ARCHIVES OF CIVIL AND MECHANICAL ENGINEERING 18 (2018) 973–982

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://www.elsevier.com/locate/acme

Original Research Article

Prediction of site overhead costs with the use of artificial neural network based model

Agnieszka Leśniak *, *Michał Juszczyk*

Cracow University of Technology, Faculty of Civil Engineering, ul. Warszawska 24, 31-155 Kraków, Poland

ARTICLE INFO

Article history:
Received 5 July 2017
Accepted 22 January 2018
Available online 4 March 2018

Keywords:
Site overhead cost
Artificial neural networks
Construction cost management

ABSTRACT

Overheads, especially site overhead costs, constitute a significant component of a contractor's budget in a construction project. The estimation of site overhead costs based on traditional approach is either accurate but time consuming (in case of the use of detailed analytical methods) or fast but inaccurate (in case of the use of index methods). The aim of the research presented in this paper was to develop an alternative model which allows fast and reliable estimation of site overhead costs. The paper presents the results of the authors' work on development of a regression model, based on artificial neural networks, that enables prediction of the site overhead cost index, which used in conjunction with other cost data, allows to estimate site overhead costs. To develop the model, a database including 143 cases of completed construction projects was used. The modelling involved a number of artificial neural networks of the multilayer perceptrons type, each with varying structures, activation functions and training algorithms. The neural network selected to be the core of developed model allows the prediction of the costs' index and aids in the estimation of the site overhead costs in the early stages of a construction project with satisfactory precision.

1. Introduction

The issue of a sufficiently reliable overheads estimation is vital for the potential contractor. According to the research presented in one of the previous works by Plebankiewicz and Leśniak [33] the influence of improper calculation of the overhead costs can be significant for the financial situation of the contracting company.

Generally, the building contractor's overhead costs are divided into two categories: site (project) overhead costs and company's (general) overhead costs [32]. Site (project) overhead costs include items that can be identified with a particular job, but not materials, labour, or production equipment. Company's overhead costs are items that represent the cost of doing business and often are considered as fixed expenses that must be paid by the contractor. In literature one can find different definitions of overhead costs [1,5,26,33,36]. On the other hand, an overhead cost of a construction project can be defined as a cost that cannot be identified with or charged to a construction project or to a unit of construction production [21]. Cilensek [19] describes

* *Corresponding author.*
E-mail address: alesniak@izwbit.pk.edu.pl (A. Leśniak).
https://doi.org/10.1016/j.acme.2018.01.014

overhead costs as those that are not a component of the actual construction work but are incurred by the contractor to support the work. The overheads include expenses that cannot be charged directly to a particular branch of work but are required to construct the project [23]. According to Polish standards of cost estimating [38], site overhead costs can be defined as all the costs incurred by the contractor on the building site in connection with the works realization, excluding the direct costs.

Overhead costs are widely discussed in literature. Relevant research on overhead costs can be divided into four main research trends [36]. Some of the researchers focus on the analysis of situation and statistical research on the understanding of the overhead costs concept, analysis of construction delays vs. overhead costs, analysis of the construction company's overhead costs distribution, and allocation and analysis of fixed expenses recovering. Assaf et al. [6] investigated the overhead costs practices and showed how the unstable construction market makes it difficult for construction companies to decide on the optimum level of overhead costs. The practices of estimating overhead costs are investigated in various countries (e.g. Great Britain [14], the USA and Canada [30], Lithuania [36], Saudi Arabia [6], Poland [33]). Particular attention is paid to a detailed computation of site overheads. A number of empirical studies relate to the determination of the project overhead cost. Factors that influence project overhead cost are widely discussed in literature in various aspects [5,17,40]. Some of them emphasize that project time is an important factor affecting project overheads [11,27]. Cooke [20] highlighted that the location of the site could affect a number of project overhead items. Brook [11] indicated that the method of work was a critical factor affecting the amount spent on project overheads. A detailed overhead costs categorization and the selection of the principal parameters of the company's activity, on which the value of overhead costs depends, was presented by Šiškina et al. [36]. Apanavičienė and Daugėlienė [2] proposed a new classification of construction companies into competitiveness classes according to the relative value of overhead costs. In other work [34], it was commented that a contractor's overhead costs, though varying from trade to trade, were dependent on annual volume of work, job type, job size, local economic conditions, support staff and equipment requirements. El Sawy et al. [25] after having conducted a series of surveys, proposed a list of factors that contribute to site overhead cost in the Egyptian construction market. The researchers in their investigations on overhead costs or its elements use different tools for instance: case-based reasoning [17], neural networks [25], exploratory factor analysis [12]. Some of the authors analyze the issue taking into account principal parameters of the construction company's activity, on which the value of overhead costs depends [36]. In other work a new classification of construction companies into competitiveness classes according to the relative value of overhead costs was proposed [2].

Artificial neural networks (ANN) refer to mathematical structures and their software-or hardware-based models which compute or process signals. The structure of the network and its mode of action is based on the brain and learning phenomena; however neural networks constitute a strongly simplified model [39]. The theory of neural networks is widely presented in literature (e.g. [9,28,31,39]). The main application of artificial neural networks includes the following [28,31]: prediction, approximation, control, association, classification and pattern recognition, associating data, data analysis, signal filtering and optimization.

Artificial neural networks began to be used in the management of construction projects in the early nineties of the last century [37]. Until today there have been a number of attempts to use artificial neural networks in engineering construction processes regarding such issues as implementation time analysis, efficiency and productivity in construction projects [24,35], predicting the maintenance cost of construction equipment [45], predicting the adoption potential or acceptability of a new construction technology [37], construction company management [13,16,18] and facilitating decision making processes in construction projects [4,42].

Apart from the issues mentioned above, there have been other attempts to apply artificial neural networks to the management of the costs involved in construction projects. One of the first publications on this topic, by Hegazy and Amr [29], aimed at the creation of a ANN-based cost-estimating model which would allow to estimate the costs of constructing motorways. A similar problem was described in [43,44]. In [41] authors described a new multi-stage framework based on ANN for cost-optimal analysis to support the deep renovation of buildings. The cost formulas for estimate sheet metal parts composed by applying neural networks was proposed in [3]. The application of ANN, in the field of construction cost management concerned also predicting cash flows [10], predicting cost deviations in high-risk projects including reconstruction, alteration, rebuilding projects [7], evaluating of project budget implementation [22] estimation of overheads in dam projects [25] or analysis of construction claims outcomes [15].

The aim of this paper is to present the results of the research on the development a regression model based on artificial neural networks which supports the prediction of the site overhead cost index and thus allows quick estimation of site overheads costs within an acceptable error range. The solution to the problem involves finding such a form of the model that will enable a specification of the site overhead cost index for construction projects. The authors' basic assumption was the application of artificial neural networks in the model, since their key feature and main advantage is their ability to generalize knowledge. This generalization allows the generation of appropriate solutions for data that did not appear in the training data set.

2. Concept of model and research phases

The authors' assumption was the development of a model that would allow the specification a site overhead cost index for a construction project. Such an index, on the basis of a computational formula, could enable a quick assessment of site overhead cost for a certain construction project. In their research, the authors intention was to develop a regression model implementing an artificial neural network. The term "regression" refers to a modelling function mapping a set of values of describing variables on the set of values of the

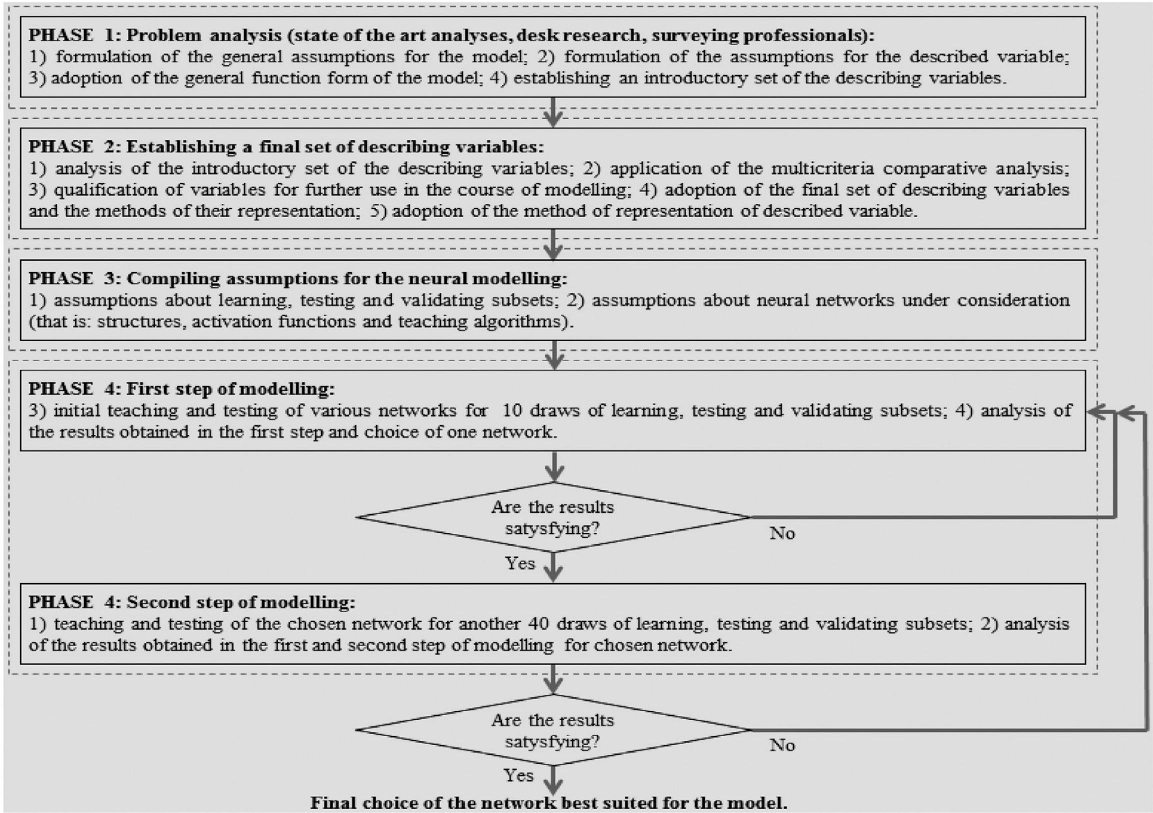

Fig. 1 – The block diagram of the research methodology.
***Source*: Own study.**

variable described. (The regression models, as it is widely accepted, can involve a mathematical equation, a set of equations or an algorithm. Here, the authors proposed implementation of neural network as an algorithm that constitutes the model.) The research comprised four phases – methodology is depicted concisely in Fig. 1.

Phases 1–3 are presented in the sections below. The fundamental research part, namely phase 4, is described in Section 3 of this paper.

2.1. *Phase 1 – problem analysis, establishing an introductory set of describing variables*

The general form of the model is illustrated by Eq. (1), while Eq. (2) depicts the general function form of the regressive model:

$$Y = F(X_j, \epsilon) \tag{1}$$

$$\hat{Y} = F(X_j) \tag{2}$$

where:

- Y – described variable of the model – site overhead cost index,
- $\hat{Y}$ – predicted value of the described variable – site overhead cost index predicted by the model,
- X_j – describing variables of the model,
- F – functional dependency connecting the describing variables with the described variable,
- ε – model error.

The fundamental assumption was an implementation of the functional dependency F implicitly by the artificial neural network. The prediction of the site overhead cost index (the value of the variable described of the model), as represented by Eq. (3):

$$\hat{Y}^i = F\left(x_j^i\right) \tag{3}$$

where:

- $\hat{y}_i$ – function F value (predicted value of the site overhead cost index) for the i-th vector of describing variables,
- F – as in Eqs. (1) and (2),
- x_{ij} – i-th vector of describing variables X_j.

The authors of this paper proposed to establish the value of the site overhead cost index, denoted hereinafter by SOC_{ind}, on the basis of three different Eqs. (4)–(6):

$$SOC_{ind1} = \frac{SOC}{LC + EC} \tag{4}$$

$$SOC_{ind2} = \frac{SOC}{LC + MC + EC} \tag{5}$$

$$SOC_{ind3} = \frac{SOC}{LC + MC + EC + SC} \tag{6}$$

where:

- SOC_{ind} – site overhead costs index (to make a distinction based on the type of calculation, the indices were additionally marked with numbers 1, 2 or 3),
- SOC –site overhead costs observed in reality,
- LC – labour costs observed in reality,
- MC – material costs observed in reality,
- EC – equipment work costs observed in reality,
- SC – subcontractors' costs observed in reality.

In order to prepare a method of determining the site overhead costs index based on artificial neural networks, an appropriate database needed to be developed. To collect a database, a survey was conducted among Polish contractors concerning the implementation of building works 400 questionnaires were sent, out of which 151 (38%) returned. After screening, 8 questionnaires were rejected. The research included quantitative studies of the factors proposed, influencing site overhead costs in relation to the construction works under analysis. The factors that were considered involved the following: the complexity of the scope of construction, localization conditions of the construction site, works implementation times, difficulties related to the implementation of works in winter, the amount of works performed by contractors themselves and the amount of works done by subcontractors. Subsequently, the real site costs that contractors carried due to the implementation of the analyzed constructions were compiled. Having included literature study and desk-research, a set of potential variables was established, describing the prediction of costs proposed for the model.

2.2. *Phase 2 – establishing a final set of describing variables, construction of a database*

An analysis of the dependencies between potential describing variables was conducted, using elements of a multi-criteria comparative analysis. The aims of the analysis were: to complement the model with descriptive variables, ensuring the form of the model as simple as possible and to reduce data redundancy, which belong to unwanted phenomena in neural modelling. All initially established potential describing variables underwent an analysis concerning the following: the relevance of the information introduced, information load, interdependencies between variables, information availability for a practical application of the model. As a result, a final set of describing variables was established and a collation of training data for the neural modelling. The variables included in the model, as well as the method of coding and their possible values, are presented in Table 1.

The database of information used in the training process of several neural networks included the known values of the described variable – Y, which could be observed in reality, as well as the related vectors of the values of the describing variables – X_j. Exemplary records are presented in Table 2.

The records of the database included coded values of describing variables and described variables in the three variants (as presented in Table 2). During the research, data for 143 construction projects under implementation in Poland in the Malopolska region were collected. For the proposed formulas of the general construction costs index *SOCind* (formulas (4)–(6)), on the basis of the construction works costs observed in reality, the values of the indices were computed. The values of the indices (three variants of the described variable) are depicted in the last three columns of Table 2.

Table 1 – Coding the input variables for the neural model (source: own study).

Variable X_j	Variable description	Method of coding	Possible values
X_1	Works type – general construction works	Binary	0 or 1
X_2	Works type – installation works	Binary	0 or 1
X_3	Works type – engineering works	Binary	0 or 1
X_4	Construction site location – in city centre	1 of *n*	1, 0, 0 or 0, 1, 0 or 0, 0, 1
X_5	Construction site location – outside the city centre		
X_6	Construction site location – non-urban spaces		
X_7	Distance between the construction site and the company's office	Pseudo-fuzzy scaling	Up to 20 km – 0.1 More than 20 km – 0.9
X_8	Works implementation time	Pseudo-fuzzy scaling	Up to 6 months – 0.1 Between 6 and 12 months – 0.5 More than 12 months – 0.9
X_9	Relations between the amount of works performed in winter to the total amount of works	Pseudo-fuzzy scaling	Up to 10%–0 Between 10% and 20%–0.1 Between 20% and 40%–0.3 Between 40% and 60%–0.5 Between 60% and 80%–0.7 Between 80% and 90%–0.9 More than 90%–1
X_{10}	Relations of the amount of works performed by subcontractors to the total amount of works	Pseudo-fuzzy scaling	Up to 20% – 0.1 Between 20% and 50%–0.5 Between 50% and 100%–0.9

Table 2 – The exemplary records of training data with the values of describing variables and described variables in three variants (source: own study).

i	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	Y		
											SOC_{ind1}	SOC_{ind2}	SOC_{ind3}
11	1	1	1	0	1	0	0.1	0.9	0.1	0.5	0.29	0.15	0.09
24	1	1	1	0	0	1	0.9	0.1	0	0.1	0.43	0.17	0.13
31	1	1	1	0	1	0	0.9	0.9	0.5	0.5	0.21	0.11	0.05
61	1	1	1	1	0	0	0.1	0.9	0.3	0.5	0.21	0.12	0.07
89	1	1	1	0	1	0	0.1	0.9	0.3	0.5	0.79	0.51	0.21
104	0	1	1	1	0	0	0.1	0.5	0.5	0.5	0.93	0.15	0.08
139	1	0	1	0	0	1	0.9	0.1	0	0.1	0.41	0.19	0.12

2.3. *Phase 3 – compiling assumptions for neural modelling*

In the process of neural modelling the authors took into account several multilayer perceptrons, as a type of neural networks, which are believed to be the best solution to the regression analysis problem [39]. The general form of the network is presented in Fig. 2. The input layer was composed of 10 neurons denoting the describing variables of the model X_j (as in Table 1). The network structure included one hidden layer in which the number of neurons ranged from two to five. The output layer consisted of one neuron indicating the described variable Y of the model (site overhead costs indexes).

A two-step procedure was assumed to establish a neural network implementing dependency F (as in Eqs. (1)–(3)).

The first step of a procedure involved training several neural networks for 10 draws of the learning subset, the validating subset and the testing subset (later referred to as *L*, *V*, *T* subsets consequently). Various network architectures, distinct activation functions, and different training algorithms were investigated. The neurons in the hidden layer employed the following activation functions: sigmoid function (7), hyperbolic tangent (8). On the other hand, the neurons in the output layer the employed: sigmoid function (7), hyperbolic tangent (8), linear function (9). These activation functions are described by the following equations, respectively:

$$g(\nu) = \frac{1}{1+\exp(-\beta\nu)} \tag{7}$$

$$g(\nu) = \tanh(\beta\nu) \tag{8}$$

$$g(\nu) = \beta\nu \tag{9}$$

where:

- $g(\nu)$ – neuron activation function,
- ν – neuron potential,
- β – activation function factor influencing its steepness.

During the neural modelling process various training algorithms were used [31,39]: conjugent gradients (CG), Levenberg–Marquardt algorithm (LM), Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS). Altogether 54 neural networks different from each other in structure, activation functions employed or training algorithms were taken into account and analyzed in the first step.

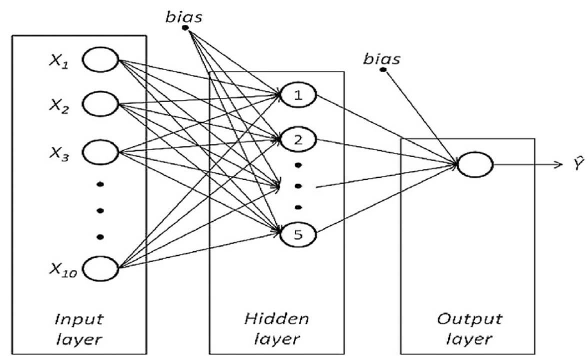

Fig. 2 – The general form of the neural network structure. *Source*: Own study.

For further investigation and the second step of modelling the authors selected the neural network which yielded the most stable training results. The quality of the network performance was established on the basis of the specified values of errors: root mean square error, *RMSE* (10) and mean average percentage error *MAPE* (11), as well as the maximum percentage error PE_{max} (12). The criteria of selection, assumed by the authors, was the acceptable performance in terms of errors range and low errors dispersion for the mentioned 10 draws of *L*, *V* and *T* subsets. The reason for this assumption was to ensure that the performance of the network is not biased by the drawing of *L*, *V* and *T* subsets.

$$\mathrm{RMSE} = \sqrt{\frac{1}{P}\sum_{p=1}^{P}\sum_{i=1}^{M}\left(y_i^{(p)} - \hat{y}_i^{(p)}\right)^2} \tag{10}$$

$$\mathrm{MAPE} = \frac{1}{P}\sum_{p=1}^{P}\sum_{i=1}^{M}\left|\frac{y_i^{(p)} - \hat{y}_i^{(p)}}{y_i^{(p)}}\right| \times 100\% \tag{11}$$

$$PE_{\max} = \max\left(\left|\frac{y_i^{(p)} - \hat{y}_i^{(p)}}{y_i^{(p)}}\right| \times 100\%\right) \tag{12}$$

where:

Table 3 – Results after the 1st step of modelling (source: own study).

ANN	ERRORS	RMSE L	RMSE V	RMSE T
MLP 10–4–1 sigmoid-linear (the chosen ANN after the 1st step of modelling)	Max	0.03895	0.04258	0.03926
	Average	0.02750	0.03526	0.03260
	Min	0.01904	0.02422	0.02422
MLP 10–4–1 sigmoid-hyperbolic tangent	Max	0.04720	0.05167	0.04926
	Average	0.03016	0.03705	0.03514
	Min	0.01970	0.02171	0.02578
MLP 10–5–1 sigmoid-sigmoid	Max	0.03391	0.04883	0.03996
	Average	0.03206	0.03854	0.03441
	Min	0.01542	0.02015	0.02750
MLP 10–5–1 hyperbolic tangent-linear	Max	0.04195	0.04521	0.04211
	Average	0.03105	0.03726	0.03903
	Min	0.01604	0.01916	0.02133
MLP 10–3–1 hyperbolic tangent-sigmoid	Max	0.03967	0.04321	0.04102
	Average	0.02860	0.03671	0.03341
	Min	0.02078	0.02712	0.02598

- p – number of the sample;
- i – number of the output layer neuron;
- y_i – known values of the costs indices being tested;
- $\hat{y}_i$ – calculated values of the costs indices being tested.

3. Research results

The analysis conducted in the first step of neural modelling involved selection of a network that obtained the most stable training results. The choice depended on the acceptable performance of the network and the dispersion of learning, testing and validation errors. The authors sought for the network, for which the differences between maximum and minimum of the errors were the smallest, in relation to the average values of the errors. The network, multilayer perceptron with a 10–4–1 structure (4 neurons in the hidden layer), was selected, with the activation functions: logistic function employed in the hidden layer and linear function employed in the output layer. Later in the paper the selected network is referred to as MLP 10–4–1. Moreover the best results in the first step of the procedure were obtained for the values of described variable Y of the model calculated with using Eq. (6). Consequently, in the further part of the paper the authors presented the results for the model where Y values were calculated with the equation mentioned above.

Table 3 presents the results of the 1st step of the modelling for five best networks. The chosen network MLP 10–4–1 is compared with four other networks in terms of RMSE errors. Maximum, average and minimum RMSE errors for learning, validation and testing (L, V, T – accordingly) that has been obtained in the training of networks in ten consecutive draws of L, V and T subsets. Information about the structure of each network (10-h-1) is given, as well as the employed activation functions for each of the networks (hidden layer activation function – output layer activation function) in the column *ANN*.

According to Table 3, one can see that for some of the investigated networks, minimum values of the RMSE errors were lower than in the case of the chosen network, however, the dispersion of the errors was greater in these cases. The criterion for the choice of the MLP 10–4–1 was the stability of the networks' training for the ten draws of o L, V and T subsets. Firstly, in case of the chosen network the differences between the minimum and maximum RMSE values were the smallest. Secondly, the differences between RMSE for learning, validation and testing were at an acceptable range.

The review of all of the networks investigated in the 1st step of modelling allowed the conclusion that in general better results were obtained for the networks with a greater number of neurons in the hidden layer, namely 10–4–1 and 10–5–1 networks (which is not surprising because more complex structures can offer a better approximation in case of nonlinear problems – compare for example [9,28]). On the other hand, neither regularity nor dependence of networks performance on the employed activation functions has been observed.

The second step of the chosen network involved further training of chosen network. In accordance with the assumed procedure of the studies for the chosen network, in the second step of neural modelling the training of selected networks, MLP 10–4–1 was performed for the subsequent 40 draws of the L, V and T subsets. Table 4 depicts the values of RMSE errors (max, average, min) of learning, validation and testing obtained both after the first and second step of modelling for the selected network MLP 10–4–1.

The final choice of the network which, was supposed to become the core of the regression model, and implement the mapping function F (as in Eqs. (1)–(3)), involved the type of MLP 10–4–1 network that had been trained on the draw number 32. For this particular network the results of RMSE errors for learning, validation and testing were the closest to the average values from all the draws of the L, V and T subsets (out of all 50 draws in both steps of modelling). The error values for the finally chosen network, namely MLP 10–4–1 trained for the 32nd draw, are illustrated in Table 5 which presents *RMSE* errors and also the *MAPE* errors in learning, validation and testing for the network. Later the final chosen network is referred to as $MLP_{(smpl32)}$ 10–4–1.

Fig. 3 presents the training results obtained for the final chosen network, $MLP_{(smpl\ 32)}$ 10–4–1, a core of the developed regression model. Learning, testing and validating results for the final chosen network are shown in the form of a scatter plot. The horizontal axis represents known and expected output values of the model Y. The vertical axis shows output

Table 4 – RMSE errors summary for chosen network (MLP 10–4-1) after step three (source: own study).

	Sampling 1–10 (1st step of modelling)			Sampling 1–50 (both 1st and 2nd step of modelling)		
	RMSE L	RMSE V	RMSE T	RMSE L	RMSE V	RMSE T
Max	0.03895	0.04258	0.03926	0.03895	0.04258	0.03926
Average	0.02750	0.03526	0.03260	0.02648	0.03200	0.03239
Min	0.01904	0.02422	0.02422	0.01822	0.02239	0.02422

Table 5 – RMSE errors summary for chosen network – $MLP_{(smpl32)}$ 10–4–1 after step three (source: own study).

RMSE L	RMSE V	RMSE T	MAPE L	MAPE V	MAPE T
0.02888	0.03252	0.03500	20.6%	17.7%	19.0%

values predicted by the model $\hat{Y}$. In the graph the points corresponding to the testing T and validating V (on the left side), and the learning L (on the right side) of the network are located mostly in the cone of error decomposing along a straight perfect fit.

4. Verification and discussion of the proposed approach

To verify the practical application of the model, it was given a task of evaluating the amount of the indirect costs on the basis of the data that was not used at the modelling stage. Thus, 5 contractors from southern Poland were asked to provide information about construction projects completed and accounted for in 2015 or 2016. The data obtained is presented in Table 6. The information includes the following: values for 10 descriptive variables X_j (as presented in Table 1), model's input data, values of the real site overhead cost indexes Y, established on the basis of Eq. (6), values of the site overhead cost indexes predicted by the model $\hat{Y}$.

For each of the cases presented in Table 5, basic errors were calculated and set together in Table 7.

The results of the prediction of the site overhead cost indexes for the new cases were satisfactory. The smallest error of the model, amounting to 3.07%, appeared in project no. 3. The highest error was obtained in case of project no. 2. The results generated by the model were presented to the contractors who provided the data. Their opinions about the index assessment for the preliminary estimation of site overhead cost were as follows: projects 1, 3 and 5 were considered highly satisfactory; project 4 was thought satisfactory; project 2 was acceptable. Therefore, it may be concluded that the application of the model gave satisfactory results.

For the purposes of assessment and verification, the authors compared the proposed neural network based model with a model built on the classical multivariate regression analysis and least squares method [8]. The linear multivariate regression model was built using of the same data that was used to train the neural networks. The classical linear model including estimates of the parameters of the model and standard errors of the parameters' estimates is given by the following formula (13):

$$\hat{Y} = \underset{(0.0338)}{0.0702} + \underset{(0.0891)}{0.0308}*X_1 - \underset{(0.1080)}{0.2202}*X_2 + \underset{(0.1220)}{0.1024}*X_3 + \underset{(0.1477)}{0.2989}*X_4 + \underset{(0.1680)}{0.4472}*X_5 + \underset{(0.1353)}{0.4719}*X_6 - \underset{(0.0696)}{0.1367}*X_7 - \underset{(0.0720)}{0.1543}*X_8 + \underset{(0.0779)}{0.6766}*X_9 - \underset{(0.0683)}{0.3253}*X_{10} \quad (13)$$

where:

- $\hat{Y}$ – predicted value of the described variable – site overhead cost index predicted by the model,
- X_j – describing variables of the model for $j = 1, \ldots, 10$, as presented in Table 1.

The performance comparison of the both the classical statistical model and the proposed neural network based model was made with the use of root mean square error – *RMSE*, as in Eq. (10) and chosen measures of descriptive statistics, the correlation coefficient R, given by Eq. (14), coefficient of determination R^2, given by Eq. (15) and coefficient of convergence φ^2, given by Eq. (16).

$$R = \frac{cov(Y;\hat{Y})}{\sigma_Y \sigma_{\hat{Y}}} \quad (14)$$

$$R^2 = \left. \frac{cov(Y;\hat{Y})}{\sigma_Y \sigma_{\hat{Y}}} \right)^2 \quad (15)$$

$$\varphi^2 = 1 - \left. \frac{cov(Y;\hat{Y})}{\sigma_Y \sigma_{\hat{Y}}} \right)^2 \quad (16)$$

where:

- $cov(Y;\hat{Y})$ – covariance of real life values and predicted values of the described variable,
- σ_Y, $\sigma_{\hat{Y}}$ – standard deviation of Y and $\hat{Y}$ respectively.

All the mentioned measures of performance have been calculated for all of the 143 cases together and compiled in Table 8.

As presented in Table 8, all calculated measures reveal that the neural network based model performance is better than the model built on the classical approach. A comparison of the two models allows one to conclude that prediction of the site overhead cost index, based on the final chosen neural network, namely $MLP_{(smpl32)}$ 10–4–1, is more reliable than in the case of employment of the classical linear model built on the multivariate regression analysis.

5. Summary and conclusions

This research resulted in development of a novel estimation method of site overhead cost index. The approach proposed by

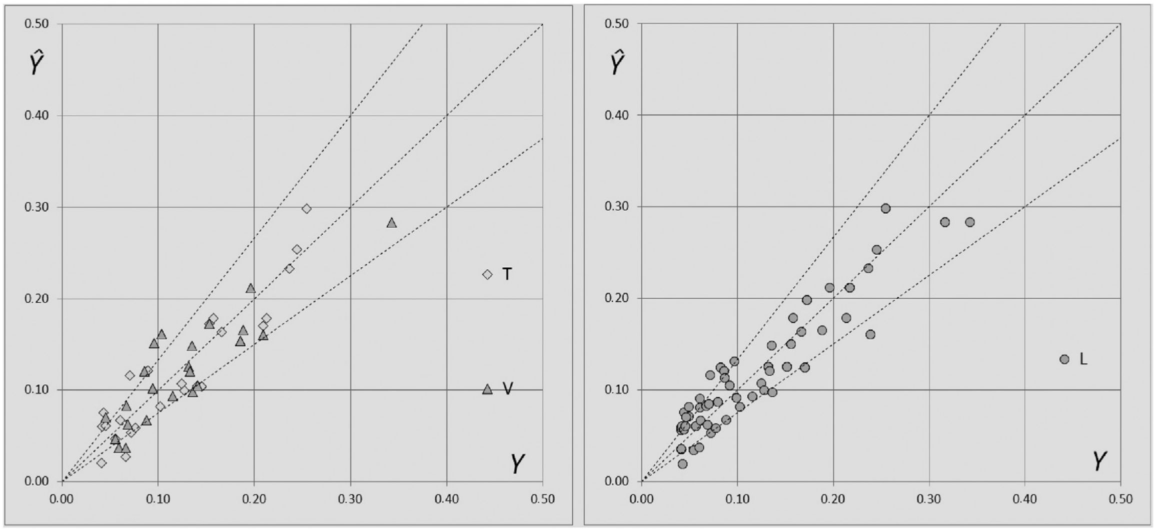

Fig. 3 – Scatter plots of training results for $MLP_{(smpl32)}$ 10–4–1 – "T" – testing, "V" – validating, "L" – learning.
***Source*: Own study.**

Table 6 – New cases introduced to the model (source: own study).

i	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	Y	$\hat{Y}$
Project 1	1	1	1	0	1	0	0.9	0.9	0.5	0.5	5.43%	5.74%
Project 2	1	0	0	1	0	0	0.1	0.9	0.3	0.9	13.63%	10.89%
Project 3	1	1	1	0	1	0	0.1	0.9	0.3	0.5	15.81%	16.29%
Project 4	1	0	0	0	1	0	0.1	0.5	0.1	0.9	4.16%	3.74%
Project 5	1	0	0	1	0	0	0.1	0.1	0.9	0.5	31.68%	29.79%

Table 7 – Basic error measures of the model predictions for the five new cases (source: own study).

i	$Y - \hat{Y}$	$\lvert Y - \hat{Y} \rvert$	$(Y - \hat{Y})^2$	$(Y - \hat{Y}) \times 100\%/Y$
Project 1	−0.003080	0.003080	0.000009	5.67%
Project 2	0.027469	0.027469	0.000754	20.15%
Project 3	−0.004858	0.004858	0.000023	3.07%
Project 4	0.004248	0.004248	0.000018	10.21%
Project 5	0.018922	0.018922	0.000358	5.97%

Table 8 – Comparison of the ANN based model and linear multivariate regression model.

Chosen measures of models' performance	Symbol	Calculated values	
		Classical linear model	ANN based model
Root mean square error	RMSE	0.05359	0.03086
Correlation coeefficient	R	0.73580	0.91725
Coefficient of determination	R^2	0.54140	0.84135
Coefficient of convergence	φ^2	0.45860	0.15865

Source: Own study.

the authors of the paper is based on artificial intelligence tools namely neural networks. A regressive model which employs artificial neural network chosen from a number of investigated networks has been proposed. The model is capable of mapping nonlinear relationships between a set of values of describing variables (which are features that characterize the construction site overheads for a project) onto a set of values of described variable which constituted the site overhead cost index. The describing variables of the model included characteristics of a construction project in relation to the type of works, the location of the construction site, the time of works completion, as well as the organizational assumptions for the construction process. The advantage of using neural networks approach instead of a classical multivariate regression approach is that there is no need to assume a priori functional relationships. The ANN, chosen to be the core of the proposed model, was fitted to the data (values of ten describing variables and one described variable) during the training process. The proposed neural networks based approach revealed its superiority over a classical multivariate linear regression approach. On the other hand, when compared to the traditional method of site overhead cost estimation, which is a preliminary detailed analysis of all the cost components, the use of the developed novel model is significantly faster and offers variant analysis of several sets of values of describing variables at a glance.

This research included the training of several types of artificial neural networks, namely multilayer perceptrons, in which various combinations of activation functions and different training algorithms were used. The networks under consideration possessed structures differentiated by the number of neurons in the hidden layer. In the first step of modelling 54 network types were considered. The results obtained in the first step helped to select the network which supported regression in the model. Then the chosen network underwent the second step of modelling. The final chosen network was the $MLP_{(smpl\ 32)}$ 10–4–1 (multilayer perceptron, with 10 neurons in the input layer, 4 neurons in the hidden layer and 1 neuron in the output layer, trained with the use of *BFGS* algorithm) selected after analysis of training results obtained for 50 draws of the learning, validating and testing subsets.

The analysis presented in this paper led to the following conclusions: the method proposed can allow to assess overhead construction costs at the early stage of the construction investment process with satisfactory precision; the results of the research confirmed the validity of using artificial neural networks in the assessment of overhead construction costs, on the basis of the proposed set of parameters characterizing the construction; the results obtained validate the application of the proposed model; the application of the model gave satisfying results.

Further research will involve the implementation of the proposed model in the form of a computer program which will allow the use of the model in practice and the exploration of possibilities of applying artificial neural networks to the problem presented with the aim of improving the model, including the use of committee machines.

Ethical statement

Authors state that the research was conducted according the ethical standards

REFERENCES

[1] B. Adey, E. Bruehwiler, R. Hydan, Effect of common cause failures on indirect costs, J. Bridge Eng. 9 (2) (2004) 200–208.

[2] R. Apanavičienė, A. Daugėlienė, New classification of construction companies: overhead costs aspect, J. Civ. Eng. Manag. 17 (4) (2011) 457–466.

[3] F. Ascione, N. Bianco, C. De Stasio, G.M. Mauro, G.P. Vanoli, CASA, cost-optimal analysis by multi-objective optimisation and artificial neural networks: a new framework for the robust assessment of cost-optimal energy retrofit, feasible for any building, Energy Build. 146 (2017) 200–219.

[4] M.E. Ashraf, Classifying construction contractors using unsupervised-learning neural networks, J. Constr. Eng. Manag. 132 (12) (2006) 1242–1253, art. no. 001612QCO.

[5] S.A. Assaf, A.A. Bubshait, S. Atiyah, M. Al-Shahri, Project overhead costs in Saudi Arabia, Cost Eng. 41 (4) (1999) 33–37.

[6] S.A. Assaf, A.A. Bubshait, S. Atiyah, M. Al-Shahri, The management of construction company overhead costs, Int. J. Proj. Manag. 19 (2001) 295–303.

[7] M. Attala, T. Hegazy, Predicting cost deviation in reconstruction projects: artificial neural networks versus regression, J. Constr. Eng. Manag. 129 (4) (2003) 405–411.

[8] D.M. Bates, D.G. Watts, Nonlinear Regression Analysis and Its Applications, vol. 2, Wiley, New York, 1988.

[9] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.

[10] A.H. Boussabaine, A.P. Kaka, A neural networks approach for cost flow forecasting, Constr. Manag. Econ. 16 (4) (1998) 471–479.

[11] .M. Brook, Preliminaries in Estimating and Tendering for Construction Work, Butterworth-Heinemann, Oxford, 1998, pp. 189–203.

[12] C.T.W. Chan, The principal factors affecting construction project overhead expenses: an exploratory factor analysis approach, Constr. Manag. Econ. 30 (10) (2012) 903–914.

[13] C.T.W. Chan, Pasquire C., A decision-making matrix model for estimating construction project overheads, in: Procs., PMI Research Conference 2006, New Directions in Project Management, Project Management Institute, Montreal, Canada, 17–19 July, (2006) 1–11.

[14] Chartered Institute of Building, Project overheads, in: Code of Estimating Practice, 7th ed., Wiley-Blackwell, 2009.

[15] K.W. Chau, Application of a PSO-based neural network in analysis of outcomes of construction claims, Autom. Constr. 16 (5) (2007) 642–646.

[16] J.H. Chen, M.C. Su, C.Y. Chen, F.H. Hsu, C.C. Wu, Application of neural networks for detecting erroneous tax reports from construction companies, Autom. Constr. 20 (7) (2011) 935–939.

[17] J.H. Chen, L.R. Yang, W.H. Chen, C.K. Chang, Case-based allocation of onsite supervisory manpower, Constr. Manag. Econ. 26 (8) (2008) 805–814.

[18] M.Y. Cheng, H.C. Tsai, E. Sudjono, Evolutionary fuzzy hybrid neural network for dynamic project success assessment in construction industry, Autom. Constr. 21 (2012) 46–51.

[19] R. Cilensek, Understanding contractor overhead, Cost Eng. (AACE) 33 (12) (1991) 21–30.

[20] B. Cooke, Contract Planning and Contractual Procedures, Macmillan, Basingstoke, 1981.

[21] W.E. Coombs, W.J. Palmer, Construction Accounting and Financial Management, 4th ed., McGraw-Hill, New York, 1989.

[22] D.K.H. Chua, Y.C. Kog, P.K. Loh, E.J. Jaselskis, Model for construction budget performance – neural network approach, J. Constr. Eng. Manag. 12 (3) (1997) 214–222.

[23] F.R. Dagostino, L. Feigenbaum, Estimating in Building Construction, 6th ed., Pearson Education, Upper Saddle River, NJ, 2003.

[24] S.U. Dikmen, M. Sonmez, An artificial neural networks model for the estimation of formwork labour, J. Civ. Eng. Manag. 17 (3) (2011) 340–347.

[25] I.Y. El-Sawy, H.E. Hosny, M.A. Razek, A neural network model for construction projects site overhead cost estimating in Egypt, Int. J. Comput. Sci. Iss. 8 (3) (2011) 273–283.

[26] A. Enshassi, A.R. Aziz, A. Karriri, Investigating the overhead costs in construction projects in Palestine, J. Finan. Manag. Prop. Constr. 13 (1) (2008) 35–47.

[27] S. Geddes, Preliminaries in Estimating for Building and Civil Engineering Works, 9th ed., Butterworth-Heinemann, Oxford, 1996.
[28] S. Haykin, Neural Networks. A Comprehensive Foundation, Prentice Hall, 1994.
[29] T. Hegazy, A. Amr, Neural network model for parametric cost estimation of highway projects, J. Constr. Eng. Manag. 124 (3) (1998) 210–218.
[30] T. Hegazy, O. Moselhi, Elements of cost estimation: a survey in Canada and United States, Cost Eng. 37 (5) (1995) 27–30.
[31] S. Osowski, Sieci Neuronowe w Ujęciu Algorytmicznym, Wydawnictwa Naukowo-Techniczne, Warszawa, 1996.
[32] R.L. Peurifoy, G.D. Oberlender, Estimating Construction Costs, 4th ed., McGraw Hill, New York, 1989.
[33] E. Plebankiewicz, A. Leśniak, Overhead costs and profit calculation by Polish contractors, Technol. Econ. Dev. Econ. 19 (1) (2013) 141–161.
[34] Rsmeans Company, RSMeans Cost Data: Student Edition, John Wiley & Sons, Hoboken, NJ, 2012.
[35] K. Schabowicz, B. Hoła, Application of artificial neural networks in predicting earthmoving machinery effectiveness ratios, Arch. Civ. Mech. Eng. 8 (4) (2008) 73–84.
[36] A. Šiškina, A. Juodis, R. Apanavičienė, Evaluation of the competitiveness of construction company overhead costs, J. Civ. Eng. Manag. 15 (2) (2009) 215–224.
[37] M. Skibniewski, L.-C. Chao, Neural network method of estimating construction technology acceptability, J. Constr. Eng. Manag. 121 (1) (1995) 130–142.
[38] Stowarzyszenie Kosztorysantów Budowlanych, Polskie Standardy Kosztorysowania Robót Budowlanych, Warszawa, 2005.
[39] R. Tadeusiewicz, Sieci Neuronowe, Akademicka Oficyna Wydawnicza RM, Warszawa, 1993.
[40] H.I. Tebin, Pricing tenders for construction, Cost Eng. 51 (1) (2009) 7–14.
[41] B. Verlinden, J.R. Duflou, P. Collin, D. Cattrysse, Cost estimation for sheet metal parts using multiple regression and artificial neural networks: a case study, Int. J. Prod. Econ. 111 (2) (2008) 484–492.
[42] M. Wanous, H.A. Boussabaine, J. Lewis, A neural network bid/no bid model: the case for contractors in Syria, Constr. Manag. Econ. 21 (7) (2003) 737–744.
[43] T.P. Williams, Predicting completed project cost using bidding data, Constr. Manag. Econ. 20 (3) (2002) 225–235.
[44] C.G. Wilmot, B. Mei, Neural network modeling of highway construction costs, J. Constr. Eng. Manag. 31 (7) (2005) 765–771.
[45] H.L. Yip, H. Fan, Y.H. Chiang, Predicting the maintenance cost of construction equipment: comparison between general regression neural network and Box–Jenkins time series models, Autom. Constr. 38 (2014) 30–38.