Support Vector Machine Based Hybrid Model for Prediction of Road Structures Construction Costs

S Petrusheva 1, D Car-Pušić 2 and V Zileska-Pancovska 1

1 University “Ss Cyril and Methodius”, Faculty of Civil Engineering, Skopje 1000 R. Macedonia
2 University of Rijeka, Faculty of Civil Engineering, Rijeka 51 000, Croatia

silvana@gf.ukim.edu.mk

Abstract. Cost prediction in early stages of construction projects is one of the crucial problems of project sustainability. Previous research has been aimed at process based and data driven model development by using various techniques, e.g. regression analysis, support vector machine (SVM), neural networks etc. According to the research results, neither of the techniques can be considered the best for all circumstances. Therefore, the research has been redirected towards hybrid modelling, i.e. combination of different techniques. In this research, for prediction of the target variable – real construction cost of road structures, available variables: contracted construction cost, contracted construction time and real construction time and cost, hybrid model – combination of SVM technique (data-driven model) and Bromilow time-cost model (TCM) (process-based model) have been used. Five hybrid models have been built for comparison purposes: SVM-Bromilow TCM, LR-Bromilow TCM, RBFNN-Bromilow TCM, MLP-Bromilow TCM and GRNN-Bromilow TCM, combining Bromilow TCM with SVM, LR (linear regression), RBFNN (radial basis function neural network), MLP (Multilayer perceptron) and GRNN (general regression neural network), respectively. The highest accuracy has been obtained with SVM-Bromilow TCM with mean absolute percentage error (MAPE) 1.01% and coefficient of determination ($R^2$) 97.61%.

1. Introduction

Problem of cost prediction in early stages of construction projects is one of the crucial problems in the life cycle of a project. Namely, it is well known that construction project cost is one of the key contract data and incorrectly estimated and contracted cost can lead to various problems, which can cause additional costs during the project implementation. The main and very frequent problem is the cost overrun in many construction projects [1–8]. For example, Baloi and Andrew [7] have presented the results of one former research [8] where cost overruns in 63% of 1,778 investigated construction projects have been detected. Flyvbjerg et al. [2] have researched cost underestimation in public transportation projects. They have researched 258 transportation projects, which have resulted in statistically valid conclusion regarding cost underestimation in infrastructure projects. That is the reason why the cost prediction, as well as cost overruns, is one of the frequent research topics. However, these issues, as the part of the economic aspect of sustainability, need to be considered more widely. This is not just about higher costs than planned. This is about sustainable and timely cost planning and its adequate allocation to all stages of the project. Overrun by items may disrupt the planned cost allocation balance with possible negative impact on other economic aspects of the project.
and quality of some works, as well. Exceeding the total project costs can also generate additional risks, e.g. infrastructural, industrial, social etc. and very negative impact on the environmental aspect of sustainability.

Cost predicting in early project stages is a demanding task. Therefore, many researchers have considered this problem in order to establish a more accurate predicting model. A lot of methods and techniques have been used for modelling, e.g. regression techniques, neural networks, supporting vectors etc. Neither of the techniques could be considered the best for all circumstances. It is recommended to compare results of several methods applied in the particular case and to choose the model with the highest accuracy. However, there could be great potential in the combination of the techniques. Therefore, the idea and the main objective of this research are to investigate some model combinations in order to develop a hybrid model with the best accuracy for considered experimental data. In this research road structure projects have been considered and the developed hybrid model can be applied for similar structures.

2. Literature review
A lot of research, as well as used methods and techniques to solve these problems have been carried out. In this research, use of Bromilow time-cost model (TCM) [9] and support vector machine (SVM) as a soft computing technique are considered. As for the BTC model, it should be emphasized that the Australian Bromilow [9] was the first to research financial execution in relation with construction time. The original model is about the dependency between time and cost, where the cost is the predictive variable. In this research, there is the opposite case in which the cost is the target variable and the time is a predictive variable. Of course, there is a lot of similar research, which has proved that there is a dependency between cost and time in construction projects [1–5, 10]. In many studies [1, 3, 5] BTC has served as an original model template. The simple regression analysis is the mathematical method that has been used to develop regression predictive process based models.

Another direction of modelling is aimed at soft computing. Developed models are data driven models. There are a lot of studies that used different types of soft computing methods in order to solve the cost prediction in construction projects [11–16]. Moreover, some of them have developed models for different types of structures. For example Ahiaga-Dagbui and Smith [17] have developed ANN based cost prediction models for hydro-engineering structures. El Sawy et al. [11] have developed ANN supported cost estimating model for site overhead costs. Developed model allows estimation of site overhead costs as a percentage of the total project cost.

Luu and Kim [14] have built an ANN model for prediction of construction cost for apartment projects in Vietnam. In comparison to MLR (Multiple Linear Regression) and Genetic algorithm model, ANN has shown the highest accuracy.

Alshamrani [15] has built a regression model for predicting initial cost of college buildings in North America where input variables were building area, floor height, number of floors, and structure and envelope type.

Petruseva et al. [18] have used SVM prediction model for construction cost and obtained very satisfactory results.

Some authors [12] developed neural network architecture that allows tunnel cost modelling in the early project phases.

Another group of authors [13] has researched costs in sports field construction projects in Poland by using several types of NNs, comparing the accuracy of obtained models.

Kim, An and Kang [19] have researched and compared results of construction costs estimation of implemented Korean residential building projects by using multiple regression analysis (MRA), neural networks (NNs) and case based reasoning (CBR). The authors have analyzed and discussed advantages and disadvantages of the models.
Petruseva et al. [18] have investigated and compared the accuracy of cost prediction results by using linear based regression models and support vector machines models (SVMs) for building projects in the Federation of Bosnia and Herzegovina.

From all of the aforementioned possibilities none of the techniques can be considered absolutely the best. Therefore, it seems expected that the research would redirect towards combination of different techniques used for modelling, e.g. regression techniques, neural networks, supporting vectors etc. The idea and the main objective of this research are to investigate some model combinations in order to develop hybrid model with the best accuracy for considered experimental data.

3. Methods

The first research step has been to establish adequate experimental database. Road structures construction projects realized over the past two decades in Croatia have been considered. Database has been established using the part of data collected through surveys during the research for Doctoral Thesis [20]. For this research, the database of 49 road structures has been established. The database consists of 32 road sections and 17 different road structures as follows: tunnels, overpasses, bridges, underpasses and animal crossings. Contracted and real construction time, contracted and real construction costs, type of object and construction year have been included in the database.

For modelling the prediction of the target variable – real construction cost, using other available variables: contracted construction cost, contracted construction time and real construction time, a hybrid model - combination of SVM technique (data-driven model) and Bromilow TCM (process-based model) has been developed.

Recently, hybrid modelling has demonstrated very promising results, using the best characteristics of two or more models and obtaining a new efficient model. SVM algorithm and Bromilow TCM are explained briefly below.

3.1. Support Vector Machine (SVM)

Solving the task of modelling from experimental data is the field of research of soft computing. Neural networks (NN) and support vector machines (SVM) have implemented this idea in the last decades, solving many different problems in modern science, specifically in engineering [21].

Soft computing algorithms implement the intelligence of nature beings, particularly learning from experience (experimental data) into the models, or other artificial systems and their most important characteristic is generalization from approximation and producing output from unknown data, using previously learned data [21].

3.2. SVM algorithm

Support vector machine (SVM) is very promising machine learning method for classification and regression tasks, demonstrating excellent results for most of the engineering problems as regression, computer vision, or recognition of objects [22]. In many applications, SVM has shown better predictive accuracy than neural networks (NN).

In order to solve the predicting problem (regression or classification), SVM should be supplied with input data, called training data, which should be the most representative for the process that should be modelled. The training data set D is composed of l n-dimensional vectors \(x_i\), and continuous variables \(y_i\), called responses from the system (Equation 1):

\[
D = \{(x_i, y_i) \in \mathbb{R}^n \times \mathbb{R}, \quad i = 1, ... l\}
\]  

(1)

The vectors \(x_i\) is called independent input variables, or predictors, and \(y_i\) is called output or target variable.
In the process of training in the regression task, SVM tries to find the dependency between \( x \) and \( y \), obtaining some approximation prediction function \( f(x, w) \). In classification tasks, this approximation function will separate the data, so it is called separation function. This function is obtained in a way to minimize the error of the approximation, called error function [21]. The function \( f(x, w) \) is usually nonlinear function of the weights \( w \), which are being adjusted in the process of learning.

Most widely used error function is the so called, Vapnik’s linear function with \( \varepsilon \) - insensitivity, (Equation 2):

\[
|y - f(x, w)|_\varepsilon = \begin{cases} 
0 & \text{if } |y - f(x, w)| \leq \varepsilon \\
|y - f(x, w)| - \varepsilon & \text{otherwise}
\end{cases}
\]  

(2)

Usually, SVM algorithm for regression task is explained by linear approximation function (Equation 3):

\[
f(x, w) = w^T x + b
\]

(3)

New empirical risk function is defined in the following way (Equation 4):

\[
R_{emp}^\varepsilon(w, b) = \frac{1}{l} \sum_{i=1}^{l} |y_i - w^T x_i - b|_\varepsilon
\]

(4)

In the case of solving the regression problem with SVM when the approximation function is linear (Equation 3), the function \( f(x, w) \) should be found, so that it approximates the pairs \((x_i, y_i)\) with \( \varepsilon \) precision. This is shown in Figure 1 [21].

Because the width of the tube is \( \frac{2\varepsilon}{\|w\|} \), for obtaining maximal accuracy of the approximation of the pairs \((x_i, y_i)\), \( \|w\| \) should be minimized.

**Figure 1.** \( \varepsilon \) - tube when approximation function is linear [21].
For simplification of the computational process, without changing the solution, \( \frac{1}{2} \|w\|^2 \) is used instead of \( \|w\| \). So, the SVM algorithm is reduced to find approximation function \( f(x, w) = w^T x + b \).

\( R_{\text{emp}} \) (eq. 4) and \( \frac{1}{2} \|w\|^2 \) are minimized, i.e. the real error, defined by Equation 5 is minimized.

\[
R = \frac{1}{2} \|w\|^2 + C \left( \sum_{i=1}^{l} y_i - f(x_i, w) \right)_{\varepsilon}
\] (5)

After solving the optimization task, the optimal values \( w_0 \) and \( b_0 \) are obtained and the required approximation function will be (Equation 6):

\[
z = f(x, w) = w_0^T x + b_0
\] (6)

When SVM should solve regression problem, which is not linear, as most of the contemporary problems, then, for finding nonlinear approximation function in the input space, mapping of the vectors \( x \) from the input space to vectors from the new multidimensional space \( F \), is used. This is accomplished with a mapping function \( \Phi \), called kernel function. This kernel function is selected so that the optimal approximation linear function is obtained in the new space \( F \). After obtaining the approximation function in the new space \( F \), the appropriate nonlinear approximation function in the input space is found by simple computation. [21]

3.2.1. Bromilow time-cost model

Bromilow “time-cost” model (TCM) expresses the relationship between the construction price and construction time (Equation 7) [9].

\[
X = a \cdot Y^b
\] (7)

\( X \) is construction time, \( Y \) is construction price, \( a \) and \( b \) are parameters, \( a \) gives the average construction time needed for construction of a monetary value and \( b \) - the dependence of the construction time on changing the price [9].

The data available for building the model was: contracted time for construction, real construction time, contracted price and real construction price. All available variables were chosen for building of the model. The Bromilow TCM (Equation 7) was applied for real price and real time of construction (Equation 8) and for contracted time and contracted price of construction (Equation 9):

\[
X_1 = a_1 \cdot Y_1^b_1
\] (8)

\[
X_2 = a_2 \cdot Y_2^b_2
\] (9)

where \( Y_1 \) and \( Y_2 \) are real and contracted construction costs respectively, and \( X_1 \) and \( X_2 \) are real and contracted construction time respectively.

In order to obtain the dependence of the real construction cost \( (Y_1) \) on other variables \( Y_2, X_1 \) and \( X_2 \) in one equation, the Equations (8) and (9) are logarithmized ((10) and (11)) and after that summarized (Equation 12):
\[ \ln(X_1) = \ln(a_1) + b_1 \ln(Y_1) \] (10)

\[ \ln(X_2) = \ln(a_2) + b_2 \ln(Y_2) \] (11)

\[ \ln(X_1) + \ln(X_2) = \ln(a_1) + b_1 \ln(Y_1) + \ln(a_2) + b_2 \ln(Y_2) \] (12)

The dependence of \( Y_1 \) on other variables is given in Equation 13:

\[ \ln(Y_1) = \frac{1}{b_1} [\ln(X_1) + \ln(X_2) - \ln(a_1) - \ln(a_2) - b_2 \ln(Y_2)] \] (13)

Equation 13 was used for implementation of the Bromilow TCM in the hybrid model proposed in this paper. As input variables, for the support vector machine predictive model, their actual values \( Y_1, Y_2, X_1 \), and \( X_2 \) were not used, but \( \ln Y_1, \ln Y_2, \ln X_1 \), and \( \ln X_2 \).

This implementation of the Bromilow model contributed significantly to the improvement of the model accuracy.

4. Results

The predictive modelling software DTREG was used for the prediction of the real cost of the road structures construction [23–24]. DTREG offers several types of predictive models and for our predictive modelling SVM was used, and Bromilow TCM was applied to its input variables. Four types of SVM models are offered in DTREG software, 2 for classification tasks and 2 for regression tasks and there are 4 kernel functions offered for each chosen SVM model: linear, polynomial, sigmoid and radial basis function (RBF). Linear kernel function was used for the considered model.

As the first, the SVM model was built without using implementation of the Bromilow TCM, knowing that the relation between construction time and cost is nonlinear. The percentage error of the model was around 40%; so that the hybrid model was build combining SVM and Bromilow TCM. After obtaining the hybrid model SVM-Bromilow TCM, with very high accuracy, another four hybrid models were built for comparison purposes: LR-Bromilow TCM, RBF NN- Bromilow TCM, MLP- Bromilow TCM and GRNN-Bromilow TCM, combining Bromilow TCM with LR (linear regression), RBF NN (radial basis function neural network), MLP (Multilayer perceptron) and GRNN (general regression neural network), respectively. The highest accuracy was obtained with SVM-Bromilow TCM model, which is discussed below.

The four input variables, which were available, are real and contracted cost of construction and real and contracted time of construction. These variables were used for model building. Real cost of construction was chosen as a target variable and the rest were used as predictors. It should be emphasized here that, because of the implementation of the Bromilow TCM, as was discussed in the previous section, the logarithm of the actual values of variables, and not the actual values themselves, were used as input.

The accuracy of the hybrid model, combining Bromilow TCM and SVM, estimated with the most used model’s estimators, MAPE (mean absolute percentage error) and R² - the coefficient of determination (expressing the general fit of the model), is: MAPE = 1.01% , R² = 97.61%. The coefficient of correlation between actual and predicted target values is 0.988 (Table 1, [24]). The validation method, used for testing the model, is 10 fold cross validation.
Table 1. The accuracy of the hybrid model (SVM-Bromilow) for validation data.

| Validation Data                      |       |
|--------------------------------------|-------|
| Mean target value for input data     | 15.817445 |
| Mean target value for predicted values | 15.792618 |
| Variance in input data               | 2.039863 |
| Residual (unexplained) variance after model fit | 0.048780 |
| Proportion of variance explained by model (R^2) | 0.97609 |
| Coefficient of variation (CV)        | 0.013963 |
| Normalized mean square error (NMSE)  | 0.023913 |
| Correlation between actual and predicted values | 0.988136 |
| Maximum error                        | 1.0255885 |
| RMSE (Root Mean Squared Error)       | 0.2208623 |
| MSE (Mean Squared Error)             | 0.0487802 |
| MAE (Mean Absolute Error)            | 0.1594287 |
| MAPE (Mean Absolute Percentage Error)| 1.0114172 |

By using sensitivity analysis and computations, DTREG reports the relative importance of the predictors for the model (Table 2, [24]).

Table 2. Overall importance of variables (DTREG software).

| Variable                  | Importance |
|---------------------------|------------|
| ln (contracted price)     | 100,000    |
| ln (contracted time)      | 0.164      |
| ln (real time)            | 0.147      |

The dependency of the actual and predicted values for the target variable is presented in Figure 2 [24].

Figure 2. Dependency of predicted and actual target values.
Minimal, maximal, mean value and standard deviation for each numerical variable are presented in Table 3 (by DTREG [24]).

**Table 3.** Minimal, maximal, mean value and standard deviation for every numerical variable.

| Variable                  | Rows | Minimum | Maximum | Mean    | Std. Dev. |
|---------------------------|------|---------|---------|---------|-----------|
| ln (contracted time)      | 49   | 3.40120 | 6.99851 | 5.58017 | 0.92660   |
| ln (real time)            | 49   | 4.29046 | 7.69166 | 5.88174 | 0.78703   |
| ln (contracted price)     | 49   | 12.27852| 18.79224| 15.61443| 1.45247   |
| ln (real price)           | 49   | 12.41654| 18.79224| 15.81745| 1.42824   |

5. Discussion

The importance of the hybrid model is going to be discussed below in more detail. When using only SVM model without implementation of the Bromilow model, i.e. when the input variables (predictors and target) for the SVM model have their actual values, the accuracy of the model for validation data, by using 10 fold cross validation, is significantly lower: MAPE = 39.43%, R²=84.96%, and the correlation coefficient is 0.940 (Table 4) [24].

By using only Bromilow model, the parameters of the model should be computed. The computation of the accurate parameters requires considerable effort, i.e. research and computing.

Before choosing SVM model as part of the hybrid model, other predictive models were tried, as part of the hybrid model using Bromilow TCM: LR (linear regression), GRNN (general regression neural network), RBF NN (radial basis function neural network) and MLP (multilayer perceptron). In fact, other four hybrid models were developed: LR–Bromilow, GRNN–Bromilow, RBF NN–Bromilow and MLP–Bromilow. The most accurate was SVM–Bromilow hybrid model. The results of the accuracy of all these five new hybrid models are shown in Table 5.

**Table 4.** Accuracy of the SVM model without implementation of Bromilow TCM (for validation data).

| Validation Data                                                                 |   |
|---------------------------------------------------------------------------------|---|
| Mean target value for input data                                               | 1.9387263e+07 |
| Mean target value for predicted values                                         | 1.6979715e+07 |
| Variance in input data                                                         | 8.8954e+014    |
| Residual (unexplained) variance after model fit                                 | 1.338e+014     |
| Proportion of variance explained by model (R²) = 0.84958                        | 84.958%        |
| Coefficient of variation (CV)                                                   | 0.596638       |
| Normalized mean square error (NMSE)                                             | 0.150415       |
| Correlation between actual and predicted values                                | 0.940120       |
| Maximum error                                                                  | 7.3843244 e+07 |
| RMSE (Root Mean Squared Error)                                                  | 1.1567187 e+07 |
| MSE (Mean Squared Error)                                                        | 1.338e+014     |
| MAE (Mean Absolute Error)                                                       | 4.1580264e+06  |
| MAPE (Mean Absolute Percentage Error)                                           | 39.432182      |
Table 5. Comparison of the accuracy for the five new hybrid models.

| Type of predictive model | MAPE (%) | $R^2$ (%) | Coefficient of correlation between actual and predicted target values |
|--------------------------|----------|-----------|---------------------------------------------------------------|
| SVM-Bromilow             | 1.01     | 97.61     | 0.988                                                         |
| LR-Bromilow              | 1.16     | 97.14     | 0.986                                                         |
| RBF NN-Bromilow          | 1.52     | 95.43     | 0.978                                                         |
| MLP NN-Bromilow          | 1.17     | 96.78     | 0.984                                                         |
| GRNN-Bromilow            | 1.38     | 95.61     | 0.979                                                         |

Luu and Kim [14] have built an ANN model for prediction construction cost for apartment projects in Vietnam and have obtained model with accuracy around 10% for MAPE. In comparison with MLR (Multiple Linear Regression) and Genetic algorithm model, ANN has shown the highest accuracy.

Alshamrani [15] has built a regression model for predicting initial cost of college buildings in North America with accuracy 94.3%, i.e. percentage error was 5.7%. The input variables were building area, floor height, number of floors, and structure and envelope type.

Authors in [16] have used SVM prediction model for construction cost and obtained very satisfactory model with accuracy of 92%, i.e. with 8% error of prediction. As input to the model, the data for early planning of the construction was used.

Authors in [25] have proposed hybrid model for predicting financial time series data combining BPNN (back propagation NN), SVM and SVR (support vector regression) and the results have shown that the new hybrid model outperforms all three single models.

6. Conclusion

As the problem of cost prediction, already in early project stages, is one of the crucial problems regarding successful project management, the research focus of this paper was to establish an adequate model with high accuracy to predict project costs.

The paper analyses the cost estimation model development using concrete database of road structures applying a hybrid model, which is a combination of process based Bromilow model, and data driven SVM model. DTREG software has been used. Accuracy with MAPE of 1.01% was obtained, with coefficient of determination $R^2$ of 97.61% and correlation coefficient of 0.998 between actual and predicted target values. The results were compared with those obtained by using other prognostic models, applying the same software.

The proposed model can be considered as a useful tool for all participants in construction project for early cost prediction, when numerous factors, which determine costs, are unknown. The future work will be focused on exploring new hybrid models and their characteristics.

It is possible to avoid cost overrun that is very often a problem in construction projects by using adequate cost predictive model. Cost overruns may disrupt the planned cost allocation balance and work quality and cause risks. Risks can cause possible negative impacts, particularly on economic, environmental and social aspects of sustainability. Therefore, it is of utmost importance to continue in research of this issue in order to improve the accuracy of cost prediction.

Statement about data
The authors declare that data supporting the results can be found in the authors databases.

Acknowledgements
The paper has been financed from the financial support funds of the University of Rijeka (project 13.05.1.3.10).
References
[1] Žujo V, Car-Pušić D and Brkan-Vejzović A 2010 Tehn. vjesnik: znanstveno-stručni časopis tehničkih fakulteta Sveučilišta u Osijeku. Contracted price overrun as contracted construction time overrun function. 17(1) pp 23–9
[2] Flyvbjerg B, Holm S M and Buhl L S 2002 Underestimating costs in public works projects: Error or lie? J. of the Americ. Plann. Assoc. 68(3) pp 279–95
[3] Le-Hoai L, Lee D Y and Lee Y J 2008 Delay and cost overruns in Vietnam large construction projects: a comparison with other selected countries J. of Civil Eng. 12(6) pp 367–77
[4] Le-Hoai L and Lee D Y 2009 Time-cost relationships of building construction project in Korea Facilities 27(13/14) pp 549–59
[5] Car-Pušić D and Radužković M 2009 Construction time-cost model in Croatia Int. J. for Eng. Modelling 22(1–4) pp 63–70
[6] Žujo V, Car-Pušić D, Žileska-Pancovska V and Ćečez M 2015 Time and cost interdependence in water supply system construction projects Technol. and Econ. Develop. of Econ. 21 pp 1–20
[7] Baloi D and Andrew F D 2002 Price Modelling global risk factors affecting construction cost performance Int. J. of Project Manag. 21(4) pp 261–7
[8] Morris P and Hough G 1987 The anatomy of major projects (New York: John Wiley)
[9] Bromilow F J 1969 Contract time performance expectations and the reality Building forum 1(3) pp 70–80
[10] Hu W and He X 2014 An innovative time-cost-quality tradeoff modeling of building construction project based on resource allocation The Sci. World J. ID 673248
[11] ElSawy I, Hosny H and Abdel Razek M 2011 A neural network model for construction project site overhead cost estimating in Egypt IJCSI Int. J. of Comp. Sci.Issues 8(3.1) pp 273–83
[12] Petroutsatou K, Georgopoulos E, Lambropoulos S and Pantouvakis P J 2012 Early cost estimating of road tunnel construction using neural networks J. of Constr. Eng. and Manag. 138(6) pp 679–87
[13] Juszczyk M, Leśniak A and Zima K 2018 ANN based approach for estimation of construction costs of sports fields Complexity 2018 7952434
[14] Luu V T and Kim S Y 2009 Neural network model for construction cost prediction of apartment projects in Vietnam Korean J. of Constr. Eng. and Manag. 10(3) pp 139–47
[15] Alshamrani S 2017 Construction cost prediction model for conventional and sustainable college buildings in North America J. of Taibah Univ. of Sci. 11(2) pp 315–23
[16] Wang Y-R, Yu C-Y and Chan H-H 2017 Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models Int.J. of Project Manag. 30(4) pp 470–8
[17] Ahiaga-Dagbui D and Smith S 2012 Neural networks for modelling the final target cost of water projects. Proc. 28th Annual ARCOM Conf. (Edinburgh, UK: Association of Researchers in Construction Management) 307–16
[18] Petruševa S, Zileska-Pancovska V, Žujo V and Brkan-Vejzović A 2017 Construction costs forecasting: comparison of the accuracy of linear regression and support vector machine models Tehn. Vjesnik 24(5) pp 1431–8
[19] Kim H G, An H S and Kang I K 2004 Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning Build. and Environ. 39(10) pp 1235–42
[20] Car-Pušić D 2004 Građevinski fakultet Sveučilišta u Zagrebu [Faculty of Civil Engineering, University of Zagreb], Metodologija planiranja održivogvremena građenja [Methodology of planning the sustainable construction time]. Dissertation. Zagreb, Croatia. (in Croatian),
[21] Kecman V 2001 Learning and Soft Computing: support vector machines, neural networks and fuzzy logic models. (Cambridge, Massachusetts: A Bradford Book. The MIT Press)
[22] Vapnik V N, Golowich S and Smola A 1997 Support Vector method for function approximation, regression, estimation and signal processing Advanc. in neural inform. Process. Syst. (Cambridge, MA: MIT Press.) 9

[23] Sherrod P 2013a Tutorial DTREG Predictive Modelling Software www.dtreg.com, access 2018.01.15

[24] Sherrod P 2013b. DTREG Predictive Modelling Software. www.dtreg.com access 2018.01.15

[25] Al-hnaity B and Abbod M 2015 Predicting Financial Time Series Data Using Hybrid Model Intell.Syst.and Applic., Studies in Computational Intelligence 650 pp 19-41