SENSITIVITY OF SIMULATED WHOLE-LIFE COSTS TO INPUT PROBABILITY DISTRIBUTIONS

Mohammed Kishk[1] , Robert Pollock[1], and Eghosa Okhuahesogie[2]

[1]*The Scott Sutherland School, The Robert Gordon University, Garthdee Road, Aberdeen AB10 7QB, UK*
[2]*School of Engineering, The Robert Gordon University, Schoolhill, Aberdeen AB10 1FR, UK*

In a typical Monte Carlo simulation (MCS), it is required to assign probability distribution functions (PDFs) for uncertain parameters. Due to lack of or irrelevance of historical whole-life costing (WLC) data, these functions are usually assumed. The research work that underpins this paper aimed to identify the significance of errors in various input parameters in a WLC exercise and to identify critical parameters in assigning a probability distribution. In carrying out the study, a number of case studies have been identified and analysed using a risk assessment software. Besides, a number of metrics to assess the impact have been identified and employed. Typical results show that varying the type and range of input PDFs only affect the certainty level of simulated WLC results while variations in the standard deviation affected the risk ranking of uncertain input variables. The most significant impact was that due to variations in the mean of input PDFs, which does not only affect the certainty of the decision, but can even change the decision altogether.

Keywords: Decision-making, Monte Carlo simulation, probability distributions, sensitivity analysis, whole-life costing.

INTRODUCTION

Whole-life costing (WLC) is highly dependent on an extensive amount of data. Assumptions and estimates are normally made whilst collecting data and predicting the behaviour of future events. As a result, the input data of WLC (life cycle, discount and inflation rates, initial cost, annual operation and maintenance costs etc.), which are based on estimates rather than known quantities is normally characterised with so `much uncertainty. Consequently, uncertainty assessment becomes crucial in WLC exercises.

Several techniques are available for risk and uncertainty assessment in a WLC exercise. The two most commonly used ones are the sensitivity analysis (SA) and the Monte Carlo simulation (MCS) techniques. The SA technique is a simple straightforward technique that is used to identify the impact on a change in a risky variable, keeping all other variables constant. However, in practise, this is not always the case as all risky and uncertain parameters can be expected to vary simultaneously. The MCS can be seen as an extension of sensitivity analysis as it takes explicit account of the fact that all risky and uncertain parameters can be expected to vary simultaneously. MCS is a means of examining problems for which unique solutions cannot be obtained and it has been used in WLC modelling by many authors (e.g. Flanagan et al., 1987, 1989; Ko et al., 1998; Goumas et al., 1999). However, it makes

[1] m.kishk@rgu.ac.uk

Kishk, M, Pollock, R and Okhuahesogie, E (2004) Sensitivity of simulated Whole-Life Costs to input probability distributions. *In:* Khosrowshahi, F (Ed.), *20th Annual ARCOM Conference*, 1-3 September 2004, Heriot Watt University. Association of Researchers in Construction Management, Vol. 1, 115-24.

the assumption that parameters subject to risk and uncertainty can be described by probability distributions. These PDFs are then used by the technique to generate the probability distribution of the dependant variable, usually the net present value in the WLC case.

To complete a MCS, therefore, it is required to assign PDFs for uncertain variables. Such functions are best derived from statistical analysis of significant historical data, which is not always available or inadequately defined. In such cases, PDFs are usually assumed. There is therefore the possibility that different WLC analysts will describe the uncertainties of the same project using different PDFs. Will these analyses come out with very different outputs that may change the decision taken? It therefore becomes necessary to assess the impact of varying the assumed input PDFs to WLC results. It is believed that this would help identify the significance of the errors or differences in various input variables in a WLC exercise and identify those parameters that require greater care in assigning a probability distribution. Sensitivity Analysis can be used to assess the impact of varying the PDFs.

The objective of the research work that underpins this paper was to study the effect of varying the parameters defining assumed PDFs of input variables in a WLC study on the simulated NPVs through a number of case studies. The rest of the paper is organized as follows. In the next section, the research methodology is outlined and a typical case study is introduced. Then, typical results are presented and analyzed. Finally, conclusions are drawn.

METHODOLOGY

Analysis Procedure

Crystal Ball 2000® (Decisioneering, 2000) was selected and used to analyse a number of case studies. This software offers an adequate PDF library, a user-friendly interface, and a good output scheme. The analysis procedure adopted in this study included the following six steps:

Step 1: an NPV Excel model is developed to simulate the mathematical WLC model.

Step 2: the preferences for running the simulation are set. Such preferences include the maximum number of trials, confidence level, and sampling method.

Step 3: assumptions are defined for uncertain input variables for the case at hand.

Step 4: the simulation process is run and the frequency and cumulative frequency count data and other relevant statistics are extracted from each simulation.

Step 5: steps 3 and 4 are repeated as required.

Step 6: results are analysed as detailed in the following subsection.

The Mathematical Model

The NPV model developed by Kishk (2001) was employed because it has been developed such that calculations are both automated and optimised. This is mainly facilitated by the derivation of automatic expressions for calculating the number of occurrences on non-annual recurring costs. In this way, these costs can be dealt with directly without the need to express each cost to a number of equivalent cash flows. Besides, compact expressions are formulated for various discount and annualisation factors. Furthermore, whole-life cost contributions of each cost can easily be followed.

It calculates the whole life cost of an alternative i, as

$$NPV_i = \sum_{l=1}^{nic_i} IC_{il} + \sum_{m=1}^{nno_i} PWO_{im} F_{im} + PWA \sum_{j=1}^{nar_i} A_{ij} + \sum_{k=1}^{nnr_i} PWN_{ik} C_{ik} - \sum_{v=1}^{nrv_i} PWS\, RV_{iv} \quad (1)$$

where NPV_i is the net present value of the alternative, $IC_{il}, F_{im}, A_{ij}, C_{ik}$ and RV_{iv} are initial, future one-off, annual-recurring, non-annual recurring costs and resale values of alternative i, and $PWO_{im}, PWA_{ij}, PWN_{ik}$ and PWS are discount factors given by

$$PWO_{im} = (1+r)^{-t_{im}} \quad (2)$$

$$PWA = \frac{1}{r}\left(1-(1+r)^{-T}\right) \quad (3)$$

$$PWS = (1+r)^{-T} \quad (4)$$

$$PWN_{ik} = \frac{1-(1+r)^{-n_{ik} f_{ik}}}{(1+r)^{f_{ik}} - 1} \quad (5)$$

$$n_{ik} = \begin{cases} \mathrm{int}\left(\frac{T}{f_{ik}}\right), & \text{provided that rem}\left(\frac{T}{f_{ik}}\right) \neq 0 \\ \frac{T}{f_{ik}} - 1, & \text{elsewhere} \end{cases} \quad (6)$$

where r is the discount rate and T is the analysis period.

Case Study

A number of case studies were used in this study. Because of lack of space, only one case study is reported. Input parameters for the selected case study are summarised in table 1. As shown, all the cost variables in equation (1) are included to ensure a wide scope for analysis. In this case, it is required to determine the whole life costs of a design alternative for an analysis period of 30 years. As shown, most cost variables are included to ensure a wide enough scope for analysis.

Table 1: Input variables for the case study

Input variables	Mean	Range
Initial cost (£'000)	480	432 - 528
Non-annual recurring costs (£'000)	250	175 - 325
Annual operating costs (£'000)	40	36 - 44
Discount rate (%)	4	2.8 - 5.2
Analysis period (years)	30	27 - 33
Frequency of repair	8	7 - 9
Salvage value (£'000)	200	150 - 250

Analysis Metrics

PDFs may be discrete or continuous. A discrete probability distribution describes distinct, finite; commonly integer values while a continuous distribution assumes all values in the range are possible. Because most construction industry activities are continuous (Flanagan and Norman, 1993), only continuous PDFs are considered in this study. These distributions are smooth solid curves that depend on one or more parameters. There are three basic types of parameters (Evans and Olson, 2002):

- A shape parameter controls the basic shape of the distribution. A distribution can have two shape parameters, e.g. the beta distribution;
- A scale parameter controls the unit of measurement or spread within the range of the distribution. Changing the scale parameter either contracts or expands the distribution along the horizontal axis; and
- A location parameter that specifies the location of the distribution relative to zero on the horizontal axis, e.g. the midpoint or the lower endpoint of the range of the distribution.

These parameters are defined by statistical terms such as mean, standard deviation, minimum, and maximum. All distributions will not have all three parameters; some may have more than one shape parameter. Understanding the effects of these parameters is important in selecting distributions as inputs to simulation models (Evans and Olson, 2002).

In addition to the mean, standard deviation, two extra metrics have been employed: the confidence interval, and the certainty level. A confidence interval is a bound calculated around a statistic that attempts to measure this error with a given level of probability (Decisioneering, 2000). Confidence intervals are important for determining the accuracy of the statistics, and hence the accuracy of the simulation. Typically, MCS software uses confidence intervals to determine when a specified accuracy of statistics has been reached, and then stops accordingly (Decisioneering, 2000).

On the other hand, certainty describes the percentage of simulation results that fall within a range. A certainty level shows the certainty of achieving the values within a specific range and it equals the shaded area bounded by that range (figure 1). Figure 2 illustrates the selection process using NPVs for the case of two competing alternatives. When comparing two competing options using their NPVs, the area of overlap, A_O, (figure 2) indicates a decision uncertainty region. In other words, an increase in the area of overlap means less certainty of decision for the best alternative and vice versa. In the limit, the certainty of the decision is absolute when there is no overlap.

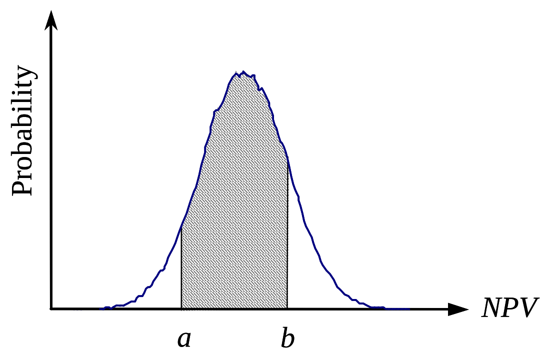

Figure 1: The concept of certainty level.

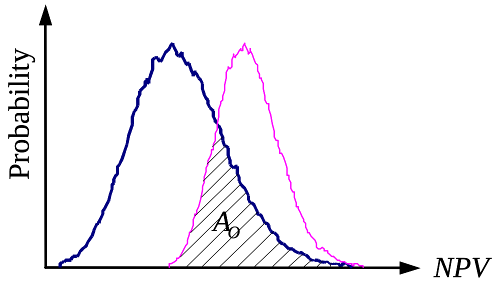

Figure 2: Comparing two competing alternatives using their NPVs.

Base Case

For the purpose of analysis and assessment of the impact of varying the input PDFs, a base case situation was considered where all input parameters are assumed to be certain, i.e. parameters like range, variance and standard deviation are equal to zero, and the mean, median and mode all have the same value. In this way, the resulting NPV distribution will reflect only the effect of the variation of the PDF of the input parameter under consideration.

RESULTS AND ANALYSIS

Effect of the Type of PDF

The effect of varying the type of probability distribution function of input variables on WLC results has been investigated. Five distribution types were considered for this study; they include: uniform, triangular, normal, lognormal, and beta distributions. These distributions are used to describe uncertainty in each input variable given in Table 1. Steps 3 to 6 of the analysis procedure were repeated, each time, changing the type of probability distribution function of an input variable while the other input variables were kept constant (certain).

Figures 3 to 6 show frequency distributions of NPVs resulting from varying the type of PDFs representing input cost variables including initial cost, annual operating cost; non-annual maintenance cost and salvage value, respectively. Although there are some shifts in the output distributions, there are no significant differences between the mean values/modes for various types of PDFs. The difference is greatest for annual costs, then initial cos. This may be attributed to the fact that annual costs contribute most to the NPV, followed by initial costs.

In general, the type of distribution has an effect on the certainty levels of the output net present values (NPVs). As expected, input uniform distributions resulted in the lowest certainty levels as indicated by the wide short output frequency distributions. In some cases, e.g. figure 3, the output frequency distributions for some distributions were practically the same; may be due to the small range estimation used in the analysis.

Therefore, it can be deduced that the type of distribution used to simulate input cost variables has no obvious effect on the output NPV mean; and hence on the ranking of competing alternatives. However, this might have an effect on the certainty of ranking.

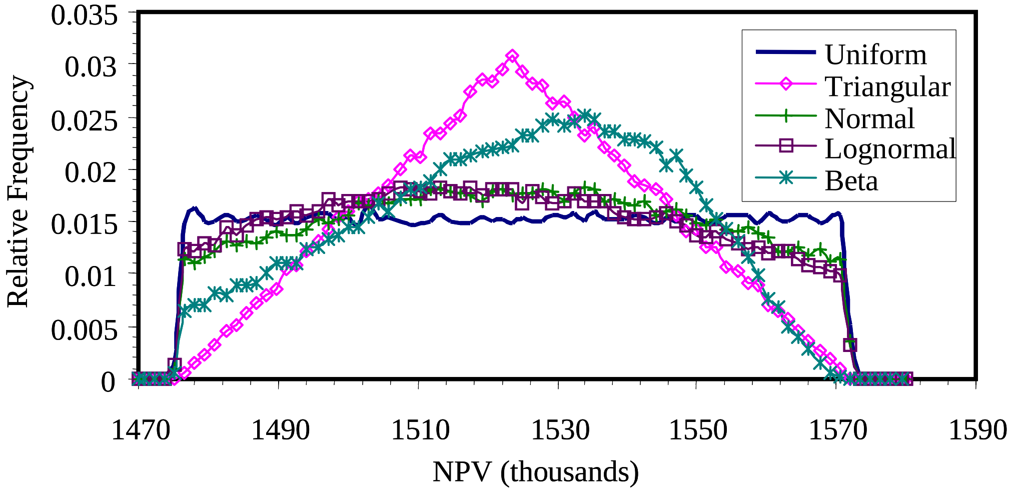

Figure 3: NPV distributions for various types of PDFs used to simulate the initial cost

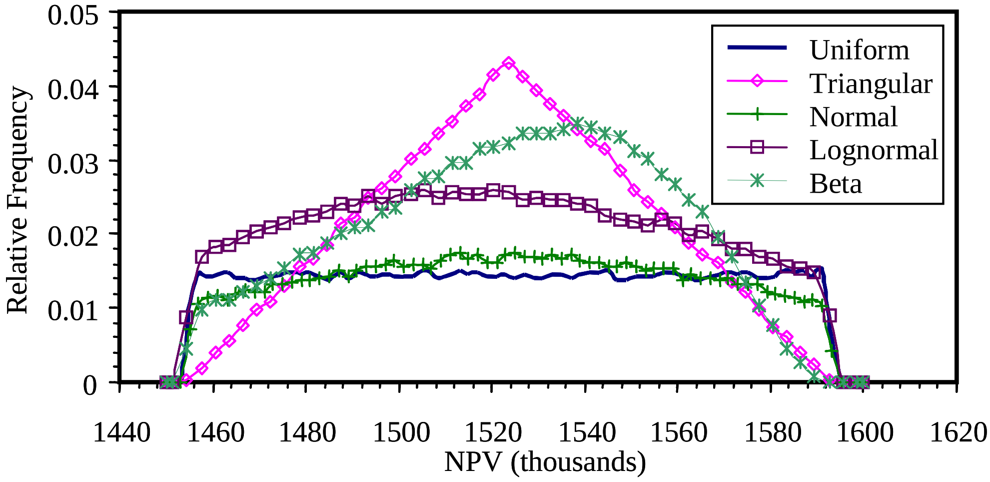

Figure 4: NPV distributions for various types of PDFs used to simulate annual operating costs

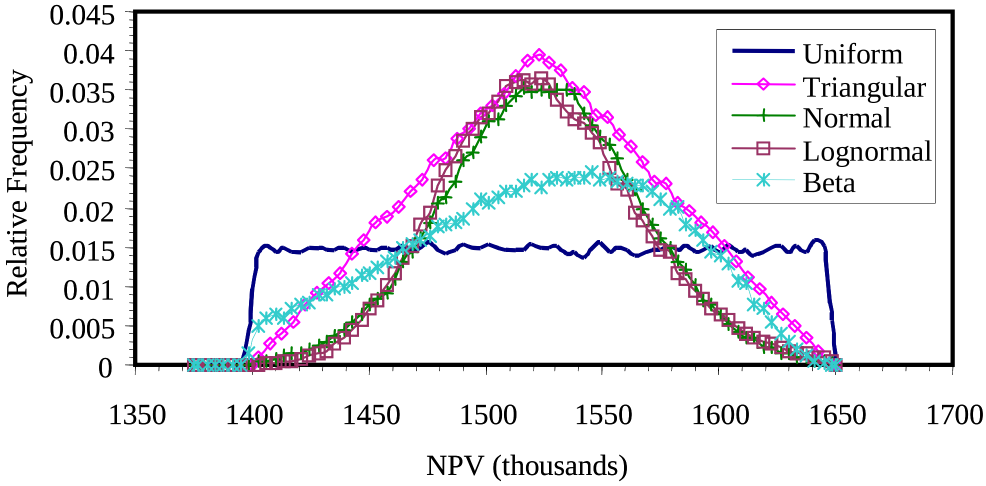

Figure 5: NPV distributions for various types of PDFs used to simulate non-annual costs

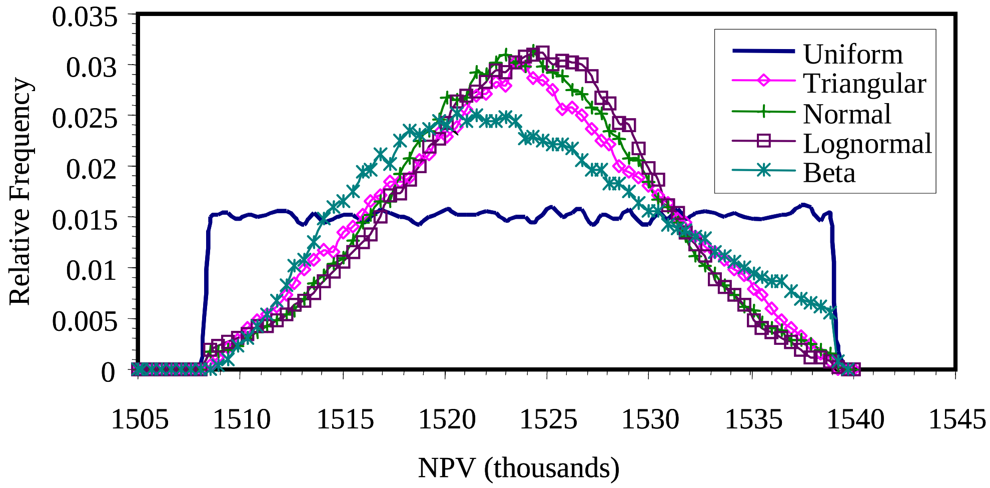

Figure 6: NPV distributions for various types of PDFs used to simulate the salvage value

The effect of the assumed type of PDF was also investigated for discounting input variables. Figure 7 shows the resulting NPVs frequency distributions for various PDFs used to simulate the discount rate. As shown, the resulting distributions are mostly right-skewed due to the discounting process; but with no significant change on the mean value or the output range can be observed. In general, a similar effect to that of input cost variables can be concluded.

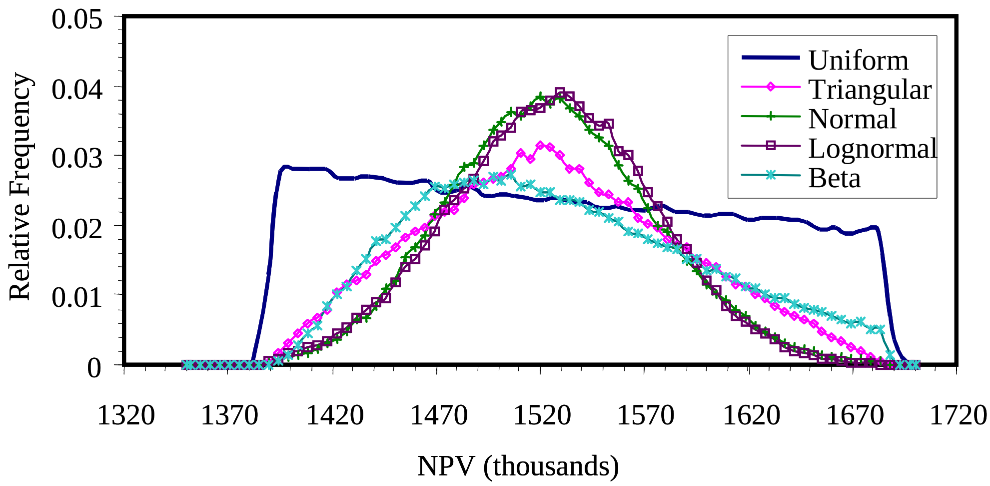

Figure 7: NPV distributions for various types of PDFs used to simulate the discount rate

Effect of the Assumed Mean

The normal distribution is usually regarded as the most important distribution in probability theory because it models many naturally occurring phenomena (Flanagan and Norman, 1993; Hayter, 2002). Therefore, it was selected and used for investigating the effect of errors in the assumption of the mean values and scatter of input cost and discounting data. Figure 8 shows the effect of varying the mean value of the non-annual recurring costs as a ratio of the certain base case denoted as the mean ratio (MR) in the figure. As expected, varying the mean value affected mainly the resulting mean outcomes as indicated by the regular shift of NPV frequency distributions. This means that a significant error in the mean value of a cost variable may not only affect the certainty of the decision, but can even change the decision altogether.

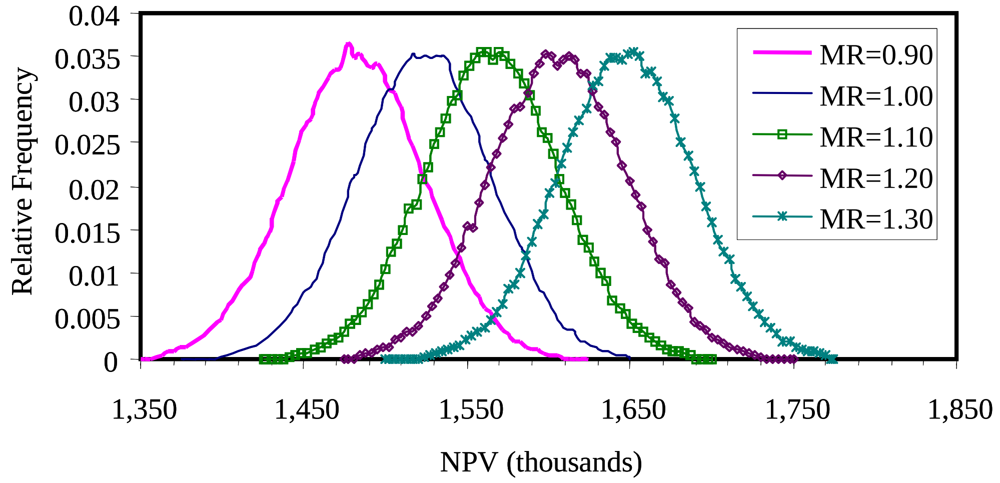

Figure 8: Effect of mean variation of non-annual recurring costs on NPVs

Equations 2 to 5 show an inverse relationship between the discount rate r and the discounting factors. This is reflected in figure 9 where an increase in the discount rate mean value lead to lower values for discounting factors which cause decreases in the WLC contributions of their associated cost variables; i.e. a corresponding reduction in the NPV mean value. It can be seen also from figure 9 that the shift the resulting NPV frequency distributions is less regular compared to the case of cost variables. This may be attributed to the non-linear discounting process as indicated by equations 2 to 5.

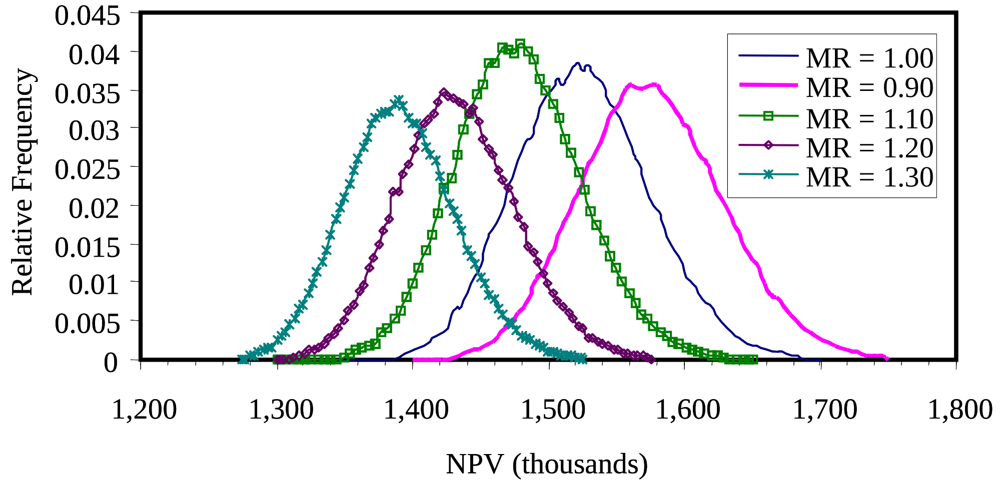

Figure 9: effect of mean variation of the discount rate on NPVs

Effect of the Assumed Standard Deviation

The standard deviation is a measure of the dispersion of the outcomes about the mean value and it is an indication of the uncertainty of a distribution. The effect of errors in the assumption of the standard deviation of input cost and discounting data was also investigated. Figures 10 and 11 show the effect of varying the standard deviation to mean (SDTM) ratio of non-annual costs and the discount rate, respectively. As shown, a change in the standard deviation only changes the shape of the probability distribution curve. It either becomes narrower or flatter. As shown, the smaller the standard deviation, the narrower the distribution graphs and the larger the standard deviation, the flatter the curve. Because the input range was fixed, the same range of the resulting NPVs can be noticed. An increase in the standard deviation would normally cause the graph to become wider. It can also be that there is almost no effect on the mean NPVs. Thus, errors in the assumption of standard deviations of input variables can only affect the certainty of the decision.

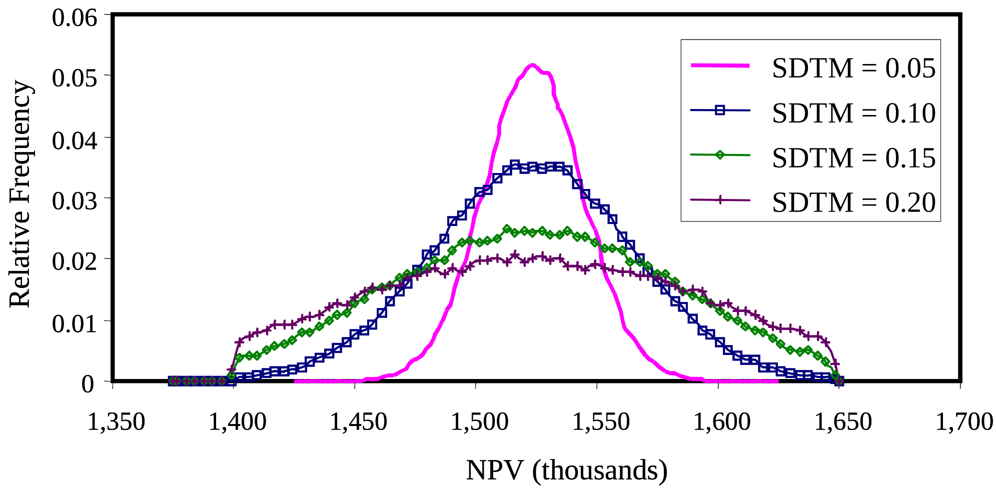

Figure 10: Effect of the assumed standard deviation of non-annual costs on NPVs

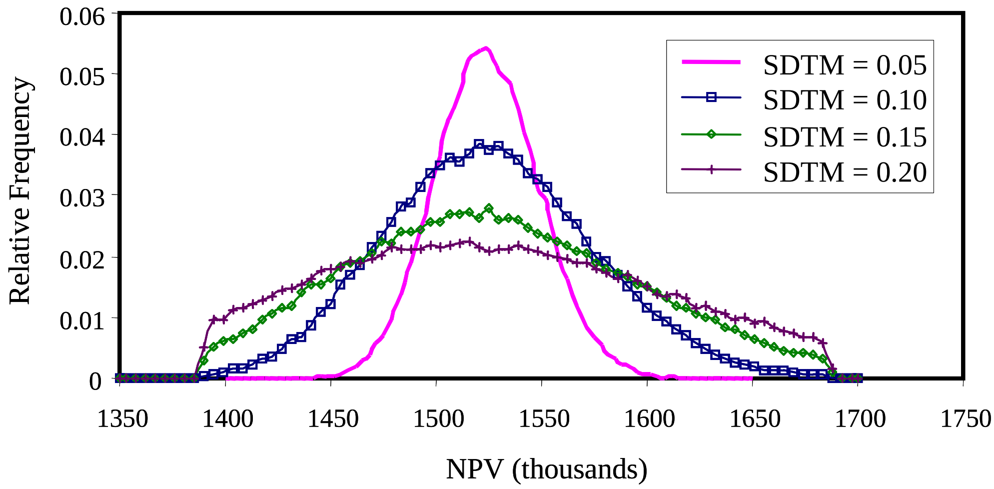

Figure 11: Effect of the assumed standard deviation of the discount rate on NPVs

CONCLUSIONS

Based on the research work that underpins this paper, the following conclusions can be drawn.

- The type of PDF used to describe uncertainty in input variables in a WLC analysis has no significant impact on the simulated output mean or range. In other words, it does not have a significant effect on WLC-based decisions. However, it can alter the certainty level of the decision.

- Errors in the assumption of standard deviations of normally distributed input variables can only affect the certainty of the decision.
- Variations in the mean value of input variables have the most significant impact on the output. It does not only affect the certainty of decision, but can change the decision altogether. In representing uncertain variables with PDFs, attention must be given to the mean value of PDF being used.

ACKNOWLEDGEMENT

The third author would like to acknowledge the financial support of the Petroleum Technology Development Fund (PTDF), a body mandated by the Nigerian government to provide scholarships for the training of Nigerian graduates.

REFERENCES

Decisioneering (2000) *Crystal Ball® 2000 user manual*. Denver: Decisioneering.

Evans, J R and Olson, D L (2002). *Introduction to simulation and risk analysis*. Second edition. New Jersey: Pearson Education Inc.

Flanagan, R, Norman, G and Furbur, D (1983) *Life cycle costing for construction*. London: Surveyors Publications.

Flanagan, R, Kendell, A, Norman, G and Robinson, G (1987) life cycle costing and risk management, *Construction Management and Economics*, **5**, 53-71.

Flanagan, R, and Norman, G (1993) *Risk management and construction*. Oxford: Blackwell Scientific.

Goumas, M G, Lygerou, V A and Papayannakis, L E (1999) Computational methods for planning and evaluating geothermal energy projects. *Energy Policy*, **27**, 147-154.

Hayter, A J (2002). *Probability and statistics for Engineers and Scientists*. Second edition. Pacific Grove, CA: Thomson Learning Incorporated.

Kishk, M (2001) *An integrated fuzzy approach to whole-life costing based decision making*. Unpublished PhD Thesis. The Scott Sutherland School, The Robert Gordon University, Aberdeen, UK.

Ko, W I, Choi, J W, Kang, C H and Lee, J S (1998) Nuclear fuel cycle cost analysis using a probabilistic simulation technique. *Annals of Nuclear Energy*, **25**, 771-789.

Metzger, J. N., Fjeld, R A, Hammonds, J S, and Owen, F H (1998). Software review: Evaluation of software for propagating uncertainty through risk assessment models. *Human and Ecological Risk Assessment*, **4**(2), 263 – 290.