WARSAW UNIVERSITY OF TECHNOLOGY Index 351733
FACULTY OF CIVIL ENGINEERING
COMMITTEE FOR CIVIL AND WATER ENGINEERING
POLISH ACADEMY OF SCIENCES ISSN 1230-2945

DOI: 10.24425/ace.2025.155100

ARCHIVES OF CIVIL ENGINEERING

Vol. LXXI ISSUE 3 2025

pp. 295 –309

Research paper

Time series analysis of hazardous events based on data recorded in a polish construction company

Zuzanna Woźniak[1], Tomasz Nowobilski[2], Bożena Hoła[3]

Abstract: The construction sector records a significant number of occupational accidents (A) and near-misses (NM), making it one of the most dangerous in the economy. In recent years, interest in near-miss events has been growing among researchers and practicing engineers, as they are considered precursors to occupational accidents. Based on a review of the literature on the subject and their own experience, the authors of the article conclude that there is a significant gap in research on near misses in the Polish construction industry. The authors believe that such studies are necessary in the context of accident reduction. The purpose of this article is to analyze the time series of near misses and accidents at work. The data used in the study come from the system of registration of hazardous events implemented in one of the Polish construction companies, recorded in 2015–2022. Due to the specific nature of construction work and the circumstances of the event, 8 categories of hazardous events were specified. For each category, a time series was built to inform about the dynamics of the changes taking place. Box plots were developed for random variables representing the time intervals between consecutive events, informing about the statistical characteristics of a given set of events (SHE_i). This research makes it possible to predict the occurrence of specific events over time and to introduce preventive measures in construction practice.

Keywords: accident at work, construction, near miss, safety management, time series

[1]MSc., Wrocław University of Science and Technology, plac Grunwaldzki 11, 50-384 Wroclaw, Poland, e-mail: zuzanna.wozniak@pwr.edu.pl, ORCID: 0000-0002-0168-5684
[2]PhD., Eng., Wrocław University of Science and Technology, plac Grunwaldzki 11, 50-384 Wroclaw, Poland, e-mail: tomasz.nowobilski@pwr.edu.pl, ORCID: 0000-0002-0599-7108
[3]Prof., PhD., Eng., Wrocław University of Science and Technology, plac Grunwaldzki 11, 50-384 Wroclaw, Poland, e-mail: bozena.hola@pwr.edu.pl, ORCID: 0000-0001-6630-8065

1. Introduction

The high accident rate in the construction industry has been confirmed in many scientific papers [1–7]. Despite the use of numerous collective and individual protective measures in construction, accidents at work still occur, causing significant injuries to workers [8]. For example, in 2022, 3703 people were injured in work accidents in Poland, including 41 fatalities and 48 serious accidents [9]. Work carried out at height, among other things, on construction scaffolding, is particularly dangerous [10, 11]. Since accidents at work happen all the time, the study of this negative phenomenon in terms of searching for its causes does not lose its relevance. The use of multivariate statistical analysis can significantly improve the level of safety in the construction industry, and precise data analysis is crucial in identifying risk factors [12].

Based on a review of the literature on the subject, the authors of the article found a research gap in the area of research on so-called near misses. To clarify the relationship between near-misses and an accidents, it is necessary to define both terms. An accident at work is considered to be a sudden event, caused by an external cause, resulting in injury or death, which occurred in connection with work [13]. On the other hand, a near-miss event is considered a precursor to a work accident [3, 14–16]. A near miss according to PN-ISO 45001:2018-06 [17] is an incident that does not result in injury and health complaints. A similar definition was formulated by Thoroman et al. [18], who defined these incidents as those that do not result in health damage, but in some cases may cause minor property damage. Near-miss incidents, when analyzed in detail, are a valuable source of information to help reduce occupational accidents and improve safety in the construction sector [5, 19]. Both occupational accidents and near-misses are hazardous events that occur due to violations of applicable labor standards and regulations, leading to disruptions in the proper course of the process [1, 10].

Near-miss incidents in the construction industry are an important area of interest for both human and economic reasons [20, 21]. They inform about the possibility of future accidents (caused by similar causes) that directly affect the health and lives of workers. The knowledge gained from the study of near-misses should be used to properly prepare workplaces on the construction site. Ignoring near-misses can lead to a workplace accident, which in turn, generates costs in terms of lost work time, stoppage of the construction process, material and image losses [1, 3]. By analyzing near misses and implementing the conclusions into construction practice, it is possible to effectively reduce the number of occupational accidents, which is a key factor in ensuring the health and safety of workers on construction sites.

Dangerous incidents on a construction site occur at random moments in time and form a stochastic process, which can be represented as a time series [14]. The purpose of the research presented in this article is to analyze the characteristics of the time series of near-misses and accidents at work and to determine the attributes of the random variable which is the time interval between consecutive incidents. The data used in the study comes from the hazardous event recording system implemented in one of the Polish construction companies. Information on the random variable of the time interval between consecutive incidents was presented in the form of box plots. The analysis made it possible to identify characteristic time patterns of incidents, which can help predict the timing of subsequent incidents (especially accidents), and more effective prevention and risk management in the construction sector. Ultimately, the results of the analysis can form the basis for developing occupational safety strategies to reduce accidents and improve working conditions on construction sites.

2. Materials and methods

The data used in the study concerned hazardous incidents recorded at one of the leading construction companies registered in Poland. The recorded incidents occurred between 2015 and 2022. The methodology of the study, proposed by the authors, included 6 stages.

Phase 1. Create a database. The data used in the study were extracted from a hazardous event registration system implemented at a leading Polish construction company. They were stored in two-dimensional tables, containing such data as the date of the incident, location of the incident, area/region, description of the incident, qualification, category, and direct cause of the incident. Each row of the table contained data on one event. The size of the analyzed set was 2396 hazardous events, including 716 occupational accidents (A) and 1680 near-misses (NM).

Phase 2. Analysis of the structure of hazardous events collected in the database. The set of all hazardous events (set of hazardous events – SHE) was divided into a set of near misses (SNM) and a set of accidents (SA):

$$SHE = SA \cup SNM \tag{2.1}$$

where: SHE – set of all events

Each of these collections was categorized by the immediate cause of the incident. The subsets were divided into 8 categories:

$$SHE_i; \quad i = 1, \ldots, 8 \tag{2.2}$$

where: SHE_1 – being hit by objects, SHE_2 – being run over by a vehicle/being hit by a vehicle, SHE_3 – human environment, SHE_4 – fall of a person SHE_5 – electricity, SHE_6 – fire/explosion/finding of unexploded munition, SHE_7 – collapse/burial/being trapped, SHE_8 – contact with moving machine parts.

The structure of the defined sets is shown in Fig. 1.

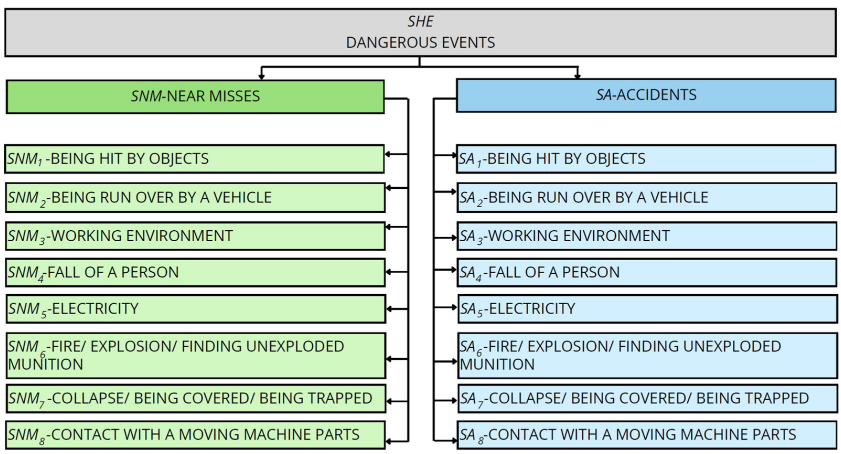


Fig. 1. Structure of hazardous events in the studied set

The set of occupational accidents (SA) is the sum of elements A_i ($i = 1, \ldots, 8$):

$$SA = \sum_{i=1}^{8} A_i \tag{2.3}$$

The set of near misses (SNM) is the sum of elements NM_i ($i = 1, \ldots, 8$):

$$SNM = \sum_{i=1}^{8} NM_i \tag{2.4}$$

Stage. 3. Build a time series of hazardous events and analyze them. The time series illustrates changes in the number of near-misses and occupational accidents over the months and years covered by the analysis. A time interval of 1 month was assumed in the study. For subsequent years, the average number of incidents was determined.

Stage. 4. Defining the time variable of the interval between events. A new variable was introduced to represent the time between consecutive events. This variable defines the number of days between the dates of consecutive events in the sequence:

$$\Delta t_j = t_{j+1} - t_j \tag{2.5}$$

where: Δt_j – the time between the jth and $(j + 1)$ events, t_j – time of occurrence of the jth event, t_{j+1}– time of occurrence of $(j + 1)$ event.

The procedure for determining intervals was performed for eight categories of incidents, for occupational accidents as well as near misses.

Phase 5. Develop box plots for the defined variables. Box plots were prepared for sets of random variables created in each category of hazardous events. They present key statistics describing the variables, such as medians, deviations from them, mean values, outliers and extremes. Figure 2 shows the characteristics of these data using box plots.

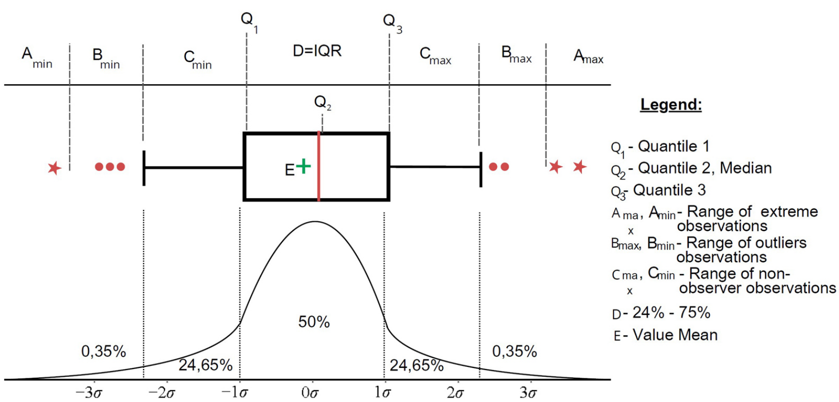


Fig. 2. Characteristics of the box plot with the comparison of probability density functions for normal distribution according to [22, 23]

Variables belonging to a specific category of hazardous events should be ordered in ascending order from minimum to maximum value.

The first quartile (Q_1) is the value of the random variable for which 25% of the random variables belonging to the category have values equal to or less:

$$Q_1 = \text{percentile}(25) \tag{2.6}$$

The median (Q_2) is the value of a random variable for which 50% of the random variables belonging to a given category have values equal to or less. The median divides an ordered data set into two parts from an equal number of observations. If the number of observations in the set is odd, the median is the value of the middle observation; if even, it is the average of the two middle values.

$$Q_2 = \text{percentile}(50) \tag{2.7}$$

The third quartile (Q_3) is the value of a random variable for which it is found 75% of the random variables belonging to the category have values equal to or less:

$$Q_3 = \text{percentile}(75) \tag{2.8}$$

The box expressed by the *IQR* formula covers the range from Q_1 to Q_3, which means that 50% of the data is inside the box range. The longer the graph, the more scattered the data. The range of the box data is:

$$IQR = Q_3 - Q_1 \tag{2.9}$$

Narrows labeled $C_{\min}$ and $C_{\max}$ also indicate the dispersion of the data. These are determined by quartiles, where the left whisker:

$$C_{\min} = Q_1 - 1,5 \cdot IQR \tag{2.10}$$

The right whisker is defined as:

$$C_{\max} = Q_3 + 1,5 \cdot IQR \tag{2.11}$$

Outliers are represented by points, and extreme values are represented by asterisks – these are observations that deviate from the majority of values in the dataset. Outlier observations mean values greater than the length of the mustache, so $B_{\min}$ and $B_{\max}$, and extreme values of $A_{\min}$ and $A_{\max}$ can be the result of measurement errors, unusual situations, or real anomalies in the data.

$$B_{\min} \leq Q_1 - 1,5 \cdot IQR \tag{2.12}$$

$$B_{\max} \geq Q_3 + 1,5 \cdot IQR \tag{2.13}$$

$$A_{\min} \leq Q_1 - 3 \cdot IQR \tag{2.14}$$

$$A_{\max} \geq Q_3 + 3 \cdot IQR \tag{2.15}$$

The average (E), marked with a green plus, symbolizes average values including all observations:

$$E = \text{medium}(data) \tag{2.16}$$

Phase. 6. Summary of study results and final villages.

3. Results and discussion

Figure 3 shows the monthly distribution of the number of occupational accidents (*A*) and near-misses (*NM*) recorded during the study period. The red colour represents *SA*, and the green colour refers to *SNM*. The dashed lines indicate the average value for *SA*-the red dashed line and *SNM*-the green dashed line for each year, respectively, which allows a quick assessment of changes over time and identification of trends in both categories of incidents.

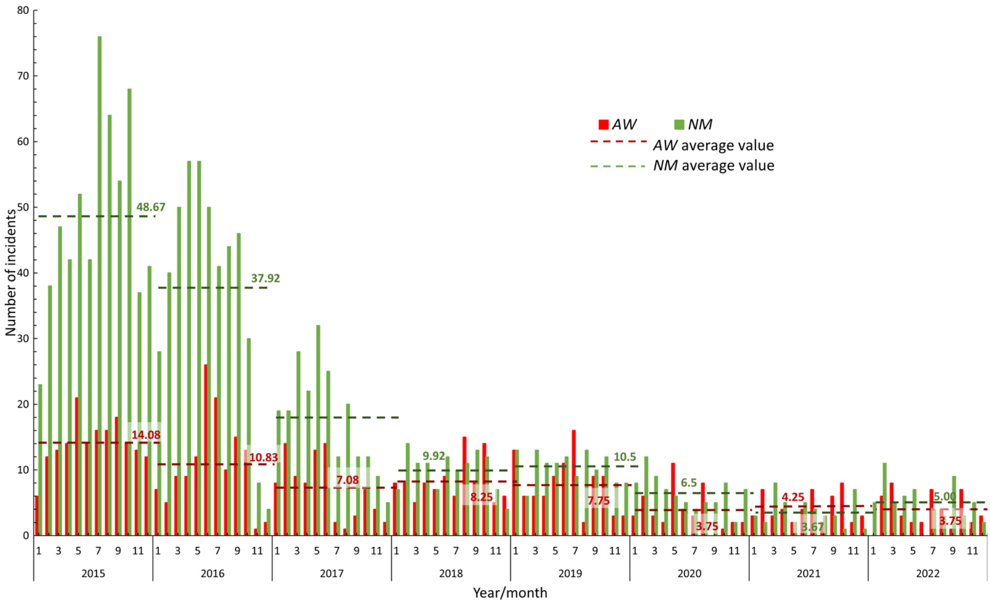
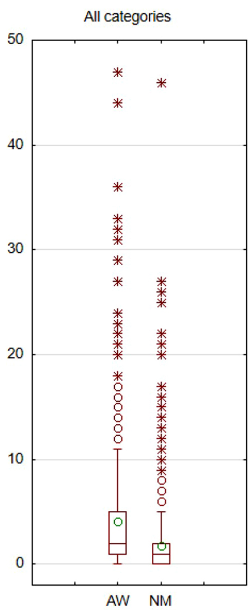


Fig. 3. Time series of *NM* and *A* events with average values by year

Both *NM* and *A* events show large fluctuations in the number of events in the time period studied and a downward trend in both. The variability in monthly data may be related to specific project conditions or seasonal changes in construction activity. In 2015, the average number of *NM* events for one month was 48.67, while the average number of *A* events for one month was 37.92. 2016 saw a 23.08% decrease in the average number of *A* events to 10.83 events/month and a 22.09% reduction in the number of *NM* events to 37.92 events/month compared to the previous year. The year 2017 is characterized by a further decline. The number of *A* decreased by 34.62% to an average of 7.08 events/month, and *NM* by 52.75% to 17.92 events/month. In 2018, there is a reversal of the trend for *A*, with an increase in the average number of accidents by 16.47% to 8.25 incidents/month, while the number of *NM* continues the downward trend, decreasing by 44.65% to an average of 9.92 incidents/month. 2019 registers a slight decrease in the number of *A* to 7.75 events/month and an increase in *NM* to 10.50 events/month, a change of 6.06% for *A* and 5.85% for *NM*, respectively, compared to the previous year. The year 2020 brings a marked decrease in *A* by 51.61% to 3.75 events/month and *NM* by 38.10% to 6.50 events/month. 2021 sees a 13.33% increase in the number of *A* over the previous year to a value of 4.25 events/month, while the number of *NM* again drops by 43.59% to 3.67 events/month. In 2022, there was a further decrease in the average number of accidents to 3.75 events per month, but a noticeable increase in the number of near-misses by 36.36% to an average of 5.00 events per month. In the first two years of the study interval, there is a noticeable difference between the number of *NM*s recorded and the number of *A* events. The results of the study, concerning the years 2015–2021, confirm the regularity previously found by Heinrich [24], Bird [25], Zimmerman and Bauer [26], according to which, before an accident occurs, there will be at least several near misses with similar causes that do not result in injury to employees. However, in subsequent years, the gap between the number of near misses recorded and the number of accidents decreases. The question can be asked: what are the reasons for such a phenomenon? The authors of the study note several factors that may have influenced such shaping of the near-misses process, namely: improvement of working conditions on construction sites and increased awareness of risks among workers, reduced enthusiasm among workers for reporting near-misses due to, for example, fear of sanctions or the time-consuming nature of reporting activities [27] or a reduction in hours worked and people employed due to the COVID-19 pandemic. The pandemic has significantly disrupted work in the construction industry, leading to material shortages, project delays, and changes in workforce dynamics due to safety concerns and worker availability [28, 29]. According to the authors, the predominant factor contributing to the decline in the number of recorded events was human influence, which is inherently unpredictable.

Figure 4 shows the box plot and characteristics of the random variable interval between events for all categories.



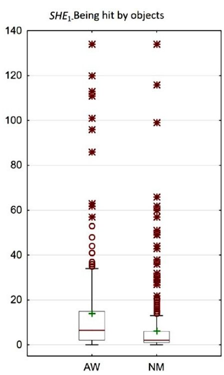
All categories	*SNM*	*SA*
Median, Q_2	1	2
Value Mean, E	2	4
Standard deviation	3	6
Min	0	0
Max	46	47
Q_1; Quartile 1	0	1
Q_3; Quartile 3	2	5
Outliers observations, B_{max}, B_{min}	6; 7; 8	54; 56; 62; 69
Extreme obseravtions, A_{max}, A_{min}	9; 10; 11; 12; 13; 14; 15; 16; 17; 20; 21; 22; 25; 26; 27; 46	79; 115

Fig. 4. Box plot of *NM* and *A* events for all categories combined

In the dataset analysed, at all construction sites operated by the subject company, the average *NM* event interval was 2 days, while the average *A* event interval was 4 days. 50% of *NM* events occurred in an interval of 1 day or less, while for *A*, 50% of events occurred in an interval of 2 days or less. The above figures indicate that the incidence of *NM* is twice as high as that of *A*, confirming the aforementioned regularity in which an occupational accident is preceded by several occurrences of near misses [24, 25]. In addition, it is also important to note that of all recorded *NM* events, 50% of them will happen in the interval from 0 to 2 days and the case of *A* in the interval from 1 to 5 days.

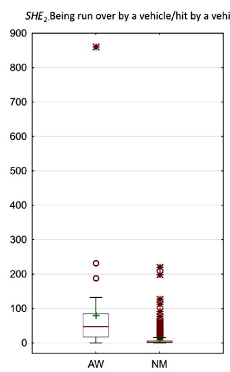
Figure 4 also shows outlier observations marked with circles and extreme observations marked with crosses. These can be due to various reasons such as delays in entering information into the database or mistakes. In addition, they can also be influenced by long interruptions due to the holiday calendar or seasonality of work. The analysis of the entire set of *SHE* events carried out is not very precise, as it does not provide a picture of how individual circumstances and types of work performed affect the initiation of hazardous events. In the next stage of the study, statistical analyses were carried out covering observations belonging to individual subsets of SHE_i; $i = 1, \ldots, 8$.

Figure 6 shows box plots illustrating the distribution of the time interval between consecutive *A* events and consecutive *NM* events in each of the defined categories.



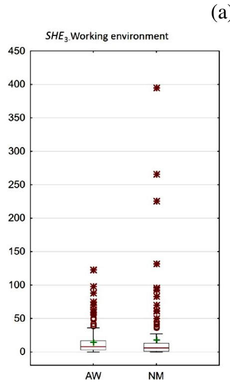
SHE_1-Being hit by objects	SNM_1	SA_1
Median, Q_2	2	7
Value Mean, E	6	14
Standard deviation	13	22
Min	0	0
Max	134	134
Q_1; Quartile 1	1	2
Q_3; Quartile 3	6	15
Outliers observations, B_{max}, B_{min}	14; 15; 16; 17; 18; 19; 20; 21	35; 36; 37; 41; 44; 48; 53
Extreme obseravtions, A_{max}, A_{min}	22; 27; 28; 29; 30; 32; 36; 38; 43; 44; 49; 51; 57; 60; 62; 66; 99; 116; 134	57; 62; 63; 86; 96; 101; 111; 113; 120; 134

(a)



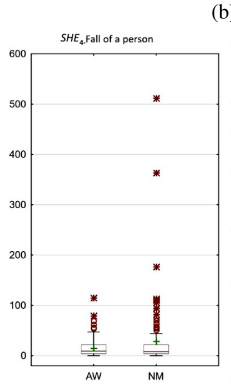
SHE_2-Being run over by a vehicle/hit by a vehicle	SNM_2	SA_2
Median, Q_2	3	47
Value Mean, E	9	80
Standard deviation	21	146
Min	0	0
Max	220	861
Q_1; Quartile 1	1	19
Q_3; Quartile 3	7	78
Outliers observations, B_{max}, B_{min}	17; 19; 20; 21; 22; 24; 25	188; 232
Extreme obseravtions, A_{max}, A_{min}	28; 31; 33; 35; 39; 43; 50; 55; 56; 60; 64; 77; 92; 113; 128; 199; 220	861

(b)



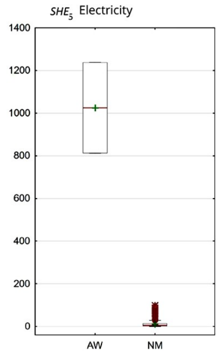
SHE_3-Working environment	SNM_3	SA_3
Median, Q_2	6	8
Value Mean, E	18	15
Standard deviation	47	20
Min	0	0
Max	395	123
Q_1; Quartile 1	2	3
Q_3; Quartile 3	13	17
Outliers observations, B_{max}, B_{min}	36; 37; 38; 41; 42; 45	39; 42; 49; 51; 56; 57; 59
Extreme obseravtions, A_{max}, A_{min}	50; 61; 62; 70; 83; 93; 96; 132; 226; 266; 395	60; 63; 70; 75; 88; 98; 123

(c)



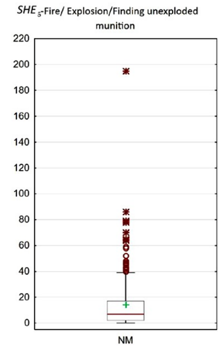
SHE_4-Fall of a person	SNM_4	SA_4
Median, Q_2	8	9
Value Mean, E	29	14
Standard deviation	67	16
Min	0	0
Max	512	115
Q_1; Quartile 1	4	4
Q_3; Quartile 3	21	22
Outliers observations, B_{max}, B_{min}	51; 55; 57; 64; 70; 76	54; 56; 62; 69
Extreme obseravtions, A_{max}, A_{min}	82; 94; 95; 106; 110; 113; 177; 364; 512	79; 115

(d)



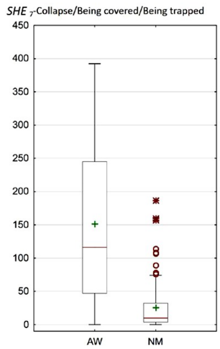
SHE_5-Electricity	SNM_5	SA_5
Median, Q_2	6	1026
Value Mean, E	11	1026
Standard deviation	15	301
Min	0	813
Max	101	1238
Q_1; Quartile 1	2	919
Q_3; Quartile 3	13	1132
Outliers observations, B_{max}, B_{min}	30; 33; 34; 35; 38; 39; 41; 43; 45; 46	-
Extreme obseravtions, A_{max}, A_{min}	48; 61; 62; 63; 66; 73; 77; 82; 101	-

(e)



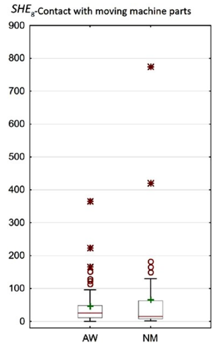
SHE_6-Fire/ Explosion/Finding unexploded munition	SNM_6	SA_6
Median, Q_2	7	-
Value Mean, E	14	-
Standard deviation	21	-
Min	0	-
Max	195	-
Q_1; Quartile 1	2	-
Q_3; Quartile 3	17	-
Outliers observations, B_{max}, B_{min}	40; 41; 42; 43; 44; 45; 46; 47; 48; 52; 58; 59	-
Extreme obseravtions, A_{max}, A_{min}	64; 65; 70; 78; 79; 86; 195	-

(f)



SHE_7-Collapse/Being covered/Being trapped	SNM_7	SA_7
Median, Q_2	10	116
Value Mean, E	25	151
Standard deviation	36	126
Min	0	0
Max	187	392
Q_1; Quartile 1	4	51
Q_3; Quartile 3	32	236
Outliers observations, B_{max}, B_{min}	76; 78; 89; 107; 114	-
Extreme obseravtions, A_{max}, A_{min}	157; 160; 187	-

(g)



SHE_8-Contact with moving machine parts	SNM_8	SA_8
Median, Q_2	15	25
Value Mean, E	66	46
Standard deviation	134	63
Min	1	0
Max	774	365
Q_1; Quartile 1	8	11
Q_3; Quartile 3	55	48
Outliers observations, B_{max}, B_{min}	149; 165; 182	166; 223; 365
Extreme obseravtions, A_{max}, A_{min}	420; 774	113; 123; 129; 151

(h)

Fig. 6. Box plots of *NM* and *A*

The analysis of the obtained results shows that depending on the type of construction work and the circumstances under which it is carried out, the incidence of near-misses and accidents at work, which is described by the interval between consecutive events, varies.

Comparing the values of the average intervals between consecutive *NM* and *A* events in the SHE_1, SHE_2, SHE_5, and SHE_6 categories, it should be noted that *NM* events occur at shorter intervals than *A* events. Thus, the occurrence of *NM* in these categories may be an indicator signalling safety problems that, if ignored, could lead to an *A* event. The significant differences in medians, especially for SA_i, indicate the need to approach each category individually.

In the SHE_1 – being hit by objects category, *NM*s occur in short intervals (median: 2 days), while *A* events have a larger median of 7 days. 50% of NM_1s occur at 3.5 times shorter intervals compared to A_1. Therefore, it can be concluded that near misses precede the occurrence of occupational accidents in this category. The average time interval *E* between the occurrence of the next event in the SNM_1 subset is 6 days, while in the SA_1 subset, it is more than twice as long at 14 days. 50% of the observations in the SNM_1 subset are between 1 and 6 days, while in the SA_1 subset of events, the range is between 2 and 15 days.

In the SHE_2 – being run over by a vehicle/being hit by a vehicle category, NM_2 occur much more frequently (median: 3 days) than A_2 (median: 47 days). *E* between the occurrence of the next event in the SNM_2 subset is 9 days, and in the SA_2 subset, it is 80 days. 50% of the events in the SNM_2 subset occurred between 1 and 7 days apart, and 50% of A_2 of this category occurred between 19 and 78 days apart. The obtained average values and time intervals, especially for occupational accidents, were very much influenced by outliers and extremes. Eliminating these values from the study population would have changed the results significantly.

The slight differences in medians in the work environment – SHE_3 category suggest a similar frequency of *NM* and *A* events. The median time interval for the SNM_3 subset is 6 days, and for the SA_3 subset is 8 days. The median event interval for the SNM_3 subset is 18 days and for the SA_3 subset is 15 days. 50% of *NM* events occurred with an interval of 2 to 8 days and 50% of occupational accidents in this category occurred with an interval of 3 to 17 days.

In the SHE_4 – fall of a person, similar medians were found in the SNM_4 and SA_4 subsets and are 8 and 9 days, respectively The average interval between incident occurrence differs by 15 days and is 29 for the SNM_4 subset and 14 for the SA_4 subset. 50% of near-miss incidents occurred with an interval of 4 to 21 days and 50% of occupational accidents in this category occurred with an interval of 4 to 22 days.

The standout category is SHE_5 – electricity. The SNM_5 subset analyzed included 79 near-misses and only 3 workplace accidents. The median interval of SNM_5 is 6 days. The average interval between events in the SNM_5 subset is 18 days. 50% of the events occurred between 2 and 13 days apart. The small number of events in the SA_5 subset means that the results obtained are not statistically reliable. In the case of electric current, the effect of electrocution depends mainly on the amount of energy released [30]. A large dose of energy can cause significant injury to a worker, while a small dose will not cause damage and such an event will not qualify as a workplace accident. It is likely that in the incidents analyzed, in many cases, workers were not injured due to too small an electrical discharge.

In the SHE_6 category of fire/explosion/ finding of unexploded munition, only near misses were recorded. In the SNM_6 subset, the median is 7 days, while the average value of the event interval is 17 days. 50% of the observations fall between 2 and 17 days. The majority of NM_6 involved finding unexploded ordnance and unexploded bombs. The authors surmise that due to strict procedures for dealing with unexploded ordnance/unexploded ordnance [31] accidents very rarely occur.

There is a large difference in the values of the individual statistical characteristics in the SHE_7 – collapse/burial/being trapped category. The median in the SA_7 subset is 116 days and in the SNM_7 subset is 10 days. These values indicate a less frequent occurrence of SA_7 subset events in this category compared to SA_1, SA_2, SA_3, SA_4, SA_8. Near-miss incidents in SHE_7 occur an average of 14.64 times per year, while occupational accidents occur 2.42 times.

In the SHE_8-contact with moving machine parts category, NM_8 occurred more frequently than occupational accidents. The median in the SNM_8 subset is 15 days, and in the SA_8 subset is 25 days. The median interval in the SNM_8 subset of 66 days is significantly higher than in the SA_8 subset of 46 days, indicating that *NM* in this category occurs an average of 5.53 times per year, while *A* averages 7.93 times per year.

4. Conclusions

The results presented in the article provide a picture of the activities of the construction company in question in the context of occupational safety in the period from 2015 to 2022. A database containing information on registered near-misses and occupational accidents was constructed, their structure was examined, and a time series of events was developed. A random variable was created, being the time elapsed between consecutive events, which allowed for their in-depth analysis and visualization of their dynamics by means of box plots. The conclusions of the study are as follows:

- The number of near-misses in the subsequent years of the study interval is greater than the number of occupational accidents. The exception is the year 2021, in which the opposite situation was observed.
- The constructed time series for both types of incidents show a definite downward trend in 2015–2019. In 2020–2022, the average number of reported incidents of both types is at a similar level.
- The decreasing ratio of near-misses to occupational accidents may indicate the presence of various factors influencing the development of both phenomena. These include improved working conditions at construction sites, increased awareness of risks among workers, reduced enthusiasm among workers for reporting near-misses due to, for example, fear of sanctions or the time consuming nature of reporting activities, and a reduction in the number of hours worked and people employed due to the COVID-19 pandemic. This aspect will be the subject of future research and analysis. Since the number of recorded potential accident events is decreasing, research should be undertaken to explain this phenomenon.
- The box plots illustrate the distribution of intervals between near misses and occupational accidents. Analysis of the obtained results indicates that depending on the circumstances under which construction work is carried out and the type of work, the incidence of hazardous events varies The smallest average interval between incidents was observed for the SNM_1 subset – being hit by objects of 6 days.
- Similarly, in the case of occupational accidents, the smallest average time interval between incidents was observed in the subsets SA_1 – being hit by objects and SA_4 – fall of a person, which was 14 days. In the SHE_6 – fire/explosion/finding of unexploded munition category, only one accident was recorded. The small number of events in the subset of occupational accidents in the electricity category means that the results obtained are not statistically reliable.
- Thanks to the intervals defined in the article between the occurrence of consecutive near-misses and occupational accidents, it is possible to predict the occurrence of an occupational accident. Implementation of the conclusions of the study will contribute to more accurate accident forecasting, improved safety management practices, better risk assessment.

The research conducted also has its limitations. On the one hand, a weakness of the research is that it was conducted only in one construction company, and this makes it impossible to generalize the results to the entire construction industry. On the other hand, the limitation of the research to only one company is advantageous, since the noted errors and deficiencies point to specific directions for preventive measures in the surveyed company. Extending the study to a larger number of enterprises is difficult due to the very limited access to data on near misses and accidents at work. This is due to the fact that their registration is not common in Poland, which makes it much more difficult to conduct research on a larger scale. In addition, companies are reluctant to agree to access event registration systems. This limits the ability to compare results between companies.

The authors plan further research to develop mathematical models, including those based on linear regression, artificial intelligence and machine learning, which, based on data on near misses, could predict the timing of workplace accidents in, for example, real time. They are also considering gaining access to the records of hazardous events at other construction companies so that they can compare their findings.

Acknowledgements

The authors received funding from the project supported by the National Centre of Science, Poland [grant no. 2021/43/O/ST8/00724 "Modelling the impact of near misses on occupational accidents in the construction industry (SAFCON)"]

References

[1] Z. Woźniak, B. Hoła, J. Ciszewski, and M. Stelmaszczyk, "Znaczenie zdarzeń potencjalnie wypadkowych dla bezpieczeństwa pracy w budownictwie (Significance of near misses for work safety in construction)", *Przegląd Budowlany*, vol. 93, no. 9-10, 2022 (in Polish).

[2] R. Irumba, "Spatial analysis of construction accidents in Kampala, Uganda", *Safety Science*, vol. 64, pp. 109–120, 2014, doi: 10.1016/J.SSCI.2013.11.024.

[3] Z. Woźniak and B. Hoła, "The structure of near misses and occupational accidents in the polish construction industry", *Heliyon*, vol. 10, no. 4, 2024, doi: 10.1016/J.HELIYON.2024.E26410.

[4] B.H.C. Lee, J.C. Chen, and K.W. Fo, "Accidents in Construction Sites: A Study on the Causes and Preventive Approaches to Mitigate Accident Rate", *Inti Journal*, vol. 1, 2018.

[5] F. Hashmi, M.U. Hassan, M.U. Zubair, K. Ahmed, T. Aziz, and R.M. Choudhry, "Near-Miss Detection Metrics: An Approach to Enable Sensing Technologies for Proactive Construction Safety Management", *Buildings,* vol. 14, no. 4, 2024, doi: 10.3390/BUILDINGS14041005.

[6] Z. Woźniak, B. Hoła, J. Ciszewski, and M. Stelmaszczyk, "Wskaźnik predykcji wypadków na podstawie liczby zdarzeń niebezpiecznych w polskim przedsiębiorstwie budowlanym (Accident prevention index based on the number of dangerous events in a Polish construction company)", *Przegląd Budowlany*, vol. 94, no. 9–10, pp. 184–189, 2023, doi: 10.5604/01.3001.0053.9396 (in Polish).

[7] M. Kowalik and W. Drozd, "Evaluation of the influence of selected factors on hazardous events in construction", *Archives of Civil Engineering*, 2022, vol. 68, no. 2, pp. 377–390, doi: 10.24425/ace.2022.140648.

[8] B. Hoła and M. Szóstak, "An Occupational Profile of People Injured in Accidents at Work in the Polish Construction Industry", *Procedia Enineering*, vol. 208, pp. 43–51, 2017, doi: 10.1016/J.PROENG.2017.11.019.

[9] Central Statistical Office, "Accidents at Work in 2022". [Online]. Available: https://stat.gov.pl/obszary-tematyczne/rynek-pracy/warunki-pracy-wypadki-przy-pracy/wypadki-przy-pracy-w-2022-roku,4,16.html. [Accessed: 25. Jul. 2024].

[10] M. Szóstak, B. Hoła, and P. Bogusławski, "Identification of accident scenarios involving scaffolding", *Automation in Construction*, vol. 126, art. no. 103690, 2021, doi: 10.1016/J.AUTCON.2021.103690.

[11] T. Nowobilski and B. Hoła, "The Qualitative and Quantitative Structure of the Causes of Occupational Accidents on Construction Scaffolding", *Archives of Civil Engineering*, vol. 65, no. 2, 2019, doi: 10.2478/ace-2019-0023.

[12] B. Hola and T. Nowobilski, "Classification of Economic Regions with Regards to Selected Factors Characterizing the Construction Industry", *Sustainability,* 2018, vol. 10, no. 5, art. no. 1637, 2018, doi: 10.3390/SU10051637.

[13] "Ustawa z dnia 30 października 2002 r. o ubezpieczeniu społecznym z tytułu wypadków przy pracy i chorób zawodowych (Act of October 30, 2002 on Social Insurance for Occupational Accidents and Occupational Diseases)", *Journal of Laws*, 2002, no. 199, item 1673, as amended (in Polish).

[14] C. Zhou, L. Ding, M.J. Skibniewski, H. Luo, and S. Jiang, "Characterizing time series of near-miss accidents in metro construction via complex network theory", *Safety Science*, vol. 98, pp. 145–158, 2017, doi: 10.1016/J.SSCI.2017.06.012.

[15] F. Chen, C. Wang, J. Wang, Y. Zhi, and Z. Wang, "Risk assessment of chemical process considering dynamic probability of near misses based on Bayesian theory and event tree analysis", *Journal of Loss Prevention in the Process Industries*, vol. 68, art. no. 104280, 2020, doi: 10.1016/J.JLP.2020.104280.

[16] M.G. Gnoni and J.H. Saleh, "Near-miss management systems and observability-in-depth: Handling safety incidents and accident precursors in light of safety principles", *Safety Science*, vol. 91, pp. 154–167, 2017, doi: 10.1016/J.SSCI.2016.08.012.

[17] PN-ISO 45001:2018-06 Systemy zarządzania bezpieczeństwem i higieną pracy. Wymagania i wytyczne stosowania. PKN, 2019.

[18] B. Thoroman, N. Goode, and P. Salmon, "System thinking applied to near misses: a review of industry-wide near miss reporting systems", *Theoretical Issues in Ergonomics Science,* vol. 19, no. 6, pp. 712–737, 2018, doi: 10.1080/1463922X.2018.1484527.

[19] W. Wu, A.G.F. Gibb, and Q. Li, "Accident precursors and near misses on construction sites: An investigative tool to derive information from accident databases", *Safety Science*, vol. 48, no. 7, pp. 845–858, 2010, doi: 10.1016/J.SSCI.2010.04.009.

[20] Z. Woźniak and B. Hoła, "Zdarzenia potencjalnie wypadkowe jako prekursory wypadków przy pracy w budownictwie (Near misses incidents as precursors of accidents at work in the construction industry)", *Materiały Budowlane*, no. 1, pp. 74–76, 2022, doi: 10.15199/33.2022.01.10 (in Polish).

[21] U.B. Celebi and F. Yilmaz, "The Importance of Safety in Construction Sector: Costs of Occupational Accidents in Construction Sites", *Business and Economic Research Journal*, vol. 6, no. 2, pp. 25–37, 2015.

[22] A.S. Perera, "Defining fuzzy membership function using box plot", *International Journal of Research in Computer Applications and Rrobotics*, vol. 5, pp. 1–10, 2017.

[23] Atlassian, "Box Plot: A Complete Guide". [Online]. Available: https://www.atlassian.com/data/charts/box-plot-complete-guide. [Accessed: 26. Jul. 2024].

[24] H.W. Heinrich, Industrial Accident Prevention. A Scientific Approach. New York, London: McGraw-Hill Insurance Series, 1931.

[25] F. Bird, *Management Guide to Loss Control*. Houston, TX: International Safety Academy, 1974.

[26] N. Zimmermann and R. Bauer, *Injuries in European Union. Statistics summary 2002-2004*. Austria: Eurostat Pocket Books, 2006.

[27] P. Duan and J. Zhou, "A science mapping approach-based review of near-miss research in construction", *Engineering Construction and Architectural Management*, vol. 30, no. 6, pp. 2582–2601, 2022, doi: 10.1108/ECAM-09-2021-0797.

[28] M. Bou Hatoum, A. Faisal, H. Nassereddine, and H. Sarvari, "Analysis of COVID-19 Concerns Raised by the Construction Workforce and Development of Mitigation Practices", *Frontiers in Built Environment*, vol. 7, art. no. 688495, 2021, doi: 10.3389/FBUIL.2021.688495/BIBTEX.

[29] H.A. Rani, A.M. Farouk, K.S. Anandh, S. Almutairi, and R.A. Rahman, "Impact of COVID-19 on Construction Projects: The Case of India", *Buildings,* vol. 12, no. 6, art. no. 762, 2022, doi: 10.3390/BUILDINGS12060762.

[30] F.B. Cambraia, T.A. Saurin, and C.T. Formoso, "Identification, analysis and dissemination of information on near misses: A case study in the construction industry", *Safety Science*, vol. 48, no. 1, pp. 91–99, 2010, 2010, doi: 10.1016/J.SSCI.2009.06.006.

[31] A. Fik and M. Puchała, "Funkcjonowanie patroli saperskich jako przykład współdziałania logistyki cywilnej (The operation of sapper patrols as an example of cooperation with civilian logistics)", *Zarządzanie Innowacyjne w Gospodarce i Biznesie: Czasopismo Naukowe o Problemach Współczesnego Zarządzania*, no. 2, pp. 49–60, 2015 (in Polish).

Analiza szeregów czasowych zdarzeń niebezpiecznych na podstawie danych zarejestrowanych w polskiej firmie budowlanej

Słowa kluczowe: budownictwo, szeregi czasowe, wypadki przy pracy, zarządzanie bezpieczeństwem, zdarzenia potencjalnie wypadkowe

Streszczenie:

W sektorze budowlanym notuje się znaczną liczbę wypadków przy pracy (*A*) oraz zdarzeń potencjalnie wypadkowych (*NM*), co sprawia, że jest on uznawany za jeden z najbardziej niebezpiecznych w gospodarce. W ostatnich latach zainteresowanie zdarzeniami potencjalnie wypadkowymi wśród naukowców i inżynierów praktyków wzrasta, gdyż są one uważane za prekursory wypadków przy pracy. Na podstawie przeprowadzonego przeglądu literatury przedmiotu oraz doświadczeń własnych, autorzy artykułu stwierdzają, że istnieje istotna luka w badaniach dotyczących zdarzeń potencjalnie wypadkowych w polskim budownictwie. Autorzy uważają, że takie badania są niezbędne w kontekście redukcji liczby wypadków. Celem niniejszego artykułu jest analiza szeregów czasowych zdarzeń potencjalnie wypadkowych i wypadków przy pracy. Dane wykorzystane w badaniach pochodzą z systemu rejestracji zdarzeń niebezpiecznych zaimplementowanego w jednej z polskich firm budowlanych, zarejestrowanych w latach 2015–2022. Ze względu na specyfikę robót budowlanych i okoliczności zdarzenia, wyszczególniono 8 kategorii zdarzeń niebezpiecznych. Dla każdej kategorii zbudowano szereg czasowy informujący o dynamice zachodzących zmian. Opracowano wykresy skrzynkowe, dla zmiennych losowych odstępu czasu między kolejnymi zdarzeniami, informujące o charakterystykach statystycznych danego zbioru zdarzeń (SHE_i). Badania te pozwalają na prognozowanie wystąpienia określonych zdarzeń w czasie oraz wprowadzanie środków zapobiegawczych w praktyce budowlanej. Zbiór wszystkich zdarzeń niebezpiecznych (*SHE*) został podzielony na zbiór zdarzeń potencjalnie wypadkowych (*SNM*) i zbiór wypadków (*SA*). Każdy z tych zbiorów został poddany kategoryzacji ze względu na bezpośrednią przyczynę zdarzenia. Podzbiory zostały podzielone na 8 kategorii: SHE_1 – uderzenie przedmiotami, SHE_2 – najechanie / potrącenie, SHE_3 – środowisko pracy, SHE_4 – upadek człowieka, SHE_5 – elektryczność, SHE_6 – pożar / wybuch / odnalezienie niewybuchu, SHE_7 – zawalenie / przysypanie / uwięzienie, SHE_8 – kontakt z ruchomymi elementami maszyn. Wprowadzono nową zmienną reprezentującą czas między kolejnymi zdarzeniami. Zmienna ta określa liczbę dni pomiędzy datami następujących po sobie zdarzeń w sekwencji. Procedura wyznaczania odstępów czasowych została wykonana dla ośmiu kategorii zdarzeń, w odniesieniu do wypadków przy pracy jak i zdarzeń potencjalnie wypadkowych. Opracowano wykresy skrzynkowe dla zestawów zmiennych losowych utworzonych w każdej kategorii zdarzeń niebezpiecznych. Prezentują one kluczowe statystyki opisujące zmienne, takie jak mediany, odchylenia od nich, wartości średnie, wartości odstające oraz ekstremalne. Zarówno w przypadku zdarzeń *NM* jak i *A*, widoczne są duże fluktuacje liczby zdarzeń w badanym przedziale czasu oraz tendencja spadkowa obu zjawisk. Zmienność danych miesięcznych może być związana ze specyficznymi warunkami projektów lub sezonowymi zmianami w działalności budowlanej. W pierwszych dwóch latach badanego przedziału czasu zauważalna jest duża różnica między liczbą rejestrowanych *NM* a liczbą zdarzeń *A*. Zbudowane szeregi czasowe dla obu typów zdarzeń wykazują zdecydowaną tendencję spadkową w latach 2015–2019. W latach 2020–2022 średnia liczba zgłaszanych zdarzeń obu rodzajów, jest na podobnym poziomie. Malejący stosunek liczby zdarzeń potencjalnie wypadkowych do liczby wypadków przy pracy może wskazywać na obecność różnych czynników wpływających na kształtowanie się obu zjawisk. Są to m.in.: poprawa warunków pracy na budowach, podniesienie świadomości o zagrożeniach wśród pracowników, zmniejszony wśród pracowników zapałdo zgłaszania zdarzeń potencjalnie wypadkowych spowodowany np. strachem przed sankcjami lub czasochłonnością czynności związanych ze zgłaszaniem zdarzeń, a także zmniejszenie liczby przepracowanych godzin i zatrudnionych osób spowodowane pandemią COVID-19. Dzięki zdefiniowanym w artykule odstępach

czasowych między wystąpieniem kolejnych po sobie zdarzeń potencjalnie wypadkowych i wypadków przy pracy możliwa jest prognoza wystąpienia wypadku przy pracy. Wdrożenie wniosków z przeprowadzonych badań przyczyni się do dokładniejszego prognozowania wypadków, poprawy praktyk zarządzania bezpieczeństwem, lepszą oceną ryzyka. Przeprowadzone badania mają również swoje ograniczenia. Główną słabością jest ich przeprowadzenie tylko w jednej firmie budowlanej, co utrudnia uogólnienie wyników na całą branżę budowlaną. Niestety istnieje ograniczony dostęp do danych o zdarzeniach potencjalnie wypadkowych, ponieważ ich rejestracja nie jest powszechna w Polsce, co znacznie utrudnia analizę na większą skalę. Dodatkowo, trudności w uzyskaniu zgody na dostęp do systemów rejestracji zdarzeń w innych przedsiębiorstwach ograniczają możliwości porównawcze z przeprowadzonymi badaniami.

Received: 2024-05-12. Revised: 2024-07-15