CONCRETE PLACING PRODUCTIVITY USING A NOVEL NEURAL NETWORK DESIGN

Doug R. Forbes, Harry McGurnaghan, L. Darren Graham[1] and Simon D. Smith

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

For planning purposes an accurate estimate of the productivity for insitu concreting operations is desirable. Various methods have been investigated to model such productivity; this paper will consider a novel neural network architecture the *Twin Nested Recurrent Network* (TNRN), in the search for increasingly accurate predictions. Neural networks are trained by providing them with a set of data from historic projects. Providing a larger training data set generally increases the accuracy of the output. In conventional neural network modelling, this training set is based upon the outcomes of a series of individual operations and thus is limited. The approach to be used here would expand this training set by considering the productivity of each individual concrete delivery used in each operation thus creating a much larger training set of data from which the neural network can learn. Further, by preparing the data in this way, an original architecture can be developed for the model. This architecture is a twin-loop nested recurrent network. The network is presented with each delivery for a given operation, and then with the productivity of the operation itself. Initial results are encouraging: this approach allows a much greater flexibility and a significant reduction in the sensitivity of changes to the network training functions.

Keywords: concreting, neural networks, planning, estimation, modelling

INTRODUCTION

Concrete placing is a process that is common to many construction projects and as such it is imperative that a planner/ manager of construction can provide accurate estimates of the productivity, cost and duration of this activity. Currently most construction contractors rely on the experience of their staff to produce such estimates. More formal techniques have been suggested by the construction research community such as regression analysis (Smith, 1999), discrete-event simulation (Smith et. al, 1995), case-based reasoning (Graham et al., 2004) and neural networks (NN). Here we shall focus on the latter, as it is advantageous to create an analytical model that will allow cost and duration estimates to be investigated to determine the response if variables are different; NNs permit this type of investigation. NNs are a means of forming an approximation that have been used to model the concrete placing process by Sonmez and Rowings (1998) and Adeli and Wu (1998).

NNs learn from experience and consist of neurons, or processing elements, that are interconnected. The position and strength of the interconnection weights determine the predictive capabilities of the network. In training a NN it is these interconnections that are altered to best reproduce an estimate of the problem. Thus, training a NN is vital in

[1] Corresponding author: email- darren.graham@ed.ac.uk; telephone- 0131 650 7207.

Forbes, D R, McGurnaghan, H, Graham, L D and Smith, S D (2004) Concrete placing productivity using a novel neural network design. *In:* Khosrowshahi, F (Ed.), *20th Annual ARCOM Conference*, 1-3 September 2004, Heriot Watt University. Association of Researchers in Construction Management, Vol. 2, 1053-62.

efficient operation. In general, by providing a larger training set, more accurate solutions can be obtained.

This study examines the use of the standard NN architectures of feedforward (FF) and Elman networks. Following this, a novel network architecture, the *Twin Nested Recurrent Network* (TNRN) is developed which makes use of the concrete delivery process by providing a larger training data set for the network.

DATA ANALYSIS

Factors effecting the data

Many factors affect the concrete placing process. Crombie (1999) listed 44 separate factors that have an effect. Factors can be classified as quantitative or qualitative. Quantitative factors are numerical values. Qualitative factors are descriptions. Portas and AbouRizk (1997) applied a weighting to qualitative subjective factors depending on the influence of the factor. Rowings and Sonmez (1996) produced a NN model of the concreting operation using ten factors: quantities completed, job type, number of workers, temperature, humidity, precipitation, percentage labourer, percentage overtime, cumulative quantities and concrete pump.

The use of a NN to model this operation is due to its complex nature and the ability of the model to incorporate all the factors. However, the only variables that should be used in the model are those that can be established by the estimator, or the model user, before the project begins (Smith, 1999). Since the data for this model development was collected by others the data to be used was limited to that available.

Data Collection

It is important for a neural network to be presented with a set of data for training that goes right to the extremes of the problem domain (Flood and Kartam, 1994) to prevent the model extrapolating the results (Hecht-Nielsen, 1990).

While collecting the data, it is important to note the size of the sample. For this study, the sample size was 249. Rowings and Sonmez (1996) used a sample size of 112, whereas, AbouRizk et al. (2001) use a sample of 39. Arditi and Tokdemir (1999) used a sample size of only 12.

The data for the study was taken from observation. These observations were carried out by others. The contractors and the corresponding clients for the sources of original data are shown in Table 1.

Table 1: Sources of data

Contractor	Client
A	Scottish Water
B	Scottish Water
C	British Waterways
D	Highways Agency (HA)

The data for the Scottish Water projects relates to the construction of wastewater treatment works (WwTWs) and the associated tanks, chambers and bases. The data for British Waterways came from the construction of the Falkirk Wheel and for the Highways Agency the data was from the concrete pours involving bridges and underpasses during motorway construction.

Since the data was not recorded specifically for this study, there are certain factors that are not required, and certain factors that could be useful. The additional factors that could be useful will be discussed later. The main aspect of the data that was not required for this study was the identity of each individual truck mixer, and the corresponding cycle times. The remaining data allows a model to be created which can be used to experiment with neural networks in the concrete placing process.

From the data available, eight variables were used to model the productivity. This response was chosen as the output since cost and duration can be calculated from it. The eight factors were: month, location on site, average truck volume, total volume, average interarrival time, number of loads, number of accepted loads and number of rejected loads.

Delivery/Operation Data Analysis

In order to make full use of the data, a value of the productivity for each individual delivery was calculated. This was based upon the arrival time at the site of the delivery, the time at the end of the discharge and the volume of the delivery. Thus, the number of productivities equalled to the number of deliveries in an operation. The overall productivity of the operation was also calculated.

For the eight explanatory variables the data is clustered. The issue that arises is that the neural network will be trained with the same input sets for different values of productivity. The reason for this difference is that the factors that affect the productivity have not been measured, or are un-measurable. To resolve this problem, the standard deviations of the productivities for each operation were calculated. From this, a network could be trained to map to two outputs: the productivity and the expected standard deviation from this value. This data is then used to train and test the TNRN approach. By doing this a range of values can be given in a prediction, and is thus more trusted by users of the system than models that solely produce a point prediction value (Portas and AbouRizk, 1997; Lu et al., 2000; Shi, 1999)

Validation and Test Data

Test Data

The validation method was the train-and-test method (Twomey and Smith, 1997). From the 249 original operations, 210 were used in the development of the networks. This represents a 75:25 split of the original data. In order to ensure that a random sample was used for testing, a random number identifier was given to each of the data points. The data set was then sorted by the identifier and the first 158 points were used for training purposed. The remainder were used for testing.

Validation Set

The additional data from the HA project provided a validation set of 39 points. This data was not independent of the training data, having been taken from the same project, however it had not been used for training purposes (Shi, 1999).

THE NEURAL NETWORKS

Feedforward and Elman Networks

Feedforward and Elman networks were created to establish the ability of NNs to model the concrete placing process. The feedforward networks are a standard approach to NN modelling. The Elman network is a simple approach to a recurrent neural network (The MathWorks, 2003). Recurrent neural networks make use of a feedback loop. This loop allows the network to form temporal patterns, and hence provide better generalisations to the problem (Elman, 1990; Liu, 2001).

Network creation is a heuristic process for both these types of network (Hegazy et al. 1994; Adeli and Wu, 1998). It is necessary to alter the number of layers in the networks and the number of neurons within each layer, the training algorithms used, and the activation functions at each layer (Hecht-Nielsen, 1989; The MathWorks, 2003).

From the work carried out using the train-and-test validation method, it was discovered that the networks were capable of predicting an average test MSE of 16 m^3/hr for the feedforward networks, and 22 m^3/hr for the Elman networks.

Twin Nested Recurrent Network Approach
The TNRN was specifically designed for this study and as far as could be determined similar architectures have not been previously developed. It combines feedforward networks with recurrent loops and makes use of the arrival of deliveries at a site to determine the productivity and the likely range the value lies within.

A diagram showing the network architecture is given in fig. 1. This approach uses two recurrent loops. For operation one, the first delivery is used to train the network and adjust the weights. This is then repeated for delivery 2, up to Q deliveries in operation 1. After each delivery the weights set by the previous delivery are used as the initial weight for the training. The network is then trained with the data for the first operation. The weights set by the delivery at this point are then used as the initial weights in operation 2, up to operation M. Once the data for operation M has trained the network, the weights set at this point are used for the final model.

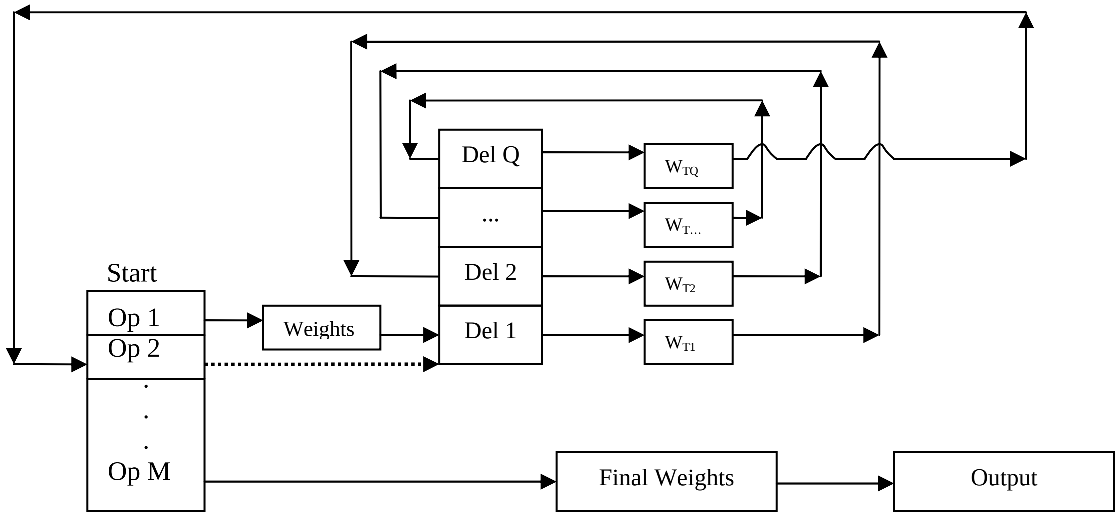

Figure 1: TNRN approach

This network was built with one hidden layer, and one output layer. Tests on the network showed that the output was independent of the number of neurons in each layer. It was also discovered that the network could not be trained without a linear activation function in the output layer.

From this, the network was created using 9 Matlab training algorithms and 3 combinations of activation function – Tan-sigmoid (T), log-sigmoid (L) and linear (P) in the first layer and linear in the second. The initial weights were set as the four combinations of 1 and 0. This created 108 networks. The networks used the same training set as the FF and Elman. The networks were trained with two outputs: productivity and standard deviation.

Results from TNRN
Training time for each network combination was approximately 100 minutes. This was not affected by the training algorithm or activation functions.

The results showed that for all four combinations of initial weights, for activation functions T-P and L-P, the productivity MSE was a constant value of $1060(m^3/hr)^2$ for all the training algorithms. However, for a P-P activation function, the MSE of the productivity varied from $611(m^3/hr)^2$ to $588,860,369(m^3/hr)^2$. The combination with the optimum value was:

- Initial weights: Input weights – 0; Hidden Layer Weights – 1
- Activation function combination: Linear-Linear
- Training algorithm: One Step Secant Algorithm

VALIDATION

Validation of TNRN Approach

The validation data results for the two outputs from this network are given in table 2 for the correlation co-efficient and the root mean squared proportional error, this is given for the productivity and standard deviation response. Figures 2 and 3 show the actual productivity and that modelled by the networks, for the productivity and standard deviation response respectively. As can be seen the relationship between the actual and the fitted data is very poor. All of the values in the table indicate that the two sets of data are not similar.

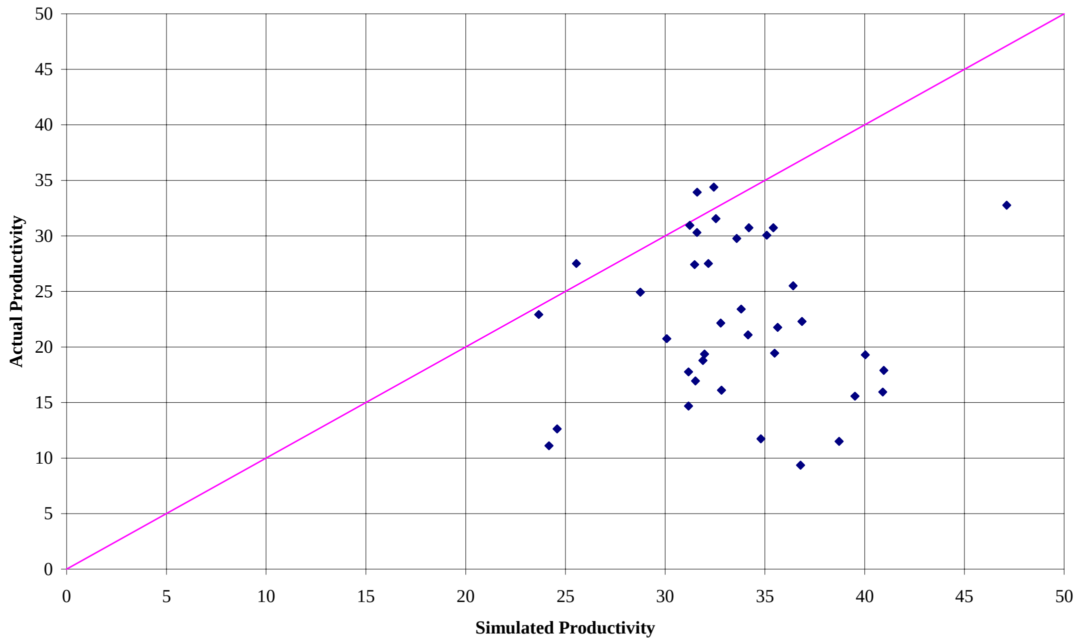

Figure 2: Simulated -v- Actual Productivity for Validation of TNRN

Table 2: Goodness-of-fit for validation data – TNRN

Response	Correlation Co-efficient, R^2	RMS Proportional
Productivity	0.016	0.967
Standard Deviation	0.042	1.043

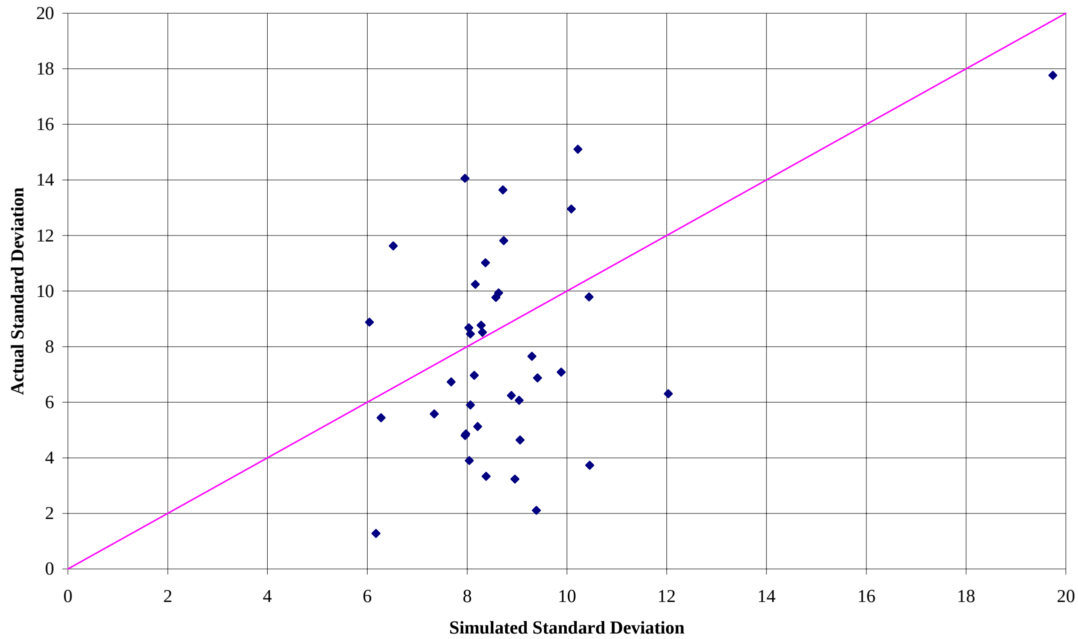

Figure 3: Simulated -v- Actual Standard Deviation for Validation of TNRN

Standard Deviation Output

Since the TNRN does not produce point predictions, it is useful to see the predicted value of the productivity with the predicted range. In fig. 4, for each of the 39 validation points, the actual productivity is plotted along the dots. The fitted productivity is marked with squares. The lines above and below show the range of the predicted value. The first tick on each indicates one standard deviation, and the second two.

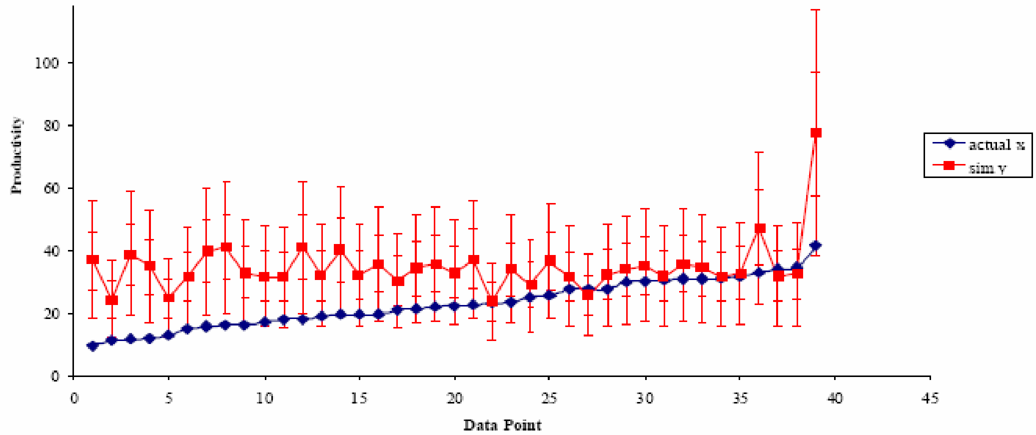

Figure 4: Actual and predicted outputs for TNRN, including standard deviation bars of predicted outputs

From fig. 4, it can be seen that for most of the predictions, the productivity is higher than required. However, all but four of the actual productivities lie within 2 standard deviations of the predicted value. It can also be seen in fig. 2 the use of the range of outputs from the TNRN. Better predictions are within the productivity range 22-38 m^3/hr. This is likely due to a larger amount of data in this range, and so the network can create a better generalisation pattern for these sets of variables.

Network Appraisal

On the face of it, the TNRN does not seem to provide good predictions to the problem. The generalisations that it provides tend to be larger than required, thus it would not be suitable for use as it does not give the same confidence as some of the FF and Elman networks. However the range of predictions does allow a value to be obtained. Two main advantages of this network are that it makes use of all the possible data within the concrete placing system, and provides a range of values as opposed to a point prediction.

DISCUSSION

Data Discussion

The data that was used for this study was not ideal, since it had not been collected for the purposes of this study. Some of the factors were incomplete, and some were not measured at all. Eight explanatory variables were used in the neural network models. It is important to note that these explanatory variables are all measurable by an estimator before a project starts.

These eight are significantly shorter than the list produced by Crombie (1999), of 44 factors. It also did not include three of the factors listed by Rowings and Sonmez (1996) all of which related to the weather conditions. The weather cannot be predicted before the start of the project, but for the use of the models to predict productivity, had the factors been different, would be superior with weather included, and a comparison would be useful of a model with weather included and one without.

The list by Crombie includes many factors that are not measurable, thus the inclusion in an analytical model would not be appropriate. However, in this study, several improvements could have been made if a better knowledge of each individual operation had been available.

Type of Pour: The reference given to each type of pour (1-6) is all that could be ascertained based upon the data. However, this is not wholly accurate. For example, at the GPO Building in Edinburgh the pour was a simply supported floor slab. This did not reflect the pump configuration that involved four 90^{o} bends, thus increasing the risk of blockages. The type of pour factor should be altered to reflect this.

Average Interarrival Time: This is an interesting factor to model, since it is dependent upon the delivery process. The interarrival time can be specified by the contractor before an operation, but it is unlikely that this will be followed exactly. For the data that was used in the model the specified interarrival times were not known. Thus, the actual interarrival times were use to construct the model. The arrival times of the deliveries are a factor in the wait times for the deliveries and the pump. This can have significant effects upon the production rate (Christian and Hachey, 1995). A more appropriate use of the data would have been to include a model with the number of late and early deliveries.

These two points show that the data is lacking. This is compounded by the analysis carried out to create the data set for the TNRN. For this, it was discovered that in each operation the levels of the eight factors were the same, but that each delivery had a different productivity based on the time it took to discharge. This meant that the eight factors used as explanatory variables must not be exclusive and that there are other factors involved in the model, which are not included. These factors could have been measurable as discussed above, or un-measurable, and therefore could not have been successfully included in the model. However, based on the modelling results the factors used provided a good basis from which a model can be created.

The data is also limited by the ranges. For the month factor, the training and testing set only range from January to October. For the validation set, this is for all twelve months. The other factors have a similar range to each other. Thus, the networks can have problems in predicting data that is outwith the training.

Another limitation of the data that was used was that it only involved the concrete pouring system. This limits the data to one aspect of the system. It has been seen that the delivery process has an effect upon the wait times. Another important aspect of the system is the finishing of the concrete. This can vary in time, for example, a wall pour will have a much smaller area requiring finishing than a floor slab, and will thus take a shorter time. The data used in the model commenced with the time that the first delivery arrived, until the last delivery left. This made no account of any delay in the arrival of the first delivery, which could result in reduced productivity if the concreting operatives are waiting. However, this is a function of the delivery process, which can be modelled as a system in its own.

Feedforward and Elman Networks

The feedforward networks gave, on the whole better results than the Elman networks. This is possibly due to the Elman network being a more complex network architecture. This complex architecture means that the network is more susceptible to fluctuations in the data, and missing variables. It has already been noted that there are missing parts in the data that could make the Elman network operate more effectively. One of the major drawbacks of these networks is the point prediction that is made, that provides no confidence intervals to guide a user.

TNRN Recurrent Network

The network was based upon the research that has been carried out using a hybrid collection of standard networks, connected in novel ways to form different architectures (Liu, 2001; de Jesus et al., 2001; Ninomiya and Sasaki, 2002).

In the training and testing of the networks, it was discovered that the TNRN was not affected by the number of hidden neurons. This is likely to be due to the small amount of data that was used in each presentation. This is a major advantage, since standard Elman and FF networks require heuristic methods to define the number of neurons, which is time-consuming and cumbersome.

The transfer functions and training algorithms (TAs) used in the development of this model showed results that were not expected. Firstly, the transfer function required a linear function in the output layer, otherwise the network would not train. Thus, with a linear transfer in the output layer, when the hidden layer function was tansig or logsig, the values of MSE for both outputs were constant for all TAs. With a linear transfer function in the hidden layer, however, the network produced a different result for each TA. The range of these was much larger than those of the standard networks.

The validation results from this network show poorer values than that of the standard approaches. The correlation co-efficient of the data is very poor for both the productivity and the standard deviation responses. However, the aim of the TNRN is not to produce a network that creates only a point prediction, but a network that will give a range of values for the given conditions. Most of the actual values lie within the two standard deviation range. However, from this plot, there is a trend for the network to be unable to make accurate predictions at the lower end of the productivity range. This is likely because training data does not extend fully to this area.

Overall, the TNRN network is likely to produce less accurate results, for similar reasons to the Elman NN. That is, that as the network architecture is more complex a better quality of data is required for training, as the network is more sensitive to fluctuations or errors. The TNRN provides promising results, and further development would be advantageous, perhaps with data from simulation to provide a larger data set.

CONCLUSIONS

In this study different types of neural networks were shown to be capable of modelling the concrete placing process with an acceptable degree of accuracy. Specifically, FF networks were shown to provide the most accurate estimates of the productivity, cost and duration of concrete placing activities. These NNs, however, had the disadvantage of taking a long time to develop, and only providing a point prediction.

It was found that the more complex the model created, the more important it is to use a set of training data that is of a high quality. Thus, for the Elman and TNRN approach, the errors in the predictions can be attributed to the limitations of the data. If more appropriate data is available for training, it may be possible to obtain better predictions with these methods.

The TNRN produced promising estimates and could be used to model similar construction processes such as earthmoving, steel erection or road paving.

Neural network modelling is a heuristic tool that can be time-consuming to define the most appropriate solutions. The TNRN, if developed to obtain more accurate estimates, would be advantageous since it cuts down on the number of component variables. Thus, the network development at this stage is not solely concerned with the end-user, but also with the developers. More work is required in this area, to examine and define the pitfalls for modellers using NNs in modelling construction processes.

REFERENCES

AbouRizk, S., Knowles, P. and Hermann, U. R. (2001) Estimating Labor Production Rates for Industrial Construction Activities. *Journal of Construction Engineering and Management*, **127**(6), 502-511

Adeli, H. and Wu, M. (1998) Regularization Neural Network for Construction Cost Estimation. *Journal of Construction Engineering and Management*, **124(**1), 18-24

Arditi, D. and Tokdemir, O. B. (1999) Comparison for Case-Based Reasoning and Artificial Neural Networks. *Journal of Computing in Civil Engineering*, **13**(3), 162-169

Christian, J. and Hachey, D. (1995) Effects of Delay Times on Production Rates in Construction. *Journal of Construction Engineering and Management*, **121**(1), 20-26

Crombie, R. (1999) *Efficiency, Quality and Waste in Concrete Operations.* BEng Thesis, University of Edinburgh

De Jesús, O., Horn, J. M. and Hagan, M. T. (2001) Analysis of Recurrent Network Training and Suggestions for Improvements, *In: Proceedings of International Joint Conference on Neural Networks, 15-19 July 2001*, **4**, 2632-2637

Elman, J. L. (1990) Finding Structure in Time. *Cognitive Science*, **14**, 179-211

Flood, I. and Kartam, N. (1994) Neural Networks in Civil Engineering I: Principles and Understanding. *Journal of Computing in Civil Engineering*, **8**(2), 131-148

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.

Hecht-Nielsen, R. (1990) *Neurocomputing*, Wokingham: Addison-Wesley

Hegazy T., Moselhi, O. and Fazio, P. (1994) Developing Practical Neural Network Application Using Backpropagation. *Journal of Microcomputers in Civil Engineering*, **9**(2), 145-159

Liu, D. (2001) Open-Loop Training of Recurrent Neural Networks for Nonlinear Dynamical System Identification, *In: Proceedings of International Joint Conference on Neural Networks, 15-19 July 2001*, **2**, 1215-1220

Lu, M., AbouRizk, S. M. and Hermann, U. H. (2000) Estimating Labor Productivity Using Probability Inference Neural Network. *Journal of Computing in Civil Engineering*, **14**(4), 241-248

Ninomiya, H. and Sasaki, A. (2002) A Study on Generalization ability of 3-Layer Recurrent Neural Networks, *In: Proceedings of the 2002 International Joint Conference on Neural Networks, 12-17 May 2002*, **2,** 1063-1068

Portas, J. and AbouRizk, S. (1997) Neural Network Model for Estimating Construction Productivity. *Journal of Construction Engineering and Management*, **123**(4), 399-410

Rowings, J. E. and Sonmez, R. (1996) Labor Productivity Modeling With Neural Networks, *In: AACE International Transactions, PROD.01*, Morgantown, WV, USA, 1-3

Shi, J. J. (1999) A neural network based system for predicting earthmoving production. *Construction management and Economics*, **17**(4), 463-471

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.

Smith, S. D. (1999) Earthmoving Productivity Estimation Using Linear Regression Techniques. *Journal of Construction Engineering and Management*, **125**(3), 133-141

Sonmez, R. and Rowings, J. E. (1998) Construction Labour Productivity Modeling with Neural Networks. *Journal of Construction Engineering and Management*, **124**(6), 498-504

The MathWorks (2003) *Neural Network Toolbox for use with MATLAB®: User Guide* [online]. Massachusetts: The MathWorks. Available from : http://www.mathworks.com/access/helpdesk/help/pdf_doc/nnet/nnet.pdf [Accessed 28 October 2003], pp. xviii, 5-1 – 5-50, 9-1 – 9-7

Twomey, J. M. and Smith, A. E. (1997) Validation and Verification. *In:* Kartam, Nabil, Flood, Ian & Garrett, James H Jr, (eds.) *Artificial Neural Networks for Civil Engineers: Fundamentals and Applications.* New York: ASCE, 44-64