MODELING SCHEDULING UNCERTAINTY
IN CAPITAL CONSTRUCTION PROJECTS

Nathan D. Boskers

Development Manager
Springwood Developments Inc.
12415 Jasper Avenue
Edmonton, AB T5N 3X9, Canada

S. M. AbouRizk

Hole School of Construction
3-133 Markin/CNRL Natural Resources Engineering
Facility
University of Alberta
Edmonton, AB T6G 2W2, Canada

ABSTRACT

Capital infrastructure projects with long-term implementation time frames are generally uncertain in nature. Engineers and planners attempting to estimate the costs of such projects often resort to using contingencies based on their experience without proper modeling of the uncertainty of costs, durations, or economic conditions. This paper presents a simulation-based model for assessing uncertainty associated with these projects. In particular the model accounts for expected fluctuations in the costs and durations of various work packages and, most significantly, it accounts for the inflation of costs over time based on when the work packages occur. The model uses Monte Carlo simulation techniques to account for time and cost and uses non-stationary time series modeling techniques to predict inflation rates. The model is implemented as a special purpose simulation template available in the public domain.

1 INTRODUCTION

The analysis of a project in its early stages is done for various reasons. An organization may have capital readily available to invest and want to determine which project, if any, should be selected to pursue. This analysis is often referred to as project selection analysis. Another reason for analyzing a project in its early stages is to determine the optimal procurement strategy to use in carrying out the project's work packages and activities. Whatever the reasons are for capital investment analysis or project procurement analysis, analysis done during the early stages of a project is extremely important because the decisions made based on this analysis will have significant impacts on the outcome of the project. Thus it is crucial for the organization to use vigorous project analysis techniques.

The Internal Rate of Return (IRR) method is typically endorsed for the purpose of analyzing projects for selection, as it is the most common and understood method in industry (Chen 1998). However, all project analysts do not support this suggestion. Others have concluded that the most appropriate method for project selection is the Net Present Value (NPV) method because the IRR method cannot take size differences in the investments into account, and it assumes that all funds received during the project can be immediately reinvested at the same interest rate as the IRR. In fact, as the IRR increases the chance that additional projects are available for reinvestment at that rate decreases (Winston 1995). Another study on project analysis methods for the purposes of determining the most appropriate analysis method for build-operate-transfer type projects broke the project analysis methods into three categories: the return methods (i.e. payback period, IRR, NPV), the risk methods (i.e. risk rating systems), and the risk-return methods (i.e. utility theory, decision trees, NPV at discount rate determined by the weighted average cost of capital). This study concluded that all of the current methods had inherent problems with them, and developed the NPV-at-risk method (Ye et al. 2000). A similar study analyzing build-operate-transfer type projects combined the project-at-risk methodology with the NPV methodology, then applied the Monte Carlo analysis to the project parameters and produced a NPV distribution as the output (Lu et al. 2000). Yet another study applied Monte Carlo analysis using the spreadsheet add-in *Crystal Ball* to project costs for the analysis of a solid waste facility. The optimal solution was the minimization of risk, which was achieved when the expected NPV was maximized and the NPV variance and standard deviation for the NPV distribution were minimized (Wajs et al. 2000). A study analyzing construction costs used Monte Carlo analysis to reduce the risk due to uncertainty in estimated construction activity costs using historical data to create cost ranges, which were then used in the Monte Carlo analysis (Cochrane 1992). This study used the spreadsheet add-in *@Risk* to perform the analysis.

Many of the commonly used analysis techniques require the forecasting of future project cash flows as input into the analysis. However, a problem arises when using these techniques in the forecasted future cash flow because the analysts assign fixed values to the activity durations. The cash flows are affected by the duration of the activities (assuming that the costs are spread evenly over the duration) and therefore are directly affected by uncertainty in the activity durations. Some analysts apply Monte Carlo methods to the costs to account for some of this uncertainty. However, although the prediction of future durations is as uncertain as the prediction of costs, neither prediction's degree of uncertainty is accounted for. Another problem arises because the analysts do not properly account for the scheduling of the project's work packages and activities. This represents another shortcoming in the current techniques, because the time at which the activity occurs in the future will have an impact on the cash flow, which in turn will affect the analysis output. A possible reason that the analysts have not accounted for the uncertainty in the duration of activities as well as the scheduling of activities could be due to the limitations created by modeling the project costs in spreadsheet programs.

This paper presents a simulation based-model for assessing uncertainty associated with capital infrastructure projects with long implementation time frames. In particular, the model accounts for expected fluctuations in the costs of various work packages, fluctuations associated with the durations of various work packages, and the timing of activity occurrences, and incorporates the inflation of costs over time based on when the work packages occur. The model uses Monte Carlo simulation techniques to account for time and cost while it uses non-stationary time series modeling techniques to predict inflation rates. The model is implemented as a special purpose simulation (SPS) template available in the public domain.

2   SUMMARY OF THE MODEL

A methodology that incorporates uncertainty in estimating capital construction projects spanning multiple years (e.g. interchange conceptualization in 2002, to be constructed in 2008) is shown in Figure 1. The methodology makes use of Monte Carlo simulation principles, within the context of a continuous simulation model where increments in time are uniform (e.g. one period is a year or a month), and inflation applied at a given period is calculated using the Box-Jenkins approach. The remainder of this section discusses the components of the developed methodology.

2.1 Break Project into Work Packages and Activities

The first two steps of the developed methodology are similar to creating a project schedule. In the first step, the analyst should break the project into work packages. By

breaking the project into work packages, the analyst is defining the major stages or phases of the work that is to be done (phases of a project lifecycle). Most projects have similar phases to their lifecycle, no matter the project type. In general, a project begins with conceptual planning and then moves on to design, construction, operation, and maintenance. Although other work packages may be defined, these are the main components on a very large scale.
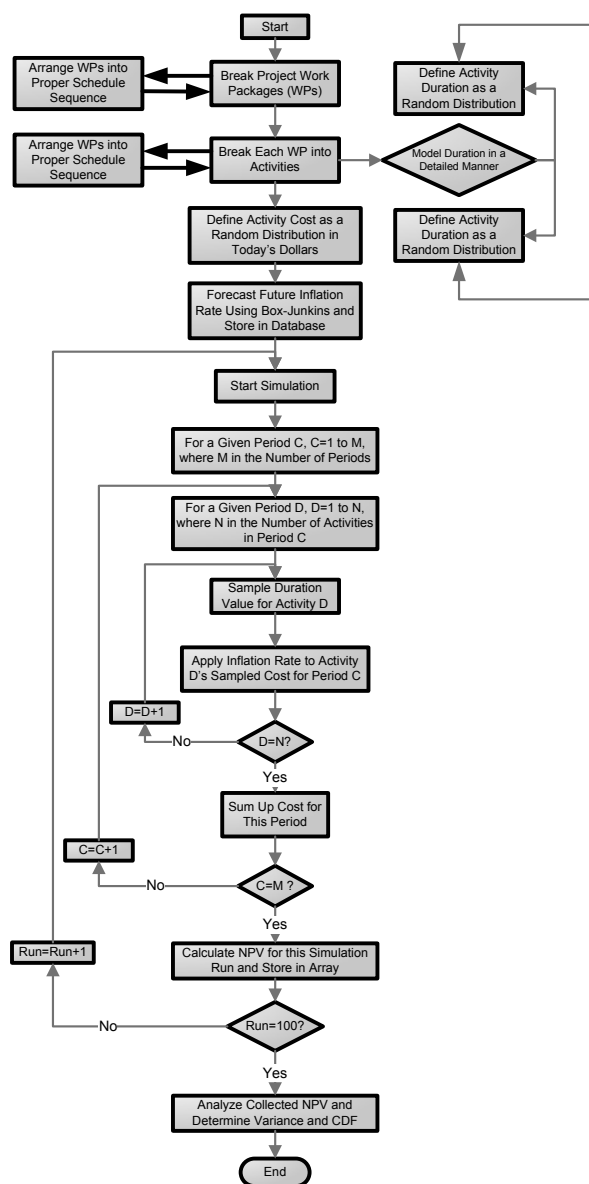

Figure 2. Project Hierarchical Modeling

Once the work packages have been defined, they should be arranged in such a way as to represent their lifecycle occurrences. This is done by developing a project schedule for the work packages, creating a network type schedule. This

establishes a proper network so that network analysis can be done on the project.

The second step of the developed methodology is also similar to creating a project schedule. The analyst should break each work package into further work packages, or into activities. The purpose of this step is to allow the analyst to incorporate as much detail into the analysis as desired. For example, with a construction work package, the analyst could break it into sub-work packages such as earthwork, foundations, superstructure, and so on. The analyst could then break these sub-work packages into activities. For example, the foundations sub-work package could be broken down into forming, pouring, and finishing activities. This concept is known as hierarchical modeling. Hierarchical modeling allows the analyst to create a model with as much or as little detail as desired. Figure 2 gives an illustration of how hierarchical modeling works. The diagram shown in Figure 2 corresponds to the first two steps in Figure 1. It should be noted that in hierarchical modeling, the number of levels that exist are endless; Figure 2 shows only three levels for the sake of brevity.
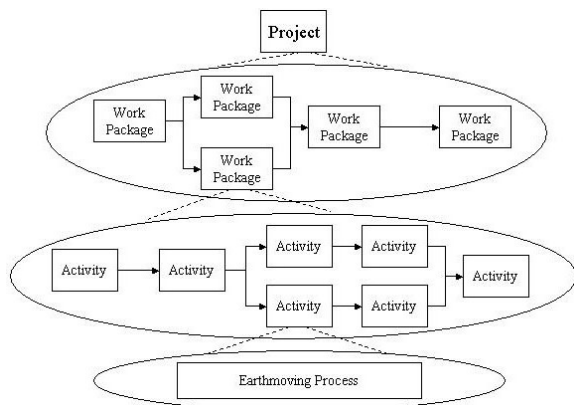

Figure 2. Project Hierarchical Modeling

The analyst is expected to arrange the activities in such a way as to represent their proper sequencing within the work package to which they belong. This can be accomplished by creating a simple network diagram (i.e. CPM schedule) and by incorporating proper network analysis into the overall project analysis.

The final task in the second step is to decide if the activity's processes should be modeled. The purpose of modeling an activity's processes would be to get more detailed and accurate output in the analysis. By modeling an activity's processes, the duration of the activity is determined by the finish time of the process, instead of using a user-defined duration value. This accounts for additional uncertainty that exists in the duration value of future activity times. If the analyst does not want to model the processes of an activity, a probability distribution may be assigned to the duration value.

2.2 Define Activity Costs

The third step in the developed methodology is to define probability distributions for the activity's costs. Later, the activity cost will be determined by sampling from the assigned distributions. By doing this, some of the uncertainty associated with the future unknown costs of the activities is accounted for.

It should be noted at this point that the values used in this step are in current dollars, where a cash flow in current dollars is defined as one in which the escalation of costs due to inflation is not incorporated, and a cash flow in actual dollars is one in which the escalation of costs due to inflation is incorporated. The escalation of costs due to inflation will be incorporated into the analysis at a later point. The values used in this step come from cost estimates performed by expert estimators, and are usually based on historical data, such as productivity rates.

2.3 Forecast Future Inflation Rate

The fourth step requires the analyst to forecast the future inflation rates of the project. Here the future inflation rate is forecasted using the Box-Jenkins time-series analysis technique. The future inflation rate is required to calculate the future cash flow in actual dollars, which is performed later in the methodology.

2.4 Calculate Time Period Cash Flows

The next steps are where the calculations in the analysis begin. Up to this point, the scheduling of activities has been defined. The cost and duration for each of the activities has also been defined as distributions. As well, inflation rates have been forecasted for each future time period. The purpose of this step is to calculate the cost of each activity in actual dollars and then sum up the total activity costs for each given time period. Basic engineering economic formulas are used to convert the activity's current cost into an actual cost.

For the purposes of project analysis, which will be described below, the timing of the actual costs needs to be determined. Therefore, network analysis is performed on the defined work packages and activities. Here, the activity durations are determined by randomly sampling the value from the probability distribution, or by modeling the activity's processes and then determining the start and finish times of each activity. The start and finish time of each activity is used to break the current costs of the activity into parts corresponding to a specific time period. The partial activity costs for each time period then have their corresponding inflation rate applied to them to convert them to actual dollars. Figure 3 shows an illustration of an example project's activities and their relationships to the predefined project time periods.

The final task in this step is the summing up of all the expenses and revenues of the project in actual dollars for each time period. The purpose of summing up the project expenses and revenues for each time period is to allow cash flow analysis to be performed, which is the project analysis technique. See Figure 3 for an illustration of how the activity costs are summed up for each time period.
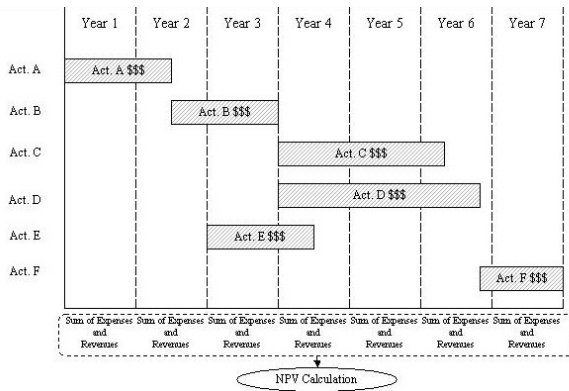

Figure 3. Bar Chart Schedule with Cash Flows

2.5 Calculate and Collect the Project NPV

The next step is to generate the output, which will be used for the project analysis. The method performed to generate the output is traditional discounted cash flow analysis, which is a standard project analysis technique.

The term project cash flow describes the amount of money in a given time period after a project's revenues, expenses, and taxes have been summed. This dollar amount may be zero, positive, or negative. Using cash flow analysis is important when project financing costs are affected by interest rates and project costs and revenues are affected by inflation rates. Both interest and inflation rates affect the overall total project cash flow. Additionally, the value of an asset is the extent that it may be expected to generate cash in the future. This value can only be determined through cash flow analysis by estimating the future cash and discounting it to the present (Smith 1991).

This task involves the calculation of the NPV on all of the project expenses and revenues for each time period. Figure 3 shows which values are used in the NPV calculation. Once the NPV is calculated, the value is collected, or stored in an array. It is here where Monte Carlo analysis is incorporated into the developed methodology. The NPV calculation is repeated many times so that a large number of NPVs can be collected.

Figure 1 shows the question 'Run = 100?' being asked to illustrate that the NPV needs to be collected many times. This collection of NPVs is used in the final step of the methodology.

2.6 Perform Statistical Analysis on Collected Data

The last step of the developed methodology requires the analyst to perform statistical calculations of mean, minimum, maximum, and variance on the collected data. A cumulative distribution curve (CDF) is also created for the collected data.

The purpose of calculating these statistics is to give the analyst information needed to analyze the project, the most important of which is the CDF. By viewing the CDF of the data, the analyst can see the associated levels of confidence versus the NPVs for the project. This concept was used when developing the NPV-at-risk methodology (Ye et al. 2000). The basic rule to this methodology is that if the NPV at a desired level of confidence is greater than zero, then the project is worth investing in, and if it is less than zero, then it is not worth investing.

The analyst may not wish to use a discount rate in the analysis. The output of the collected data in this case would just be the total project cost in actual dollars. This is significant if the analyst is analyzing a project that will not generate any revenue but will rather be just an expense, as in the case of many public works projects. The analyst would be able to use the CDF and determine the cost of the project at any level of confidence, where 100% is the maximum available value and 0% is the minimum confidence value available to be calculated.

3 DEVELOPED SPECIAL PURPOSE SIMULATION TOOL

The developed methodology described above is set up in such a way that it can be incorporated into a computer simulation model. This will reduce the possibility of calculation errors and speed up the calculation process. Performing the above methodology without the aid of a computer would be tedious and take up enormous amounts of time. The simulation environment Simphony was used to develop a Special Purpose Simulation (SPS) tool. The steps that incorporate schedule network analysis and activity duration process modeling could not be done if a simple spreadsheet software program was used. It is here that much of the current methodologies failed, partly due to the lack of an advanced simulation environment where the models could be created.

SPS tools developed in Simphony are called templates. The template created is called the planning template and it is a SPS tool developed for the purpose of analyzing and optimizing capital investment projects. The planning template:

- Has a graphical user interface that allows for the easy creation of capital investment project models;

- Can be reused over and over again for a countless number of projects to be analyzed; and
- Fully incorporates the developed methodology discussed previously.

3.1 Template Structure

Every template created in Simphony uses modeling elements for the purpose of building project-specific simulation models. The planning template has seven modeling elements. Each modeling element contains all or a combination of the following: input parameters, outputs, and statistics. In some cases, modeling elements can obtain input values from other elements on different hierarchical levels, as well as report outputs to other modeling elements on different hierarchical levels.

The planning template utilizes some of the functional features built into the Simphony environment. The most important feature is the statistics and charts that Simphony automatically produces. Using the collected NPV data, Simphony performs all the statistical calculations on the data including the CDF.

3.1.1 Modeling Elements – Input Parameters

The input parameters in the modeling elements require the template user to manually enter a value. The inputted values will be used by the template to calculate the output. All the modeling elements and their respective inputs are listed in Table 1.

Table 1: Modeling Elements with Input Variables

Modeling Element	Description of the Input Variable	Deterministic	Stochastic	Subjective Variable
Project Planner	Discount Rate (%)	X		
Collector	Overhead (%)	X		
Activity	Model Duration?	N/A	N/A	Yes/No
	Forecast Type	N/A	N/A	Select from list
	Activity Duration		X	
	Activity Direct Cost ($)		X	
	Activity Indirect Cost ($)		X	

The *project planner* element is the highest element on the hierarchical structure of elements. The input parameters required in this element are related to the overall project analysis. For example, the discount rate value is entered as a percentage in this element. As well, outputs generated from the modeling elements on lower hierarchical levels are reported to this element and used as input. The *project planner* has the following input parameter for simulation: Discount Rate (see Table 1). The discount rate value is used in the computation of the overall project's NPV.

The *collector* element represents a project's work package. This element is below the *project planner* element in the hierarchical structure, but above the *activity* element, described below. The input parameters in this element relate to the sub-work packages and activities inside the given work package. The *collector* element has the following input parameters for simulation: Collector Description, and Overhead (see Table 1). The overhead input parameter is entered as a percentage and is multiplied by the total costs obtained from the outputs of the elements in the lower hierarchical levels. This parameter is useful because the overhead cost estimate (project office staff and other office resources) at the conceptual planning stages of a project is typically given as a percentage of the direct project work package cost.

The *activity* element is the element used to represent a project's activities. This element is below the *collector* element in the hierarchical structure. The input parameters for this element relate to the costs and durations of the activities. The input parameters entered here affect the overall project analysis output. The input parameter duration affects the timing of all subsequent activities and work packages, which affect the total project NPV. The input parameter cost affects the overall time period cash flow, which also affects the total project NPV. The *activity* element has the following input parameters for simulation: Model Duration, Forecast Type, Activity Description, Activity Duration, Activity Quantity, Activity Direct Cost per Quantity, and Activity Indirect Cost (see Table 1).

The *start*, *finish*, *in*, and *out* elements are all used for the purpose of guiding the path of the simulation model. None of these elements have any input parameters.

3.1.2 Modeling Elements – Output and Statistics

The planning template performs calculations and computations to produce the output and statistics an analyst would use for project analysis. Simphony allows the template designer to create template specific output and statistics if desired and provides built in functional features.

The only output in the planning template that is seen by the analyst is found in the *finish* element. The template designer specifically incorporated the output into the planning template so that the analyst can determine the overall project finish time for each simulation run. The other output in the planning template is produced by Simphony and used as input for other elements, where it is needed for further computations. This output is irrelevant to the analyst, so it is not given in visual form.

The statistics represented in graphical form are calculated in the *project planner* element. The first statistic is the NPV, which represents steps 8 and 9 of the developed methodology discussed previously. The NPV is collected

for each simulation run and the statistics are calculated on the collected data. The analyst uses the given statistics—the project's maximum, minimum, and mean NPV, a CDF, and duration—on the collected NPVs to analyze the project. As discussed previously, the CDF is an indicator at a specific level of confidence to determine if the project is worth investing in or not. The final statistic, project duration, is used to determine the possible range of project length so that planning can be made accordingly.

3.2 Creating a Model

Creating a project specific model in the planning template is quite simple, and requires little simulation knowledge from the user. The Simphony environment allows for the creation of models through the use of a graphical interface.

Using the planning template specifically, the user only needs to create a project in the form of a CPM network diagram. The model is created is such a way as to model the hierarchical modeling discussed previously. The graphical user interface allows for the creation of the models that look similar to the diagram shown in Figure 2. Once the desired model is created, all the user needs to do is enter in the required input values and run the simulation. As discussed earlier, all the outputs and statistics will be automatically generated.

3.3 Construction Operation Modeling Features in the Developed SPS Tool

In second step of the methodology previously discussed, the analyst has the option of modeling an activity's operation processes or of entering in a value for the activity's duration. The planning template was developed to incorporate this very simply. If the template user decides to model the operation, the user can use other templates previously developed in the Simphony simulation environment to provide these inputs. The user can also simply enter a duration value. An example is given below to illustrate how this feature works.

3.4 Example Description

An example earthwork construction operation is used to represent the operations of the 'Cut and Fill' activity in the project network diagram. The operation is a basic excavation task, consisting of a scraper loading with fill, traveling loaded to the dump site, dumping the fill, and traveling back to the load site empty. The details of the operation were assumed and are outlined as follows:

- Required volume of fill: $500m^3$
- Number of earthmovers: 4 ea.
- Earthmover carrying capacity: $5m^3$
- Loading time: 5 min.

- Travel to dump site time: 20 min.
- Dump time: 3 min.
- Return to cut site time: 17 min.
- Number of push dozers: 2 ea.

Figure 4 shows how the excavation operation model was created in the planning template. The earthwork operations are part of the 'Cut and Fill' activity, which is part of the 'Earthworks' work package. The graphical image is the user interface that the user would encounter when using the Simphony SPS tool. The logic of the operation in relation to the network diagram is that when the required amount of fill has been dumped at the dump site, the activity is complete, and the next activity in the network diagram can proceed.

After running the simulation model, the output duration for the earthwork operation was 2.42 working days. This value would then be used as the duration for the 'Cut and Fill' activity in the network diagram instead of a manually entered value in the *activity* elements' parameters.

4   MODEL IMPLEMENTATION AND CASE STUDY

An interchange project by the City of Edmonton was used as a case study to test the developed SPS tool. At the time of modeling, the project was at the conceptual design phase. The purpose of the study was twofold: to test the developed SPS tool and to perform risk analysis for the project owner to provide a more reliable project cost estimate. A more accurate cost estimate would enable the owner to properly assign financial resources to the project and not tie up any additional capital.

In the following section, the term 'range estimate' will be used to describe the application of Monte Carlo techniques to project costs and revenues.
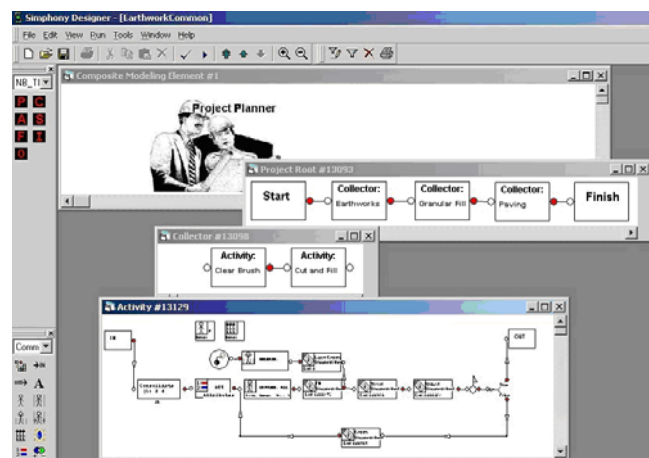

Figure 4. Earthwork Operations Modeled in Network Diagram

4.1 Cost Estimate Models Developed

Four models were developed for the analysis process. The first model was the fixed cost estimate provided from the City of Edmonton. This model did not incorporate any uncertainty, nor did it incorporate any future costs due to inflation. This model was used as the base model. The second model was the fixed estimate plus the incorporation of an inflation rate of 3%. The third model was a range estimate of the costs, without any inflation costs incorporated, but with network analysis incorporated. The fourth model was a range estimate and network analysis model, with an inflation rate of 3%.

It should be noted that the owner provided all the input data used for all the developed models. The data included durations, quantities, and costs, given as parameters for triangular and uniform distributions.

4.2 Analysis Procedure

Upon determining the project network sequencing, the project network was inputted into the developed SPS tool in Simphony for the discussed models. Once the network was inputted, all the project activity parameters, such as duration and cost, were defined. Once all the parameters were defined, all the developed simulation models were run for 100 simulation runs, and statistics were gathered for each run. Figure 5 shows the CDF for the range estimate plus forecasted inflation costs.

4.3 Discussion

As indicated earlier, the purpose of the case study was two-fold. The first reason was to test the developed SPS tool. Some observations were made upon reviewing the output of the four models. The first observation is that the mean range estimate costs for all cases were higher than the fixed estimate costs. This means that if the fixed estimate was used for budgeting purposes, the actual costs of the project would most likely be overrun. Here it can be seen that by incorporating some uncertainty into the analysis, the project cost is actually higher than the fixed cost. The second observation is that the range estimates that incorporated the network analysis produced mean outputs that were higher than the range estimates without the network analysis. This shows that the uncertainty of duration times has an impact on the project cash flow, and should therefore be accounted for. In this case, the project would not have enough capital if a budget were created based on the use of Monte Carlo techniques for cost uncertainty alone.

The second reason for the case study was to perform risk analysis for the project's owner. The results obtained indicated that the data provided by the owner was not ideal in the sense that the fixed estimate totals were too similar to the range estimate totals. The main reason for this is the
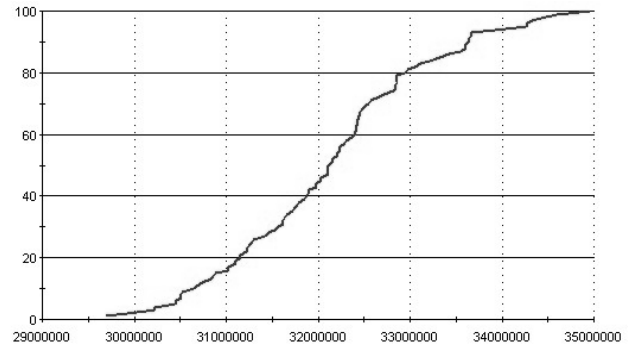

Figure 5. CDF for Range Estimate with Network Analysis

input data provided by the owner. For example, the largest work package in value was a fixed number, with no variance at all. Other large work package items had relatively small variance, which could also contribute to the small overall project cost variance. According to a study of infrastructure projects (AbouRizk et al. 2002), the suggested variance for preliminary design estimates for road construction should be $\pm$ 25%. The results of the analysis show that the variance for this project is about $4 million, which is about 9.4%. This actual variance determined from the analysis is much less than the value above. This implies that the input data provided by the owner was not as generous in the breadth of ranges as it should have been, considering the phase of the design. In order to ensure that the data provided for the input parameters are accurate, the activities that make up 80% of the total project cost should be checked, as they are the activities that most contribute to the project's overall cost.

5 CONCLUSION

The decisions made with regard to capital investment prior to the commitment of large amounts of capital and the decisions made before much of the project's life has occurred, are the most important decisions an analyst can face. The decisions made at the early stages of a project have a large impact on the future outcome of the project. This is especially true when considering the economics of the project. Incorrect decisions made at the early stages of a project could result in large amounts of economic losses in the future. Therefore, vigorous, detailed, and accurate analysis needs to be conducted for potential capital investment projects, as well as committed projects, in the early stages of their lives in order to ensure that risk is minimized, due to the impacts that wrong decisions can make. Current analysis methods fail to properly incorporate network analysis and activity duration uncertainty into the overall project analysis steps. This paper presented a methodology that incorporates network analysis and duration uncertainty in a project analysis technique. This methodology was used as the simulation framework for a

developed SPS tool created in the simulation environment Simphony.

REFERENCES

AbouRizk, S. M., Babey, G. M., and Karumanasseri, G. (2002). Estimating the cost of capital projects: an empirical study of accuracy levels for municipal government projects. *Canadian Journal of Civil Engineering*, 29(5): 653-661.

Chen, M.T. (1998). Simplified project economic analysis. *Cost Engineering*, 40(1), 31-35.

Cochrane, R.E. (1992). Using @risk to predict project costs. *Proceedings of the 36th Annual Transactions of the American Association of Cost Engineers*, Morgantown, WV., F.3.1-F.3.5.

Lu, Y.C., Wu, S., Chen, D.H., and Lin, Y.Y. (2000). BOT projects in Taiwan: financial modeling risk, term structure of net cash flows, and project at risk analysis. *The Journal of Project Finance*, 5(4) , 53-63.

Smith, J. (1991). How companies value properties. *CIM Bulletin*, 84(953), 50-52.

Wajs, W., Bieda, B., and Tadeusiewicz, R. (2000). Project cost analysis for niepolomice municipal solid waste using the Monte Carlo simulation. *Risk Analysis II*, 225-234.

Winston, R.E. (1995). Rapid method for capital investment decisions. *Cost Engineering*, 37(4), 41-45.

Ye, S., and Tiong, R.L.K. (2000). NPV-at-risk method in infrastructure project investment evaluation. *Journal of Construction Engineering and Management*, ASCE, 126(3), 227-233.

AUTHOR BIOGRAPHIES

NATHAN BOSKERS is a development manager at Springwood Developments in Edmonton, Alberta, Canada. He completed his MSc at the University of Alberta in 2004. His e-mail address is <nboskers@telus.net>.

S. M. ABOURIZK is a professor of construction engineering and management in the Hole School of Construction at the University of Alberta. He also holds the NSERC/Alberta Construction Industry Research Chair and the Canada Research Chair in Operation Simulation. His e-mail address is <abourizk@ualberta.ca> and his Web address is <www.construction.ualberta.ca>.