffic (inbound, outbound)

Figures 12 to 13 are graphs match schedules for the two months of data (summer and winter) on day and night time (0800-2000 and 2000-0800) for inbound and outbound vessels. In all the verification tests performed at a level of significance $\alpha = 0.05$.

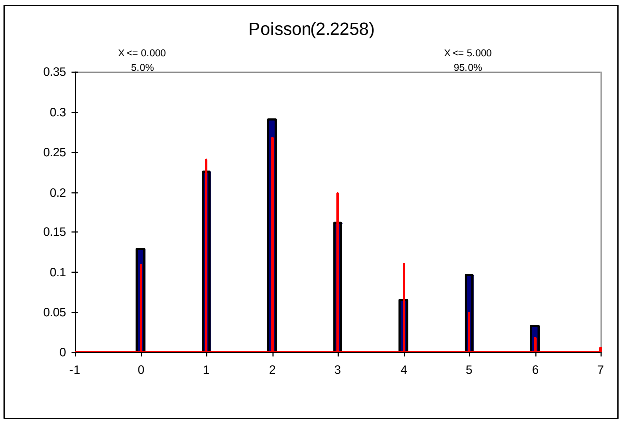

Figure 12: Histogram with fitted distribution for gate S1, inbound traffic, summer

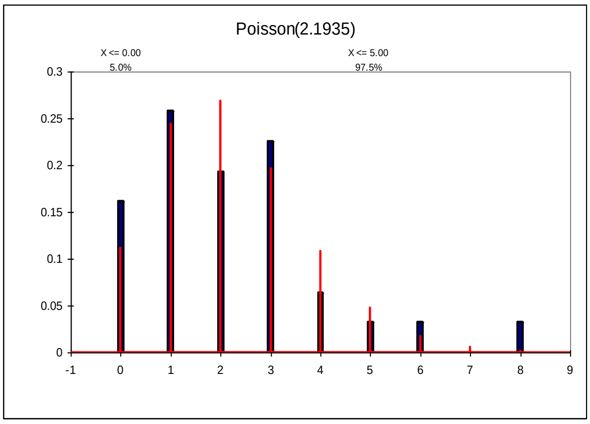

Figure 13: Histogram with fitted distribution for gate N1, outbound traffic, summer

Table 1: Tests results for chosen gates in GDANREP

Month	hrs	Inbound traffic		
		Intensity	Test Value	P Value
January	0800-2000	1.806	1.279	0.8650
	2000-0800	1.613	1.476	0.6878
July	0800-2000	2.226	0.290	0.9619
	2000-0800	2.193	1.189	0.7557
		Outbound traffic		
		Intensity	Test Value	P Value
January	0800-2000	1.742	3.364	0.4988
	2000-0800	1.484	0.634	0.8887
July	0800-2000	2.806	2.229	0.6937
	2000-0800	2.000	1.281	0.7336

5 CONCLUSIONS

The methodology presented in the paper is the basis for the development of a mathematical model of traffic. Creation of precise model that presents results should be verified by studying on traffic flows and its distributions on different water areas.

Using this approach it is possible to obtain generic rules that describe the vessel path and corresponding speed in many different areas. To do so, the case study area (Southern Baltic) will be split into several characteristic waterway and segments and the location specific results will be generalized to their specific segments.

Can be concluded that by using an analysis of historical AIS data, clearly more insight is obtained in the detailed individual vessel behavior. This understanding of the behavior can be formulated in generic rules. These rules can be implemented in maritime model, which improves the simulation of the individual vessel path and vessel speed.

It should be noted that the distribution applies solely to the AIS registered ship traffic. Thus, the leisure boats and fishing boats will be modeled separately to form a complete ship traffic distribution. Leisure boat traffic is only to a limited extent present in winter, spring and fall. Thus, seasonal variations will be included in the model to account for pleasure boat traffic in the total ship traffic volume and in the ship traffic distributions.

Further work should be focused on improving the model that will involve the examination of traffic flows on the open and limited water, depending on the hydro meteorological conditions, ship traffic, and depending on the existing dangers (e.g. shoal).

ACKNOWLEDGEMENT

This paper was created with support of EfficienSea project; partially EU founded Baltic Sea Region Programme 2007-2013.

REFERENCES

1. IMO NAV 51/3/X March 2005 SUB-COMMITTEE ON SAFETY OF NAVIGATION 51st session. Overview of the ships' traffic in the Baltic Sea – HELCOM 2008.
2. Overview of the ships' traffic in the Baltic Sea. HELCOM 2008.
3. J. Guziewicz, W. Ślączka, "Methods for determining the maneuvering area of the vessel used in navigating simulation studies", VII MTE Conference, Szczecin 1997.

4. L. Gucma “The study of probabilistic characteristics of traffic flow on the fairway Szczecin- Świnoujście”, Szczecin 2005
5. Guidelines for the installation of a shipborne automatic identification system (AIS), SC/Circ.227, IMO, January 2003