FORECASTING INDICES FOR CONSTRUCTION ACTIVITIES IN THE HEAVY CIVIL ENGINEERING INDUSTRY: A JOINT APPROACH METHODOLOGY

N Dawood and W Bates

School of Science and Technology, University of Teesside, Middlesbrough, TS1 3BA, UK

There are several forms of technique available for use in the prediction of time series data and not one method can be applied to all time series. Essentially, there are only two forms of approach to forecasting: Subjective, which incorporates the judgement of 'experts' (persons who work closely with the industry or product they are set to forecast), and objective, which apply a scientific process to the analysis of previous data using a statistical approach to create 'a fit' to the historic data then applying the 'model' to predict future occurrences. The objective techniques vary in complexity. However, it does not follow that the more complex the approach the better the forecast. The author would like to state that where accuracy is required, a combination of objective statistical analyses with subjective evaluation can be shown to be the most efficient.

Keywords: Forecasting, heavy engineering, indices.

INTRODUCTION

Forecasting performs most effectively as a two-stage process with trends extracted from historical data and subjectively analysed by an expert estimator (Bates and Dawood 1996). There have been several attempts to see whether analytical mapping can indeed outperform the judgement of a so-called expert and vice versa (Edmundson, et al. 1988). However, numerous publications have shown that combination methods are undoubtedly the most comprehensible (Bunn and Wright 1991, Dawood and Neale 1993, Pereira et al. 1989, Bates and Dawood 1996). The authors are not striving to say that this is the answer to all our forecasting problems but the most understandable. Estimators in the heavy civil engineering industry (process, water, railways, bridges) tend not to be statisticians and like the analytical approaches to be kept simple, so they can apply some of their knowledge of the market to the prediction.

Although expert judgement tends to be used predominantly in the early stages of projects, it does not forego the need for analysis of historical data if it exists. To develop an accurate prognosis, as much information as possible is required whether numerate or opinionated. Trends and projections tend to aid the estimator with his prediction as a visual understanding of the past can help users develop the future. To make the process of forecasting simpler, indices can be developed from antecedent data and predicted using a variety of models. Explanations of these modelling techniques appear in several publications (Bates and Dawood 1996) and are outlined in a subsequent paragraph.

Dawood, N N and Bates, W (1998) Forecasting indices for construction activities in the heavy civil engineering industry: a joint approach methodology. *In:* Hughes, W (Ed.), *14th Annual ARCOM Conference*, 9-11 September 1998, University of Reading. Association of Researchers in Construction Management, Vol. 1, 200-9.

Considering the information supplied by the respondents in a questionnaire survey (Bates 1997a) of the heavy engineering industry and surveys of other industries (Sparkes and McHugh 1984), there seems to be a lack of knowledge and training regarding forecasting in the UK. Developing skills in forecasting can be an arduous task and the estimator may never find the time. In today's world of cost estimation, experts are not selling their time but their 'knowledge'. Therefore, this knowledge must be accurate and if not the inaccuracies must be justified by reasons.

To prevent loss or errors, the knowledge mentioned should be harnessed in an expert system. Forecasting time series information has been made simple with the development of expert forecasting software such as "Forecast Pro" and generally by the world of IT (Woodward et al. 1994). Complex statistical methods are carried out by the processor without the intervention of the user. This may all seem a little 'black box'. Nonetheless, as the estimator is only interested in a visual or tabular representation of the underlying trends and not the analysis that goes into producing these trends, this only aids the outcome.

The mathematical analysis can quite easily be developed into an expert system. However, expert systems giving estimators' gut feeling are not so simple. Expertise is inherent with an experienced estimator as market feel. However, this feel is difficult to grasp and to put into a logical format. Nevertheless, if it was to be controlled and verified within the heavy civil engineering industry it could be an extremely useful tool in:

1. Speeding up the expert's response to the client
2. Accelerating the training of junior estimators
3. Providing a means for encapsulating expert knowledge from experienced personnel.

All of these provide confidence, not only in the estimator's figures but also for the client with an improved and guaranteed estimate and possible reasons for error.

The difficulty in producing reasoned information from a computer processor has gone with the simplicity of knowledge based shells (KBS). The only doubtfulness is extracting the information from the expert. In the case of the mathematical analysis, codes and formulas are used with confidence. However, in the case of market feel, there are no formulae but reasoning on different sets of circumstances at that time. The authors plan to carry out a series of workshops that will indeed produce the required information in a structured format, for use within a prototype model. This may be brought to light in future publications.

As part of the overall research into the development of the indices for the heavy civil engineering industry, a forecasting strategy for these indices will have to be introduced and that is the prime objective of this paper. The remainder of this paper discusses and introduces forecasting techniques and a methodology for integrating the subjective and objective styles. The subjective refinement utilised as part of the joint approach is also reviewed and detailed in other publications by the author.

FORECASTING TECHNIQUES

As aforementioned there are several techniques available for forecasting time series data. The approaches tend to fall into two overall categories, either analytical or subjective. In this section the author will be reviewing the most common of the techniques available in the business management field.

Review of the subjective techniques

Subjective techniques are those based on the judgement of 'experts'**.** There are several different forms of subjective forecast and they are reviewed in a proceeding paragraph. The techniques are based on the knowledge of the persons working closely with the element they are forecasting. The predictions are usually fairly accurate and should not be considered unusable because there is little or no mathematical background to the theories. In fact a lot of the companies in the UK only use judgement in their predictions (Bates 1997) as they found statistical methods to be far more cumbersome, expensive and no more accurate (Sparkes and McHugh 1984).

Normally the forecasts are produced by people who are directly involved with the item they are forecasting, and have a sixth sense about how they are going to move in the future. The familiar techniques can be found in any judgmental forecasting text (Wolfe and Flores 1990) and are summarised below. Table 1 (shows the judgmental techniques):

Table 1: Subjective forecasting techniques

Subjective method	**Description**
Sales force estimates	Sales force estimating is the method where the salesperson is required to produce an estimate of future movements in the market.
The survey method	This method is extremely simple to carry out with very little expertise required. Basically it asks customers their opinions.
The average judgement method	Based on the judgement of personnel concerned with the business
The Delphi method	Based on a questionnaire approach given to a panel of experts, where the answers from the first round of the questionnaire are applied to develop the next.
Identification methods	Based on identifying who the experts are.
Graph and table methods	Based on eyeball judgements of previous data, either by graph or by tables of results

Review of the objective techniques

As the prime intent of this paper is to produce a joint approach to forecasting in the heavy engineering industry, it is best to review the mathematically-based methods also. The statistical approaches are the backbone to the conjunctive method with the subjective evaluation being the refining element. The judgmental methods are normally applied in a polishing capacity to the analytically forecast data. In later publications a more detailed structure for applying the subjective refinement will be introduced. Nonetheless, as we are trying to produce a statistically sound model a prose of the most common analytical methods is surveyed in Table 2.

JOINT APPROACH METHODOLOGY

What the authors plan to show here is a systematic methodology for applying subjective intervention within a statistical procedure. The model is made up of an objective semi-complex method which allows the use of judgmental intervention, so as to refine the outcome. The authors will be looking in great detail at its procedure and a simple application to the heavy engineering industry.

The semi-complex approach utilised is the Classical Time Series Decomposition (Proportionate Model). Decomposition refers to the breaking down of the data into workable components, then aggregating the prediction of each component to form an

Table 2: Objective forecasting techniques

Objective method	**Equations**	**Description**
Moving average models	$\hat{Y}_{t+1} = \frac{(\sum_{i=0}^{N-1} Y_{t-i})}{N}$	This approach is ideal for finding seasonal effects, by smoothing the peaks in the actual data
Smoothing models	$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$	This group of models is based on the equation(left) that can be used to smooth the actual series, the trend and the seasonality depending on which model.
Decomposition models	Y = T * S * C * I	This model breaks up the data into workable components, which are easier to forecast, however requires some expert intervention
Arima models / Box Jenkins	$z_t = f_1 z_{t-1} + \ldots + f_p z_{t-p} + a_t - t_1 a_{t-1} - \ldots - t_q a_{t-q}$	This is a very complex set of models which are made up of auto-regressive components and moving average components based on the generic equation(left)
Causal models	$Y_i = b_0 + b_1 X_i + b_2 X_{2i} + b_3 X_{3i} + \ldots + b_k X_{ki}$	These models are moving away from the projectionist type models by considering the factors that influence the overall factors movement
Learning models	$\hat{Y}_{t+1} = \sum_{i=1}^{N} w_i x_{t-i+1}$ $Y = \sum_{i=1}^{k} \alpha_i f(\sum_{j=1}^{n} \gamma_j X_j)$	These models fall into two categories where the weights applied to the actual series for learning or as weights to the influencing factors. (Neural Networks)

overall prognosis using Equation 1 and can be carried out by following the steps discussed (vide infra).

$$\mathbf{Y = T * S * C * I}$$

Equation 1: Classical Time Series Decomposition (Y actual series; T trend; S seasonality; C cyclicality; I randomness.)

Extracting the trend

The trend may be extracted by using least squares regression to create the ***b*** coefficients from the top line of 1. The proceeding lines in the matrix represent simultaneous equations which are used to formulate these coefficients. Normally in this industry the series tend to be linear therefore, $\boldsymbol{b_{N>1} = 0}$.

	Y =	$\mathbf{b_0}$	$\mathbf{b_1X}$	$\mathbf{b_2X^2}$	$\mathbf{b_3X^3}$	$\mathbf{b_kX^k}$
1	ΣY	$b_0 * n$	$B_1\Sigma X$	$B_2\Sigma X^2$	$b_3\Sigma X^3$	$b_k\Sigma X^k$
X	ΣXY	$b_0\Sigma X$	$B_1\Sigma X^2$	$B_2\Sigma X^3$	$b_3\Sigma X^4$	$b_k\Sigma X^{k+1}$
$\mathbf{X^2}$	$\Sigma X^2 Y$	$b_0\Sigma X^2$	$B_1\Sigma X^3$	$B_2\Sigma X^4$	$b_3\Sigma X^5$	$b_k\Sigma X^{k+2}$
$\mathbf{X^3}$	$\Sigma X^3 Y$	$b_0\Sigma X^3$	$B_1\Sigma X^4$	$B_2\Sigma X^5$	$b_3\Sigma X^6$	$b_k\Sigma X^{k+3}$
$\mathbf{X^k}$	$\Sigma X^k Y$	$b_0\Sigma X^k$	$b_1\Sigma X^{k+1}$	$b_2\Sigma X^{k+2}$	$b_3\Sigma X^{k+3}$	$b_k\Sigma X^{2k}$

Matrix 1: The matrix for applying the best fit to a series of points based on polynomial equations.

Developing seasonality
This component produces seasonal variations about a unity value and also reduces some of the randomness of the series. The stages required to achieve the values are shown next:

1. Find the moving average of the data (using a suitable seasonal period such as twelve months)
2. Divide the actual historical indices by the moving average values for the same period
3. Find the average seasonal factor for each period over the whole term
4. Normalise the seasonal factors (by multiplying them by the reciprocal of the mean of the factors calculated in 3.)

Reviewing cyclicality
The cyclicality (still including some randomness) may be valued by simply rearranging Equation 1 to give Equation 2.

Cyclicality = actual indices ÷ [trend x seasonality]

Equation 2: Rearranged Decomposition

Forming the forecast
Aggregating the components to form the forecast is extremely simple. However, the cyclicality component which maybe refined subjectively with varying levels of complexity is not. An overview of the subjective approaches that maybe applied will be shown with reference being made to internal papers by the author. The analytically forecast model has three fundamental components:

1. The trend line which should be extrapolated into the future (Figure 1)
2. The seasonality component, based on the repeated seasonal factors continued into the future (Figure 2)
3. The cyclicality which should be values that follow the general movement of the business cycle (Figure 3) An element of judgement is applied to this element.

Subjective intervention
Firstly, we require expert opinion no matter what level of complexity. Nevertheless, how do we find an expert? One person may be accurate in some cases but not all. Therefore, a group consensus should be considered as the generic opinion (Moore and Miles 1991).

Bates (1997b), relating an initial estimate to an out-turn cost, used subjective influences to improve the estimate. The improvement is reached through the development and understanding of how some risk factors create a percentage addition to project cost, based on expert opinion. There are three techniques available:

1. Number of Factors
2. Weighted Number of Factors
3. Cumulative Factor Distributions

Each of the above methods increases in complexity from top to bottom and any of them may be applied as part of the refinement to our analytically produced index.

However, the factors used in Bates (1997b) are related to the industry in which the estimate is being made and the specific project influences. In this element of the prediction, the influences are related to the market movements and will be known as market relationships (MR). There are several market relationship factors, some of which are outlined below:

- Government variations and changes
- Inflation changes
- Interest rate changes
- Legislation

The market refinement influences have distributions related to the cyclicality found through knowledge elicitation. Therefore, cyclicality may be related by Equation 3.

$$\mathbf{Cyclicality} = \mathbf{1} + \left[\sum_{\mathbf{K}=\mathbf{1}}^{\mathbf{M}} \phi_{\mathbf{K}} \times \mathbf{E}\left[\mathbf{MR}\right]_{\mathbf{K}} \right]$$

Equation 3: Cyclicality Movement (ϕ_K Represents the weightings of each of the factors)

$E[MR]_K$ Represents the expected addition value for each factor found through knowledge elicitation.

The component inside the square braces represents the variation about the unity value for cyclicality as an addition or subtraction. Furthermore, until these factors have been investigated, a detailed evaluation such as this cannot be carried out. However, some of the simpler approaches may still be applied. The example later is based entirely on the judgement of the author, thus showing with even simple judgement an increase in accuracy may be achieved.

APPLICATION TO THE HEAVY CIVIL ENGINEERING INDUSTRY

Forecasts of heavy engineering indices tend to be sparse. Their existence, apart from the very general indices in the Cost Engineer or those produced for mechanical items by the Department of the Environment, is very limited and their accuracy has also been questioned. The author wishes to put forward some example data from the heavy civil engineering industry and forecast the information using the joint approach introduced above. Moreover, as a comparison, predictions were made using some of the objective techniques above and a subjective forecast was supplied by an expert in the field of heavy civil engineering estimating. The information regarding the indices can be broken into its constituent components so that a forecast can be produced. A visual representation of this decomposition is shown subsequently.

Figure 1 shows the actual data series, the extracted ‘linear’ trend and the twelve-month moving average used to formulate the seasonality factors as explained above. Figure 2 shows the seasonal variation. Cyclicality is the final component and where the subjective evaluation may be introduced. The historical variations are shown in Figure 3.

Figure 3 is the most important graph of all the above due to it being the graph to which experts should refer for their background data so as to make an informed prediction for the future values. The cyclicality values by their variation can be

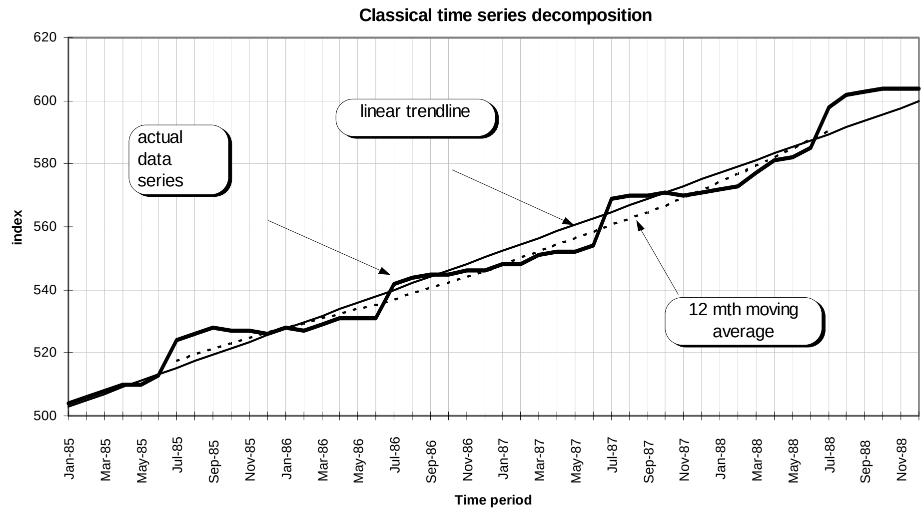

Figure 1: An overall breakdown of the data series.

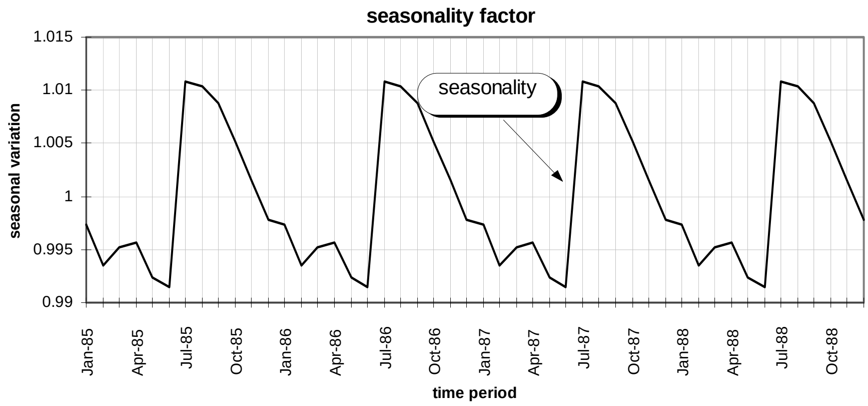

Figure 2: Seasonal variation repeating over the term

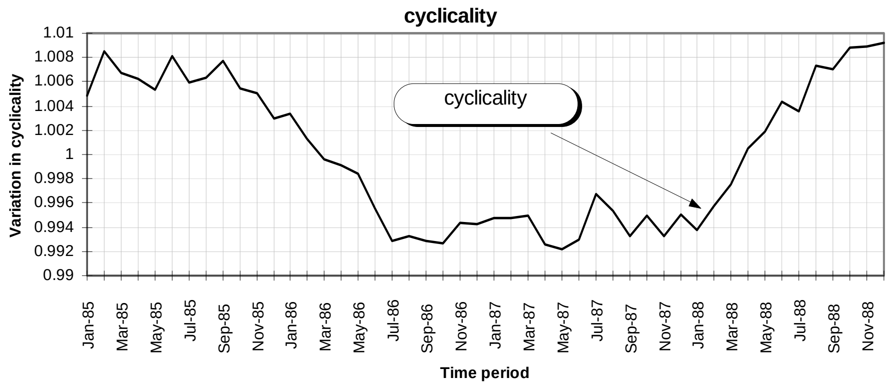

Figure 3: Variation in cyclicality (including noise)

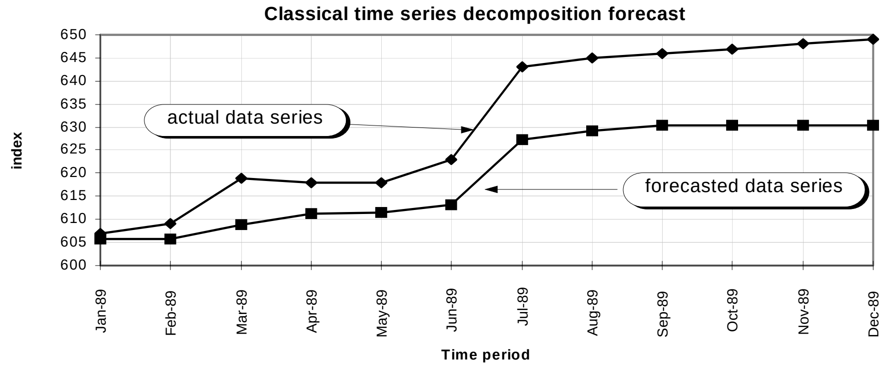

Figure 4: Actual and forecast indices based on CTSD historical model

shown in Figure 5 to cause noticeable differences to the forecast of the indices. Figure 4 shows a forecast where the chosen values of cyclicality had no real subjective evaluation applied and were assumed to be unity.

Figure 4 shows that the forecast plot follows well the actual data series, however slightly underestimated. This underestimation is due to the prediction of the values for cyclicality. Using any of the above subjective evaluation methods an improved value for cyclicality can be predicted. The data shown in Table 3 is represented in graphical form in Figure 6. Figure 6 shows how each of the methods compares in forecast-ability. Figure 5 shows how a slight variation on cyclicality through simple evaluation can improve the forecast significantly. If a more structured approach were to be applied a more improved forecast would be achieved. Figure 6 proves that complexity does not improve accuracy. Furthermore, in visual terms Figure 5 shows the difference in changing the decomposed cyclicality component from unity to the last value estimated from the model in Figure 3. This shows that with simple evaluation the forecast can be improved. In addition, with expert judgement of the market this may be improved again.

CONCLUSIONS AND FUTURE WORK

From visual inspection of the graphs and table above it is obvious that accuracy is not directly related to complexity with regard to forecasting, with our joint approach comparing well with some of the complex methods but with greater understanding. Further, combination forecasts leave the necessary room for improvement and show that a little judgement can improve the accuracy significantly. The expert opinion alone tended to underestimate the situation especially the seasonal jump in the data.

The heavy engineering industry is going through cultural changes in contract set-up and execution strategies. Alliances and incentives are the buzz-words. Initiatives like CRINE and ACTIVE hope to achieve cost reductions on project cost, related to the budget. However, if the budget is not accurate then cost reductions cannot be measured as there is no benchmark. This is where indices are most useful. Therefore, this area of change shows the need for accurate forecasting and model development.

As part of the overall research further aspects of the field of cost engineering will be looked into in an attempt to develop an index and a forecasting strategy for the heavy engineering industry. The indices and forecasting approach will form part of the makeup for an out-turn cost development from an initial estimate.

Date index produced	Actual index value	Expert estimate	Double moving average	Double exponential smoothing	Winters seasonal	Box Jenkins	CSTD Cyc=1	CSTD Cyc=last
Jan-89	607	604	608	608	603	605	600	606
Feb-89	609	604	610	610	605	606	600	606
Mar-89	619	607	613	612	606	609	603	609
Apr-89	618	608	616	614	609	612	605	611
May-89	618	610	618	616	611	612	605	611
Jun-89	623	611	621	618	613	614	607	613
Jul-89	643	620			618	627	621	627
Aug-89	645	625			628	629	623	628
Sep-89	646	625			630	630	624	629
Oct-89	647	626			631	631	624	629
Nov-89	648	627			632	631	623	629
Dec-89	649	629			632	631	623	629

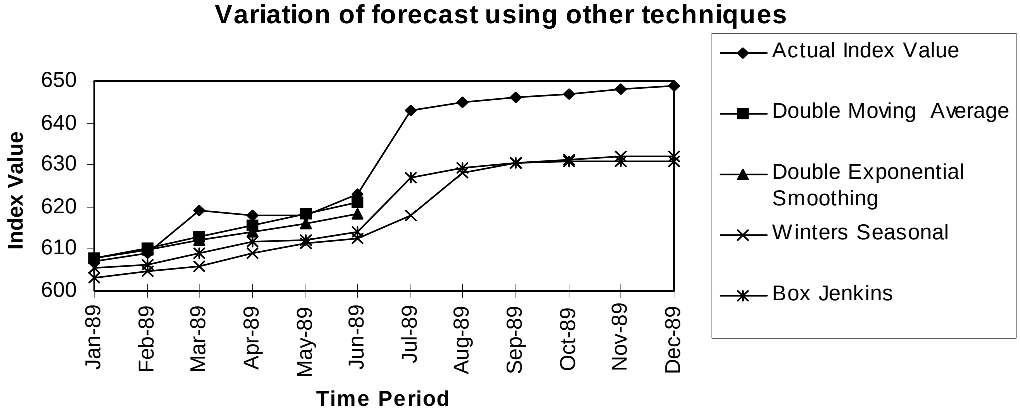

Figure 5: Graphical Representation of Tabulated Data

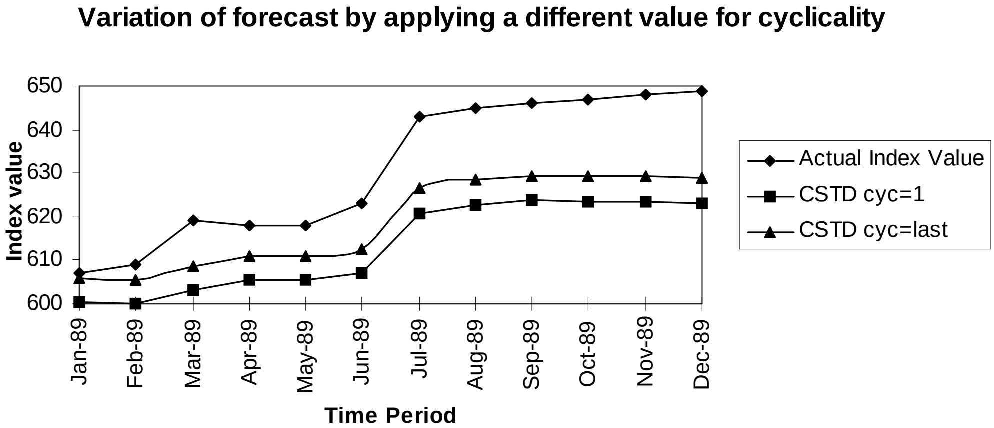

Figure 6: Comparison in forecast through alteration of the cyclicality value

All of the methods outlined in this paper will be harnessed as part of an expert system known as TAROT which will utilise a statistical forecasting package known as Runes for Excel and MSACCESS databases to formulate the forecasts mentioned. Further publications on the development of out-turn cost figures, expert systems and risk allocation will be produced in the coming months.

REFERENCES

Bates, W (1996) *Internal report on statistical forecasting methods and their application to the Heavy Civil Engineering*, University of Teesside, Middlesbrough.

Bates, W (1997a) *Internal report on the estimating and forecasting practices in the UK.* University of Teesside, Middlesbrough.

Bates, W (1997b) *A risk allocation strategy for construction activities in the process industry,* University of Teesside, Middlesbrough.

Box, G.E.P., and Jenkins, G.M. (1970) *Time series analysis forecasting and control,* San Francisco; Holden Day.

Bunn, .D and Wright, G. (1991) Interaction of judgmental and statistical forecasting methods: issues and analysis. *Management Science*, **37**(5), 501-518.

Dawood, N.N. and Neale, R.H. (1993) Forecasting the sales of pre-cast concrete building products, *Construction Management and Economics*, **11**(2), 81-98.

Edmundson, B., Lawrence, M. and O'Connor, M. (1988) The use of non-time series information in sales forecasting: a case study. *Journal of Forecasting*, **7**, 201-211.

Fitzgerald, E. and Akintoye, A. (1995) The accuracy and optimal linear correction of UK construction tender price index forecasts. *Construction Management and Economics*, **13**(6), 493-500.

Moore, C.J. and Miles J.C. (1991) Knowledge elicitation using more than one expert to cover the same domain. *Artificial Intelligence Review*, **5**(4), 255-271.

Pereira, B.D. et al. (1989) Experience in combining subjective and quantitative forecasts. *Journal of Forecasting*, **8**(3), 343-348.

Sparkes, J.R. and McHugh, A.K. (1984) Awareness and use of forecasting techniques in British industry. *Journal of Forecasting*, **3**, 37-42.

Wolfe, C. and Flores, B. (1990) Judgmental adjustment of earnings forecasts. *Journal of Forecasting*, **9**(4), 389-405.

Woodward, S., et al. (1994) Cost Management: The Software Solution, *Construction Papers 31*, Ascot; CIOB.

Wright, G. and Ayton, P., (1987) *Judgmental forecasting*. London; Wiley.