Developing a Parametric Model for Estimating Process Control Costs

Larry R. Dysert, CCC

ABSTRACT:
This article discusses the development of a parametric model used to prepare conceptual estimates for process control costs on capital projects. This discussion is based on a specific parametric model used by the Eastman Kodak Company, which is probably unique to its capital projects. However, the specific algorithms and the model are not the focus —this article describes how the model was developed, and how it is being used.

KEY WORDS: parametric model, conceptual estimates, process control

A parametric cost model is an extremely useful tool for preparing early conceptual estimates, when there is little technical data or engineering deliverables to provide a basis for using more detailed estimating methods. This article discusses the development of a parametric model used to prepare conceptual estimates for process control costs on capital projects. This discussion is based on a specific parametric model used by the Eastman Kodak Company, which is probably unique to its capital projects. However, the specific algorithms and the model are not the focus —this article describes how the model was developed, and how it is being used.

This article reviews the types of data collected to support the model, the analysis of the data, the creation of a spreadsheet model based on this analysis, and the various forms of output and presentation data available from the model. What is learned from this particular parametric model can be applied to other circumstances and used to create other forms of parametric models applicable in other situations.

A PROCESS CONTROL ESTIMATING SYSTEM (PROCEP)

Cost estimators are continually searching for effective ways to prepare conceptual and early strategic estimates for projects when minimal project scope information is available. At Kodak, the capital estimating department identified the process controls discipline as an area that was difficult to estimate accurately during the strategic estimating process. To address this need, the estimating department decided to develop a parametric estimating model based upon actual cost history and on key relationships that could be identified between costs and specific parameters of a process control system. The estimating application developed is an Excel-based spreadsheet called the process control estimating program (PROCEP).

PROCEP allows estimators to create process control estimates with the following characteristics.

- Efficient—it produces conceptual estimates in a matter of minutes when given the appropriate design parameters.
- Consistent—it produces a consistent estimate format using documented key design parameters as the estimate basis.
- Accurate—it produces estimates with a predicted accuracy range of -30/+50 percent, applicable for class 5 and class 4 estimates using AACE International's estimate classifications.
- Flexible—it can be used to develop process control estimates as a part of a total project estimate, or where the process controls work comprises the entire project (a process controls upgrade project).
- Defensible—it clearly highlights the design parameters used and provides key ratios and metrics for comparison with other projects.

The PROCEP estimating system has been in use since 1990, and has been a valuable tool. It has gained the confidence of the process control engineering and design community at the company, and is often used to help determine the most cost-effective process control schemes by comparing various alternatives very early in the project's life cycle.

DEVELOPING A PARAMETRIC MODEL

The development of a parametric estimating model can appear to be a daunting task; however, the use of modern computer technology (including popular spreadsheet programs) can make the process tolerable, and much easier than it would have been many years ago. The process of developing a parametric model should generally involve the following steps:

- cost model scope determination;
- data collection;
- data normalization;
- data analysis;
- data application;
- testing; and
- documentation.

The following discussion focuses on how these steps were accomplished in regards to the PROCEP parametric estimating system.

Cost Model Scope Determination

The first step in developing a parametric model is to establish its scope. This includes defining the end use of the model, the physical characteristics of the model, the cost basis of the model, and its critical components and cost drivers.

For the PROCEP system, the end use of the system is to prepare strategic and conceptual estimates of process control costs, including process control hardware, instruments, installation, software, engineering and design, and debugging. It is intended to be a reasonably simple spreadsheet model that could be used with minimal project scope information to produce estimates with an intended accuracy range of up to -30/+50 percent. It should produce a standard estimate report and also display key metrics for comparisons with other projects.

The model should be based on actual costs from completed projects and reflect an organization's engineering practices and technology. It should be applicable for the variety of different process technologies that are involved with Kodak projects (machine process, chemical process, building automation, etc.). The model should be based on current-year costs or provide a mechanism to escalate to current-year costs.

The model should be based upon key process control parameters that can be defined with reasonable accuracy early in project scope development, and should allow an estimator to easily adjust the derived costs for specific complexity, or other factors affecting a specific project.

Data Collection

Data collection and development for a parametric model require a significant effort. The quality of the resulting parametric model can be no better than the quality of the data it is based upon. Both cost and scope information must be identified and collected. For the PROCEP model, it was decided to collect data from recent projects where the cost and scope information for process controls could be identified readily and clearly. The collection involved both complete projects (where process controls is a part of the overall project scope) and projects that were primarily process control projects (upgrades or replacements of process control systems).

The type of data to be collected was decided upon in cooperation with the process control engineering community. A data input form was developed to aid in collecting the data. Over time, this form has been revised as the data needs were better identified, and the parametric model was revised.

The original data collection study began in 1990, and collected information on 40 projects dating back to 1987. Since its inception, the PROCEP model has been updated with additional project data in 1994, 1995, and 1998. The database of project information currently holds data on approximately 100 projects.

The data obtained by the input forms are entered into a Excel spreadsheet that includes the following cost information:

- major hardware and purchased software costs;
- instrumentation purchase costs;
- installation costs (labor and bulk materials—separated for hardware and instrumentation, if available);
- engineering and design costs;
- project administration costs;
- software development costs; and
- debugging costs.

The following scope information and process control system characteristics are also collected:

- type of process control system (general type as well as hardware manufacturer/model information);
- type of process controlled (chemical, machine, HVAC, etc.);
- installation conditions (hazardous, corrosive, etc.);
- Input/output (I/O) count (analog and digital or new and existing); and
- types of I/O (pressure, temperature, analytic, etc.).

Data Normalization

After the data have been collected, the next step in the process of developing a parametric model is to normalize the data before the data analysis stage. Normalizing the data refers to making adjustments to the data to account for differences between the actual basis for each project and a desired standard basis of data to be used for the parametric model. Typically, data normalization implies making adjustments for:

- escalation;
- general location;
- site conditions;
- system specifications; and
- cost scope.

For the PROCEP model, the data were required to be normalized for escalation to bring all costs up to the same base year. All of the project data was from the Kodak Park Rochester site, so no adjustments were required for location. Site conditions and system specifications were elements identified on the data input sheets, and were not normalized. Instead, these items were analyzed to determine their effect and used as key parameters of the resulting parametric model. Normalizing for cost scope implies making adjustments for unusual costs that a specific project may incur but that a "typical" project would not. A comments section on the data input forms was used to identify these occurrences, and the data were adjusted to exclude the costs for unusual circumstances.

Data Analysis

The next step in the development of a parametric model is data analysis. There are many diverse methods and techniques that can be used in data analysis, and they are too complex to delve into here. Typically, data analysis consists of performing regression analysis of costs versus selected design parameters to determine the key costs drivers for the model. Most spreadsheet applications now provide regression analysis and simulation functions, which are reasonably simple to use.

For PROCEP, a number of regression analysis cases (linear and nonlinear) were run against the data to determine the algorithms that eventually comprised the parametric model. The various relationships (cost versus design parameters) were examined for "best fit" by looking for the highest R-squared value. R^2 has the technical sounding statistical name of *coefficient of determination*, and is commonly used as a measure of the goodness of fit of a regression equation. In simple terms, it is one measure of how well the regression equation explains the variability of the data. The algorithms resulting from the regression analysis are applied to the input data sets to determine on a project-by-project basis how well the regression algorithm predicted the actual costs.

Regression analysis can be a time-consuming process (especially with the simple regression tools of a spreadsheet program), as experiments are made to discover the best-fit algorithms. As an algorithm is discovered that appears to provide good results, it must be tested to ensure that it properly explains the data. Advanced statistical tools can quicken the process but can be more difficult to use. Sometimes an estimator will find that erratic or outlying data points need to be removed from the regression runs in order to avoid distortions in the results. It's also important to recognize that many cost relationships are nonlinear, and therefore one or more input variables will be raised to a power. Experimentation is needed with both with the variables you are testing against and the exponential powers used for the variables. Regression analysis tends to be a continuing trial-and-error process until the proper results are obtained that appear to explain the data.

For the PROCEP model, the following data relationships resulted in regression relationships that were used in the parametric model:

- hardware costs versus I/O count addressed by the process control system;
- field instrumentation costs versus I/O count purchased;
- field installation costs versus I/O count installed;
- software development costs versus I/O count addressed by the system;
- engineering costs versus field capital cost; and
- debugging costs versus field capital cost.

Data Application

The data application stage involves establishing the user interface and presentation form for the parametric cost model. Using the mathematical and statistical algorithms developed in the data analysis stage, the various inputs to the cost model are identified, and an interface is developed to provide the user with an easy and straightforward way to enter this information. Electronic spreadsheets provide an excellent mechanism to accept user input, calculate costs based upon algorithms, and display the resulting output.

As mentioned previously, PROCEP is an Excel spreadsheet application. A data entry form was generated to accept the user input that includes:

- estimate title information;
- total digital and analog I/O to be addressed by the controller hardware;
- total digital and analog I/O to be purchased by the project;
- total control valves to be purchased;
- total digital and analog I/O to be installed;
- hardware type (distributed, programmable logic controller, etc.);
- redundant hardware/spare capacity information; and
- type of process being controlled.

In addition, there are a number of complexity adjustments that a user can make based on his/her understanding of the specific project requirements, cost model, and experience. The spreadsheet uses built-in Excel protection features to allow the user to only enter information in the data entry cells. All other cells, including calculations and lookup table information, are protected from being modified or overwritten.

The spreadsheet contains macros to allow users to easily print the results of an estimate in a consistent format. The model also calculates and displays various key ratios between costs that can be used to benchmark the results and to compare cost trends over time.

Testing

One of the most important steps in developing a cost model is to test its accuracy and validity. When using regression analysis, one of the key indicators of how well a resulting algorithm explains the data is a term called R^2. This is the coefficient of determination, and it provides a measure of how well the resulting algorithm predicts the calculated costs. An R^2 value of 1 indicates a perfect fit, while an R^2 value of .89 indicates an 89 percent confidence that the regression equation explains the variability in cost (based on the data used in the development of the model). However, a high R^2 value by itself does not imply that the relationships between the data inputs and the resulting cost are statistically significant.

Once a regression analysis has been performed, and an algorithm with a reasonably high R^2 value has been obtained, an estimator still needs to examine the algorithm to ensure that it makes common sense. In other words, perform a cursory examination of the model to look for the obvious relationships that you expect to see. If the relationships from the model appear to be reasonable, run additional tests for statistical significance and to verify that the model is providing answers within an acceptable range of error.

One quick check to run is to test the regression results directly against the input data to see the percent error for each of the inputs. This lets you quickly determine the range of error, and interpreting the results can help you to determine problems with the algorithm.

After all of the individual algorithms have been developed and assembled into a complete cost model, it is important to test the model as a whole against new data (data not used in the development of the model). Consult a statistical text for more information about testing regression results and cost models.

Documentation

The resulting cost model and estimating application must be documented thoroughly. A user manual for the estimating application should be prepared that shows the steps involved in preparing an estimate using the cost model, and clearly describes the required inputs to the cost model. The data used to create the cost model should be documented, including a discussion of how the data were adjusted or normalized for use in the data analysis stage. It is usually desirable to make the actual regression data sets available, along with the resulting regression and test results. All assumptions and allowances designed into the cost model should be documented, as should any exclusions. The range of applicable input values, and the limitations of the model's algorithms, also should be explained.

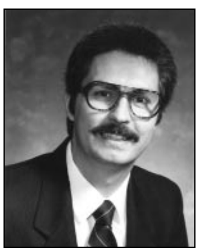
For the PROCEP model, the instructions for use are contained in a user guide. Additional information is contained in a separate resource manual and includes the following:

- original data collection forms;
- tables showing the regressed data sets and regression results;
- lists of the resulting algorithms used in the PROCEP cost model;
- graphical representations of the algorithms plotted along with the results

can help you to determine problems with the algorithm; and

- summary charts that show the overall relationships derived from the cost model.

Parametric cost models can be a valuable resource when preparing early conceptual estimates. Our experience at Kodak has demonstrated that effective parametric estimating systems can be developed using basic skills in estimating, mathematics, statistics, and spreadsheet software. It is important to understand that the quality of the results can be no better than the quality of the input data, and great care should be taken during the data collection stage to gather appropriate and accurate project scope and cost data. Using the statistical analysis and modeling capabilities of modern spreadsheet (and/or statistical) software, cost models can be created that provide a consistent and accurate basis for conceptual estimates.



Larry R. Dysert, CCC, is a project estimator for the Eastman Kodak Company in Rochester, NY. Larry has been an active member of AACE International, and is currently serving as Chair of the Technical Board. In the past, Larry has held several positions in the local Genesee Valley Section of AACE, including serving as President, Vice President, and Programs Director. Prior to joining Kodak, Larry managed his own consulting firm for 3 years, and worked for many years with a large A&E firm. Larry has 21 years of experience in project estimating and project management for a wide variety of industries. He has spoken at many national [US] and international conferences relating to estimating and project management topics.◆