Time-Cost Relationship for Predicting Construction Duration

Agata Czarnigowska[1] and Anna Sobotka[2]

Abstract

Estimates of the construction time and cost are of key importance in the early phases of the project – they serve as a basis for the decision whether to commence with planning or not, and are used as input for budgets and programmes. The project's success depends on reliability of these estimates. It is thus crucial to answer the question: what do the project time and cost depend on? or easier to answer: correlated with? The answer can be based only on experience – personal as well as recorded in databases or mathematical models. Tools facilitating construction project planning on the basis of past experience are the object of research for many years. The paper discusses the time-cost relationship proposed by Bromilow in nineteen-sixties and adopted in many later multifactor models of construction duration. On the basis of a sample of 100 Polish public road projects, the statistical validity of the Bromilow's model was confirmed. The model was compared with multifactor models based on statistical regression and regression trees. Applicability of the models was discussed with respect to their errors and confidence intervals.

Keywords: time-cost relationship, forecasting, regression trees.

Introduction

The existence of a relationship between construction time and cost has been considered obvious: the time-cost-performance triangle appears in practically all project management handbooks (e.g. Kerzner, 1984). However, it would be interesting to find out if this relationship could be described quantitatively, and if so, if it might find any practical application. Cost is a generalized measure of any project's scale and complexity. Assuming that a reliable estimate of the project cost is possible to be made at early stages of planning, the cost may be considered known at the moment when project duration is to be decided. This rather optimistic assumption was the foundation for numerous attempts aimed at establishing a time-cost model that could be used for predicting project duration on the basis of project cost.

Statistical time-cost models

Literature review

Application of models using systematically recorded experience to planning and managing new project has been an object of interest of many researchers (Lai et al., 2008; Lee et al., 2008; Kaplinski, 1997). Among the models considered, regression-based ones are reported

[1] Assistant Professor, Faculty of Civil Engineering and Architecture, Lublin University of Technology, 20-618 Lublin, Nadbystrzycka 40, Poland, Tel.: +48 815384441, fax: +48 815384648; e-mail: a.czarnigowska@pollub.pl

[2] Professor, Department of Geomechanics, Civil Engineering and Geotechnics, AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow, Poland, Tel: +48 126174767, Fax: +48 126172059, e-mail: sobotka@agh.edu.pl

to provide a useful tool in cost planning (e.g. Chou and Tseng, 2011). Though project cost databases tend to be more popular than those of project schedules, there have been a number of attempts to construct models of project duration based on historical data.

The first time-cost regression model of construction projects is attributed to Australian researchers who, having analyzed cost and duration of a sample of construction projects completed during late 1960ies, proposed the following model, later referred to as the Bromilow's time-cost model (Kaka and Price, 1991):

$$L = K \cdot C^{B}, \tag{1}$$

or its equivalent:

$$\ln L = \ln K + B \cdot \ln C, \tag{2}$$

where L is the number of working days from the contractor's possession of the building site to the completion of works; C – actual value of works as paid by the client, expressed in A$ million; K and B – constants.

Bromilow's findings were checked by other researchers on the basis of new samples (Kaka and Price, 1991; Chan, 2001; Yousef and Baccarini, 2001; Ogunsemi and Jagboro, 2006). The form of the time-cost function (1) was confirmed to match sample data better than other function types tried, though determination coefficients obtained by the authors were not high (for large samples of non-uniform projects below 0.75). Large yearly fluctuations of the constants B and K were reported, though without any particular trend (Skitmore and Ng, 2001)

Statistical significance of the time-cost relationship gave rise to numerous attempts to create a multifactor regression model of construction duration that would incorporate project qualities other than cost and provide a better fit than the Bromilow's model. Table 1 compares selected models presented in the literature, where L stands for construction duration expressed in days, and b_i represent constants.

With few exceptions (Skitmore and Ng, 2003, Love et al., 2005, Stoy et al., 2007), cost was usually considered the most important independent variable present in multifactor models. Generally, there is no agreement on what factors should be the basis for estimating the duration. Despite the fact that the models presented in the literature were claimed to be statistically correct and significant, the authors often came to contradictory conclusions: some found that e.g. the client sector (public/private), building function or size strongly affected construction duration, others excluded them as insignificant. The initial selection of factors considered was also a matter of assumption, as the models were not always aimed at duration predictions – some were by-products of search for factors correlated with duration, some were used for measuring time performance. Thus, there were researchers who focused on management factors, whereas other preferred more "tangible" qualities, either known well ahead of commencement with works, or possible to be determined only after the project was finished.

The log-log relationship between time and cost in these multifactor models was widely argued to provide best fit, though some different functions were also proposed (Stoy et al., 2007; BCIS, 2004a; Martin et al., 2006; BCIS 2009). The authors were rarely specific about the quality measures of their models. The prediction and confidence intervals for the estimates can be found only in Stoy et al. (2007), BCIS (2004a) and BCIS (2009). Naturally, the larger and more diversified the samples, the greater errors were observed.

Most researchers analysed projects related with construction of buildings, so there are only a few works devoted to civil engineering projects. Kaka and Price (2001) analysed 140 UK road projects and found that the form of contract (fixed price vs. adjusted price) affects strongly the Bromilow's time-cost model parameters. Yousef and Baccarini (2001) conducted similar work on the basis of 46 sewerage projects in Australia, but did not considered factors other than costs. Irfan et al. (2011), disposing of large samples, focused on highway projects and created separate regression models for different project types (maintenance, resurfacing, construction, bridge construction, traffic infrastructure) that used planned cost and contract type as predictors of duration.

Table 1. Overview of multifactor models (selected works),
L – construction duration, b_i – constants

Autor, sample	Factors found significant	Regression function, quality measures
Walker, 1995 33 non-residential buildings, new-built	– cost (C) – ratio of time extension (x_1) – scope (fit-out/other) (x_2) – level of quality requirements (x_3) – management style (x_4) – design and construction teams communication quality (x_5) – efficiency of IT use (x_6)	$\log L = \log C + b_1 x_1 + b_2 x_2 + ... + b_6 x_6$ determination coefficient R^2=0.9987 and percentage errors of estimate (for particular observations, not summarised)
Chan and Kumaraswamy, 1999, 56 blocks of flats of the same standardised design system, the same public client,	different set of factors describe as-planned and actual duration; for actual duration L: – cost (C) – project type (flats for sale / rent) (x_1) – facade type (prefab or other) (x_2) – volume of the building (x_3) – gross floor area (x_4) – number of storeys (x_5)	$\ln L = b_0 + b_1 \ln C - x_1 + x_2 + b_3 \cdot x_3 + - b_4 \frac{x_4}{x_5}$ percentage errors of the estimate of L for each observation, maximum errors of +/-7%, the model's *MAPE*=2,51%
Skitmore i Ng, 2003 Australia, 93 buildings, new	– contractual time (CT) – contract type (lump sum / other) (x_1) – procurement method (x_2) cost excluded from the analysis	$\log L = b_0 + b_1 \log CL + x_2 + x_3$ Adjusted determination coefficient R^2=0.938
Love et al., 2005 Australia. 126 buildings, new or refurbished	– usable floor area (x_1) – number of storeys (x_2) cost excluded from the analysis, as cost not known until the project is finished	$\log L = b_0 + b_1 \log x_1 + b_2 \log x_2$ Adjusted R^2=0.96 *MAPE*=50%
Stoy et al. ,2007, Germany, 200 buildings from BKI database, 16 buildings for verification	– gross floor area (x_1) – number of winters (x_2) – project planning time in months (x_3) dependent variable is logarithm of construction speed, cost excluded	$\ln \frac{x_1}{L} = b_0 + b_1 \ln x_1 - b_2 x_2 - b_3 \ln x_3$ Adjusted R^2=0.915, *MAPE*=20%, Errors of estimate for test sample projects range (-29%;9%)
Hoffman et al., 2007, USA, 616 military buildings, new or refurbished	– cost (C) – client type (x_2-x_4) – management (own / contr. out) (x_5) – design type (own / contr. out) (x_6)	$\ln L = b_0 + b_1 \ln C - b_2 x_2 - b_3 x_3 + - b_4 x_4 - b_5 x_5 - b_6 x_6$ Adjusted R^2=0.374
BCIS, 2004a UK, 1500 new buildings, from KPI database	– cost (C) – procurement system (x_3) – contractor selection (x_4) – client type (x_5) – function (x_6) – region (x_7)	$\sqrt{L} = b_0 - b_1 \log C + b_2 \log^2 C + b_3 x_3 + + b_4 x_4 + b_5 x_5 + b_6 x_6 + b_7 x_7$
BCIS, 2009, UK, 4500 buildings, new or refurbished, BCIS database	factors as above Separate regression functions for new-builts and refurbishments Winter period does not affect duration significantly	equation as above the "calculator" provides prediction and confidence intervals for the estimates; separate equations for new-built and refurbishment projects

Discussion on the literature review

Most authors claim that it is possible to apply regression models to estimating construction time on the basis of cost, so they assume that it is easier to estimate construction cost than time, and that cost estimates are accurate enough to provide a basis for time estimates. There is an abundance of evidence on discrepancies between early budgets and costs at completion. Case studies focus on most striking examples (Potts, 2005; Polonski, 2006; Magnussen and Olsson, 2006), but there are statistical overviews of the scale and frequency of cost miscalculations: quite alarming by Flyvbjerg et al (2002), and a number of less pessimistic (Ng et al. 2001; KPI UK, 2003; BCIS, 2004b; KPI New Zealand, 2005).

Another issue is the reliability of the winning-bid price as a measure of the project scope and scale. There is evidence that contractors' bids are sensitive to intensity of competition, subjective risk perception and even season of the year when a call for tenders is announced. This can be observed in bid spreads in public procurement procedures. In Poland, they are expressed by an average bid dispersion factor W_z (Borowicz, 2005):

$$W_z = \frac{1}{n}\sum_{i=1}^{n}\frac{C_{i\max}}{C_{i\min}}, \tag{3}$$

where n is the number of tender procedures investigated, and C_{imax} , C_{imin} are, respectively, the highest and the lowest bid in each procedure. For instance, for the years 2000-2007, the average bid dispersion factors of public projects in Poland were from 1.23 to 1.43, and bid dispersion in particular cases reached even 250% (Borowicz, 2005 and 2008). Thus, the relationship between the contract price and actual value of works may be rather loose.

Under these circumstances, there are reasons to question both the "as-planned" and "contractual" cost in the role of independent variable for planning construction time. Moreover, it occurs that predictability of cost is generally no better than predictability of time (Martin et al., 2006; KPI UK, 2003). This provides more arguments against using cost as a predictor of time. However, before it is rejected, it would be interesting to check the model's sensitivity to cost miscalculation. In general, values of the constant B in the models presented in the literature range from 0.2 to 0.5, and the smaller B, the smaller the effect of cost on the value of the time estimate (see formula 1). The time-cost models are thus not very sensitive to cost estimate errors.

Investigation of time-cost relationship of Polish road projects

Research methodology

The aim of research was to create a model of road project construction duration based on relationships among the project qualities. These relationships were to be determined while analysing project qualities that were likely to be known or estimated at early planning stage, without considering details on organisation of works or construction methods. The model was (potentially) to find application in assuming construction durations at the stage of feasibility checks. For the purpose of this study, a project was defined as a scope of works contracted in one public procurement procedure and covered by one contract, supplemented with change orders and contract annexes, if applicable.

Stages of research were following: literature review, interviews with the construction clients – to determine their approach to estimating construction durations and budgets at the planning stage (outside the scope of this paper), data collection (by analysing project records – no ready-made databases were available), preselection of project qualities correlated with duration by means of the regression tree, construction of a multifactor regression model, and finally comparing the models obtained underway:

1. a simple regression (least squares method) between functions of actual duration (L) and actual cost (C): $f(L) = b_0 + b_1 f(C)$, where b_0 and b_1 are model parameters,
2. a *CART* regression tree based on 25 project qualities recorded at data collection stage and likely to be known at early stage of project planning;
3. multiple linear regression model (least squares method) relating a function of actual duration (L) and functions of predictors selected from the same project qualities as used to construct Model 2, $f(L) = b_0 + b_1 f(x_1) + b_2 f(x_2) + ... + b_n f(x_n)$, where $b_0 - b_n$ are model parameters, and $x_1 - x_n$ are predictors.

Calculations were conducted by means of *Statistica* 8.0.

The sample

The sample comprised 100 public road projects, completed between 2003 and 2008 in three neighbouring regions in south-eastern Poland. The projects considered differed in scope and type (Figure 1), and their cost ("as planned", including VAT) ranged from PLN 800 thousand to PLN 500 Million. The sample was considered representative of road projects from the analyzed period and location, and its size was at least 15% of the size of the population (imprecise due to non-uniform reporting methods used by the public clients).

One of the early assumptions of the research was analyzing similar projects, as new-build circular roads. As occurred during the data collection process, the number of such projects was too small to be used for statistical analyses, and the majority of works contracted in the analyzed period consisted in modernization of the existing infrastructure.

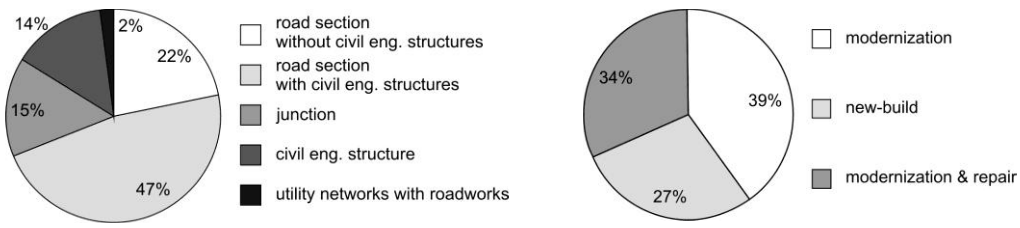

Figure 1. Sample structure according to object of works and project type

Therefore, projects varying in scope and type were included in the sample. Their similarity lay in overall conditions: the clients were from public sector, acting under similar budgetary constraints, the regions were similar in terms of natural and economic environment and level of infrastructure development, the works were contracted according to the public procurement law, the only criterion of contractor selection was the lowest price, and contract duration was enforced by the client. A diversified sample implies that the model derived from the data would be a far going generalization.

Simple linear regression model

Analysis of scatter diagrams (Figure 2) and experiments with several functions confirmed that, among simple regression models considered, Bromilow's model (Formula 2) provides best fit for the analysed sample, and is statistically correct. The model (Formula 7) is significant (F-test) and of significant parameters (t-tests), residuals are normally distributed, with constant variance and expected value of 0. Normality was checked by analysing residual histograms, scatter diagrams, and by normality tests: Kolmogorov-Smirnov/Lilliefors' and Shapiro-Wilk's. Homoscedasticity of residuals was checked by analyzing residual scatter diagrams, and by Lagrange test (Stanisz 2007). The Bromilow's model for the sample (Model 1) is described by the following equation:

$$\ln L = 1.2067 + 0.4649 \ln C \tag{4}$$

the model's adjusted determination coefficient $\overline{R}^2 = 0{,}636$, and standard error *SEE=0.504*. The mean absolute percentage error *MAPE*=44,84% is comparable with the scale of errors of time-cost models using logarithm transformation of duration, presented in the literature.

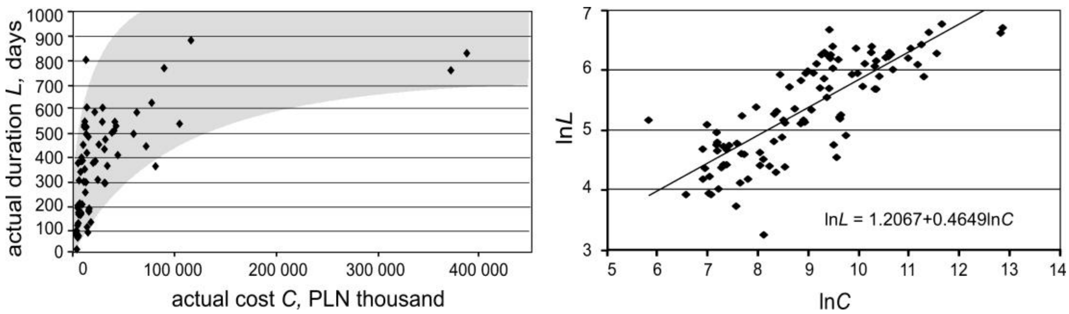

Fig. 2. Scatter diagram of actual construction duration *L* against actual cost *C*, and scatter diagram of log-log values with regression line

Regression tree

While collecting input, data on 25 project qualities were collected. Those qualities were considered likely to be known at early stages of project planning, and were of various types: categorical and quantitative, related with geometric parameters of the road, scope of works, road class, location, number of bridges and many more, and are listed in Figure 4. All these potential predictors were used to construct *CART* models.

The method consists in recursive division of the set of observations into subsets (two subsets at a time), according to one quality at a time, to obtain the greatest possible reduction of heterogenity of observations in the subset (Gatnar, 2001). Here, the heterogeneity was measured by the variance of durations of projects in the subset. The best tree was selected according to Breiman's procedure (Gatnar, 2001).

The best-fit model (Model 2), presented in Figure 3, uses seven predictors: assumed number of winters during construction, construction cost, number of culverts along the route, client type (either national or regional road office), total length of civil engineering structures, number of intersections, number of parking/bus bays. If applied in practice, it would assign a project one of ten durations: 91, 161, 166, 291, 396, 487, 490, 573, and 810 days. Some of them differ by only a few days.

The model's adjusted determination coefficient $\overline{R}^2 = 0{,}924$ (Gatnar, 2001) indicates that the model is well fitted to the sample. Considering the relatively small number of observations used to create the model, and their being diverse, this is not automatically an advantage when it comes to using the model for predictions.

There may be doubts about using the number of winters as a predictor of construction duration, as hard to estimate as the duration itself. However, interviews with the client's representatives indicated that the clients decided to fit a project in a certain number of years at the beginning of the project planning process, which arises from budgetary constraints and long-term planning of public organisations. Therefore, the number of winters can be considered defined in advance. Selection of this variable was prompted also by other research (Stoy et al 2007; BCIS 2004a) – where it was to allow for seasonal changes in speed of works.

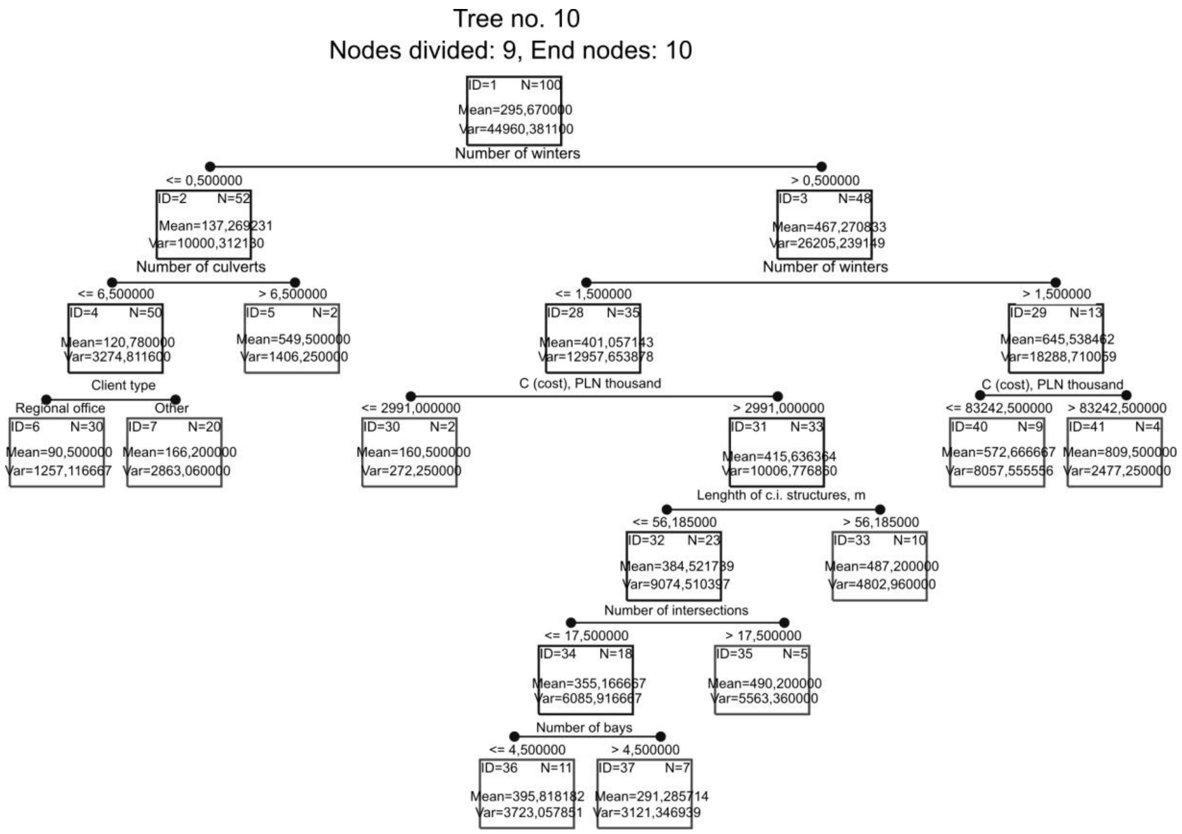

Figure 3. The best regression tree – Model 2

Multiple linear regression model

It was assumed that a linear multiple regression model would be looked for, and its parameters were to be determined by the least squares method. With only 100 cases in the sample, using stepwise regression to select most suitable of 25 potential predictors (or actually over 40, as categorical variables were converted into binary variables) was considered inefficient. However, while constructing regression trees, one can identify variables that are strongly correlated with the predicted variable. These are not necessarily present in the best regression tree (Gatnar, 2001). Figure 4 presents the relative importance of potential predictors, determined in the procedure of constructing regression trees. For further investigations, nine potential predictors were selected arbitrarily: six of the "most important" defined in *CART* analysis (Figure 4), and additionally those present in the best regression tree. These nine factors were then used for constructing regression models by means of stepwise regression (forward selection and backward elimination):

- construction cost,
- number of winters,
- length of civil engineering structures in the scope of a project,
- number of civil engineering structures,
- total length of roads covered by the project,
- number of culverts along the route,
- number of bays,
- number of intersections,
- client type (either regional or national road agency)

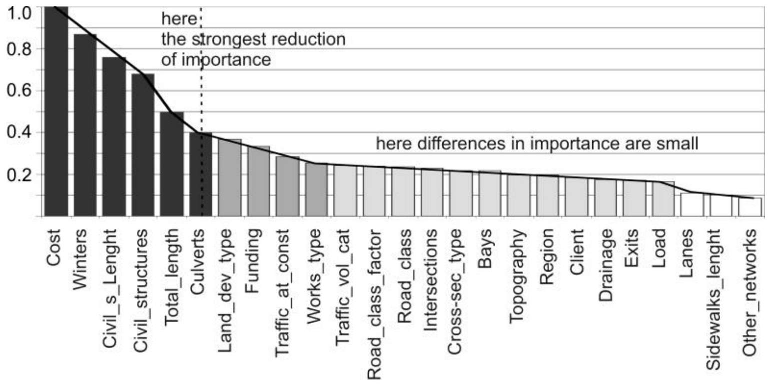

Figure 4. Relative importance of variables defined while constructing regression trees

Several models were tried, differing in transformations of variables. The best fitted was Model 3, with four predictors: cost (*C*), civil engineering structures length (C*ivil_s_length*), number of winters (*Winters*), and number of civil engineering structures (*Civil_s*):

$$\sqrt{L} = -4.43 + 1.89 \ln C + 0.56 \ln Civil\_s\_length + 4.44 Winters - 0.28 Civil\_s\,, \tag{5}$$

The model fulfils the assumptions of the least squares method, its adjusted determination coefficient is $\overline{R}^2 = 0.867$, mean absolute percentage error *MAPE*=13,17%, and standard error *SEE*=2.28. Considering parameters of Equation 8, one can observe that the estimate is very sensitive to the number of winters that can be hard to asses.

Quality of the models

The models use different transformation of predicted value (ln*L* in the case of simple regression, *L* for regression tree, and $\sqrt{L}$ for multifactor regression), so the statistics of adjusted determination coefficient $\overline{R}^2$, standard error *SEE*, or mean absolute percentage error *MAPE* cannot be directly compared. As the models are meant to be used for predicting duration expressed in days (*L*), errors expressed in days were calculated (L_i is the observed duration, and $\hat{L}_i$ – duration calculated on the basis of the model):

$$e_i^{days} = L_i - \hat{L}_i\,, \tag{6}$$

Analysing them, one can see the scale of dispersion between expected vs. observed – for the set of observations used to build the models. Values of these errors, and the mean absolute percentage error, $MAPE^{days}$ (Figure 10):

$$MAPE^{days} = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{e_i^{days}}{L_i} \right|. \tag{7}$$

are directly comparable, though not normally distributed. Figure 5 compares the scale of errors in days for all models considered. For practical applications, the best model would be the one of lowest dispersion. In this case, it is the regression tree. Its $MAPE^{days}$ is 23%. Model 1 has $MAPE^{days}$ of 45%, and Model 3 – 28%.

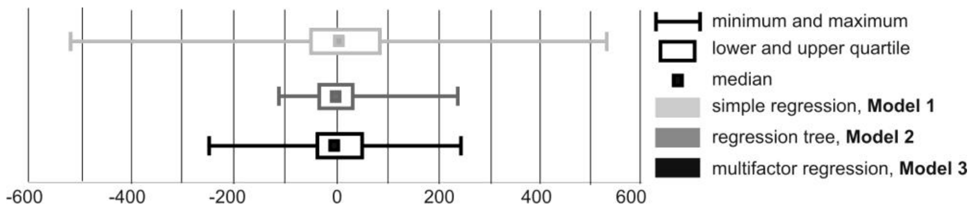

Figure 5. Comparison of model errors expressed in days

To compare the models' predictive ability, duration estimates (in days) were calculated for seven projects not included in the initial sample. Their qualities stayed within the ranges covered by the models. Figure 6 shows predicted vs. observed values of project durations. Again, for practical applications, the best model would be the one of lowest error. The test sample is small, which affects reliability of the conclusions to follow.

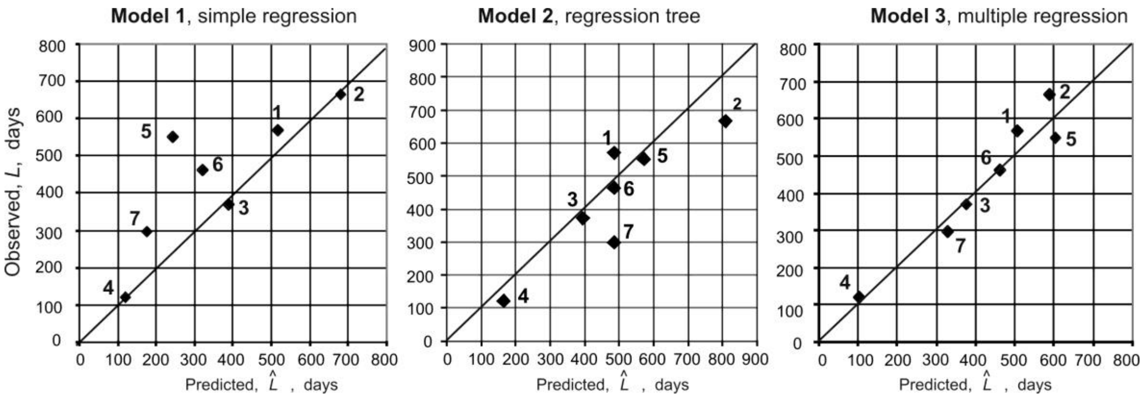

Figure 6. Predicted against observed durations (days) of test sample

In the case of these particular projects, now it is Model 3 that seems to provide most precise predictions, as the observed values are quite close to predicted values. To express it in numbers, one can calculate mean absolute percentage errors in days for the test sample: the "best-looking" prediction model (Model 3) has $MAPE_{test}^{days} = 9\%$, the second-best is Model 1 (Bromilow's) with $MAPE_{test}^{days}$=21%. The regression tree (Model 2) provides the least accurate estimates with $MAPE_{test}^{days}$=22%. This is due to the fact that it is too well adjusted to the initial sample, and the test sample simply does not follow the same pattern.

The quality of predictions based on regression functions cannot be judged without prediction and confidence limits for the estimates of durations. However, in the case of non-parametric Model 2, there are no grounds to calculate prediction and confidence intervals in a way that could be compared with Models 1 and 3.

Figure 7 presents these data for parametric models (expressed in months for better readability), assuming 95% confidence level.

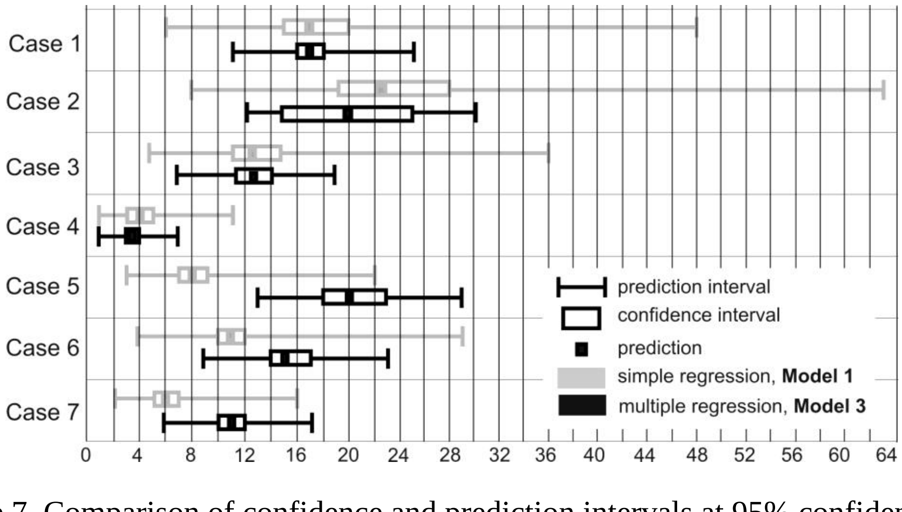

Figure 7. Comparison of confidence and prediction intervals at 95% confidence for Model 1 and Model 3, test sample projects, durations expressed in months

Model 1 seems too inaccurate to find any practical application. This may be illustrated by Case 3 from test sample: the user can be 95% sure that the expected duration for project of such cost is between 11 and 15 months (confidence interval for regression), but can be also 95% sure that a particular project of such cost may take between 5 and 36 months (prediction interval).

The multiple regression Model 3 is certainly more accurate, but confidence and prediction intervals are still quite broad: for Case 3, the confidence limits for the regression are 11 and 14 months, and prediction limits – 7 and 19 months.

Summary and conclusions

The sample considered in the paper was small and diversified. However, some statistically significant relationships between construction duration and other project qualities have been found.

Calculations confirmed the universal character of the Bromilow's time-cost model. Certainly, the model has some advantages: it is simple and, at least for the considered sample, statistically correct. However, its errors are high (mean absolute percentage error in days is 45%) , and the prediction and confidence intervals impractically broad.

Using a non-parametric method of regression trees, 25 qualities of the analysed project were checked with regard to their relationship with construction duration. Only four of them (the most important according to the non-parametric analysis) stayed in the final multifactor regression model: construction cost, number of winters within the construction period, total length of civil engineering structures (as bridges) in the project, and the number of these civil engineering structures. The predictors are different than these presented in the literature – this is of course specific to the type of projects analysed (the literature focuses mostly ob buildings, not roads) and initial assumptions on what factors to consider.

What is interesting, the form of a multifactor regression equation most frequent in the literature,

$$\ln L = b_0 + b_1 \ln C + b_i \ln x_i + b_j x_j \,, \tag{8}$$

where L is duration, C is cost, x_i is a continuous variable, x_j represents a discreet variable, and b_i – parameters, did not provide the best fit for the analysed sample. The following equation proved more appropriate:

$$\sqrt{L} = b_0 + b_1 \ln C + b_i \ln x_i + b_j x_j \,. \tag{9}$$

It is statistically correct, immune to outliers, of lower errors and narrower predictions and confidence intervals, and also not very sensitive to errors of the predictors' estimates, with the exception of the number of winters.

A non-parametric model of regression trees (*CART* type) also provides a good fit, though predictions less accurate, than the classic regression models. Interestingly, it uses a different set of predictors than the multifactor regression model: instead of number of civil engineering structures, there appeared: client type, number of culverts, number of intersections and number of bays. However, with the sample being small (100 used to construct the model and 7 to validate it) and diversified, such models are not reliable.

Time-cost regression models for repeatable projects (e.g. buildings of the same function, structure type, similar layout and location) could be more precise. Similarly, if more independent variables were considered, and samples were larger, better models could be provided. A number of researchers report their achievement in this field (Irfan et al. 2011) and there exists at least one commercial regression-based duration "calculator" (BCIS 2009). This may serve as evidence of practical applicability of parametric models in planning construction duration. Such models have some advantage over other models based on experience, such as "black box" expert systems, or neural networks – they are portable: regression models are expressed as equations, and to use them, one does not need to dispose of the whole database or software. Moreover, the reasoning process behind the model is quite obvious. This may be the reason why, in the time of quick development of artificial intelligence methods, statistical analyses do not loose on popularity.

Further research in the field may include: investigation on other factors, constructing other model types (here, specification of regression functions was based on results presented in the literature and scatter diagram analyses, and the simplest approach of least squares method was used), and applying artificial intelligence tools create better models. This however requires expanding the database.

Acknowledgments

The authors would like to thank Mr Joe Martin, Executive Director, BCIS, for help and information on the Guide to Building Construction Duration, and the General Directorate for National Roads and Motorways, Divisions in Lublin, Krakow and Rzeszow, the Lublin, Krakow and Rzeszow Directorates of Regional Roads for providing access to project documentation.

References

BCIS, 2004a. *Guide to Building Construction Duration*, BCIS, London.

BCIS, 2004b. *Time and Cost Predictability of Construction Projects*. *Analysis of UK Performance.* BCIS, London.

BCIS, 2009. *Construction Duration Calculator*, www.bcis.co.uk/ construction_duration_ calculator, accessed in April 2009

Borowicz A., 2005. Public procurement of construction works (Statistical analysis). *Konferencja naukowo-techniczna Budownictwo Polskie w rok po wstapieniu do Unii Europejskiej*, Gdansk 9-11 czerwca, 41-47 (in Polish)

Borowicz A., 2008. Tendencje i uchybienia w kalkulacji cen zamowien publicznych na roboty budowlane. *Materiały XIV Konferencji Naukowo-Technicznej*, Ciechocinek, 15-17 października, 9-29 (in Polish)

Chan A.P.C., 2001. Time-cost relationship of public sector projects in Malaysia. *International Journal of Project Management*, 19, 223-229.

Chan D.W.M., Kumaraswamy M.M., 1999. Modelling and predicting construction durations in Hong Kong public housing. Construction Management and Economics, 17, 351-362.

Chou, J.-S., Tseng, H.-S., 2011. Establishing expert system for prediction based on the project-oriented data warehouse. *Expert Systems with Applications*, 38, 640-651.

Flyvbjerg B., Holm M.S., Buhl S., 2002. Underestimating Cost in Public Works Projects: Error or Lie? *Journal of the American Planning Association*, Vol. 68 No. 3, 279-295.

Gatnar E., 2001. Nieparametryczna metoda dyskryminacji i regresji. PWN, Warszawa.

Hoffman G.J., Thal A.E., Webb T.S., Weir J.D., 2007. Estimating Performance Time for Construction Projects. *Journal of Management in Engineering*, ASCE, Vol. 23 No 4., 193-199.

Irfan, M., Khursid, M.B., Anastasopoulos, P., Labi, S., Moavenzadeh, F., 2011. Planning-stage estimation of highway project duration on the basis of anticipated project cost, project type, and contract type. *International Journal of Project Management*, Vol. 29 (1), pp.78-92.

Kaplinski O., 1997. *Modelling of Construction Processes: A Managerial Approach*. KILiW PAN, IPPT, Studia z Zakresu Inzynierii No. 43, Warsaw.

Kerzner H., 1984. *Project management. A systems approach to planning, scheduling and controlling*. Van Nostrand Reinhold, New York.
KPI UK, 2003. All *Construction Industry Key Performance Indicators Wallchart 2003*. Construction Best Practice, Watford
KPI New Zealand, 2005. *The New Zealand Construction Industry National Key Performance Indicators Wallchart 2005 Data*. Centre for Advanced Engineering Christchurch, New Zealand.
Kaka A., Price A.D.F., 1991. Relationship between value and duration of construction projects. *Construction Management and Economics*, 9, 383-400.
Lai Y-T., Wang W-C., Wang H-H., 2008. AHP- and simulation-based budget determination procedure for public building construction projects. Automation in Construction, 17, 623-632.
Lee J-R, Hsueh S-L, Tseng H-P., 2008. Utilizing data mining to discover knowledge in construction enterprise performance records. *Journal of Civil Engineering and Management*, 14(2), 79-84.
Love P. E. D., Tse R. Y. C., Edwards D.J., 2005. Time–Cost Relationships in Australian Building Construction Projects. *Journal of Construction Engineering and Management*, ASCE, 131(2), 187-194.
Magnussen O.M., Olsson N.O.E, 2006. Comparative analysis of cost estimates of major public investment projects. *International Journal of Project Management* 24, 281–288
Martin J., Burrows T.K., Pegg I., 2006. Predicting Construction Duration of Building Projects. *Shaping the Change XXIII FIG Congress,* Munich, Germany, October 8-13.
Ng S.T., Mak M.M.Y., Skitmore R.M., Lam K.C., Varnam M., 2001. The predictive ability of Bromilow's time-cost model. *Construction Management and Economics*, 19: 165-173.
Ogunsemi D.R., Jagboro G.O., 2006. Time-cost model for building projects in Nigeria. *Construction Management and Economics*, 24: 253-258.
Polonski M., 2006. The analysis of the reliability of realization costs and investment's time limits in Warsaw. *Electronic Journal of Polish Agricultural Universities, Topic Civil Engineering* 9 (4) www.ejpau.media.pl/volume9/issue4/art-10.html (accessed in May 2010).
Potts K., 2005. The New Scottish Parliament Building – A Critical Examination Of The Lessons To Be Learned. *Third International Conference on Construction in the 21st Century (CITC-III) Advancing Engineering, Management and Technology*, 15-17 September, Athens, 534-539.
Skitmore R.M., Ng S.T., 2001. Note. Australian project time-cost analysis: statistical analysis of intertemporal trends. *Construction Management and Economics*, 19, 455-458.
Skitmore R.M., Ng S.T., 2003. Forecast models of actual construction duration and cost. *Building and Environment*, 38(8), 1075-1083.
Stanisz A., 2007. *Przystepny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny*. Tom 2. Modele liniowe i nieliniowe. Statsoft Polska, Krakow.
Stoy Ch., Polalis S., 2007. Early estimation of building construction speed in Germany. *International Journal of Project Management*, 25, 283-289.
Walker D.H.T., 1995. An investigation into construction time performance. *Construction Management and Economics*, 13, 263-274.
Yousef G., Baccarini D., 2001. Developing a cost-time model for estimating construction durations of sewerage projects in Perth, Western Australia. *Journal of Construction Research*, 2, 213-220.