EST.01

So You Think You're an Estimator?

Mr. Larry R. Dysert, CCC

This paper focuses on many of the issues and problems associated with rough order of magnitude estimating. How do you prepare an estimate when there is very little information on which to base the estimate? While teaching at the University of Chicago, physicist Enrico Fermi had a reputation for asking his students, without any warning, seemingly impossible questions, such as, "how many piano tuners are there in Chicago?"

We face similar conceptual estimating problems everyday in the world of projects. Our management asks for an estimate for a new project using a revolutionary technology never implemented before, and with a larger capacity than has ever been built; and oh, by the way, they need the estimate by tomorrow. This paper is intended to discuss some of the techniques that can be used, as well as some of the problems faced when we are placed in such situations.

Conceptual Estimating Techniques

There are many order-of-magnitude or conceptual estimating techniques that have been developed over the years and each is useful in certain situations. Often a single estimate will rely on using a combination of estimating techniques for different portions of the project. The conceptual estimating techniques that will be briefly discussed in this paper are:

- capacity factoring;
- parametric modeling;
- end-product units method;
- analogy; and
- expert judgment.

This paper will only present a summary of the estimating techniques, however the references identified at the end of the paper can point you towards a more detailed explanation.

Capacity Factoring

A capacity factored estimate (CFE) is one in which the cost of a new proposed project is derived from the cost of a similar project of a known capacity. The basic estimating algorithm relies on the typical non-linear relationship between capacity and cost shown in the following equation:

$$\$B/\$A = (Cap_B/Cap_A)^e$$

(equation 1)

where $\$A$ and $\$B$ are the costs of the two similar projects, Cap_A and Cap_B are the capacities of the two projects, and "e" is the exponent (or capacity factor) that drives the non-linear relationship.

The exponent "e" (or the capacity factor) is actually the slope of the log-curve that is drawn to reflect the relationship between actual costs and capacities of two or more completed projects. When the exponent has a value less than 1, it reflects the typical "economy of scale" cost relationship that we expect from a change in capacity of a project.

For example, if we were to estimate the cost of a new refinery that is 25 percent larger than the last one we built, we would expect that the costs to build the larger refinery would increase by less than 25 percent. Thus, this estimating technique is sometimes known as the "scale of operations" technique.

Capacity factored estimating techniques can be applied to a wide range of industries and projects to prepare quick feasibility and project screening estimates. This technique is very common in the process industries where the exponent "e" typically has a value between 0.5 and 0.85, depending on the type of plant; and in fact yet another name for this estimating technique is the "six-tenths rule" because of a common reliance on using an exponent value of 0.6 if no better information is available. With an exponent of 0.6, doubling the capacity of a project or plant increases the project costs by 50 percent.

It is important to realize, however, that as project capacities increase, the exponent also tends to increase in value, as illustrated in Figure 1. The capacity factor exponent between projects A and B may have a value of 0.6; between projects B and C, the exponent may have a value of 0.65; and between projects C and D, the exponent may have risen to 0.72. As project capacities increase to the limits of current technology, the exponent approaches a value of 1. At this point (or as the value of the exponent becomes larger than 1), it becomes more economical to build two projects of a smaller size, rather than one large project.

In applying the capacity factoring cost estimating technique, we convert the algorithm used to explain the relationship between cost and capacity to the following cost estimating relationship:
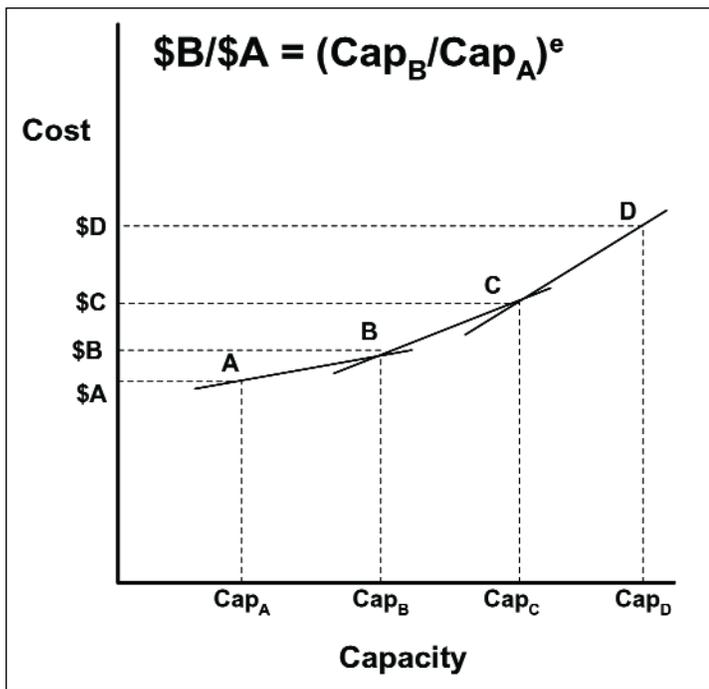
Figure 1—Exponents differ in value across capacity ranges. Cap$_A$ is the capacity of Plant A, and so on.

$$\$B = \$A \times (Cap_B/Cap_A)^e$$

(equation 2)

Where $B is the estimated cost for a new project, $A is the actual cost of a similar project, Cap$_B$ is the capacity of the new project, Cap$_A$ is the capacity of the similar completed project, and "e" is the capacity factor exponent.

Let's examine a typical CFE situation where we need to estimate the costs of a 100,000 BBL/Day Hydrogen Peroxide unit to be built in Philadelphia and completed in 2007. We have recently completed a 150,000 BBL/Day plant in Malaysia with a final cost of $50 Million in 2004. Our recent project cost history shows a capacity factor of .75 is appropriate. The simple approach is to just use our capacity factor algorithm:

$$\$B = \$50M \times (100/150).75 = \$36.9M$$

(equation 3)

However, this would be too simple and incorrect! A better approach is to adjust for the differences in scope, location, and time. The plant in Malaysia included piling, tankage, and owner costs that will not need to be included in the proposed plant for Philadelphia. Construction in Philadelphia is expected to cost 1.25 times the construction costs in Malaysia (location adjustment). Escalation will be included as a 1.06 multiplier from 2004 to 2007. There are costs for additional pollution requirements in Philadelphia that were not included in the cost of the Malaysian plant. Taking these into account, the estimate now appears like this:

150,000 BBL/Day Plant in Malaysia	$50M
Deduct Piling. Tankage, Owner Costs	-$10M

Adjusted Cost for Scope	= $40M
Malaysia to Philadelphia Adjustment (X 1.25)	= $50M
Escalate to 2007 (X 1.06)	= $53M
Factor = $53M X (100/150).75	= $39M
Add Pollution Requirements (+$5M)	= $44M

The example presented here is applicable to the process industries, but the basic capacity factoring technique is appropriate in many other industries as well. I have seen similar techniques used in the commercial building industry (using building square foot area as the unit of capacity), and even in the software development industry (using expected lines of code to be written as a unit of capacity).

This method is most effective when the new (to be estimated) and completed projects are near-duplicates, and are reasonably close in size. If the capacity factor used in the estimating algorithm is reasonable, and the project being estimated is relatively close to the size of a similar project of known cost, then the potential error from the a CFE is well within the level of accuracy of an order-of-magnitude estimate.

A key is to account for differences in scope, location, and time. The recommended way to do this is to deduct costs from the completed project (the known base case) that are not applicable to the new project. Apply location and time adjustments to normalize the costs, and then use the capacity factor estimating algorithm to adjust for project size. Finally, add any additional costs that are required for the new project, but were not required for the base case project.

Parametric Modeling

A parametric cost model can be an extremely useful tool for the preparation of early conceptual estimates. A parametric estimating model is a mathematical representation of one or more cost estimating relationships (CER's) that provide a logical and predictable correlation between the functional or physical characteristics of a project and its costs. A capacity factored estimate can be thought of as a simple parametric model (using capacity as a single independent variable); however, sophisticated parametric models will often involve several independent variables and cost drivers.

Derivation of a parametric estimating model can be a daunting and complicated undertaking. The model should be based on the collection and analysis of actual cost data from completed projects, along with key engineering and design data. The key is to identify the significant project design parameters that can be defined with reasonable accuracy early in project scope development, and that are correlated with statistical significance to project costs. The model should also provide the capability for the estimator to make adjustments for specific factors affecting a particular project.

The data collection efforts for developing the model require significant effort. Both cost and design scope information must be identified and collected. It is best to collect the information at as low a level of detail as possible, as it can always be summarized later if an aggregate level of cost information provides a better cost model. After the data has been collected, it should be normalized for time, location, site conditions, project specifications, and cost scope.

Cooling Range (Deg F)	Approach (Deg F)	Flow Rate (GPM)	Actual Cost	Predicted Cost	% Error
30	15	50,000	$1,040,200	$1,014,000	-2.5%
30	15	40,000	$787,100	$843,000	7.1%
40	15	50,000	$1,129,550	$1,173,000	3.8%
40	20	50,000	$868,200	$830,000	-4.4%
25	10	30,000	$926,400	$914,000	-1.3%
35	8	35,000	$1,332,400	$1,314,000	-1.4%

Table 1—Actual Costs Versus Predicted Costs With Parametric Equation

This then leads to the next step of data analysis that involves the regression of cost versus selected design parameters to identify the key cost drivers. Regression analysis often requires many iterative trials to develop the best-fit CERs (or estimating algorithms) that will form the parametric model.

Usually, a CER will take one of the following forms:

$$Cost = a + bV1 + cV2 + \ldots \quad \text{(linear relationship)}$$
(equation 4)

or

$$Cost = a + bV1x + cV2y + \ldots \text{(non-linear relationship)}$$
(equation 5)

where V1 and V2 represent the values of input design variables; a, b, and c are constants derived from the regression analysis; and x and y are exponents (also derived from the regression analysis). Often a single estimating algorithm will involve both linear and non-linear cost relationships.

The cost estimating algorithms derived from the regression analysis are then examined to ensure that they provide reasonable and expected relationships between costs and the key design parameters, as well as tested for statistical significance and to verify that the model is providing results with an acceptable range of error.

An example of a fairly simple parametric estimating model is the following equation that uses three design parameters to calculate the estimated costs of an induced-draft cooling tower. An induced-draft cooling tower is typically used in process plants to provide a recycle cooling water loop. The units are generally prefabricated, and often installed on a turnkey basis by the equipment vendor. Key design parameters affecting costs are the cooling range, the temperature approach, and the water flowrate. The cooling range is the temperature difference between the water entering the cooling tower and the water leaving it; and the approach is the temperature difference between the cold water leaving the tower and the wet-bulb temperature of the ambient air.

The parametric estimating algorithm was developed from the regression analysis of design and cost information for recently completed units and normalized (adjusted for location and time) to a Northeast US, year-2000 timeframe:

Estimated Cost = $86,600 + $84,500 X (Cooling Range, Deg F)$^{.65}$ - $68,600 X (Approach, Deg F) + $76,700 X (Flowrate, 1000 gal/min)$^{0.7}$

(equation 6)

The algorithm shows that the cooling range and flowrate affect costs in a non-linear (exponential) fashion, while the approach affects costs in a linear manner. Increasing the approach results in a less-costly cooling tower as it increases the heat efficiency of the heat transfer taking place.

Table 1 shows the actual costs of six induced-draft cooling towers along with the predicted costs (all costs in Year 2000 $) from the parametric estimating equation. The percent error is well within an acceptable level of accuracy for an order-of-magnitude estimate.

As with any estimate, adjustments for location and time will need to be applied to the costs derived from a parametric model, as well as adjustments for additional or modified scope from that assumed in the model.

Parametric models can be much more complex than the single CER shown in the above example. In addition to several CER's, a complex parametric model may include an extensive database of technical and cost history and require extensive documentation to communicate the assumptions, ground rules, and logic incorporated in the model. Parametric models have been created to prepare estimates for everything from commercial construction projects to the space shuttle to software development.

Parametric models can be a valuable resource in the preparation of early, order-of-magnitude estimates. Effective parametric models can be developed using basic skills in estimating, mathematics, and statistical analysis; and implemented using sophisticated programming application or simple spreadsheets. The quality of the results from a parametric model are obviously no better than the quality and analysis of the input data used in creation of the model. Great care should be taken during the data collection stage to gather appropriate and accurate project scope and cost data, and the model should be thoroughly tested to ensure that the results are logical, consistent, and meet the expected accuracy levels.

End-Units Method

This conceptual estimating methodology is generally used when enough historical data exists to from similar projects in order to relate the end-product (capacity units) of a project to its costs. This techniques allows an estimate to be prepared relatively quickly, requiring only the end-product units of the proposed project. Examples of the relationship between costs and end-product units are:

- the cost of building an electric generating plant and the plant's capacity in kilowatts;

- the construction cost of a hospital and the number of patient beds;
- the development cost of a software program and the number of function points (screens, reports, calculations, etc.) to be included in the application; and
- the construction cost of a parking lot and the number of parking spaces required.

For illustration purposes, consider the construction of a 1500 room luxury hotel, and assume a similar hotel has recently been completed at a nearby location. The hotel just completed included 1000 guest rooms, as well as a lobby, restaurants, meeting rooms, parking garage, and swimming pool. The total construction cost for the 1000 room hotel was $67,500,000. The resulting cost per room is $67,500.

We can then calculate the cost of the new 1500 room hotel of comparable design and features as $101,250,000 ($67,500/room x 1500 rooms). This simple calculation has, however, ignored several factors that may impact costs. For example, it has ignored any economies-of-scale (capacity factors) that may result from constructing a larger hotel, and it has assumed that the cost of the common facilities (lobby, restaurants, pool, etc.) vary directly with the increase in the number of guest rooms. If cost data exists to understand the cost impact of these differences, then further adjustments to the estimated costs should be made to account for these influences. Similarly, if the location or the timing of the proposed hotel differs significantly from the known cost data point, then cost adjustments should be made to account for these differences.

Very similar in concept to the end-product units estimating methodology is the physical dimensions method. This estimating technique uses the physical dimensions (length, area, volume, etc.) of the item being estimated as the driver of costs. For example, the estimate for constructing a building may be based on the square meters or cubic volume of the building, and similarly the cost of an oil pipeline or a highway may be based on a linear basis.

As with the end-product units method, this technique also depends on historical information from comparable facilities. Let's consider the need to construct a 3,600-m^2 warehouse. Again, a recently completed warehouse of 2,900 m^2 in a nearby location was recently constructed at a cost of $623,500 (or $215/$m^2$). The completed warehouse usd a 4.25-m wall height, thus enveloping a volume of 12,325 m^3 (or a cost of $50.50/$m^3$ on a volume basis).

In determining the cost of the new 3,600-m^2 warehouse, we estimate the costs on a m^2 basis at $774,000 ($215/m^2 x 3600 m^2); however, since the new warehouse will utilize a wall height of 5.5m, we may decide that estimating on a volume basis is more appropriate. The volume of the new warehouse will be 19,800 m^3 (3600 m^2 x 5.5m), and the estimate on a volume basis results in an estimate of $1,002,000. Again, we will still need to take into account other estimating adjustments for location, time, economies-of scale, etc. based on information available to us.

Analogy

An analogy estimate is typically characterized by the use of a single historical data point serving as the basis for the estimate. Analogy estimating methods are often used when a parametric model or other estimating algorithms (capacity factors, equipment factors, etc.) cannot be applied. This may be because of a lack of adequate historical data to support the development of conceptual estimating algorithms, or perhaps because the proposed project differs significantly from those projects that existing estimating algorithms can address. In any case, an analogy estimate is typically prepared by selecting a completed project as a base case, and then adjusting the historical costs for the technical, performance, complexity, physical, and other differences between the new project and the base case.

Because of its typical reliance on a single data point, the process to compare the characteristics of the new and base case project and the extrapolation process used to derive new costs are critical to the accuracy of an analogy estimate. Generally, technical experts are used to help assess a quantitative difference between the base case project and the project to be estimate. It is the estimator's task to develop the cost impact of the quantified differences. This involves both objective and subjective judgments. Some differences, such as differences in size can be calculated using fairly deterministic methods such as capacity factors (as described above), and other differences such as metallurgy, or other physical characteristics, can also be calculated using proven or recognized adjustment factors. However, some quantified variances such as complexity or performance factors require much more subjectivity in establishing the cost impact due to the differences between projects. This is part of the "art" of estimating, and often requires extensive experience (i.e., the "school of hard knocks") to develop an appropriate feel for the adjustments required.

Luckily, most new projects (even those considered revolutionary) typically can be broken down into sub-systems, of which only a portion will involve significantly new technology. Thus, some subsystems can be estimated with relative high accuracy, and only those subsystems that involve significant changes in complexity or technology advances are subject to the greater estimating uncertainty requiring a large degree of subjectivity in assessing the cost impacts of differences to base case historical costs.

As with most estimating methods, analogy estimating tends to be both easier to apply and result in improved accuracy if a systematic process is applied. First, the new project should be as clearly defined as possible, especially in reference to the characteristics (capacity, size, design, complexity, etc.) that may be applicable in locating or determining a comparable base case project upon which to establish a starting point for estimating costs.

If possible, the project should be broken down into logical subsystems or components. Those components that are very similar to existing components for which reliable historical cost data (or cost factors) exist can be estimated by appropriate estimating techniques. The components that involve significant new technology, or for which reliable historical cost information does not exist, will need to be evaluated to determine the characteristics that can best be used to determine corresponding base case components.

When identifying the characteristics used to determine comparable base cases, it is important to focus on characteristics that drive significant cost impacts. For example, metallurgy may be much more important in determining comparable components than color.

The next step is to address the differences between the new components and the base case components, focusing on the characteristics that drive costs. This is where assistance from a technical specialist may be required. The technical specialist (at this stage of the design process) may still need to rely on subjective assessments (such as the new widget is 20 percent more complex than the old widget, or the new widget likely requires 30 percent more moving parts that the old widget). The key is to attempt to have the technical specialist quantify as much as possible in an objective fashion, and to provide subjective assessments only where absolutely required.

The estimator must then collect, analyze, and normalize the costs from the base case components before determining the cost adjustments (or factors) to be applied to account for the technological differences. After adjusting costs both objectively and subjectively, the costs for the various components are then combined into the aggregate total cost estimate.

Reliance on Historical Information

The conceptual estimating methods described thus far are very reliant on having relevant historical cost information upon which to base the estimates, whether that information is encompassed as capacity factors, parametric estimating models, end-product unit costs, or historical project costs to be used as a base case in the derivation of an analogy estimate. For the most part, conceptual estimating methods are characterized by requiring significant effort in data gathering, data analysis, and estimating methods development before estimate preparation ever begins. There's obviously a large effort in historical cost analysis to develop accurate estimating factors and estimating algorithms to support conceptual estimating methods. Preparing the conceptual estimate itself takes relatively little time, sometimes less than an hour.

There are still times, however, when you simply have no reliable historical information or estimating algorithms upon which to base an estimate. You might be asked to estimate the cost of a project involving an entirely new technology never used by your organization before, or you might simple have failed in the past to collect and analyze actual project cost and technical information in order to develop conceptual estimating tools. In these cases, you may be forced to rely on "expert judgment."

Expert Judgment

As its name implies, expert judgment (or expert opinion) is an estimating technique that relies almost solely on the experience, knowledge and assessment of one or more experts. When you have no objective information on which to base an estimate, you may be forced to simply ask the opinion of a person that is knowledgeable of the project to be estimated and the costs of (hopefully) similar projects.

The expert may be acquainted with the project costs of other companies in the same industry, or otherwise have some useful information on which to base his judgment, but in the end that is what it is—a somewhat subjective judgment that lacks the objectivity of a mathematically derived calculation.

Obviously, the more objective knowledge and personal experience that the expert can apply to the specific estimating situation, the better the result should be. A problem, however is that any single expert may be subject to biases that are difficult to discern. To avoid this inherent bias when using a single expert to provide an estimated cost, a group of experts will often be used to develop an expert judgment estimate. A common technique applied to reaching group consensus is called the "Delphi Method."

Originally conceived by the Rand Corporation in 1948, the Delphi Method allows a group of subject matter experts to reach a group consensus using a disciplined and systematic approach. Generally, the basic approach follows these steps:

- the teams of subject matter experts is assembled, but told not to discuss their work (or any pre-conceived ideas) with one another.
- a facilitator provides each of the subject matter experts with the project information, and asks each expert to provide an estimated value based on his knowledge and experiences.
- the facilitator then distributes all estimates (usually anonymously) to the team, allowing each expert to see all of the estimated values.
- each expert then revises his estimate, and the process continues until the collection of estimated values reaches a consensus value.

Typically, a group consensus for the estimate is reached after only a few cycles of the process. Along with the estimate, the subject matter experts will often provide information about their assumptions, risk issues, etc. that they developed while compiling their estimate. This information would also be distributed in the round-robin review of all team member estimates, allowing each expert to see some of the thought process that went into each of the estimates.

Generally, as each review round takes place, the experts start developing a rough agreement on the assumptions, and the individual estimates get closer and closer. Eventually, the estimates are all within a narrow range and a particular value is selected (often the average of the individual estimates).

There are several variants to this basic technique, but the basic concept is to eliminate individual biases in the "expert" opinions, and to reach group consensus in a non-confrontational manner. Sometimes, it can still be difficult to reach a consensus. For example, there may a situation where three out of four experts have settled on a value of $50 million for an estimated value, and one expert remains at a value of $80 million. There are of course many different ways in which to address this situation, but the most common would probably be to accept the $50 million value as the estimate, but with a stated risk that a member of the team considers $80 million as a more accurate value.

Sometimes, rather than having the experts simply review the team estimates and backup information before submitting a new estimate, the facilitator will prompt the team to engage in open discussion of all the issues. After the open discussion of opinions, each expert then provides his own estimate in a similar fashion to that described above.

As with any estimating technique, the desire is that the estimators involved are adequately assessing and making adjustments for all project characteristics that affect project costs. At the end of the process, the consensus assumptions, risk issues, and other per-

tinent information should be documented and accompany the project estimate.

Common Problems with OOM Estimating Methods

When properly applied, order-of-magnitude or conceptual estimating methods can provide quick, and sufficiently accurate estimates for feasibility studies, and other early project decisions. When based upon good, historical cost information the techniques described above can be used very effectively and with reasonable accuracy.

One of the largest problems faced with conceptual estimating techniques is obtaining a clear understanding of the project scope. Clearly, the level of scope definition is low compared to that which will be available for later estimates. For early estimates, the estimator is often working directly with the business unit in gaining alignment on the project scope to be estimated. Early communication must exist between the estimator and the project team or business unit on the expectations for the estimate, and the estimator's abilities to meet those expectations.

The estimator must be clear to identify the level of accuracy that can be expected from the level of scope information available, and the available cost information and estimating tools and techniques available to support the estimate. Alignment also needs to take place to establish the boundaries for the estimate—what is supposed to be included in the estimated costs, and what is to be excluded. Early communication helps to avoid misunderstandings and failed expectations at a later date.

The estimator needs to be aware that the business unit or project team may have a preconceived cost value for a project even at these earliest stages of scope definition. The estimator must ensure that he prepares an unbiased and realistic estimate based on the scope of work to be accomplished, and does not become prejudiced by any preconceived estimate values.

Another of the biggest problems in using conceptual estimating techniques is the reliance on the basic estimating calculation (or algorithm) to produce an estimated value, and then not adequately adjusting the calculated costs for the unique peculiarities of the project being estimated. For example, when using the capacity factor technique, the estimator may fail to adequately normalize the costs of the base case project, and fail to properly identify and quantify the scope differences between the base case and proposed project. If the proposed project contains $10 million of additional scope items that were not included in the base case project, then the capacity factor algorithm is not going to account for those costs.

Often the estimator fails to fully understand the basis of the historical cost information available. If a historical average end-product unit cost value of $100,000 per hospital bed was normalized to cover only the hospital costs, and not the associated costs for parking structures and related infrastructure, then the estimator needs to be aware of this and adjust estimates accordingly if his proposed project includes these additional items.

Lastly, estimators often fail to adequately document early estimates. The basis of estimate document is often even more important for conceptual estimates than for later estimates because of the tendency for management to "cast into stone" the first estimated cost they receive for a project. Later, when the capacity of the project has doubled, the implemented technology is completely different, and the project was constructed in a different location and two years later than originally planned, management wonders why their order-of-magnitude estimates are never very accurate. In fact, the conceptual estimate may have been very accurate—it was just for a different project than ended up being constructed. Having a comprehensive basis of estimate that documents the scope of the project being estimating, the project location, time, and any other assumptions and costs data used in developing the estimate can help to refute the notion that all early estimates are bad.

*T*his paper summarizes many of the order-of-magnitude estimating methodologies that can be used when preparing early, conceptual estimates—the "Fermi type" problem when management wants an estimate by tomorrow for a new, radical, never-been tried process. Using a sound and disciplined approach, and well-documented historical cost data and estimating factors, these conceptual estimating techniques can be used to prepare sufficiently accurate estimates to support early decision making. The paper also addresses some of the common problems and pitfalls encountered with early estimates. Well prepared conceptual estimates enable management to make sound business and financial decisions at the early stages of a project. If we get that right, we can then be prepared to achieve success throughout the project.

REFERENCES

1. Black, Dr. J.H., "*Application of Parametric Estimating to Cost Engineering*," **AACE Transactions**, AACE International, West Virginia, 1984.
2. Chilton, C. H., "*Six-Tenths Factor Applies to Complete Plant Costs*," **Chemical Engineering**, April 1950.
3. Dysert, L. R., "*Developing a Parametric Model for Estimating Process Control Costs*," **AACE Transactions**, AACE International, West Virginia, 1999.
4. Dysert, L. R., "*Estimating*," **Skills and Knowledge of Cost Engineering**, **5th Edition**, AACE International, West Virginia, 2004.
5. FAA, **Life Cycle Cost Estimating Handbook**.
6. Miller, C. A., "*Capital Cost Estimating - A Science Rather Than An Art*," **Cost Engineer's Notebook**, AACE International, West Virginia, 1978.
7. NASA, **Parametric Estimating Handbook**.
8. Rose, A., "*An Organized Approach to Parametric Estimating*," **Transactions of the Seventh International Cost Engineering Congress**, 1982.
9. Williams, R. Jr., "*Six Tenths Factor Aids in Approximating Costs*," **Chemical Engineering**, December 1947.

Mr. Larry R. Dysert, CCC
5009 NE 292nd Ave
Camas, WA 98607

E-mail: dysert.larry@syncrude.com