A METHOD FOR EFFECTIVELY IMPLEMENTING CONSTRUCTION PROCESS PRODUCTIVITY ESTIMATION MODELS

L. Darren Graham[1] **and Simon D. Smith**

School of Engineering and Electronics, The University of Edinburgh, Edinburgh, EH9 3JN, UK

The international research community have proposed many accurate models for estimating the productivity of construction processes. Few of these models permit practical implementation at the early stages of a construction project, and therefore are of limited use to construction practitioners. For example, contractors require accurate estimates of productivity in the early stages of a project, to avoid being lumbered with an inadequate quantity of plant at the construction stage. This paper considers the earthmoving process, and examines the effect of inadequate resources on productivity. The causes of under-resourced operations are discussed and a dynamic modelling framework is proposed as a method to prevent this phenomenon. Dynamic modelling can be applied to any modelling method and constitutes: separating the operation into planning stages; choosing the variables which are known and significant at each stage; building a model based on those variables. The framework is applied and validated for a linear regression model of earthmoving productivity, providing acceptable results. The accuracy of the results from the model could be improved by the adoption of the principles of concurrent engineering. Finally, the implementation of dynamic modelling in simulation, case-based reasoning and neural networks is discussed.

Keywords: estimating, planning, modelling, productivity

INTRODUCTION

There have been numerous modelling methods, such as simulation, neural networks, regression analysis and case-based reasoning, proposed to estimate the productivity of construction processes. The productivity is normally a measure of the amount of work completed per hour, and can be used to estimate the amount of time and financial cost required to complete an individual operation of a construction process. The time and cost estimates for each operation, of every construction process are required to make a plan of a complete project, and therefore these estimates must be accurate. In the past there have been many simulation productivity models, such as: CYCLONE (Halpin, 1977), UMCYCLONE (Ioannou, 1989), STROBOSCOPE (Martinez and Ioannou, 1994) and TruckSim (Smith, 1995; Smith et. al, 1995). More recently the tendency has been to use the artificial intelligence modelling techniques neural networks (Chao and Skibniewski, 1994; Portas and Abourizk, 1997; Sonmez and Rowings, 1998) and case-based reasoning (Graham and Smith, 2003; Graham et. al, 2004; Graham and Smith, 2004) to estimate construction productivity.

Although, all of the above models have been proven capable of providing accurate estimates of productivity, there have been many associated usage difficulties,

[1] Tel: 0131 650 7207, e-mail: darren.graham@ed.ac.uk

Graham, L D and Smith, S D (2004) A method for effectively implementing construction process productivity estimation models. *In:* Khosrowshahi, F (Ed.), *20th Annual ARCOM Conference*, 1-3 September 2004, Heriot Watt University. Association of Researchers in Construction Management, Vol. 2, 1043-52.

preventing the widespread application of the models in the construction industry. One of the main difficulties lies in the need to provide a supplier (of plant and materials) with an accurate order of the requirements for a particular operation (or operations) far in advance of the day of construction. This holds for any large UK Contractor, whether the supplier is internal or external to an organisation. The productivity models proposed so far have not taken into account the need for early, accurate estimates and hence are less useful to industry than they could be.

A consequence of placing a late order with a supplier, especially of plant, is to potentially leave an operation under resourced. To highlight the effect that a lack of resource has on operational performance, a study is presented to quantify the changes in performance encountered in an operation, resulting from a reduction in resources. The consequences of recruiting too much plant (over resourced) are also discussed. Next, a method of making productivity models more useful is presented, based on the philosophy that a model should be dynamic. That is, the variables which are required to be input by the user (planning engineer) should be realistically known at pre-defined stages of a construction project; the model variables change (normally increase in number) as the time to construction draws nearer and a planning engineer has more knowledge of an operation. In this paper, the dynamic modelling framework has been applied to the productivity estimation of earthmoving operations using linear regression analysis, at two project stages.

The structure of this paper is as follows:

- Summary of the collection of data
- The effect of poor plant allocation on construction operations
- The dynamic modelling framework
- An example application of dynamic modelling: linear regression analysis of earthmoving productivity
- Re-planning the validation operations
- Implementing dynamic modelling using other modelling methods
- Conclusions and future work

GATHERING OF REAL CONSTRUCTION DATA

The data used in this study was collected during a time study of four road-building projects, undertaken by the same contractor in UK. These projects varied widely in terms of: the quantity of soil that was required to be moved, from the relatively small A52 Ashbourne relief road (120,000 m^3) to the M1/A1 link road (3,000,000 m^3); the location of the projects (South, Midlands and North of England); the soil types encountered; and numerous additional factors. In total, 141 earthmoving operations were recorded, 90 of which were deemed suitable for use in this study; the 51 others were discounted due to missing data. Two-thirds of the 90 operations were selected for use in model development, the remaining third to be used in model validation. Table 1 presents a list of the recorded variables and their corresponding determinability at: four to six weeks prior to construction (planning stage one); and, one week to one day before construction (planning stage two). If a variable could be determined with a high degree of confidence (high in the table) it was considered for use in modelling (see table 1).

THE EFFECT OF POOR PLANT ALLOCATION ON CONSTRUCTION OPERATIONS

It is important to highlight the effect that the under-provisioning (poor allocation) of plant has on the performance of construction operations. The aim of this study is to quantify this effect, and will do so by:

- Using a linear regression model to recreate the thirty validation operations which have taken place in the past, reducing the number of haulers used in the operations by increments of one, until the final recreated operation uses three haulers less than the original. Recording the performance of each new operational set-up.
- Measuring the difference (percentage) between the original performance and each modelled operational performance to gain some perspective on the effect of reducing the quantity of plant.

Table 1: Recorded variables and variables considered for modelling

Variable	Determinability		Model Variables	
	Planning Stage One	Planning Stage Two	Planning Stage One	Planning Stage Two
Haul road soil parameters	Medium	High	Haul road gradient	Haul road soil parameters
Haul road gradient	High	High	Plant specifications	Haul road gradient
Haul road rolling resistance	Low	Medium	Loader bucket volume	Haul road length
Haul road length	Medium	High	Number of haulers used	Plant specifications
Plant specifications	High	High	Total volume to be moved	Load cycle times
Plant operator ability	Low	Medium	Bucket passes per load	Hauler travel times
Load area characteristics	Medium	Medium	Project number *	Loader bucket volume
Type and load of obstructions	Low	Medium		Bucket passes per load
Weather	Low	Low - Medium		Number of haulers used
Material susceptability to weather	Low	Low - Medium		Total volume to be moved
Load cycle times	Medium	High		Month of excavation
Hauler manoeuvre times	Medium	Medium		Project number *
Hauler travel times	Medium	High		
Load dump times	Medium	Medium		
Loader bucket volume	High	High		
Bucket passes per load	High	High		
Number of haulers used	High	High		
Total volume to be moved	High	High		
Month of excavation	High	High		

* included to allow an examination of the effect that an individual project has on all the collected data. If the effect is significant, the model will not be useful for providing estimates of new projects.

The results of the poor plant allocation study are shown in table 2; they are an average over the entire validation data set. Clearly, a reduction in the number of haulers causes a steady decline, of 16.21% per truck of operational productivity and a non-linearly increasing operational cost. Therefore, it is very important to ensure that an adequate quantity of plant is available for a specific operation, and the dynamic modelling method is one way of providing this.

Table 2: Average percentage difference from the observed, for the whole validation set

Performance Measure	Number of Haulers Removed		
	One	Two	Three
Productivity	-16.21	-32.42	-48.63
Cost	7.53	12.01	23.70

THE DYNAMIC MODELLING FRAMEWORK

The study in the previous section highlighted the detrimental effect of having too few trucks in an earthmoving operation. What causes an under-resourced operation? There are a number of answers to this question: an inaccurate estimate of the probable rate of work (productivity) is made; an order for plant requirements is submitted to the supplier too late to allow fulfilment. What can be done to reduce the occurrence of an under-resourced operation? To diminish these issues, two methods could be pursued:

- Concurrent engineering principles should be adopted by construction firms. Briefly, this means bringing together all parties: designer, architect, client, construction planner and contractor to learn as much about the project at the earliest possible stage. This shall not be dealt with here.
- The dynamic modelling framework should be applied to productivity estimation models. This shall now be discussed in full.

The aim of dynamic modelling is to provide an accurate estimate of the productivity of a construction operation at the earliest possible stage in the planning of a construction project. This aim is achieved through using information about the operation that is realistically known at certain stages of the planning process. Currently, the dynamic model is split into two stages: planning stage one, four to six weeks prior to construction, with the aim of allowing a planning engineer to calculate the resource requirements; planning stage two, one week to one day prior to construction, to allow a confirmation of the order or make a slight adjustment to the order based on new information.

The framework of the dynamic modelling philosophy is:

- Make a list of all available variables and determine how if they are known at either planning stage one or two. If a variable is known it is added to a second list.
- Select the significant variables from the second list for each planning stage; principle component analysis (PCA) was used to choose these variables in the study.
- Make two models (one for each planning stage) based on the significant variables.
- Validate the model to provide confidence that it provides accurate estimates.
- At planning stage one; supply the known details of the operation to the model, and experiment with the number of resources to produce estimates of productivity and operational cost. Compare the productivity and cost estimates and note the plant resource which provides the best value for money, e.g. a compromise between speed and cost. Increase the ideal numbers of plant by one to provide an initial conservative estimate. Communicate these requirements to the plant supplier(s).
- At the second planning stage: supply the required information to the model, to produce an adjusted estimate of productivity and cost; experiment with the plant requirements to find the compromise between speed and cost. At this stage the ideal number of plant should be chosen and communicated to the plant supplier(s). It is unlikely that the plant requirements will have changed

significantly, therefore the suppliers should be able to meet the contractor's needs.

AN EXAMPLE APPLICATION OF DYNAMIC MODELLING: LINEAR REGRESSION ANALYSIS OF EARTHMOVING PRODUCTIVITY

This example demonstrates the application of linear regression to modelling earthmoving productivity through following the dynamic modelling framework outlined above. The data used in this experiment was introduced earlier in this paper. To re-cap, 90 operations were observed, 60 of which shall be used in model development, the remainder for validation.

Select the highly determinable earthmoving process variables
The highly determinable variables of the earthmoving process have been selected and are shown in table 1. Note there are 6 and 11 such variables, at planning stages one and two, respectively. In addition a project number index has been included to determine whether the models only hold true for specific projects- if this were the case the models' predictive usefulness would be limited; this results in, 7 and 12 process variables at planning stages one and two, respectively.

Select the significant variables of the earthmoving process
PCA was used to select the significant variables of the process. The numbers of variables were reduced from: 7 to 5 at planning stage one; 12 to 9 at planning stage two. For both planning stages, the project index was insignificant, thus the models will be useful for providing predictions of future operations.

Develop a model for each planning stage
A linear regression model was developed for both planning stages one and two, using the remaining data. The regression method used was the backward elimination stepwise, meaning that a variable which was found to be less significant in the linear fit than a threshold value, would be removed. The regression is then performed again on the reduced data set, and this process continues until the remaining variables are above the set threshold of significance. The planning stage one variables were reduced from 5 to 3, and stage two variables reduced from 9 to 6. The regression equations for the productivity are:

$$\text{Productivity} = 18(\text{No. Haulers}) + 10.8(\text{Hauler Type}) + 0.3(\text{Volume}) + 44.5 \qquad \text{Stage One}$$

$$\begin{aligned}\text{Productivity} = &-14.5(\text{Month}) + 9.6(\text{Loader Type}) + 25.5(\text{No. Haulers}) - 40(\text{Load Cycle Time}) \\ &+ 27.2(\text{Buckets per Load}) - 7.2(\text{Travel Time}) + 147.7 \qquad \text{Stage Two}\end{aligned}$$

How well does the model fit the data? A coefficient of variation in the model of 1 signifies a perfect fit. However, a value of 0.7 is most likely to be an acceptable fit to a complex system like the earthmoving process. The R^2 values for the planning stage one and two models were 0.755 and 0.88, respectively. Thus, these regression fits can be deemed to be tentatively acceptable. There is a need for more knowledge of the predictive capabilities of the models, and this shall be provided through model validation (see next indent).

Model validation

The details of 30 observed operations, which had not been used in the model development, were used in the model validation. Validation consisted of entering the information of each operation into both regression models and recording the estimated productivity values. These values were compared with the observed productivities and it was found that on average: the planning stage one model overestimated productivity by 5.8%, with a standard deviation in results of 27.7%; the planning stage two model overestimated productivity by 3.13%, with a standard deviation in the results of 14.3%. Thus, due to more information being known about the process at the later stage, the model of stage two is more accurate. A plot of the models' estimates against the observed productivities for the validation set is shown in figure 1. Clearly, the models are roughly following the trend of the observed data. In conclusion, both models are suitable accurate for the purposes of this study.

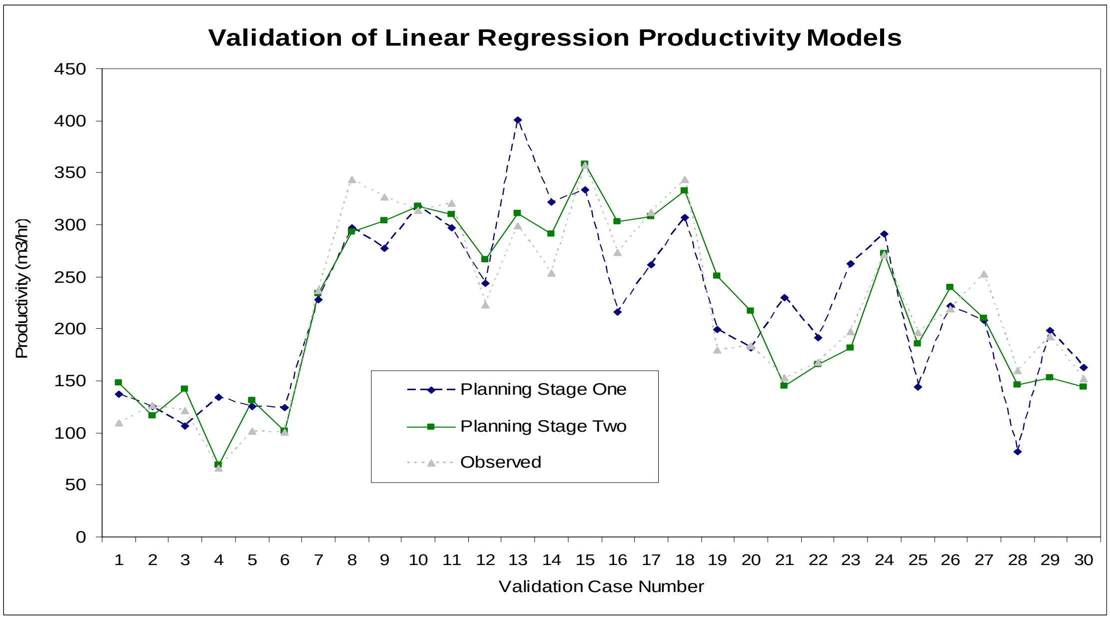

Figure 1: The results of model validation

Using the models to order plant

The known details of a future operation can be supplied to the validated models, in order to receive estimates of productivity, duration and cost. The number of plant required (number of haulers) can be experimented with (between 2 and 6 haulers in this study) and a compromise between duration and cost of an operation can be made to find the ideal planning solution. In this example the details of the operation, known at planning stage one were: Type of available hauler = Caterpillar D400D; Excavation volume = 400 m^3. At planning stage two, the known details of the operation were: Month of excavation = May; Type of available loader = Caterpillar 245 backhoe excavator; Load cycle time = 2 minutes; Buckets per load = 7; Travel time = 6 minutes. The number of haulers was experimented with, and the resulting productivity and cost estimates for each experiment, at planning stages one and two, are shown in tables 3 and 4, respectively. Note that in this example the trend of the costs observed in tables 3 and 4 are the reverse of those in table 1, the consequence being that this operation provides more value for money when only two haulers are used, and it is likely that poor forward planning would have resulted in this operation having been over-resourced, i.e. 3 haulers used instead of 2.

Table 3: Experimental results at the early planning stage (one)

	Number of Haulers				
	2	3	4	5	6
Productivity (m3/hr)	210.2	228.1	246.1	264.1	282
Cost (£)	216.65	271.98	319.24	360.06	395.67
Cost/ Volume (£/m3)	0.537	0.675	0.792	0.893	0.982

To make the decision on which operational set-up provides the best value for money a ratio of the cost divided by volume is calculated. As volume is fixed the ratio provides a basis for directly comparing the results of the two planning stages and the value for money of each resource quantity. In the cost/volume ratio, a smaller value signifies a better compromise between cost and time. Clearly, at the first planning stage, having two haulers provides the best value for money. Therefore, following the dynamic modelling framework, an order should be placed at this time, of four to six weeks to construction, with the plant supplier for three haulers (ideal plus one). Examining the experimental results at the second planning stage reveals that the situation is unchanged in terms of the ideal quantity of haulers. However, it is noteworthy that the value for money for two haulers has been reduced (cost/volume ratio increased) in the second stage, while the remaining hauler quantities provide increasing value (cost/volume ratio decreased).

Table 4: Experimental results at planning stage two

	Number of Haulers				
	2	3	4	5	6
Productivity (m3/hr)	209	234	260	285	310
Cost (£)	217.97	264.75	302.36	333.25	359.08
Cost/ Volume (£/m3)	0.541	0.657	0.75	0.83	0.89

RE-PLANNING OF VALIDATION OPERATIONS

Each of the 30 validation operations discussed in the previous section, were re-planned using the methods outlined above, to produce new estimates of cost and productivity. By comparing the cost estimates with those actually observed in the operations, it is possible to calculate the average improvement/deterioration in the cost (per cubic metre) to complete an operation, over the whole validation set, from using the dynamic modelling framework to plan the operations. It was found that an average cost saving of £0.21 per cubic metre was possible.

Assuming that the validation set of operations is an appropriate survey of the four studied projects, the cost saving per cubic metre of earth moved can be used to calculate the possible cost savings in all four projects. The total amount of earth moved in the four studied projects equalled 5.82 million cubic metres, and the actual cost of this work was calculated as £5.98 million. Through re-planning using the dynamic modelling framework the total cost could have been £4.75 million, offering a reduction in cost of 20.5% from the actual. It is noteworthy that this is only a theoretical saving to the direct cost of earthmoving- indirect costs were not considered. Clearly the saving indicates the potential impact that the dynamic modelling framework could have on a construction industry which needs to reduce waste and to improve its performance.

IMPLEMENTING DYNAMIC MODELLING USING OTHER MODELLING METHODS

The following section is a brief summary of the actions required to implement the dynamic modelling framework in other commonly used construction process modelling methods.

Case-Based Reasoning (CBR)

CBR consists of formulating a data (case) base of past examples of a problem and using several different methods to identify appropriate past examples as possible solutions to new problems. Thus, the implementation of dynamic modelling can be achieved in two ways: one, as suggested in the dynamic modelling framework, point 3- have a separate case base for each planning stage, incorporating only the significant variables of the respective stages; two, vary slightly from the framework, point 3 by building one case base holding the significant variables of both planning stages, but only consider, at any one time, the variables relevant to a particular stage of the project. The remainder of the framework can be followed using CBR.

Artificial Neural Networks (ANN)

ANNs form approximations of the trends contained within data sets and are therefore useful in modelling many non-linear or near-linear problems, such as predicting the productivity of the earthmoving process. ANNs work in a similar way to linear regression analysis and the dynamic modelling framework can be followed directly by: separating the data into two groups, planning stages one and two; splitting the data within each of these groups into two further groups- training and testing; training two models to mimic planning stages one and two; testing the models predictive capabilities (validation) and then using them to plan future construction operations.

Discrete-Event Simulation (D-E Sim)

Discrete-event simulation models are often driven by probability distributions. These distributions represent the variability contained within the process, which is being modelled. In the earthmoving process, the variability can be described by representing the stochastic elements in the process, such as the time of travelling from the loader to the dump area (travel time), using probability distributions. Probability distributions that accurately represent the stochastic elements can then be used to build up a picture of the productivity of the process, over an entire operation. The probability distributions which represent the stochastic elements will change with different operational conditions, i.e. the probability distributions representing the travel time at planning stage one, are unlikely to remain the same at planning stage two. At the moment, to understand the nature of this change would require an expert. However, Graham and Smith (2003; 2004) have produced a computer model which removes the need to have expert knowledge of the system to define its representative probability distributions, and the dynamic modelling framework shall be applied to this model in the future.

CONCLUSIONS

A by-product of poor forward planning is that a construction project can be under-resourced. This is due to a plant/materials supplier not having a sufficient stock, at late notice, to meet the demands of a construction project. This study highlighted the effect of having under-resourced operations on the earthmoving process performance. It was found that for every hauler unit an operation was under resourced by, the productivity was reduced on average by 16%, with a concurrent non-linear increase in cost. Thus, the provision of too little resources should be avoided. Intuitively, providing too many resources should also be avoided, to prevent the incursion of additional costs- an extra hauler makes a large difference to the cost of an operation, and this can propagate for a whole project.

In an attempt to improve the planning of projects, and to prevent the issues raised above, this paper proposed that construction productivity models are developed following the dynamic modelling framework. This framework consists of using all the significant knowledge possessed by a planning engineer at specific project stages, to allow accurate estimates to be made of the resources required in an operation, well in advance of construction.

The dynamic modelling framework was applied to the linear regression analysis of the earthmoving process, providing models which accurately recreated a validation set of operations. Experimentation was undertaken on how the productivity and cost of the operation changed with a change in the number of haulers. The results of this experimentation allowed the validation operations to be theoretically re-planned, with an ideal number of haulers selected for each situation. This new estimate of performance, in terms of cost and duration was compared with the observed values. This comparison was used to determine if any theoretical improvement could have been made, by applying the dynamic modelling framework to the planning of operations in the first place. It was found that a theoretical 20% saving on cost could have been made through applying the dynamic modelling framework.

REFERENCES

Chao, L-C., and Skibniewski, M. J. (1994) Estimating construction productivity: a neural network based approach, *J. Comp. In Civil. Engrg., ASCE*, 8(2), 234-251.

Graham, L. D., and Smith, S. D. (2003) A hybrid model to improve the estimation of concreting operations, *In:* Greenwood, D (Ed.), *19th Annual ARCOM Conference*, 3-5 September 2003, University of Brighton. Association of Researchers in Construction Management, Vol. 2, 505-512.

Graham, L. D., and Smith, S. D. (2004) Estimating the productivity of cyclic construction operations using case-based reasoning, *Advanced Engineering Informatics*, In Print.

Graham, L. D., Smith, S. D., and Crapper, M. (2004) Improving concrete placement simulation with a case-based reasoning input, *Civil Engineering and Environmental Systems*, Vol. 21(2), 137-150.

Halpin, D. W. (1977) CYCLONE: method of modeling job site processes, *J. Constr. Div., ASCE*, Vol. 103, 489-499.

Ioannou, P. G. (1989) UM-CYCLONE discrete-event simulation system: user's guide, *Rep. UMCE-89-12,* Dept. of Civ. Engrg., Univ. of Michigan, U.S.A.

Martinez, J., and Ioannou, P. G. (1994) General purpose simulation with Stroboscope, *Proc. 1994 Winter Simulation Conf., IEEE,* Piscataway, N.J., U.S.A.

Portas, J., and Abourizk, S. (1997) Neural network model for estimating construction productivity, *J. Constr. Engrg. and Mgmt., ASCE*, 123(4), 399-410.

Smith, S. D. (1995) Productivity estimation of earthmoving operations using a discrete-event simulation model, *PhD Thesis*, University of Edinburgh, UK.

Smith, S. D., Osborne, J. R., and Forde, M.C. (1995) Analysis of earth-moving systems using discrete-event simulation, *J. Constr. Engrg. and Mgmt.*, *ASCE*, 121(4), 388-396.

Sonmez, R., and Rowings, J. E. (1998) Construction labor productivity modeling with neural networks, *J. Constr. Engrg. and Mgmt.*, *ASCE*, 124(6), 498-504.