Multi-Attribute Regression Analysis for Concrete Pavement Productivity Estimation

Antonis Panas

Centre for Construction Innovation, Department of Construction Engineering and Management, Faculty of Civil Engineering, National Technical University of Athens, Greece
cvapanas@mail.ntua.gr

John-Paris Pantouvakis

Centre for Construction Innovation, Department of Construction Engineering and Management, Faculty of Civil Engineering, National Technical University of Athens, Greece
jpp@central.ntua.gr

THE ESTIMATION OF CONCRETE PAVEMENT PRODUCTIVITY IS INTRICATE BECAUSE OF SEVERAL FACTORS, SUCH AS THE AVAILABLE WORKING WIDTH AND LENGTH, THE CONCRETE LAYER THICKNESS, THE CONSTRUCTION METHODOLOGY, AUTOMATION CAPABILITIES AND THE WORKING CONDITIONS. This study intends to assess the productivity of the concrete pavement operations by taking into consideration the effect of the above factors using the regression analysis technique. In regression models, one dependent variable (productivity) is considered against several independent variables (width, length, working conditions, health & safety level). Direct observation, site visits and video recording of actual concrete pavement activities on a construction site over a period of four months has been used for data collection. The regression relationships are plotted, so as to derive empirical nomographs and correction coefficients which can be used to adjust actual productivity against theoretical baselines. In this sense, process-oriented estimation methodologies for specific construction operations can be developed, by which each factor's contribution to productivity can be estimated. The results indicate that a larger working area increases productivity for a given concrete layer thickness. The estimated correction factors fit the actual measurements in a statistically significant fashion, however their predictive capability is expected to improve as the study sample increases.

INTRODUCTION

Concrete pavement layering is a process that requires the combination of labor- and machine-intensive operations. As such, the implications and the interactions of the affecting factors pose a significant research interest, given that published productivity or cost data are scarce and based mainly on road-building research efforts (Yiang and Wu, 2007). This paper aims to (i) identify the relationship of concrete paving productivity for specific operational factors, (ii) explore the variation of productivity in relation to the working efficiency and the Health and Safety (H&S) status, (iii) create productivity estimation models and (iv) generate new estimation

methodologies in the form of empirical relations or nomographs showing the sensitivity of achieved productivity in relation to production variables.

The structure of the paper is as follows: First, the concrete paving operations are described along with the main factors that affect their productivity. Subsequently, the research methodology is explained followed by the presentation of the analysis' results stemming from field measurements of actual paving operations. The main inferences emerging from the study are discussed and, finally, the main conclusions and future directions of research are delineated.

Background

Concrete pavement construction within the scope of this paper regards the layering of ready-mixed concrete for the construction of heavy duty surfaces in external areas, which are intended to be operated by large scale machinery (e.g. airports, harbours etc.) (Knutson et al., 2008). The activities that have to be executed for the completion of a given surface of concrete pavements are depicted in Figure 1, accompanied by a list of the required resources that have to be deployed. A short description of the illustrated tasks follows.
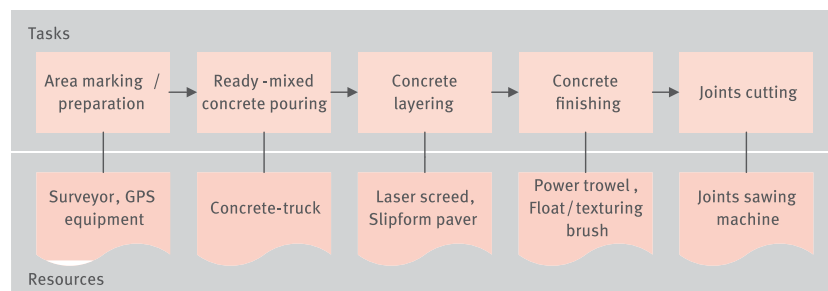
specialised equipment is utilized for concrete layering and screeding in its final position. There are three options available for concrete screeding: (a) using hand screeds, (b) using GPS-guided laser screeders and (c) utilising slipform paving equipment. After having reached the desired thickness in the pavement layer, power trowel ("helicopters") equipment is used for finishing the surface and a float/texturing brush is used in order to acquire a rough final surface with increased traction. The final task is the contraction and expansion joints cutting on a given grid (e.g. every 7-8m) which takes place after initial curing of the concrete. Each one of these tasks can be completed at a certain theoretical productivity level. However, it is evident that each project is different and, thus, deviations from theoretical values are expected, leading to the actually effective productivity achieved on site. The effective productivity of paving operations can be formalised by Equation 1 as shown below:

$$Q_{eff}^{P} = f\left(Q_{eff}^{L}, Q_{eff}^{F}, Q_{eff}^{C}\right) = f\left(Q_{th}^{L} \times \prod f_i, Q_{th}^{F} \times \prod f_i, Q_{th}^{C} \times \prod f_i\right)$$

where: $Q_{eff/h}^{P,L,F,C}$ = effective/theoretical productivity for paving (P) [m²/h], layering (L) [m²/h], finishing (F) [m²/h] and joints cutting (C) operations [m/h];

As shown in the above mentioned equation, productivity of concrete paving operations is affected by factors which may reflect environmental conditions (weather conditions, ambient temperature) (Yiang and Wu, 2007), structural features (pavement thickness, joint spacing, construction methods) (Vandenbossche et al., 2011), as well as managerial influences (management efficiency, health and safety status) (Nunally, 2007). However, pertinent research in the field has been concentrated mainly on asphalt paving operations (Miller et al., 2011), within urban areas (Lee et al., 2000) which poses challenges in terms of the lane-closing strategy and the interactions with the heavy road traffic (Hassan and Gruber, 2008). This paper adopts a different approach, by scrutinising micro-level factors which are believed to influence achieved productivity. Therefore, concrete paving operations are examined by taking into account the working width and length, the health and safety status and working efficiency. It is evident that each one

of the aforementioned factors holds a certain set of attributes, as shown in Table 1. The working width (w) holds a minimum value of 4m which represents the min working range of both the slipform paver and the laser screed. The working length (l) has no certain limitation and represents the maximum length of the lane that is worked by a crew on a given workday. The efficiency factor generally represents the ratio of the actual working time of a resource considering any operational delays to the ideal 60-min working hour (Panas and Pantouvakis, 2010b). In our case, the efficiency factor denotes the delays caused due to the abnormal shape of the working area in comparison to a pure rectangular shape. In other words, it is assumed that any deviation from the rectangular working surface causes more delays in the working cycle due to the need for

Figure 1: Concrete pavement construction tasks and resources.

First, a survey line is established in the working area to facilitate setting forms at proper grade and alignment. Consequently, ready-mixed concrete is poured from trucks, whereas

f_i = productivity factor of i variables for the adjustment of theoretical to effective productivity[-].

micro-movements of the equipment as well as the crews. The health and safety status can be evaluated by implementing both quantitative (Hallowell and Gambatese, 2009) as well as qualitative approaches (Mitropoulos et al., 2009). Within the framework of this research a qualitative approach has been adopted based on a job hazard analysis. For simplicity reasons, the main criterion that has been taken into account was the execution of concrete paving operations in the proximity or concurrently with other construction activities. The latter increases the H&S risk, since there is a danger of collision with other vehicles circulating on the site. In addition, paving operations take place adjacent to building works, which increases the risk of objects falling from a height.

site personnel as well as studying of project documentation (drawings, quantity take-offs, progress payment orders, labour hours logs). The quantitative and qualitative data have been used for the selection of the variables described in the previous chapter. In total, 24 data points have been collected representing on-site workday measurements. One data cluster has been specified for each variable as shown in Table 1. The limits of each cluster are denoted by the minimum and maximum values or the ordinal values for every quantitative and qualitative factor respectively. In this way, a valid experimental framework is established which will consequently help the categorization and the in-depth analysis of the data within the regression models (Panas and Pantouvakis, 2010a).

where: Y = dependent or response variable; $X_1, X_2, ..., X_k$ = explanatory, independent or predictor variables; b_0 = intercept; $b_1, b_2, ..., b_k$ = regression coefficients

For this study, multiple linear regression has been performed, in order to generate productivity regression models. The working width and length are the explanatory variables, whereas productivity is the dependent variable. Three types of models are created: (a) concrete layering, (b) concrete finishing and (c) joints cutting productivity models. A variety of metrics and statistical procedures has been applied in order to ensure the statistical validity and robustness of all the models and their parameters. At first, scatter diagrams have been prepared and checked visually for the detection of any possible outliers. Then, statistical checks based on the residual and standardised residual information in the regression model statistical outputs were performed. The p value of t-test for the coefficients of

Factors	Attributes (min, max)
Working width (w [m])	4,00m-32,00m
Working length (l [m])	27,00m-210,00m
Efficiency (n [min/min])	Qualitative (Adequate, Non-adequate)
Health and Safety (H&S) status (hs [-])	Qualitative (Acceptable, Non-acceptable)

Table 1: Paving operations factors and their attributes

On the basis of the variables described before, the collected data have been divided respectively according to their attributes in specific clusters which enables their computational processing. More details on the data management strategy and the research methodology are provided in the next section.

Methodology

The study methodology is depicted in Figure 2. Each phase of the research methodology is explained in the following paragraphs.

Data elicitation

Data were collected through work studies of actual paving operations within a four month period. Direct observation and video recording were used as primary data elicitation instruments. Secondary data were gathered by open interviews with senior project management staff, construction managers and
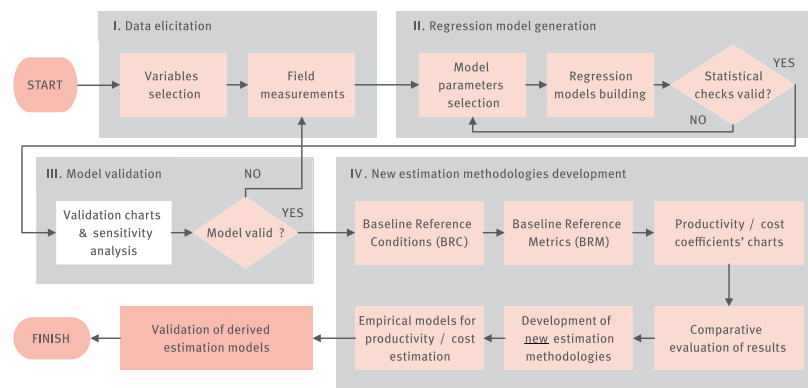
Figure 2: Research methodology.

Regression model generation

Regression models are based on a data-oriented technique, where the collected data are directly associated with each other, without considering the process behind this data. The general form of a multiple linear regression model is given in Equation 2 below:

$$Y = b_0 + b_1 \times X_1 + b_2 \times X_2 + ... + b_k \times X_k$$

the regression models has been calculated to evaluate their fit. Probabilities less than 0.05 are considered as significant evidence that the coefficients are not zero. The statistical features used to check the models' robustness as a whole were (a) R-square, (b) Adjusted R-square and (c) Root mean square er-

ror. The reader is recommended to consult Kutner et al. (2005) or any other regression references for the definition of each criterion. The JMP® version 9 software developed by SAS® has been used to build and validate all the models.

Model validation

The validation process was performed by comparing the outputs of the developed models to the actual collected data. For this reason the data sample has been divided in a model building (70%) and a validation set (30%) by random selection. At this point it should be highlighted that the variance of the estimated regression coefficients based on the model-building set is generally larger than the variance if the entire set was going to be used. Therefore, after model validation, the final regression model is calculated based on the entire set. In addition, irrespective of their significance, the models' results have to be validated and compared with real world data to ensure that the designed models are good enough for real construction practice use. Hence, the validation process includes the substitution of validation data inputs to the designed models so as to compare predicted results of productivity models to collected data. The results of the validation process are illustrated in validation charts which will be presented in the next sections.

New estimation methodologies development

This section serves the main objectives of the study, in terms of the research contribution. The first step is the definition of the "baseline reference conditions" (BRC), namely the operational conditions under which every operational coefficient can be neglected, as it is supposed that it does not affect productivity. On a theoretical basis, this means that when certain conditions are met, then $f_i=1,00, \forall i \subset N$ (see Equation 1), and, consequently, theoretical and effective productivities coincide. For further information the reader is referred to

Panas and Pantouvakis (2010b). When the BRC are established, the "Baseline Reference Metrics" (BRM) are defined, namely the initial productivity values against which actual measurements are going to be benchmarked. BRMs can be extracted from estimation handbooks or from a company's historical record. The variation of the theoretical BRMs to the actual data is visualised by the creation of charts which facilitate the comparative analysis of the studied operations and, ultimately, enable the formulation of new estimation methodologies, which are not computationally complicated, but rather simple and useful estimation tools. The latter can be achieved by estimating the values of the f_i, thus yielding empirical models in the form of $Q_{eff}=Q_{th} \times \prod f_i$. Finally, after having ensured that the produced regression models are validated statistically and in practice, they can be directly applied in the estimation process.

Analysis

This section presents the results of the analysis for all three types of productivity regression models which are built: (a) concrete layering, (b) finishing, (c) joints cutting. Productivity regression models have been developed for the last three tasks of Figure 1, (concrete layering, finishing, joints cutting). For brevity reasons, only the concrete layering regression model development process will be presented in detail, since the methodology for the formulation of the other two models is the same.

Concrete layering regression models

Working width and length are the primary variables selected for this model. The first step would be the scanning of the data for outlying values and their examination to see if they are valid observations. A scatterplot 3D graph has been plotted (Figure 3) which can detect trivariate outliers in three-dimensional space.
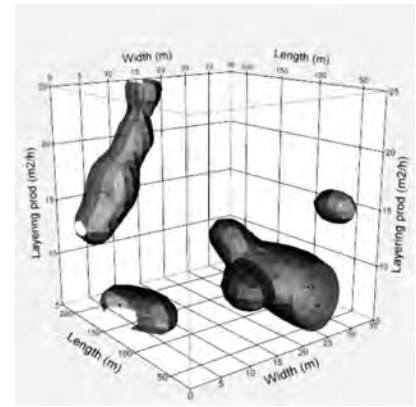

Figure 3: Scatterplot 3D graph for trivariate outliers detection.

Since all the data lie within the designated margins of Table 1, the next step is the multiple regression model generation, with working width and length being the predictor variables and layering productivity considered as the response variable. The results of the model fitting process are displayed in Figure 4 below. The first three graphs are called leverage plots because they convey the idea of the data points pulling on the lines representing the fitted model. Thus, the strength of the effect is shown by how strongly the line of fit is suspended away from the horizontal by the points. Since in all three graphs the 95% confidence curve crosses the horizontal reference line, the effect is considered to be significant. The summary of fit table shows that the model has an R-square value 0,87347, which represents the coefficient of multiple determination that measures the proportional reduction of total variation in productivity using working width and length as independent variables. In other words, it represents the total variability in productivity explained by working width and length. R-squares values of >0,80 imply that data vary little around the fitted models and, thus, are generally acceptable (Kutner et al., 2005). The analysis of variance table shows that width and length have an observed significance probability (Prob >F) of <0,0001, which is significant at the 0,05 level. The parameters estimates table shows the estimated regression

coefficients in the linear model and a t-test for the hypothesis that each parameter is zero. The Prob›|t| metric is 0,0076 for the intercept and ‹0,0001 for the working width and length respectively, which means that $b_0 \neq 0$, $b_1 \neq 0$ and $b_2 \neq 0$ at 99% confidence. Ultimately, the following productivity model is used (Equation 3):

$$Q_{eff}^{L} = -9,833424 + 0,6624123 \times \left[Width\right] + 0,0972577 \times \left[Length\right]$$

Figure 5a illustrates the plotted regression model as a function of working length, for different width values. The respective validation chart (Figure 5b) shows that the model values fit actual productivity data adequately (average accuracy ±15%). A sensitivity analysis has also been undertaken to investigate the model's variation to a change in the working efficiency level (n) as well as the Health and Safety status (hs) (see Table 1). Both factors have been modelled as binary values taking the value X=1,00 if efficiency is adequate and the H&S status is acceptable, and X=0,00 for non-efficient working conditions or non-adequate H&S status. Hence, theoretically the four combinations of the factors' values (n=1; n=0; hs=1; hs=0) may yield four different regression models. However, the application of the experimental framework described in the research methodology section concluded that valid and adequate data points were found only for the n=1 and hs=0 cases. The developed regression models as well as their statistical features are shown in Table 2.
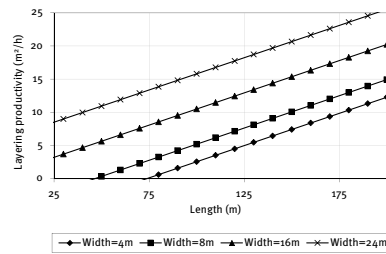

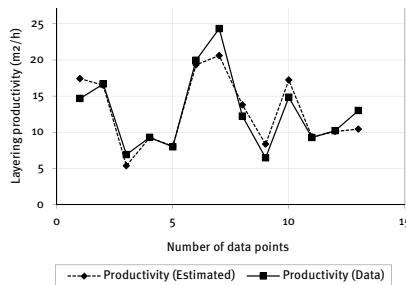**Figure 5a:** Concrete layering productivity estimation chart.

Figure 5b: Concrete layering productivity validation chart.

No.	Case	Concrete layering productivity model	Prob › F	R-square	Prob›|t| intercept	Prob›|t| width	Prob›|t| length
1	n=1	-14,95629 + 0,7800604x[Width] + 0,1155696x[Length]	0,0313	0,82	0,1261[a]	0,0269	0,0130
2	hs=0	-4,593776 + 0,4864072x[Width] + 0,0663698x[Length]	0,0013	0,93	0,0545	0,0005	0,0010

[a]This outcome indicates statistical invalidity for the intercept in these models, but the models are still valid because the p-value for the models as a whole is ‹0,05, even though there is some statistical inconsistency with the intercept.

Table 2: Sensitivity analysis regression models for working efficiency and H&S status

"GPS-guided laser screeder" technique. Therefore, for statistical validity reasons the analysis is conducted for this operation only and the data division is as follows (Table 3).

The BRC scenario is set according to the preferences of the estimator. In this case, the BRC scenario is defined as the one corresponding to a working width of 0-6m and working length of ›80m. This choice is justified by the fact that the maximum operating width of the laser screeder's blade is around 6m long whereas in most cases the pavement lane to be constructed had a length of over 80m. Therefore, the BRM for concrete layering is established as the average of data points 6, 9, 15 and 16, which is 9,27m²/h. Since two variables are taken into account (i=2), then two productivity factors are defined $f_1=f_{width}$ and

$f_2=f_{length}$. The lack of adequate amount of data points in the first row of Table 3 (no data points in the 0-55m cluster and only 1 data point in the 55-80 cluster) means that f_{length} cannot be calculated in a statistically valid fashion. This is not the case however for the f_{width} factor, whose estimation is shown in Equations 4a,b,c below.

The analysis denotes essentially that for a working length greater than 80m, effective productivity in the 9m-17m working width will be slightly more than double than the equivalent value for a 0m-6m width of operation. For example, if it is assumed that a concrete paving operation is going to take place with a working width of 9-17m, in a length of more than 80m then the effective
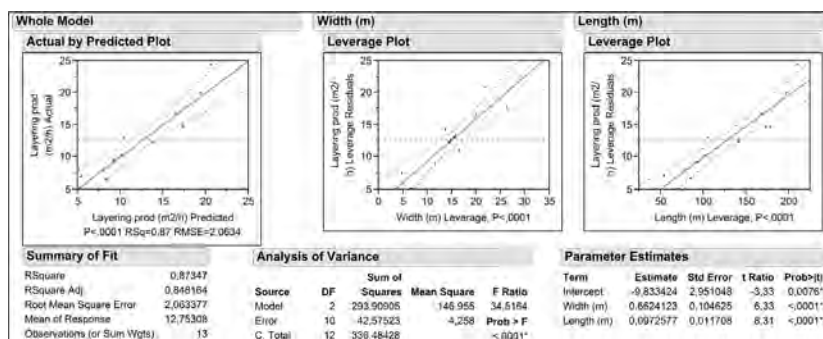

Figure 4: Concrete layering productivity regression model results.

	Length			f_{width}
Width	0-55	55-80	>80	
om-6m		5	**6, 9, 15, 16**	1,00
6m-9m			7, 8	1,69
9m-17m	10		13, 14, 21	2,06
>17m	11, 12, 18	17, 19, 22		

Table 3: Division of collected data points in the specified clusters

$$f_{width} = \frac{Q_{eff}\langle 6,9,15,16\rangle}{Q_{eff}\langle 6,9,15,16\rangle} = 1,00 \Rightarrow Q_{eff}^L = Q_{th}^L \times 1,00, w \in [0-6], l \in [>80]$$

$$f_{width} = \frac{Q_{eff}\langle 7,8\rangle}{Q_{eff}\langle 6,9,15,16\rangle} = 1,69 \Rightarrow Q_{eff}^L = Q_{th}^L \times 1,69, w \in [6-9], l \in [>80]$$

$$f_{width} = \frac{Q_{eff}\langle 13,14,21\rangle}{Q_{eff}\langle 6,9,15,16\rangle} = 2,06 \Rightarrow Q_{eff}^L = Q_{th}^L \times 2,06, w \in [9-17], l \in [>80]$$

productivity for the concrete layering task would be given by the relationship $Q_{eff} = Q_{th} \times 2,06 = 9,27 \times 2,06 = 19,10$ [m²/h]. For validation purposes, data point 13, which corresponds to the previously described scenario, yielded an actual measured productivity of 19,95m²/h. So, the estimation of the produced model seems to be a very good prediction. It should be noted, though, that the correction factors have been derived from a rather small sample. It is logical, that as more data points are being added to the sample, the predictive capability of the correction factors is going to be improved. However, the validity of the estimation process per se is independent of the sample size and should only abide with the statistical inferences of the experimental framework, as described in the research methodology.

Finishing and joints cutting regression models

This section briefly presents the analysis results for the rest of the studied tasks, namely the finishing and joints cutting operations. For brevity reasons, the final results of the produced models along with their statistical properties are going to be presented, since the applied methodology remains the same as in section 4.1. The regression models for the finishing and joints cutting operations are shown in Table 4, whereas the productivity and validation charts are illustrated on Figures 6a/b (average

accuracy ±23%) and 7a/b respectively (average accuracy ±25%). A summary of the estimated f_{width} factors for all three regression models is shown in Table 5.

Discussion

The analysis and all the aforementioned regression models has been based on the laser screed layering method. It is interesting to examine what the resulting productivity will be for a changed working method, by looking into two alternatives: (i) concrete layering by hand and (ii) slip-

No.	Case	Productivity model	Prob > F	R-square	Prob>|t| intercept	Prob>|t| width	Prob>|t| length
1	Finishing (m²/h)	-23,64306 + 2,3417823 * [Width] + 0,2037962 * [Length]	0,0001	0,85	0,0111	0,0001	0,0001
2	Joints cutting (m/h)	-5,613561 + 2,2789522 * [Width] + 0,1695937 * [Length]	0,0023	0,70[a]	0,6541[b]	0,0007	0,0099

[a,b]This outcome indicates statistical invalidity, but the models are still valid because the p-value for the models as a whole is <0,05, even though there is some statistical inconsistency with the intercept and R-square.

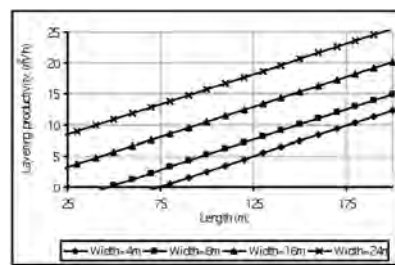Table 4: Finishing and joints cutting productivity models

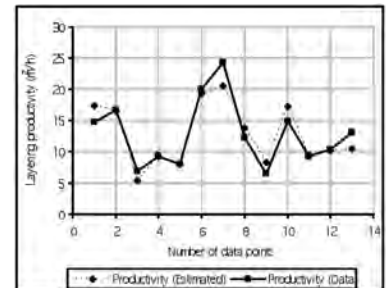Figure 6a: Pavement finishing productivity estimation chart

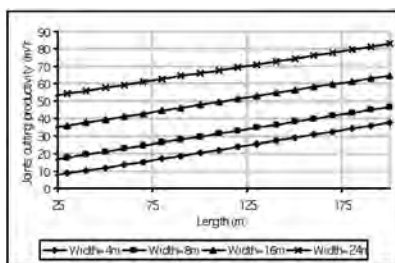Figure 6b: Pavement finishing productivity validation chart

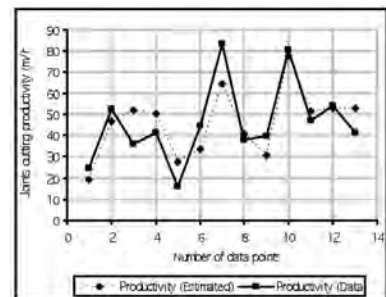Figure 7a: Joints cutting productivity estimation chart

Figure 7b: Joints cutting productivity validation chart

	Length >8om		
Width	Layering	Finishing	Joints cutting
om-6m	**1,00**	**1,00**	**1,00**
6m-9m	1,69	1,39	1,06
9m-17m	2,06	2,39	2,28

Table 5: Correction factor fwidth for paving operations

form paving technique. For concrete layering by hand, it seems that there is a significant change in productivity. Indicative field measurements indicate a decrease of more than 50%. This fact may be attributed to the increase in the required time for finishing, since the accuracy of work is severely affected in relation to the automated working method of the laser screed. Thus, it is logical that inclinations in the magnitude of 1/1000 are not easily reached by hand.

On the contrary, the productivity variation for the slipform paving technique seems to be rather small, more specifically in the region of ±5%. This can be explained as follows: The analysis takes into account the time required for the actual pouring, screeding and finishing of the concrete. In that sense, both methods are close in terms of production output, however, it should be noted that the slipform paver is limited at a working width of 4m, whereas the laser screed works freely. The main difference though lies in the preparation activities before the concrete pouring, which relates to the time that is needed by the surveyor in order to mark the layering region, indicate the required altitudes etc.

In terms of the models per se, it is notable that for higher width and length values the respective productivity models yield larger outputs. In other words, the larger the working area the more productive the deployed crews can be. A probable interpretation of this finding is the fact that in larger working areas, on-site congestion is considerable smaller, enabling a more efficient collaboration of the resources.

In total, the difference of current study with other approaches lies in that the applied methodology is not limited in the generation of forecasting models, but rather identifies the relationship of specific operational factors regarding the achieved productivity. As such, process-oriented estimation methodologies can be developed to cover for any lack of previous data or grounded theory behind specific construction operations.

Conclusions

This study has presented an integrated approach to (i) analyse construction operations, (ii) comprehend the factors that affect their productivity, (iii) establish a valid methodology for quantifying their impact, (iv) conduct sensitivity analysis in relation to selected variables, (v) develop statistically valid regression estimation models and (vi) develop new, process-oriented estimation methodologies to be applied in future projects. The research is both relevant to industry practitioners, since it comprises an estimation tool that can be easily adapted to an organisation's cost management objectives, as well as researchers, since it provides the framework upon which every construction operation, can be decomposed it its constituent tasks and, consequently, analysed in depth. Despite its limitations in scope, the research results have been validated and the achieved accuracy is satisfactory. A possible extension of the current study could be (a) on the breadth of the analysis, where cost regression models may be similarly developed to indicate variations in unit costs and (b) on the applied methodology, where a departure from deterministic methods towards the adoption of stochastic, simulation-based analysis may enhance the models' validity.

7. References

Hallowell, M.R. and Gambatese, J.A. (2009), "Activity-based safety and health risk quantification for formwork construction", *Journal of Construction Engineering and Management*, Vol. 135, No. 10, pp. 990-998.

Hassan, M.M. and Gruber, S. (2008), "Simulation of Concrete Paving Operations on Interstate-74", *Journal of Construction Engineering and Management*, Vol. 134, No. 2, pp. 2-9.

Knutson, K., Schexnayder, C.J., Fiori, C. and Mayo, R. (2008), *Construction management fundamentals*, McGraw-Hill, NewYork.

Kutner, M.H., Nachtsheim, C.J., Neter, J. and Li, W. (2005), *Applied linear statistical models*, Mc-Graw Hill, New York.

Lee, E.B., Ibbs, C.W., Harvey, J.T., and Roesler, J.R. (2000), "Construction productivity and constraints for concrete pavement rehabilitation in urban corridors", *Transportation Research Record*, Vol. 17, No. 12, pp. 13-24.

Miller, S.R., Hartmann, T. and Doree, A.G. (2011), "Measuring and visualizing hot mix asphalt concrete paving operations", *Automation in Construction*, Article in press.

Mitropoulos, P., Cupido, G. and Namboodiri, M. (2009), "Cognitive approach to construction safety: Task demand-capability model", *Journal of Construction Engineering and Management*, Vol. 135, No. 9, pp. 881-889.

Nunally, S.W. (2007), *Construction methods and management*, 7th ed., Prentice-Hall, Upper Saddle Rive, NJ.

Panas, A. and Pantouvakis, J.P. (2010a), "An empirical framework for the performance-based evaluation of health and safety's contribution to sustainable construction", *Organization, Technology and Management in Construction*, Vol. 2, No. 2, pp. 173-181.

Panas, A. and Pantouvakis, J.P. (2010b), "Comparative analysis of operational coefficients' impact on excavation operations", *Engineering Construction and Architectural Management*, Vol. 17, No. 5, pp. 461-475.

Vandenbossche, J.M., Mu, F. and Burnham, T.R. (2011), "Comparison of measured vs. predicted performance of jointed plain concrete pavements using the Mechanistic-Empirical Pavement Design Guideline", *International Journal of Pavement Engineering*, Vol. 12, No. 3, pp. 239-251.

Yiang, Y. and Wu, H. (2007), "Production rates of highway construction activities", *International Journal of Construction Education and Research*, Vol. 3, No.2, pp. 91-98.