ISSN 0554-6397
Prethodno priopćenje
(*Preliminary communication*)

Jasminka Bonato
E-mail: jasminka.bonato@pfri.uniri.hr
Biserka Draščić Ban
E-mail: biserka.drascic@pfri.uniri.hr
Ivan Tudor
E-mail: ivan.tudor@pfri.uniri.hr
University of Rijeka, Faculty of Maritime Studies, Studentska 2, 51000 Rijeka, Croatia

Linear Trend Model in the Analysis of Ship Losses and Navigation Safety

Abstract

By researching the available statistical data of the European Maritime Safety Agency (EMSA) on total losses in shipping, the most common cause, ship sinking, was analyzed in the period from 2015 to 2020. The data are useful for qualitative and quantitative assessments, and indicate factors that cause unwanted events.
The aim of this paper is to highlight the possibility of prediction, which is crucial for the assessment of risk and safety in shipping, by applying a statistical model in the analysis of losses in shipping. In the analysis and processing of the available databases, Excel was used, as part of the Microsoft Office software package, which enables statistical analysis and drawing graphs.

Keywords: statistical model, security, maritime, Excel

1. Introduction

The safety of the global shipping industry is important, therefore, based on the data reported in EMCIP on total shipping losses in the period 2011-2020, specific areas affecting the risk and safety assessment were analyzed. Areas relevant to security assessment are security awareness, environmental impact and risk assessment. This includes, for example, misinterpreting the intentions of other ships, anticipating the impact of meteorological conditions, such as wind and low visibility, and performing critical operations, for example, selecting an anchorage area.

The available European Maritime Safety Agency/Safety analysis of EMCIP data - Navigation accidents databases enable quantitative and qualitative analysis of reported events, as well as the factors that contribute to these events. Most of the reported problems relate to operational work methods, organizational factors and risk assessment[1], but data show that the most common human action is related to navigational accidents [2].

The factors that contribute to navigational accidents are the sinking of the ship, stranding, fire, damage to the machine or the ship's hull, are shown in table 1, while the graphic representation of the factors that contribute to navigational accidents is shown in figure 1.

Table 1 [1]

Samples/lost ships	∑
Ship skinging	477
Stranded ships	172
Fires	99
Machine damage	51
Hull damage	29

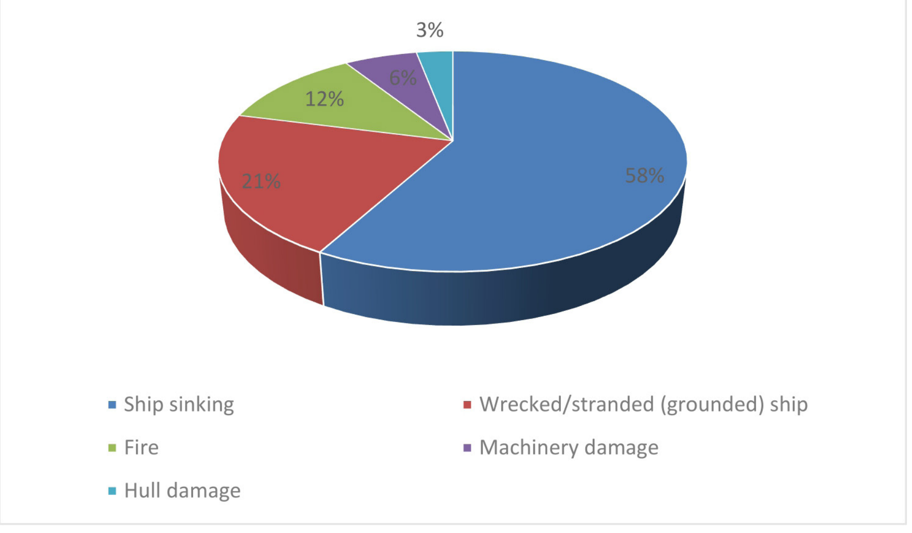

Figure 1: Factors contributing to navigational accidents [Source:Authors].

The research aims to emphasize the importance of quantitative methods of risk management, which in the long term means a contribution to an effective decision-making process [3]. Along with the introduction of new technologies, improved ship design, introduction of risk management methods, in the long run it would also mean a reduction in losses.

The linear trend method was chosen, which enables the analysis and prediction of accidents and their causes. [4,5]

2. Linear trend model

Forecasating theture state and movement od the phenomenon, based on existing variable values, demands installation of a specific model. Descriptive statistics enable tabular and graphical presentation, while the dynamics of the phenomenon can be monitored using linear trend analysis, which was applied in this research[6,7]. The most common cause of losses in shipping, ship sinking, was analyzed in the period from 2011 to 2020, data in table 2.

Table 2 [1]

Year	**Ship sinking (y_t)**	x_t	x_t^2	$x_t y_t$	$\hat{y}$
2011.	46	0	0	0	62.727
2012.	54	1	1	54	6259.3876
2013.	70	2	4	140	6256.0482
2014.	50	3	9	150	6252.7088
2015.	66	4	16	264	6249.3694
2016.	48	5	25	240	6246.03
2017.	57	6	36	342	6242.6906
2018.	31	7	49	217	6239.3512
2019.	31	8	64	248	6236.0118
2020.	24	9	81	216	6232.6724
∑	477	45	285	1871	

Graphical display of ship sinking, as most common shipping losses, for the same period, shown in figure 2.

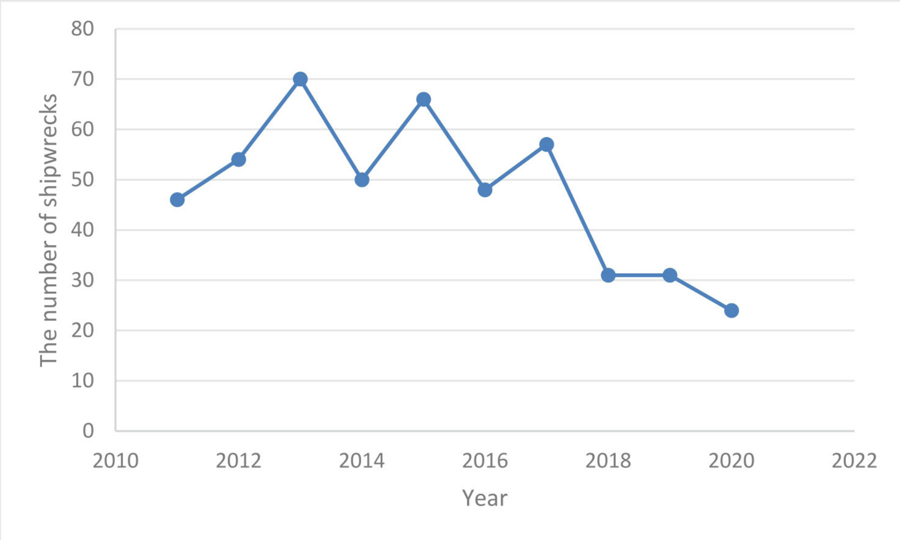

Figure 2: Disploay of losses, ship sinking in period between 2011.-2020.

2.1. Determining the trend equation

A linear trend is the direction that determines the tendency of the phenomenon in the observed period. In general, the trend model to be evaluated is: $\hat{Y} = \hat{a} + bX$, where X is time. (1)

In our case, it is years because the time series data is given by years. The value of X is determined depending on the starting point, which is given or is chosen arbitrarily. All periods after zero are 1, 2, 3, ..., table 2.

The unknown parameters are a and b. The method of least squares gives the following expression for calculating the parameter b:

$$b = \frac{\sum_{i=1}^{N} X_i Y_i - N \cdot \bar{X}\bar{Y}}{\sum_{i=1}^{N} X_i^2 - N \cdot \bar{X}^2} \quad (2)$$

The trend coefficient b shows the average change in the dependent variable when the independent variable increases by a unit of time. $\bar{X}$ and $\bar{Y}$ are simple arithmetic variable values.

$$\bar{X} = \sum_{i=1}^{N} \frac{X_i}{N} = \frac{45}{10} = 4.5 \quad (3)$$

$$\bar{Y} = \sum_{i=1}^{N} \frac{Y_i}{N} = \frac{477}{10} = 47.7 \quad (4)$$

Including the dana from table 2 into equation (2), you get the value of parameter b, equation (5).

$$b = \frac{\sum_{i=1}^{N} X_i Y_i - N \cdot \bar{X}\bar{Y}}{\sum_{i=1}^{N} X_i^2 - N \cdot \bar{X}^2} = \frac{499 - 6 \cdot \frac{15}{6} \cdot \frac{257}{6}}{55 - 6 \cdot (\frac{15}{6})^2} = -3.34 \quad (5)$$

Linear trend coefficient amounts to -3.34 which means that within the observerd period, the sinking of ships deacreased by about 3 ships yearly. Parameter a is given with expression $a = \bar{Y} - \bar{X} \cdot b$ (6)
and represents the expected value of the time series in the original period. If the origin is not in the middle of the period, then the parameter a is not the average number of units.

$$a = \bar{Y} - \bar{X} \cdot b = \frac{257}{6} - \frac{15}{6} \cdot (-8.2) \approx 67.72 \quad (7)$$

Coefficient now amounts to 67.7 which represents average ship sinking in the period of 2011. to 2020. Now the equation of the trend (8) can be displayed graphically, as show in picture 3.

$$\hat{Y} = a + b\,X = 67.7\text{-}3.34\,X \quad (8)$$

The parameters â and b are estimated so that the line Ŷ passes between the actual points of the time series and best interprets the connection between them.

In the last column of table 2 are the values obtained by calculation from the trend equation and are considered as theoretical data of the observed phenomenon.

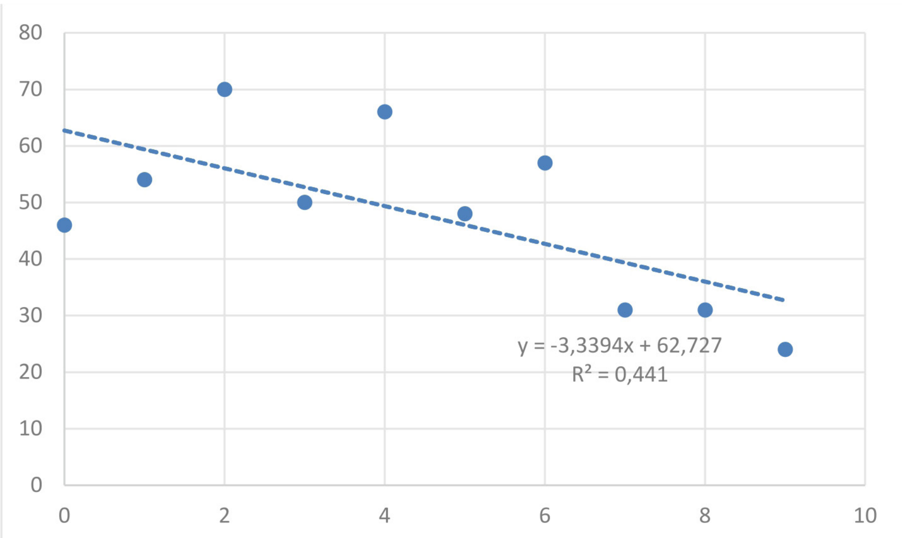

Figure 3: Graphic representation of the linear trend

The direction of the evaluated trend model has a negative slope, that is, it shows a negative trend of the observed time series.

Finally, the rated model is X=0 u 2011.godini

Value for x is 1 year

Value for y is 1 sinking of a ship

2.2. Represenativeness measures

In the paper, the representativeness of the model was evaluated using the coefficient of determination, which is based on the analysis of variance expression (9).

$$r^2 = \frac{SP}{ST}, \tag{9}$$

It is defined as the ratio of the sum of the squares of the deviations interpreted by the trend model SP and the sum of the squares of the total deviations, ST.

$$SP = a\sum_{i=1}^{N} Y_i + b\sum_{i=1}^{N} X_i Y_i - N\bar{Y}^2 \tag{10}$$

$$ST = \sum_{i=1}^{N} Y_i^2 - N\bar{Y}^2 \tag{11}$$

Inserting the values into the definition expressions for SP and ST yields:

$$SP = 62.72 \cdot 477 - 3.34 \cdot 1871 - 10 \cdot 47.7^2 = 915.4 \tag{12}$$

$$ST = 24839 - 10 \cdot 47.7^2 = 2086.1 \tag{13}$$

Value of determination coefficient:

$$r^2 = \frac{SP}{ST} = \frac{915.4}{2086.1} = 0.438 \approx 0.44 \tag{14}$$

3. Results

After evaluating the parameters of the trend model, the question of representativeness arises, i.e. the ability of the model to explain the movement of the dependent variable of the time series Y through time X.

The value of the coefficient usually ranges from 0 to 1 and the model is more representative if this value is closer to unity. Since r^2 is less than 0.6 in our analysis, the model is not representative enough. This means that 44% of the total squared deviations of the dependent variable of the time series from the arithmetic mean were interpreted by this model.

4. Discussion

The importance of the trend equation is that it can be used for estimation, i.e. to calculate the expected frequency for a given or selected time unit. This assessment is valid for a phenomenon that can be accepted to move in the future at the same pace as before [8].

Thus, based on the model, we can calculate the number of drownings in, for example, 2025, X=14, which according to the trend equation would amount to :

$$Y= 62.72- 3.34\cdot 14 = 15.96 \approx 16$$

5. Conclusion

Every phenomenon, the movement of which is presented as a time series, is influenced by several parameters, whose influence is visible from the graphic display, but sometimes a more complex statistical analysis is needed in order to establish and analyze the connections between the parameters.

The linear trend model explains the linear movement, positive or negative, of the values of the observed time series over time.

In addition to displaying the linear movement of the occurrence of the time series based on the evaluated model, the value of the occurrence for some future periods can be predicted, which is what this research wants to highlight. In this example, a short time series is shown, therefore the representativeness of the selected model is not fully confirmed. For further statistical research, it is desirable to analyze longer series for which more relevant conclusions can be drawn during the observed time period.

6. References

1. Available from: https://www.agcs.allianz.com/content/dam/onemarketing/agcs/agcs/reports/AGCS-Safety-Shipping-Review-2021.pd
2. Annual overview of marine casualties and incidents 2022, https://www.emsa.europa.eu/newsroom/latest-news/item/4867-annual-overview-of-marine-casualties-and-incidents-2021.html
3. BENŠIĆ, MIRTA;ŠUVAK, NENAD: *„An introduction to probability and statistics“*, Strossmayer, Department of Mathematics,Osijek, 2014.
4. DIZDAR, DRAŽAN: *„Quantitative methods“*, Faculty of Kinesiology, University of Zagreb, Zagreb 2006.
5. http://www.efos.unios.hr/imijoc/wp-content/uploads/sites/71/2013/09/Primjena-statisti%C4%8Dkih-metoda-s-naglaskom-na-trend-korelaciju-i-regresiju.pdf
6. https://emsa.europa.eu/mass/rbat.html
7. https://hrcak.srce.hr/file/111856
8. PEDIŠIĆ, ŽELJKO; DIZDAR, DRAŽAN: *„Handbook of Quantitative Methods “*, Faculty of Kinesiology, University of Zagreb ,Zagreb 2010.
9. PIVAC, SNJEŽANA; BOŠKO, ŠEGO: *„Textbook and collection of exercises“*, Alka script d.o.o. Zagreb, 2014.
10. ŠOŠIĆ, IVAN: *„Applied statistics “*, Školska knjiga, Zagreb 2004.
11. TOMAŠEVIĆ, MARKO: *„Statistical methods in research“*,Sveučilište u Splitu, Pomorski fakultet, Split 2007.
12. ZENZEROVIĆ, ZDENKA: *„Statistical handbook“*, Pomorski fakultet u Rijeci, Rijeka 2004.