# MODELLING DELAYS IN BRIDGE CONSTRUCTION PROJECTS BASED ON THE LOGIT AND PROBIT REGRESSION

## A. Leśniak[1], M. Juszczyk[2], G. Piskorz[3]

Construction projects, even exemplarily planned and organized, bear a risk of unforeseen events and problems which can result in completion of the works after the deadline, that is delays. The construction of bridges is an inseparable part of road and rail projects and construction and expansion of the transport network. The paper aims at finding a relationship between the independent variables characterizing bridge projects and the delays during their implementation. Two alternative models were proposed to solve the problem: logit and probit. The data set comprising road and rail bridges built in Poland in the last 12 years (2005 – 2017) was used to build the models. The evaluation, quality and accuracy parameters of proposed models were determined in the final part of the paper.

*Keywords*: delays in construction projects, delays in bridge construction, logit regression, probit regression.

## 1. INTRODUCTION

A successful completion of a construction project means that it is finished within the cost budget, in time, according to the planned quality, is delivered safely and meets the customer's requirements. There are many papers devoted to the analysis of construction projects in terms of time [16, 24], costs [18, 28, 32, 40], logistics [43], process organizational aspects [12, 13, 14, 19, 22, 26, 37, 38],

---

[1] Ass, Prof., Ph.D., Eng.,.Cracow University of Technology, Faculty of Civil Engineering, Institute of Construction Management, Warszawska 24 st., 31-155 Cracow, Poland, e-mail: alesniak@l3.pk.edu.pl

[2] Ph.D., Eng., Cracow University of Technology, Faculty of Civil Engineering, Institute of Construction Management, Warszawska 24 st., 31-155 Cracow, Poland

[3] M.Sc., Eng., doktoral student, Cracow University of Technology, Faculty of Civil Engineering, Institute of Construction Management, Warszawska 24 st., 31-155 Cracow, Poland

work safety [15], innovative solutions [19] and project contracting [23]. There are many factors affectingthe course of a construction project in time, and frequently they can be foreseen already at the works planning stage. These are, i.e. project technical parameters, applied technological and organizational solutions. There is also a group of factors that are difficult to foresee, such as weather, breakdowns, incompetence of contractors. The occurrence of such factors may cause the contractual completion date to be exceeded [27]. The delay causes can be divided into two groups [27]: justified and unjustified, and the responsible party (the party bearing the consequences) can be either the Employer or the Contractor. The unjustified delays are a separate group, and usually are attributed to the Contractor. In such case the Contractor is not authorized to extend the time for completion and to receive remuneration. The Contractor-dependent factors are related mainly to the availability of resources, supervision and experience [27]. Many studies conducted in developed and developing countries have shown that delays in construction works are a common and frequently recurring problem in the implementation of construction projects [45]. The studies on identification of delay causes have been conducted, inter alia, in the United States [6], in the United Kingdom [33], in Saudi Arabia [2], in Thailand [45], in Turkey [20] and in Slovakia [31, 34]. More than 100 factors were identified as a result of these studies. The most frequently indicated delay causes included: insufficient Contractor's experience, errors in design documentation, design modifications at the implementation stage, Employer's problems with project financing, slow decision making on the part of the Employer, unrealistic time schedules. A few attempts to identify the causes of delays in construction projects have been also made in Poland [17, 21, 27, 30]. The studies presented in [29] and [30] conducted among the contractors indicated that that errors in design documentation, slow decision making on the part of the Employer, and poor quality site management are the main reasons of delays in construction projects. In [11] the authors presented the opinions of employers according to whom the main delay causes include errors in design documentation, low quality workforce, and unfavorable weather. In [30] the authors selected groups of factors which result from contractual provisions and can be diagnosed already at the pre-contract stage. In [5] it was proved that the most important delay-generating factors are errors in design documentation, poor cooperation between the Employer and the Contractor, and unfavorable weather. The authors proposed also various methods to analyze the factors causing the delays: factor analysis [11], cluster analysis [29], Dematel method [8], flexible operation strategies [39]. These studies prove a huge interest in the problem of construction project delays among scholars, but the vast majority of the papers deals with identification and grouping of delay causes. Models and methods have been proposed in the literature to effectively estimate the construction project duration [7, 10, 23, 41, 42], but there is a lack of

proposals of models to forecast the delay duration. Among the few, one can indicate [3, 4] in which the authors use the Artificial Neural Networks to estimate a credible duration of construction projects. Construction of bridges has a significant impact on the success of road or rail projects. The bridges are necessary to establish the continuity of transport routes and cross the obstacles. In most cases, the bridges are a part of a larger project involving the construction of a road which throughput they are to ensure. They are crucial points, often milestones for the Contractor.

The purpose of this paper is to solve the classification problem. The essence of the analysis is to search for the relationship between the independent variables characterizing bridge projects and the delays during their implementation.

## 2. MODEL CONCEPT

Classification is data exploration method which aim is to assign individual cases to appropriately defines classes whose number is limited. Each object is described by means of an information vector consisting of values from the domain of certain descriptive attributes and of one selected attribute called the decision attribute [36]. Its value divides the set of objects into predefined classes consisting of objects which have the same decision attribute value. The classification process is a construction of a model involving the search for a relationship (preferably a functional relationship) between the decision attribute and the descriptive attributes.

The research involved the search for a binomial model in which the explained variable Y is quantified by a dichotomic value (taking one of the two possible values). It was assumed that variable Y takes two possible variants defined by the codes: "YES" for occurrence of delay, and "NO" for absence of delay (the labels should equate with values 1 and 0, respectively). If $p_i$ is the probability of event $Y_i=1$, then $1-p_i$ is the probability of event $Y_i =0$ and the expected value of variable $Y_i$ is:

(2.1) $$E[Y_i] = 1 \cdot p_i + 0 \cdot (1 - p_i) = p_i$$

In binomial models it is assumed that $p_i$ is a function of the explanatory variables values vector $x_i$ for the $i$th object and parameters vector $\beta$:

(2.2) $$P_i = P(y_i = 1) = F(x_i^T \beta)$$

Depending of the type of function F, there are inter alia logit models (2.3) and probit models (2.4).

$$(2.3) \qquad P_i = F(x_i^T \beta) = \frac{1}{1 + \exp(-x_i^T \beta)}$$

$$(2.4) \qquad P_i = F(x_i^T \beta) = \int_{-\infty}^{x_i^T \beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

In order to build the classification model first it is necessary to create the database of cases and define the descriptive attributes (model input variables) and the decision attribute (output variable quantified by means of a dummy variable).

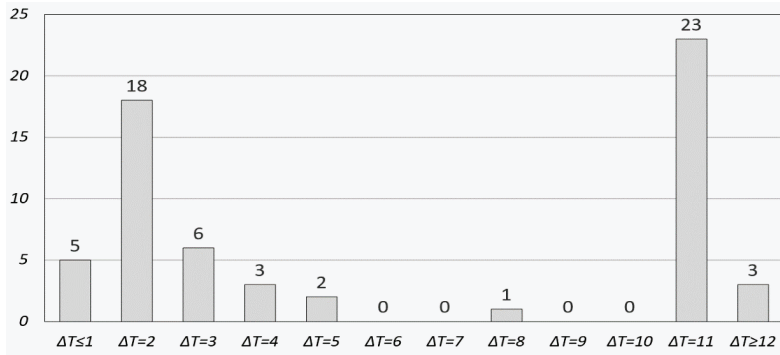## 2.1. DATABASE CONSTRUCTION AND DEFINING THE MODEL VARIABLES

The data from the database were taken from the analysis of design documentation and documents from the design and implementation stages of road and rail bridges built in Poland during 12 years (2005–2017). The information in database included the project implementation parameters, parameters of bridges, and the size of delay (if any). The data about 169 bridge projects were collected, including 48 projects without delays, and 121 with delays. In order to make the sizes of delayed and non-delayed subsets roughly equivalent, the half of cases were randomly removed from the delayed subset. Finally, 109 cases were used in the regression analysis. Delay in months $\Delta T$ was defined as follows:

$$(2.5) \qquad \Delta T = T_{act} - T_{plan}$$

where:

$T_{plan}$ – planned duration of a bridge construction,

$T_{act}$ – actual duration of a bridge construction.

Fig. 1. Number of bridge construction projects regarding delay $\Delta T$.

(source: own study)

Figure 1 presents the number of bridge projects with delay $\Delta T$ for cases included in the regression analysis after a random removal of the half of delayed cases.

Tables 1 and 2 show in total 15 parameters describing the project and the bridge, with their description, data type and original values. These parameters are independent (input) variables in the model.

Table 1. Variables related to general information about the bridge construction project (source: own study).

| Description | Symbol $X_j$ | Raw data type | original values | pseudo-fuzzy scaled values |
|---|---|---|---|---|
| Type of object | $X_1$ | categorical | Bridge<br>Viaduct<br>Wharf | 0.1<br>0.5<br>0.9 |
| Classification regarding the use of object | $X_2$ | categorical | Pedestrian/Bicycle<br>Roadway<br>Railroad | 0.1<br>0.5<br>0.9 |
| Type of a contract | $X_3$ | categorical | Build<br>Design and build | 0.1<br>0.9 |
| Total construction cost | $X_4$ | numerical | Total construction cost as-built in PLN | - |
| Planned duration of construction phase | $X_5$ | numerical | Planned duration in months | - |

Table 2. Variables related to parameters of bridge constructions (source: own study).

| Description | Symbol $X_j$ | Raw data type | original values | pseudo-fuzzy scaled values |
|---|---|---|---|---|
| Total length of a structure | $X_6$ | numerical | Total length | - |
| Width of a structure | $X_7$ | numerical | Width | - |
| Number of spans | $X_8$ | numerical | Number of spans | - |
| Type of a structural solution | $X_9$ | categorical | Beam<br>Arched/Framed/Box | 0.1<br>0.9 |

| Type of a structure regarding material | $X_{10}$ | categorical | Steel | 0.1 |
| | | | Reinforced concrete | 0.5 |
| | | | Prestressed concrete | 0.9 |
| Type of bridgehead supports | $X_{11}$ | categorical | None | 0.1 |
| | | | Sollid-walled | 0.5 |
| | | | Columns | 0.9 |
| Type of intermediate supports | $X_{12}$ | categorical | None | 0.1 |
| | | | Sollid-walled | 0.5 |
| | | | Columns/Piles | 0.9 |
| Type of suports' foundation | $X_{13}$ | categorical | Shallow | 0.1 |
| | | | Deep | 0.9 |
| Type of surface | $X_{14}$ | categorical | Roadway pavement | 0.1 |
| | | | Railway pavement | 0.9 |
| Load class | $X_{15}$ | categorical | Class A | 0.9 |
| | | | Class B/class C | 0.5 |
| | | | Class D/class E | 0.1 |

In variables whose original data were of a categorical type, a pseudo-fuzzy scaling to the interval <0.1 – 0.9> was used; the scaling method is also shown in the tables. In numerical variables, the linear scaling of original values to the interval <0.1 – 0.9> was used. Table 3 presents the basic descriptive statistics of variables used in the regression after data scaling and the values of coefficient of variation $U$ calculated according to the formula:

$$(2.6) \qquad U_j = \frac{\sigma_j}{\bar{x}_j}$$

where:

$\sigma_j$ – standard deviation of the sample for variable $X_j$,

$\bar{x}_j$ – mean value for variable $X_j$.

Table 3. Descriptive statistics for independent variables $X_j$ (source: own study).

| Symbol $X_j$ | average $\bar{x}_j$ | median $M_e$ | standard deviation $\sigma_j$ | coefficient of variation $U_j$ |
|---|---|---|---|---|
| $X_1$ | 0.554 | 0.715 | 0.249 | 0.448 |
| $X_2$ | 0.390 | 0.500 | 0.203 | 0.520 |
| $X_3$ | 0.522 | 0.500 | 0.202 | 0.388 |
| $X_4$ | 0.827 | 0.900 | 0.232 | 0.281 |
| $X_5$ | 0.174 | 0.145 | 0.102 | 0.585 |
| $X_6$ | 0.207 | 0.166 | 0.139 | 0.671 |
| $X_7$ | 0.250 | 0.219 | 0.111 | 0.444 |
| $X_8$ | 0.166 | 0.100 | 0.114 | 0.687 |
| $X_9$ | 0.283 | 0.100 | 0.338 | 1.192 |
| $X_{10}$ | 0.588 | 0.500 | 0.274 | 0.466 |
| $X_{11}$ | 0.504 | 0.500 | 0.102 | 0.202 |
| $X_{12}$ | 0.394 | 0.100 | 0.338 | 0.860 |

| | | | | |
|---|---|---|---|---|
| $X_{13}$ | 0.401 | 0.100 | 0.389 | 0.971 |
| $X_{14}$ | 0.225 | 0.100 | 0.292 | 1.297 |
| $X_{15}$ | 0.753 | 0.900 | 0.229 | 0.304 |

As indicated by Table 3 $U_j > 0.10$, for all variables $X_j$. the values of input variables proposed for the model have a sufficient variability and it is not necessary to remove quasi-constant variables. The output variable in the model is a qualitative variable representing "*YES*" for occurrence of delay, and "*NO*" for absence of delay.

# 3. CONSTRUCTION AND EVALUATION OF PROPOSED MODELS

## 3.1. LOGIT AND PROBIT REGRESSION

The logistic regression analysis and its variety, the probit regression, are indicated in the literature among the most widely used methods of solving the classification problem [1, 9]. The logistic regression model (so-called logit model) is used to explain qualitative variables depending on the level of explanatory variables (qualitative or quantitative). The explained variable $Y$ generally takes two values 1 and 0, representing two opposite events, e.g. a phenomenon occurred/ did not occur, which in the discussed case means: the delay in the bridge construction occurred/ did not occur [1, 25, 35]:

$$(2.7) \qquad Y_i^* = ln \frac{p_i}{1 - p_i} = \beta_0 \alpha_0 + \beta_1 X_{1i} + \beta_2 x X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i$$

where:

$\beta_j$ – model structural parameters,

$\varepsilon_i$ – random component,

$ln \frac{p_i}{1 - p_i}$ – logit,

$Y_i^*$ – qualitative non-observable variable [hidden],

$X_{ji}$ – values of model explanatory variables,

$p_i$ – probability of dependent variable $Y_i$ being 1 calculated based on the logistic probability density function:

$$(2.8) \qquad p_i = \frac{e^{x_i' \beta}}{1 + e^{x_i' \beta}} = \frac{1}{1 + e^{-x_i' \beta}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki})}}$$

The non-observable variable $Y_i^*$ is defined as a hidden variable because only the binary variable $Y_i$ is observed in the form: $Y_i = \begin{cases} 1; & Y_i^* > 0 \\ 0; & Y_i^* \leq 0 \end{cases}$

In the probit model, the probability value is determined by the normal distribution function:

$$(2.9) \qquad p_i = \int_{-\infty}^{z_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

where:

$z_i = \beta_0 + \beta_1 X_{1i} + \beta_2 x X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i$

## 3.2. ESTIMATION OF QUALITATIVE VARIABLE

An attempt was made to estimate the binomial qualitative variable ("YES" for occurrence of delay, and "NO" for absence of delay) and to find out which factors and with what weight and in what direction affect the construction project delay. Structural parameters of the logit and probit models $\beta j$, as given in the equations 2.6 and 2.7, were computed with the use of STATISTICA[TM] - an advanced analytics software package (Table 4).

Table 4. Structural parameters of the logit and probit models (source: own study).

| $X_j$ | Structural parameters of models $\beta_j$ | |
|---|---|---|
|  | Logit regression | Probit regression |
| $X_0$ | 2.371 | 1.015 |
| $X_1$ | -4.792 | -2.258 |
| $X_2$ | 4.613 | 2.680 |
| $X_3$ | -5.961 | -3.288 |
| $X_4$ | 4.533 | 2.531 |
| $X_5$ | -3.704 | -2.848 |
| $X_6$ | 15.352 | 9.023 |
| $X_7$ | 7.510 | 3.447 |
| $X_8$ | -23.038 | -12.545 |
| $X_9$ | -4.231 | -2.063 |
| $X_{10}$ | 1.536 | 0.745 |
| $X_{11}$ | -4.207 | -1.946 |
| $X_{12}$ | 0.030 | -0.176 |
| $X_{13}$ | 0.935 | 0.519 |
| $X_{14}$ | -0.836 | -0.304 |
| $X_{15}$ | -0.452 | -0.163 |

Positive sign of model structural parameter $\beta_j$ indicates that an increase of a variable increases the probability of variable Y being 1 (YES). The negative sign indicates that an increase of a variable increases the probability of value 0 (NO). For example, an increase of variable $X_4$ – Construction Cost, increases the probability of delay, whereas an increase of variable $X_8$ – Number of Spans increases the probability that delays will not occur.

The models were evaluated based on classification effectiveness expressed as the number of correctly identified cases. The summary, in the form of accuracy matrix, is presented in tables 5 and 6.

Table 5. Results of classification - logit model (source: own study)

| Actual | Predicted | | Total: | Correctly classified: |
|---|---|---|---|---|
| CLASS | "YES" (1) delayed project | "NO" (0) not delayed project | | |
| "YES" (1) delayed project | 51 | 10 | 61 | 84% |
| "NO" (0) not delayed project | 17 | 31 | 48 | 65% |
| Total: | 68 | 41 | 109 | |
| Odds ratio: 9.3 | | Correctly classified cases for both classes together: 75% | | |

Table 6. Results of classification - probit model (source: own study)

| Actual | Predicted | | Total: | Correctly classified cases: |
|---|---|---|---|---|
| CLASS | "YES" (1) delayed project | "NO" (0) not delayed project | | |
| "YES" (1) delayed project | 51 | 10 | 61 | 84% |
| "NO" (0) not delayed project | 19 | 29 | 48 | 60% |
| Total: | 70 | 39 | 109 | |
| Odds ratio: 7.8 | | Correctly classified cases for both classes together: 73% | | |

The comparison of both models indicates that the logit model gave slightly better results (75% of correctly identified cases). The odds ratio of probability of $Y_i = 1$ to probability $Y_i = 0$ is higher in the logit model and equals 9.3. This means that this classification is slightly more than nine times better than a classification that could be expected from chance.

The developed regression model can also be verified using the analysis of residual values. Analysing the residuals [44], one can quickly and effectively detect possible deviations from the correct regression analysis and indicate outlying observations. The first step included the plotting of residuals normality chart in order to obtain a graphical check of residuals normality for both models (Fig. 2)
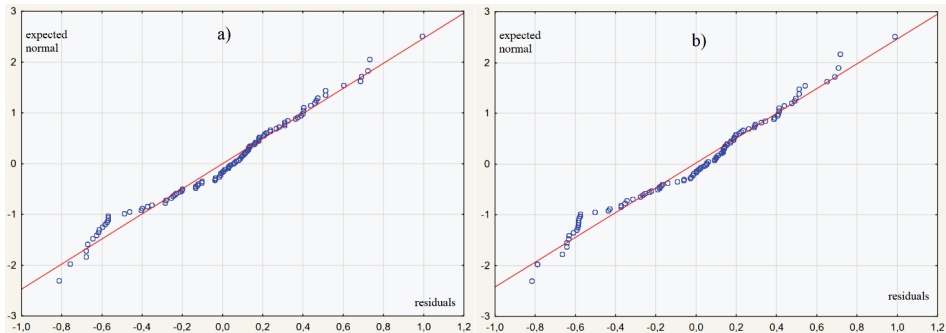
Fig. 2. Normality of the distribution of residuals: a) logit model, b) probit model

(source: own study, as obtained in STATISTICA® software suite)

The results allow a visual evaluation of residuals conformity with normal distribution. Although in the lower part the points slightly depart from the straight line, in their majority they lie along the straight line, thus confirming the normality of residuals distribution.

# 4. CONCLUSIONS

The authors have proposed the logistic regression and the probit model to search for relationships between the bridge construction parameters and construction delays. The explained variable $Y$ in the model had two possible variants described by codes "*YES*" for occurrence of project delay and "*NO*" for absence of delay. The results of models evaluation were similar, but the logit model have slightly better results (75% of correctly identified cases). Its classification correctness was better in class "1" – delay (84%) than in class "0" – no delay (65%). Please note that the proposed models allow a classification of the explained variable Y which is quantified by means of a dichotomic value (taking one of the two possible values). The authors are planning an attempt to classify the delay duration, not only the mere fact of delay occurrence. This requires using different tools, e.g. Artificial Neural Networks.

# REFERENCES

1. D. A. Aczel, Statystyka w zarządzaniu, Wydawnictwo Naukowe PWN, Warszawa 2000.

2. A. Al-Kharashi, M. Skitmore, Causes of delays in Saudi Arabian public sector construction projects, Construction Management and Economics, 27[1], 3-23, 2009.

3. H. Anysz, B. Buczkowski, The association analysis for risk evaluation of significant delay occurrence in the completion date of construction project. International Journal of Environmental Science and Technology,1-6, 2018.

4. H. Anysz, A. Zbiciak, A., N. Ibadov, Pomiar trafności prognoz opóźnień w realizacji przedsięwzięć budowlanych za pomocą sztucznych sieci neuronowych. Materiały Budowlane, [6], 3-5, 2016.

5. H. Anysz, A. Zbiciak, Przyczyny powstawania opóźnień w realizacji kontraktów budowlanych - analiza wstępnych wyników badania ankietowego, Autobusy: Technika, Eksploatacja, Systemy Transportowe, 3, 963-972, 2013,

6. J.R. Baldwin, J.M. Manthei, Causes of delays in the construction industry, ASCE Journal of the Construction Division, 97, 177-187, 1971.

7. A. Czarnigowska, A. Sobotka, Time–cost relationship for predicting construction duration. Archives of Civil and Mechanical Engineering, 13[4], 518-526, 2013.

8. M. Dytczak, G. Ginda, T. Wojtkiewicz, Identyfikacja roli czynników opóźnień realizacji złożonych przedsięwzięć budowlanych, Budownictwo i Inżynieria Środowiska, 2, 481-485, 2011.

9. E. Gatnar, Symboliczne metody klasyfikacji danych, Wydawnictwo Naukowe PWN, Warszawa 1998.

10. B. Gładysz, D. Kuchta, D. Skorupka, A. Duchaczek, Fuzzy analysis of project duration in situations of risk, AIP Conference Proceedings, 1648, 600003, 2015.

11. M. Głuszak, A. Leśniak, Construction delays in clients opinion–multivariate statistical analysis. Procedia engineering, 123, 182-189, 2015.

12. B. Grzyl, E. Miszewska-Urbaska, M. Apollo "Building Information Modelling as an Opportunity and Risk for Stakeholders Involved in Construction Investment Process", Procedia Engineering, 196, 1026-1033, 2017.

13. T. Hanak, J. Korytarova, Subsidy Risk Related to Construction Projects: Seeking Causes. Open Engineering, 8[1], 484-489, 2018.

14. B. Hoła "Identification and evaluation of processes in a construction enterprise", Archives of Civil and Mechanical Engineering, 15[2], 419-426, 2015.

15. B. Hoła, M. Szostak: "Analysis of the state of the accident rate in the construction industry in European Union countries, Archives of Civil Engineering, 61[4]:19-34, 2015.

16. N. Ibadov, J. Kulejewski, Evaluation of the project timeliness with fuzzy constraints, International Conference Of Numerical Analysis And Applied Mathematics [ICNAAM-2015], AIP Conference Proceedings Vol.1648 No.1, 600002, 2015.

17. P. Jaśkowski, S. Biruk, Analiza czynników ryzyka czasu realizacji przedsięwzięć budowlanych. Czasopismo Techniczne, Budownictwo, 1-B[2], 157-166, 2010.

18. M. Juszczyk, A. Leśniak, K. Zima, ANN Based Approach for Estimation of Construction Costs of Sports Fields, Complexity 2018, 1-11. DOI:10.1155/2018/7952434

19. O. Kapliński, Innovative solutions in construction industry. Review of 2016–2018 events and trends. Engineering Structures and Technologies, 10[1], 27-33, 2018.

20. A. Kazaz, S. Ulubeyli, N.A. Tuncbilekli, Causes of delays in construction projects in Turkey. Journal of Civil Engineering and Management, 18[3], 426-435, 2012.

21. A. Kowalczyk, W. Meszek, M. Rejment, A. Dziadosz, Delays in construction works in Polish construction industry from the contractors' point of view. In MATEC Web of Conferences, 222, 01008, 2018.

22. J. Korytarova, T. Hanak, R. Kozik, E. Radziszewska-Zielina "Exploring the contractors' qualification process in public works contracts" Creative Construction Conference, Procedia Engineering, 123, 276-283, 2015.

23. A. Kosecki, Kontraktowanie realizacji przedsięwzięć budowlanych, PWN, Warszawa, 2019.

24. M. Krzemiński, Optimization of work schedules executed using the flow shop model, assuming multitasking performed by work crews, Archives of Civil Engineering, 63[4], 3-19, 2017.

25. D.T. Larose Metody i modele eksploracji danych, Wydawnictwo Naukowe PWN, Warszawa 2008.

26. M. Lendo-Siwicka, M., Połoński, K. Pawluk, K. [2016]. Identification of the interference in the investment process during the realization of a Shopping Centre–a Case Study. Archives of Civil Engineering, 62[1], 159-172, 2016

27.  A. Leśniak, Przyczyny opóźnień budowy w opiniach wykonawców, Czasopismo Techniczne. Budownictwo, 1-B, 2, 57-68, 2012.

28.  A. Leśniak, M. Juszczyk, Prediction of site overhead costs with the use of artificial neural network based model. Archives of Civil and Mechanical Engineering 2018, 18[3]: 973-982. DOI:10.1016/j.acme.2018.01.014

29.  A. Leśniak, M. Juszczyk, G. Piskorz, Grupowanie czynników opóźnienia realizacji przedsięwzięcia budowlanego z wykorzystaniem analizy skupień. Materiały Budowlane, 10, 62-65, 2018.

30.  A. Leśniak, G. Piskorz, Potential reasons for works delays resulting from the provisions of the agreement. Technical Transactions, 6, 135-143, 2017.

31.  A. Leśniak, G. Piskorz, M. Spišáková, D. Mačková, Causes of delays in construction works resulting from the provisions of the contract in Poland and Slovakia, 27 [1], 71–81, 2018.

32.  A. Leśniak, K. Zima, Cost calculation of construction projects including sustainability factors using the Case Based Reasoning [CBR] method, Sustainability 2018, 10[5],1608. DOI: 10.3390/su10051608

33.  C.S. Lim, M.Z. Mohamed, An exploratory study into recurring construction problems, International Journal of Project Management, 18[4], 267-273, 2000.

34.  D. Mackova, M. Kozlovska, R. Baskova, M. Spisakova, K. Krajnikova, Construction-duration prediction model for residential buildings in Slovak Republic based on computer simulation. International Journal of Applied Engineering Research, 12[13], 3590-3599, 2017.

35.  G. S. Maddala, Ekonometria, Wydawnictwo Naukowe PWN, Warszawa 2008.

36.  S.H. Nguyen, Matematyka stosowana, Systemy decyzyjne. Uniwersytet Warszawski, 2011, http://www.mimuw.edu.pl/~son

37.  B. Nowogońska, J. Cibis, Technical problems of residential construction. In IOP Conference Series: Materials Science and Engineering, 245, 5, 052042, 2017.

38.  P. Nowotarski, J. Pasławski, J., Matyja, Improving Construction Processes Using Lean Management Methodologies–Cost Case Study, Procedia engineering, 161, 1037-1042, 2016.

39.  J. Pasławski, J. Drzewiecka, Analiza zakłóceń procesów budowlanych, Budownictwo i Inżynieria Środowiska, 2/2011, 475-479.

40.  E. Plebankiewicz, K. Zima, D. Wieczorek, Life cycle cost modelling of buildings with consideration of the risk, Archives of Civil Engineering, 62[2], 149-166, 2016.

41.  M. Połoński, K. Pruszyński, Impact of baseline terms on the course of critical paths and time buffers in the modified Goldratt's method. Archives of Civil Engineering, 59, 3, 313-320, 2013.

42.  D. Skorupka, The method identification and qualification of construction project risk, Archives of Civil Engineering, LI, 4, 647-662, 2005.

43.  A. Sobotka, A. Czarnigowska, Analysis of supply system models for planning construction project logistics. Journal of Civil Engineering and Management 11, 73-82, 2005.

44.  A. Stanisz, Przystępny kurs statystyki, Statsoft Polska, Kraków 2007.

45.  S.U.R Toor, S.O. Ogunlana, Problems causing delays in major construction projects in Thailand. Construction management and economics, 26[4], 395-408, 2008.

**LIST OF FIGURES AND TABLES:**

## MODELOWANIE OPÓŹNIEŃ W PRZEDSIĘWZIĘCIACH MOSTOWYCH Z WYKORZYSTANIEM REGRESJI LOGITOWEJ I PROBITOWEJ

*Słowa kluczowe*: opóźnienia w przedsięwzięciach budowlanych, opóźnienia w przedsięwzięciach mostowych, regresja logitowa, regresja probitowa

**STRESZCZENIE:**

Pomyślne zakończenie przedsięwzięcia budowlanego oznacza terminowe ukończenie budowy w ramach założonych kosztów i planowanej jakości oraz spełnieniem oczekiwań inwestora. Istotny wpływ na powodzenie inwestycji drogowych czy kolejowych ma realizacja obiektów mostowych. Są one niezbędne dla zapewnienia połączeń i pokonania przeszkód w terenie, a ich obecność gwarantuje ciągłość tras komunikacyjnych. Obiekty mostowe stanowią newralgiczne punkty, będąc nierzadko kamieniami milowymi dla wykonawcy robót.

Celem artykułu, jest rozwiązanie problemu klasyfikacyjnego dotyczącego opóźnień w realizacji obiektów mostowych. Istotą analizy jest poszukiwanie związku między zmiennymi niezależnymi charakteryzującymi przedsięwzięcia obejmujące budowę obiektów mostowych, a faktem wystąpienia opóźnienia podczas ich realizacji. W pracy poszukiwano modelu dwumianowego, w którym zmienna objaśniania $Y$ jest kwantyfikowana za pomocą wartości dychotomicznej (przyjmującej jedną z dwóch możliwych wartości). Założono, że zmienna $Y$ przyjmowała dwa możliwe warianty określone za pomocą kodów - etykiet: „*TAK*" odpowiadającego opóźnieniu przedsięwzięcia mostowego, „*NIE*" oznaczającego brak opóźnienia przedsięwzięcia mostowego (kody należy utożsamiać odpowiednio z wartościami 1 i 0).

W artykule zaproponowano rozwiązanie problemu z wykorzystaniem dwu alternatywnych modeli – logitowego i probitowego. W analizie wykorzystano dane zebrane w ramach badań własnych obejmujące parametry realizacji przedsięwzięć mostowych i parametry obiektów mostowych, a także wielkość opóźnienia (jeżeli wystąpiło). W wyniku przeprowadzonych badań zgromadzono dane dotyczące 169 obiektów mostowych, w tym 48 obiektów, w których nie wystąpiły opóźnienia i 121, w których wystąpiły opóźnienia. W analizie wykorzystano dane przedstawiające 109 przypadków.

W toku pracy badawczej zbudowano dwa alternatywne modele: logitowy i probitowy oraz przeprowadzono analizę wyników. Wyniki uzyskane dla obu modeli były zbliżone, jednak nieco lepsze wyniki klasyfikacji uzyskano w przypadku modelu logitowego, który zidentyfikował poprawnie 75% przypadków. Procent poprawności klasyfikacji tego modelu był lepszy w klasie „1" oznaczającej wystąpienie opóźnienia (84%), niż w klasie „0" brak opóźnienia (65%). Autorzy zwracają uwagę, iż zaproponowane modele pozwalają na klasyfikacje zmiennej objaśnianej Y, która jest kwantyfikowana za pomocą wartości dychotomicznej. W planach badawczych autorów jest podjęcie próby klasyfikacji wielkości opóźnień, nie tylko samego faktu jego wystąpienia. Takie założenie wymaga zatasowania innych narzędzi np. sztucznych sieci neuronowych.