CONSTRUCTION COSTS FORECASTING: COMPARISON OF THE ACCURACY OF LINEAR REGRESSION AND SUPPORT VECTOR MACHINE MODELS

Silvana Petruseva, Valentina Zileska-Pancovska, Vahida Žujo, Aida Brkan-Vejzović

Original scientific paper

Each contract for a construction project has the costs as an essential element, so the accuracy of forecasting the construction costs can have an impact on the project realization, and also, on the project participants' business. Data for structures (75) were used for modelling with two predictive models: linear regression model (LR) and support vector machine (SVM) model, using Bromilow's model for cost and time relation and predictive modelling software DTREG. The mean absolute percentage error (MAPE) for the SVM model is 0.3% and for the linear regression model is 4.79%. Comparison of the models' results pointed out that the forecasting with SVM was significantly more accurate.

Keywords: construction costs; forecasting; linear regression; support vector machine

Predviđanje troškova građenja: usporedba točnosti modela linearne regresije i modela podupirućih vektora

Izvorni znanstveni članak

Bitni element svakog ugovora za građevinski projekt su troškovi kao bitni element pa točnost predviđanja troškova može utjecati na izvršenje projekta, a također i na poslove onih koji su uključeni u projekt. Podaci za objekte (75) upotrebljeni su za modeliranje s dva modela za predviđanje: model linearne regresije (LR) i model podupirućih vektora (SVM), uz primjenu Bromilowa modela za odnos vremena i troška i softvera DTREG - za prediktivno modeliranje. Srednji apsolutni postotak greške (MAPE) za SVM model je 0.3% a za model linearne regresije 4.7%. Usporedba rezultata modela pokazala je da je predviđanje sa SVM značajno točnije.

Ključne riječi: linearna regresija; podupirući vektori; predviđanje; troškovi građenja

1 Introduction

The construction costs are an element of each contract for a construction project, so they are a matter of project participants' great interest [1].

In spite of the costs importance for the construction projects, costs overrunning is a worldwide common issue [2]. Many projects are not finished within the costs that are specified [3], i.e. they do not achieve their costs objectives [4-6]. Hence, completing the construction project within the costs that are contracted has become a challenge [7].

Construction costs are determined by numerous changeable factors, so it is impossible to acknowledge all of them in the process of construction cost forecasting. Additionally, the relationship between costs, quality and time should be respected [8], which increases the complexity and the responsibility of the construction costs forecasting process. Therefore, costs estimation is an issue of special interest for construction project participants, particularly in the cases when a fast costs prediction is needed. Also, accurate forecasting of costs is of crucial importance for the project participants' business, especially for large-scale projects with long period of construction [9].

1.1 Aim of the paper

Paper's aim is to create/design models for fast construction costs prediction using two untraditional forecasting models: linear regression model and support vector machine model, and to compare their accuracy.

Linear regression (LR) methods have many advantages such as their speed of convergence in comparison with neural networks, which usually have low speed of convergence, also their small requirement for memory space and their simplicity of use, but they cannot work well with non-linear functions as neural networks and support vector machines (SVM).

SVM are very accurate predictive models which in many cases have demonstrated better prediction than neural networks, in regression and also in classification problems and because of that they are being intensively investigated in the last several years.

There is important difference between the assumptions in the training with SVM and the classical statistical modelling. In most of the classical statistical approaches the experimental data are being modelled by a set of functions which are linear in parameters and the assumed joint probability for the data is normal, Gaussian distribution, for most of the real-life problems. But in recent years, for many modern real-life problems these assumptions turned out to be inappropriate, because most of the contemporary problems are highly dimensional (with an increasing dimensionality of the input), so that the model will need exponentially increasing number of terms, thus becoming very hard for use. Also, the distribution laws for most of the real-life data are non-Gaussian, often very far from the normal distribution. SVM models have been developed to overcome all these problems and to be suitable for the standard modern data sets, which in most of the cases are with small training data-sets and highly dimensional [10].

Starting from the above mentioned, in this paper these two models (LR and SVM) were applied to real data from the practice and more accurate predicting was obtained with the SVM model.

2 Literature review

Construction costs are among issues that are highly pronounced in construction practice and their

investigation has become a challenge for many authors worldwide. They have investigated different aspects of construction costs such as: costs estimation, costs forecasting, costs overruns, the relation cost-time etc. [11-29].

For estimating and forecasting the construction costs there are traditional models which are based on quantities, design, resources, functional elements or building operational units etc. But, many authors put particular attention on the untraditional models for costs estimation and forecasting. They use new techniques and practices such as the experimental models, regression models and simulation models [12].

Authors in [13] have tested and have refitted the Bromilow's linear regression model. They have shown that various projects need various estimates of parameters. They developed two type of models: for industrial projects and for projects that are not industrial. Similarly, other authors [14] investigated the correlation between time and cost overruns by applying the Bromilow's time-cost algorithm.

The functional relationship construction time - construction cost for highways was explored in [15]. Additionally, for forecasting cost and time regression models were created in [16]. Models depend on contract sum, contract period, contractual arrangement, knowledge of the contractor selection method, client sector, project type etc. Regarding sensitivity analyses it was shown that the errors in actual construction cost forecasting for large and small projects are virtually the same [15].

In [17] authors proposed a probabilistic model for predicting risk effects on projects costs and duration. For developing the statistical regression model and sample tests they used historical data of similar construction projects. With the 95% probability of the model the precision of the mean cost and time prediction was ±0.035% .

A parametic model for estimation of cost for highway projects was presented in [18]. A supervised neural network model optimized with Genetic Algorithms was established together with parameters of determination which notably impact the price of the highway projects. The model uses the Levenberg-Marquardt algorithm as a back-propagation rule and Hyperbolic Tangent function as a transfer function for both hidden and output layers. The model was trained and evaluated. Then, after carring out case study for testing its validity and accuracy in handling real data, a graphical user interface module was coded for the model to facilitate its usage and manipulation with future highway projects practical applications. The results showed that the developed model is reliable to be used at early stages of highway projects because its percentage error is 16% (and is lower than the allowed 20%). Similarly, [19] gives cost estimation model that is based on artificial neural network and is useful for highway projects at the conceptual phase in developing countries.

A model for cost estimation in road projects that is useful for construction managers is built in [20]. The model is developed using support vector machine. The 12 factors were identified to be the most important factors affecting the cost-estimating model. A total of 70 case studies from historical data were used for modelling and

the built model was successfully able to predict project cost with accuracy of 95%.

A parametric model for estimation of cost for railway systems in urban areas in predesign stage is given in [21]. Incorporating neural networks and regression analysis, the authors have developed powerful parametric model for selecting the parameters with big impact on the cost during early project stages.

In paper [22] authors developed predictive regression model that uses factors which affect project cost and examined their importance. From the literature review in this study it is indicated that significant effects on the overall project cost come from several factors. They are connected with: clients, experts duration, etc. Some of the six most important factors to project cost identified by the study are the levels of: design complexity, construction complexity; technological advancement etc. The contractor's experience on similar projects; the consultant and the client; the suitableness of contractor's plant and equipment used are the most important among those factors, so they were used for developing the cost predictive model.

For estimation of the construction cost the back-propagation network (BPN) model is applied in [23]. General algorithms are incorporated in BPN in order to choose the parameters of BPN's. The authors obtained very effective and accurate model for estimating construction costs.

An efficient cost estimation tool is presented in [24], useful for project managers and designers in early phase of design process for buildings, using neural network methodology for estimating the cost of square meter for 4-8 storey residential reinforced concrete-structural systems.The accuracy that was achieved was 93%.

In [25] support vector machine, an artificial intelligence technique is used for construction cost estimation that is useful for planners and owners for predicting the cost of a construction project. Through interview with experts and literature review, the factors which impact the cost estimate most, are identified. The data from 29 construction projects are used in the training process of the SVM. The average prediction error of the model that they obtained was less than 10% and the computation time was less than 5 minutes.

Paper [26] presents the development of ANN ensemble and SVM classification models for predicting the schedule and cost success, using the status of the early planning, as input to the model. The authors obtained satisfactory prediction results. For forecasting the success of the project' cost, the models of SVM produced the best result of forecasting with accuracy of 92%, and the models of ANN gave the best result of forecasting with accuracy of 80% - for schedule success forecasting. The conclusion of the paper is that good planning is needed in the early project' stage. That leads to better project' outcomes. Also, the techniques of artificial intelligence modeling are more suitable for sample data that are non-linear in comparison with the traditional statistical regressions.

The authors of the study [27] propose a new model, named as EAC-LSPIM, capable to deliver accurate predictive results using Least Squares Support Vector Machine (LS-SVM), machine learning based interval

estimation (MLIE), and Differential Evolution (DE). LS-SVM is used for regression analysis and implements supervised learning. MLIE is used for obtaining the prediction intervals and DE is used in the cross validation process for obtaining the optimal values of tuning parameters. At a certain level of confidence, the model gives the lower and upper prediction limits.

Comparison of the accuracy of regression analysis, SVM and NN (neural networks) in predicting construction costs is presented in [28]. The authors used historical data for construction cost for school building projects and they have obtained that NN gives more accurate prediction results than SVM and regression model, and they concluded that NN models are more suitable for predicting the cost of buildings for schools. For developing the model authors used 197 buildings data and 20 buildings data for testing the model. All three models produced a high correlation between the predicted cost and the cost that are actual.

Support vector machine is also used in [29] to create intelligent model for predicting the conceptual cost in construction industry during the early phases of the projects. In order to point out the performance of the SVM model, non-linear regression model and back-propagation NN model, were developed. The comparison with these two techniques illustrated that SVM performs better than these two techniques.

3 The research method

Using the questionnaires and interviews with contractors and site engineers, data for 75 structures constructed in Bosnia and Herzegovina have been obtained. Key collected data covered: contracted price, contracted time, real price, real time of construction, and the reasons for non-compliance of deadline.

The predictive modelling software DETREG [30, 31] was used for developing two predictive models: linear regression (LR) and support vector machine (SVM). The well-known 'time-cost' model proposed by Bromilow [32] was implemented in both of these models and the accuracy of the prediction of these models was compared.

3.1 Linear regression prediction model for construction project costs

The method for modelling relationship of one target (dependent) variable y with one or several independent variables (predictors) x, is called *linear regression* (LR). When there is only one predictor variable it is called *simple linear regression*, and when there are several predictors, it is called *multiple linear regression*.

The linear regression model is given by Eq. (1) [30]:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n \tag{1}$$

y is the dependent variable (target) and x_1, x_2,…, x_n, are the independent variables (predictors). B_1, β_2,…, β_n are parameters (coefficients). β_0 is a constant [30].

Because LR is used for approximation of linear functions to data points, it does not work well on real data which usually have non-linear relationships.

But LR has many practical uses. For example, it can be used for measuring the strength of the relationship between x_j and y, or to determine which x_j have no relationship with y. One of the most important features of the LR model is that it is simple and well understood and in comparison with neural networks (which can model non-linear relationships), is much faster, requiring minimum memory space. Also, by analysing the values and the sign of the regression coefficients, the affecting of the target variable by the predictors can be inferred [30].

In the case of simple linear regression, the approximation function is a straight line. In the case of two independent variables, the approximation function is a plane. Fig. 1 presents the approximated plane with two independent variables [30].

For particular observation, the difference between the actual and the predicted value of the dependent (target) variable is the error of the prediction which is also called "deviation".

The goal of every regression analysis is obtaining the β coefficients from Eq. (1), which minimize the error of the prediction.
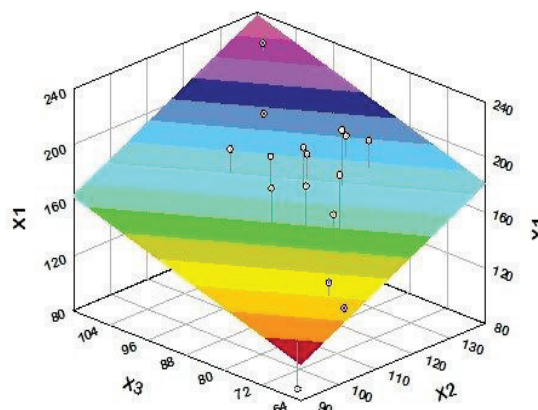

Figure 1 Fitted plane with 2 predictive variables [30]

For the prediction of the construction cost, the 'time-cost' model proposed by Bromilow [32] is implemented in our LR model. It is given in Eq. (2):

$$T = K \cdot C^B \tag{2}$$

where T is contracted time; C is contracted price; K is parameter for measuring productivity, it expresses the average time required for the construction of a monetary value; B is a parameter which expresses the dependence of the time on the change of the price [32].

The linear form of Eq. (2) can be obtained by its logaritmizing, as shown in Eq. (3):

$$\ln T = \ln K + B \cdot \ln C \tag{3}$$

and this form is used for the requirements of the linear regression model.

From Eq. (3) we shall express the cost as a function of time (Eq. (4)):

$$\ln C = \frac{1}{B} \ln T - \frac{1}{B} \cdot \ln K \tag{4}$$

Now single linear regression can be applied using this linear form and the values of the parameters K and B shall be obtained for this model.

For predicting the construction costs of the 75 structures, the above Eq. (4) was used in the linear regression predictive model from DTREG software package [31].

Tab. 1 and Tab. 2 show the results of the linear regression predictive model using DTREG software.

Table 1 LR model for predicting real price of construction depending on real time of construction

Linear Regression Parameters Computed Coefficient (Beta) Values					
Variable	Coeff.	Std. Error	t	Prob(t)	95% Conf. Interval
ln(real t)	1.22251	0.0972	12.57	< 0.00001	1.029 1.416
Const.	7.11869	0.4653	15.30	< 0.00001	6.191 8.046

Using Eq. (4), ln (real time) is used as independent (predictor) variable and ln (real price) is used as target variable, not the actual values of real time and real price of construction. This application of Bromilow' model contributed very much to the improvement of the accuracy of the model.

The first column of Tab. 1 shows the coefficients of the linear regression model. They are: $1/B = 1.22251$, the coefficient before the variable ln (real time), and the coefficient $(-1/B)\ln K = 7.11869$. From this last equation,

the parameter K can be computed: $K = e^{-\frac{7.11869}{1.22251}}$.

The model will be (Eq. (5)):

$$C = \left(\frac{T}{K}\right)^{\frac{1}{B}} \text{ i.e. } C = \left(\frac{T}{e^{-\frac{7.11869}{1.22251}}}\right)^{1.22251} \quad (5)$$

From the data set available for modelling, one part is used for training, and the rest is used for validating of the model. Tab. 2 shows the most used estimators of the accuracy of the model for validation data: the coefficient of determination R^2, the coefficient of correlation between the predicted and actual values of the target variable and *MAPE*, the mean absolute percentage error.

Table 2 Estimation of the accuracy for validation data (LR model)

Coefficient of determination $R^2 = 0.66583$ (66.583%)
Coefficient of correlation between actual and predicted target values $r = 0.816458$
MAPE = 4.7913225

3.2 Support vector machines (SVM) predictive model
3.2.1 Support vector machines

The Support vector machine (SVM) is a relatively new modelling method that has been rapidly developed in recent years, showing very good results at obtaining accurate models for many problems like pattern classification, function approximation, and regression problems.

The architecture of SVM models is very similar to that of neural networks.

There are two categories of support vector machine models, depending on the kind of the problem they solve: support vector regression (SVR) and support vector classification (SVC).

Till now SVM research has been intensively oriented toward solving practical problems, evolving in very active field of investigation. Because of its very good generalization to unknown data, SVMs have demonstrated very good performances at prediction with time series and also regression problems and in the last several years SVM methods are becoming standard methods of machine learning [33].

SVM are also called *kernel methods*, i.e. methods which use kernel functions to transform non-linear learning problem in linear by transforming the input space into output multidimensional space with kernel mapping functions and solving the problem in the output multidimensional space. This holds for solving regression problems and also classification problems. Fig. 2 shows mapping from two-dimensional into three-dimensional space [34].

This concept of mapping the input space into new multidimensional space and solving the problems in the new space is very efficient, allowing SVM models to perform very complex separations.
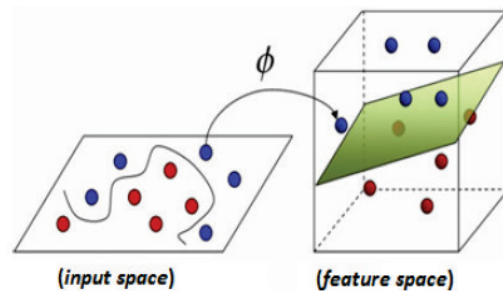

Figure 2 Mapping in multidimensional space [34]

3.2.2 Regression using SVM

The novel training concept implemented in SVM shall be explained shortly. Similarly to neural networks SVMs have the ability to approximate arbitrary multidimensional function to some desired level of accuracy, and so they are very desirable for modelling non-linear and unknown processes [10].

The only information available for learning is the input data set S

$$S = \{(x_i, y_i) \in X \times Y\}, i = 1, 2,..., m \quad (6)$$

which contains m data pairs from which SVM will be trained to learn the relationship between input predictors and output target variable, obtaining some approximation function $f(x,w)$. $x_i \in X$ are the input n-dimensional vectors, and $y_i \in Y$ are the values of the target variable, or outputs of the SVM model. w is weight vector whose components will be obtained in the process of training [10].

After the process of training support vector regression (SVR) model obtains the non-linear n-dimensional approximation function $f(x,w)$ which has some deviation

from the actual target values from the training data set, which is, in fact, the *error of the approximation*.

The accuracy of the SVR is being measured by computing the error of the approximation. The Vapnik's ε *-insensitivity* error (or loss) function is widely used. It is given in Eq. (7) [10].

$$L_\varepsilon = | y - f(x, w) |_\varepsilon = \\ = \begin{cases} 0 & \text{for } | y - f(x, w) | \le \varepsilon \\ | y - f(x, w) | - \varepsilon & \text{otherwise} \end{cases} \qquad (7)$$

ε is the maximum distance of the function f from the actual values of the target variable, i.e. the values of the function f must lie in a tube with radius ε (Fig. 3, [35]).

Solving the SVR problem will be shortly explained on a linear regression problem. In this case, the approximation function f that should be obtained is linear, and it should approximate all data pairs (x_i, y_i) with maximal deviation ε (Eq. (8) and Fig.4):

$$f(x, w) = w^{\mathrm{T}} x + a \qquad (8)$$
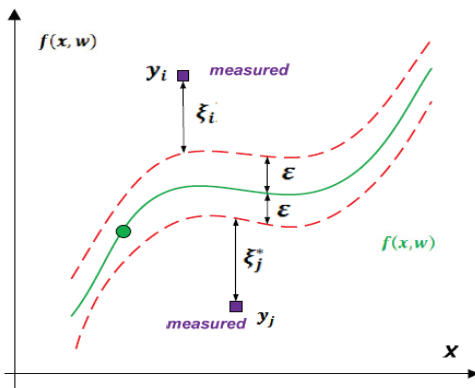

Figure 3 ε - tube [35]

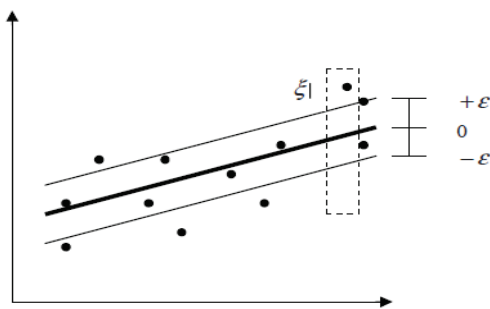**Figure 4** ε - tube when the approximation function is linear

Because the width of the tube is inversely proportional to the norm of vector w, $|| w ||$, the maximal allowed distance of the pairs (x_i, y_j) will be computed by minimization of $|| w ||$. If for computational convenience (without changing the result) this condition is defined to be minimization of $\frac{1}{2} || w ||^2$ instead of minimization of $|| w ||$ and considering Eq. (7) and Eq. (8), the solution of this problem is reduced to convex optimization problem given by the conditions (9):

Minimization of $\frac{1}{2} || w ||^2$ subject to $\begin{cases} y_i - wx_i - a \le \varepsilon \\ wx_i + a - y_i \le \varepsilon \end{cases}$ (9)

The required regression function $f(x, w) = w^{\mathrm{T}} x + a$ is being computed by minimizing the term [10]:

$$R = \frac{1}{2} || w ||^2 + C \left(\sum_{i=1}^{l} | y_i - f(x_i, w) |_\varepsilon \right) \qquad (10)$$

In most of the cases the regression problems are non-linear. In this case, SVR algorithm solves the problem by mapping the input *n*-dimensional space (defined by input *n*-dimensional vectors x_i) into some multi-dimensional, so called *feature space* with some mapping function Φ, called *kernel function*. The role of the kernel function is to transform the non-linear problem from the input space into linear problem in the new feature space. Then the problem in the new feature space is solved as linear regression problem, as was discussed above [10, 35].

3.2.3 Predicting construction cost with SVM model using DTREG software

Predicting modeling software DTREG ([30, 31]) is used for forecasting the construction cost of the objects.

SVR (support vector regression) and SVC (support vector classification) models are offered as SVM models by DTREG software.

Since different kernel functions will be suitable for different type of data, DTREG offers several kernel functions: sigmoid, radial basis function (RBF), polynomial and linear function. It is often recommended to try several kernel functions in order to choose the best one, but in most of the cases RBF function is the best one.

The parameter which controls the stopping of the optimization process of SVM algorithm is called *Epsilon value*, which is tolerance factor whose value can be changed by the user. By its reducing, a more accurate model will be generated, and by its increasing the computation time will be reduced [30].

There are different methods for validating the accuracy of the SVM model in the DTREG software.. They are: 1) random percent validation when random percent rows from the input data set are held out and used for validating of the model after the training process; 2) using variable for choosing the rows for validating; and 3) V-fold cross validation, when V-1/V proportion of the rows from the data set is being used in each SVM model during the training process. The other rows are used for validating. In our model V-fold cross validation is applied, as most recommended [30].

DTREG has two methods for obtaining optimal values for the most important parameters of the model which influence the accuracy of the model.

For obtaining optimal function value, DTREG for SVR has the only option offered: **minimization of the total error**.

The authors Chih-Chung Chang and Chih-Jen Lin [36] have contributed very much to the improvement of the SVM algorithms, for their mathematical theory and also for their practical applications and DTREG software

for SVM is based on the LIBSVM project by these authors.

Tab. 3 shows the results for the estimators of the accuracy of the DTREG's SVM model for the validation data: the coefficient of determination, coefficient of correlation and MAPE.

Table 3 Estimation of the accuracy for validation data for SVM Model

Coefficient of determination $R^2 = 0.99571$ (99.571%)
Correlation between actual and predicted target values $r = 0.997880$
$MAPE = 0.3034884$

We shall stress here the importance of choosing the relevant predictors for the target variable. Using Eq. (3) (the implementation of the Bromilow' time-cost model), the actual values of the predictors (real time, contracted price and contracted time), and the target variable (real price) are not used as input variables of the model, but logarithm of their values, which contributed very much for the accuracy of the model.

4 Analysis of the results and discussion

Usually R^2 and MAPE (mean absolute percentage error) are being used as estimators of the accuracy of the predictive models. In statistics, the coefficient of determination R^2 measures the general suitability of the predictive model, indicating how good data points match a curve or a line. The values of R^2 belong to the interval [0,1] and expressed in percent R^2 computes how much the variation of the output response is attributed to the predictors of the model. The value $R^2 = 0.666$ for our LR model may be interpreted: around 67% of the variation in the response is attributed to the chosen predictors, the rest 33% are due to unknown variability.

MAPE usually expresses accuracy as a percentage. For this model *MAPE* = 4.79 means that the percentage error of the LR model is 4.79%.

The Support Vector Machine model accuracy is shown in Tab. 3 and the results $R^2 = 0.996$ (99.6%), and *MAPE* = 0.30 indicate considerable increasing of the accuracy in comparison with LR model.

As it can be seen from the results, the SVM model gave excellent predictive result for our data and significantly more accurate in comparison with linear regression.

We should stress here that when we do not know in advance the type of the relationship between the predictors and the dependant variable, then we should try several predictive models in order to obtain the best one for our data, which will give most accurate predicting, because for different data, different type of predictive model can be suitable. For example the authors in [28] have obtained more accurate predicting with ANN than with SVM. Authors in [25] have obtained satisfying results with SVM with error less than 10% and time less than 5 min. Authors in [20] have obtained also good results with SVM predicting obtaining 95% accuracy of the model.

5 Conclusion

The construction process is a complex process that is influenced by numerous and changeable factors.

Additionally, the accurancy of construction costs forecasting can have an essential role for the process of construction and for the project participants bussiness. Therefore, the cost forecasting is a particularry difficult and responsible process.

Learning from previous projects costs experience is an important issue. For that purpose, a data base for costs of previously realized construction project was formed. Data were used for creating models for forecasting of the construction cost using linear regression (LR) and support vector machine (SVM).

With the SVM model we received a significantly more accurate forecasting. The result $R^2 = 0.996$ (99.6%), and MAPE=0.30 shows excellent predictive capabilities of the SVM, considering that these results are for real problems from the practise.

One of the weaknesness of the SVM model is its speed of convergence, in relation to the linear regression model.

The models are convenient for rapid and efficient forecasting of construction cost and they are not a substitution of detail cost estimation process. Due to that, they are applicable for the initial phase of the construction projects by project participants and by clients.

These models limitation is that they are applicable in construction projects without strong influence of physical factors (poor organisation of construction site and works, incomplete documentation, incorrect documentation, bad climate conditions etc).

Nevertheless, these research results are a good base for future investigation of some algorithms for speeding up the convergence of the SVM model like hyper parameters search space reductions algorithms or PSO (particle swarm optimisation) algorithms which can considerably speed up the convergence of SVM algorithms. Also, the approaches used may serve as an experience for developing other models for construction costs prediction.

The research discussed in our paper could be useful for construction projects participants and also for improving construction process in general.

Acknowledgements

6 References

[1] Zhang, Y.; Ng. S. T. An ant colony system based decision support system for construction time-cost optimization. // Journal of Civil Engineering and Management. 18, 4(2012), pp. 580-589. https://doi.org/10.3846/13923730.2012.704164

[2] Ahmed, S.; Azhar, S.; Castillo, M.; Kappagantula, P. Construction delays in Florida: An empirical study. pp. 1-44. https://ohlsc.fiu.edu/pdfs/Research_Reports/Delays_Project.pdf (21.10.2014)

[3] Azhar, N.; Farooqui, R. U.; Ahmed, M. S. Cost overrun factors in construction industry of Pakistan. // First international conference on construction in developing countries (ICCIDC-I), Advancing and Integrating Construction Education, Research & Practice. (2008), pp. 499-508.

[4] Moura, H. P.; Teixeira, J. C.; Pires, B. Dealing with Cost and Time in the Portuguese Construction Industry. // CIB2007-422, CIB World Building Congress 2007, pp. 1252-1265. https://www.irbnet.de/daten/iconda/ CIB5165.pdf (15.12.2016)

[5] Olawale, A. Y.; Sun, M. Cost and time control of construction projects: Inhibiting factors and mitigating measures in practice. // Construction Management and Economics, 28, 5(2010), pp. 509-526. https://doi.org/10.1080/01446191003674519

[6] Czarnigowska, A.; Sobotka, A.Tim cost relationship for predicting construction duration. // Archives of Civil and Mechanical Engineering, 13, 4(2013), pp. 518-526, http://dx.doi.org/10.1016/j.acme.2013.05.004

[7] Cantarelli, C.C.; Cost overruns in Dutch transportation infrastructure projects, Bijdrageaan het ColloqiumVarvoersplanologischSpeurwerk, 2009. URL: www.cvs-congres.nl/cvspdfdocs/cvs09_156.pdf. (14.10.2014)

[8] Yadaiah, V.; Haragopal, V. V. Multi-Objective Optimization of Time-Cost-Quality Using Hungarian Algorithm. // American Journal of Operations Research, 6, 31-35. https://doi.org/10.4236/ajor.2016.61005

[9] Hwang, S.; Park, M.; Lee, H.; Kim, H. Automated Time-Series Cost Forecasting System for Construction Materials. // Journal of Construction Engineering and Management. 138, 11(2012), pp. 1259-1269. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000536

[10] Kecman, V. Support Vector Machines – An Introduction. School of Engineering Report 616, The University of Auckland, Auckland, NZ, Springer-Verlag, Berlin Heidelberg, 2005.

[11] Mačková, D.; Bašková, R. Applicability of Bromilow´s Time-cost Model for Residential Projects in Slovakia. // SSP - Journal of Civil Engineering. 9, 2(2014), pp. 5-12.

[12] Akintoye, A.; Fitzgerald, E. A survey of current cost estimating practices in the UK. // Construction Management and Economics. 18, 2(2000), pp. 161-172. https://doi.org/10.1080/014461900370799

[13] Ng, S. T.; Mak, M. Y. M.; Skitmore, R. M.; Lam, K. C.; Varnam, M. The predictive ability of Bromilow's time–cost model. // Construction Management and Economics. 19, 2(2001), pp. 165-173. https://doi.org/10.1080/01446190150505090

[14] Žujo, V.; Car-Pušić, D.; Žileska-Pančovska, V.; Ćećez, M. Time and cost interdependence in water supply system construction projects. // Technological and Economic Development of Economy.

[15] Shr, J.-F.; Chen, W.-T. Functional model of cost and time for highway construction projects. // Journal of Marine Science and Technology. 14, 3(2006), pp. 127-138.

[16] Skitmore, R. M.; Ng, S. T. Forecast models for actual construction time and cost. // Building and Environment. 38, 8(2003), pp. 1075-1083. https://doi.org/10.1016/S0360-1323(03)00067-2

[17] Abu Hammad, A. A.; Alhaj Ali, M. S.; Sweis, J.G.; Bashir, A. Prediction model for construction cost and duration in Jordan. // Jordan Journal of Civil Engineering. 2, 3(2008), pp. 250-266.

[18] Adel, K.; Elyamany, A.; Belal, M. A.; Kotb, S. A.Developing Parametric Model for Conceptual Cost Estimate of Highway Projects. // International Journal of Engineering Science and Computing. 6, 7(2016), pp. 1728-1734.

[19] Sodikov, J. Cost Estimation of Highway Projects in Developing Countries: Artificial Neural Network Approach. // Journal of the Eastern Asia Society for Transportation Studies. 6, (2005), pp. 1036-1047.

[20] El-Sawalhi, I. N. Support Vector Machine Cost Estimation Model for Road Projects. // Journal of Civil Engineering and Architecture. 9, (2015), pp. 1115-1125

[21] Sonmez, R.; Ontepeli, B. Predesign cost estimation of urban railway projects with parametric modeling Predesign cost estimation of urban railway projects with parametric modeling. // Journal of Civil Engineering and Management. 15, 4(2009), pp. 405-409. https://doi.org/10.3846/1392-3730.2009.15.405-409

[22] Ganiyu, O. B.; Zubairu, K. I. Project Cost Prediction Model using Principal Component Regression for Public Building Projects in Nigeria. // Journal of Building Performance. 1, 1(2010), pp. 21-28.

[23] Kim, G.-H.; Yoon, J.-E.; An, S.-H.; Cho, H.-H.; Kanga, K.-I. Neural network model incorporating a genetic algorithm in estimating construction costs. // Building and Environment. 39, 11(2004), pp. 1333-1340. https://doi.org/10.1016/j.buildenv.2004.03.009

[24] Gunaydın, H. M.; Dogan, S. Z. A neural network approach for early cost estimation of structural systems of buildings. // International Journal of Project Management. 22, 7(2004), pp. 595-602. https://doi.org/10.1016/j.ijproman.2004.04.002

[25] Cheng, M.-Y.; Wu, Y.-W. Construction Conceptual Cost Estimates Using Support Vector Machine. // 22nd International Symposium on Automation and Robotics in Construction ISARC 2005 / Ferrara (Italy), 2005, pp. 1-5.

[26] Wang, Y.-R.; Yu, C.-Y.; Chan, H.-H. Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models. // International Journal of Project Management. 30, 4(2012), pp. 470-478. https://doi.org/10.1016/j.ijproman.2011.09.002

[27] Cheng, M.-Y.; Hoang, N.-D. Interval estimation of construction cost at completion using least squares support vector machine. // Journal of Civil Engineering and Management. 20, 2(2014), pp. 223-236. https://doi.org/10.3846/13923730.2013.801891

[28] Kim, G.-H.; Shin, J.-M.; Kim, S.; Shin, Y. Comparison of School Building Construction Costs Estimation Methods Using Regression Analysis, Neural Network, and Support Vector Machine. // Journal of Building Construction and Planning Research. 1, 1(2013), pp. 1-7. https://doi.org/10.4236/jbcpr.2013.11001

[29] Vahdani, B.; Mousavi, M. S.; Mousakhani, M.; Sharifi, M.; Hashemi, H. A Neural Network Model Based on Support Vector Machine for Conceptual Cost Estimation in Construction Projects. // Journal of Optimization in Industrial Engineering. 5, 10(2012), pp. 11-18.

[30] Sherrod, P. DTREG Predictive Modeling Software – tutorial. 2013. www.dtreg.com (15.09.2013)

[31] Sherrod, P. DTREG Predictive Modeling Software. 2013. www.dtreg.com (13.09.2013)

[32] Petruseva, S.; Zujo, V.; Zileska-Pancovska, V.Neural Network Prediction Model for Construction Project Duration. // International Journal of Engineering Research & Technology. 2, 11(2013), pp. 1646-1654.

[33] Haykin,S. Neural Networks: A Comprehensive Foundation, 2nd ed., Pearson Education Inc., Canada, 2005.

[34] Pesko, I. Model for estimation of costs and time of construction for urban traffic arteries. // PhD Thesis (in Serbian), Novi Sad, 2013.

[35] Kecman, V. Learning and Soft Computing, Support Vector Machines, Neural Networks and Fuzzy Logic Models, Massachusetts Institute of Technology (MIT), USA, 2001.

[36] [36] Chang, C. C.; Lin, C. J. LIBSVM – A Library for Support Vector Machines, (2005), http://www.csie.ntu.edu.tw/~cjlin/libsvm/ (10.10.2013)

Authors' addresses

Silvana Petruseva, PhD, Associate Professor
Faculty of Civil Engineering,
"Ss. Cyril and Methodius" University,
Blv. Partizanski Odredi, 24,
1000 Skopje, Republic of Macedonia
silvana@gf.ukim.edu.mk

Valentina Zileska - Pancovska, PhD, Full Professor
Faculty of Civil Engineering,
"Ss. Cyril and Methodius" University,
Blv. Partizanski Odredi, 24,
1000 Skopje, Republic of Macedonia
valentinazp@gf.ukim.edu.mk

Vahida Žujo, PhD, Associate Professor
Faculty of Civil Engineering,
University Dzemal Bijedic,
Univerzitetski Kampus,
88104 Mostar, Bosnia and Herzegovina
vahida.zujo@unmo.ba

Aida Brkan-Vejzović, PhD, Assistant Professor
Faculty of Economics,
University Dzemal Bijedic,
Univerzitetski Kampus,
88104 Mostar, Bosnia and Herzegovina
Aida.Brkan@unmo.ba