Estimate Validation and Bias Assessment: **Ratio-to-Driver Method**

BY **JOHN K. HOLLMANN, PE CCP CEP DRMP FAACE HON.LIFE**

ABSTRACT

Cost estimate validation is often mentioned in AACE International literature but not described in depth. This article describes the practice of cost estimate validation including a method called ratio-to-driver. Validation starts with the business establishing an objective in terms of a cost strategy that is captured in the basis of estimate document. The cost strategy defines the estimating approach in respect to desired base estimate bias (and every estimate is biased). Achievement of the cost strategy is the quality being assured by validation. Then, reliable, normalized metrics (cost estimating relationships in ratio form) are developed from a comparison set of projects drawn from an historical database (or obtained from some other reliable source). Database systems often do double-duty as validation tools; a precursor to the future of analytics and machine learning. The ratio-to-driver method applies the metrics in a logical, stepped sequence of comparisons that seeks to pinpoint the cause of variations. Because base estimate bias is a systemic risk, and validation measures bias, validation is also a first step in quantitative risk analysis. While a long-established practice, estimate validation is not defined in AACE Recommended Practice 10S-90, Cost Engineering Terminology, and is only superficially covered in other estimating RPs. As such, this article is intended as a basis for an RP that will be aligned with others that include validation or benchmarking. The primary effected RPs (with abbreviated titles) are: 31R-03 (estimate review), 34R-05 (basis of estimate), 35R-09 (estimate planning), draft CE-81 (estimate requirements) and 42R-08 (parametric risk analysis). This article was first presented at the 2019 AACE International Conference & Expo as EST.3184.

Introduction and Background – What is Estimate Validation?

This article will define recommended practices for cost estimate validation including a method called ratio-to-driver. The practice of validation also applies to estimating time; i.e., schedule durations. However, this article and the potential RP are limited to concise cost estimates. Cost estimate validation is often mentioned in AACE International literature in conjunction with estimate reviews (as in "estimate review and validation") or estimate benchmarking, but validation as its own practice is not described in depth. The practice is touched on in the AACE *Total Cost Management (TCM) Framework*, Chapter 7.3 [28] and the following AACE Recommended Practices (RPs with topics in parentheses): 31R-03 (estimate review) [3], 34R-05 (basis of estimate) [4], 35R-09 (estimate planning) [5], draft CE-81 (estimate requirements) [10], and 42R-08 (parametric risk analysis) [8]. The terms estimate review and estimate validation are also not included in the AACE RP 10S-90, Cost Engineering Terminology [1]. Before describing the practice, the following provides proposed definitions.

The following relevant definitions, as included in the current RP 10S-90, set the context for defining estimate review and estimate validation:

- VALIDATION – Testing to confirm that a product or service satisfies user or stakeholder needs. Note difference from verification.
- VERIFICATION – Testing to confirm that a product or service meets specifications.
- QUALITY – Conformance to established requirements (not a degree of goodness).
- BENCHMARKING – A measurement and analysis process that compares practices, processes, and relevant measures to those of a selected basis of comparison (i.e., the benchmark) with the goal of improving performance. The comparison basis includes internal or external competitive or best practices, processes or measures. Examples of measures include estimated costs, actual costs, schedule durations, resource quantities, etc.

RP 31R-03, Revising, validating, and documenting the estimate, further describes estimate review and estimate validation as unique steps in a process as shown in Figure 1, [3].

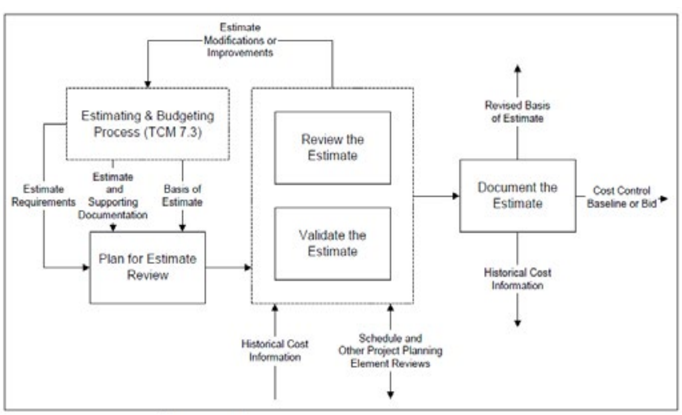

FIGURE 1 Estimate Review, Validation and Documentation Process (RP 31R-03)

Considering the descriptions of the RP 31R-03 steps and the existing 10S-90 terms for context, proposed definitions for estimate review and estimate validation are:

- ESTIMATE REVIEW – A quality assurance process, typically qualitative in nature, to test or assure that an estimate of cost or time technically conforms to *estimating requirements*.
- ESTIMATE VALIDATION – A quality assurance process, typically quantitative in nature, to test or assure that an estimate of cost or time meets the *project objectives* in regards to its appropriateness and competitiveness. A form of benchmarking that compares relevant estimate cost, time and/or resource measures (e.g., metric ratios) to those of a selected basis of comparison.

The main differentiator between these practices is that validation is to assure *project objectives* are achieved while review is to assure *conformance to technical requirements*. Estimate requirements are the topic of RP CE-81 (in technical committee review [10]). However, the objectives that validation are intended to assure have not been well defined.

PROJECT OBJECTIVES AND ESTIMATE VALIDATION

TCM Framework section 4.1.2.3 (Establish Objectives and Targets) states that, "Objectives usually reflect the general success criteria of the asset owner and/or whoever is funding [or bidding on] the investment" [26]. In respect to project cost (or time), success is usually viewed and measured in two ways; predictability and competitiveness. If improving *predictability* is an entity's success criteria, the measurement focus is on accuracy; i.e., being on budget and on forecast cash flow. If improving *competitiveness* is an entity's success criteria, the measurement is cost effectiveness; i.e., lower absolute cost (or time) for the same scope (this is similar to the concept of value). It is a challenge to balance predictability and competitiveness. For example, targeting lower cost (better performance than in the past) often means taking risks which results in more uncertainty and less accuracy [18].

In estimating, predictability and competitiveness objectives will be expressed as an explicitly planned *bias* in the base estimate (and later the control budget). For example, targeting is a typical competitive strategy wherein planned improvements in cost and performance (supported by improvements in practices) are set as goals. On the other hand, predictability strategies tend to have a financial focus on or bias toward hitting budgets by period (cash flow) and overall. Predictable strategies are common in government projects that are authorized and funded on a fiscal year basis (unfortunately, losing sight of competitiveness is often not an effective use of tax money). In any case, the goal of validation should not be to repeat history; when combined with effective quantitative risk analysis, improvement in project systems, practices and outcomes should result.

In addition to predictability or competitiveness bias, the project *execution strategy* has cost estimating implications in that it should communicate whether the project is *cost-driven* or *schedule-driven*. This strategy guides decision making when changes and risk responses are assessed during execution and earlier in project planning. One major benchmarking firm defines a schedule-driven project as "one in which the business is willing to *trade* capital cost to achieve schedule" [21]. The driver implies a bias but does not determine it; for example, while a schedule-driven project may sacrifice cost for schedule, there may still be a target objective for the cost being sacrificed. Assuring this "driver" objective is achieved is done during quality assurance of the change and risk management processes.

In summary, the first step in the practice of estimate validation is to establish and communicate the objective(s) of the estimate. The objectives should address predictability and/or competitiveness along with any other goals. Without stated, clear objectives, the value of validation is greatly diminished (just as estimate review is less valuable if there are no stated estimate requirements to assure). The place to communicate the objectives is first in the overall estimate plan (re: RP 35R-09 for buildings or 36R-08 for process) and later in the basis of estimate (re: RP 34R-05) and estimate requirements for estimates by third parties (draft RP CE-81). As of this writing, these AACE RPs do not explicitly include communicating the cost objective. However, the author recommends that a "Cost Strategy" section be added to the various estimating RPs to define how to communicate the business objective in a way that directs and facilitates the estimating process.

Strategy	Typical Basis and Bias of the Base Estimate
Targeted	The base represents a defined level of performance relative to past performance based on analysis. Usually aggressive (competitive) but may be conservative (predictable) relative to the past as defined by a documented cost strategy.
Historical Norms (or Analogy)	The average or typical performance of past projects (i.e., *realistic*.) Includes the impact of nominal risk events of the past. This is a conservative (predictable) approach as defined in a documented cost strategy or undefined but consistent approach of estimating function.
Ad-Hoc	Approach left to will of the estimator. May use a *de facto* historical norms approach, but it is less consistent and reliable.
Database Reference	Use a reference database with a defined basis, which could be any of the three above. The basis of legacy data is often murky, and bias may be inconsistent from item to item. Users apply "database markups" that may result in "historical norms" if the database and markups have consistency. If not, the strategy is similar to *ad-hoc*.

TABLE 1 Typical Cost Strategies That Guide Cost Estimating ([9] with permission)

DOCUMENT A COST STRATEGY FOR THE BASE ESTIMATE

The cost strategy is a statement in the basis of estimate and the estimate requirements describing the objective of the estimating process and the general approach to achieving that objective. (This sentence can serve as a proposed RP 10S-90 definition for *cost strategy*). The strategy is defined by the business (or tender) sponsor and the strategy statement must be agreed by them. Unfortunately, one rarely sees a clear strategy stated in basis documents and the sponsors may not even understand why the estimator is asking the question (which is a good reason to ask). Table 1 describes some typical cost strategies (often implied and not documented) that more or less guide the *base* estimating process. Each strategy has a description of its bias toward predictability (conservative) or competitiveness (aggressive) or indeterminate or random (never recommended).

The author's experience is that management usually trusts the estimating lead to apply some sort of reasonable cost strategy to the base; it is left to the estimating lead's discretion. The estimating lead in turn often repeatedly applies some institutionalized approach that is "the way we do it" but is not articulated. Because most estimators are risk averse (not anxious to take responsibility or blame for overruns), management is in effect defaulting to the historical norms approach in Table 1. This is the behavior of a predictability culture which is also often punitive in respect to cost overrun and schedule slippage. Cost competitiveness is more difficult to achieve than predictability and being predictably competitive is the most challenging to achieve. An example of a competitive cost strategy statement that one might find in a complete basis of estimate where estimating is backed by an excellent historical database is as follows:

> *The base cost and duration estimate values will reflect aggressive but reasonably achievable current pricing and performance. "Aggressive but reasonably achievable" means that the assumed performance will reflect the first quartile level (i.e., p25) of historical performance or equivalent for similar strategies and scope excluding the impact of identifiable changes and risks [18].*

The cost strategy for the base estimate should be consistently reflected elsewhere in the basis of estimate as applicable. For example, when describing the basis for equipment costs in a competitive strategy, a "the least cost, technically acceptable tender" approach might be chosen as opposed to a more conservative "mid-point of tenders" or other approach. Such a clear cost strategy statement will provide the estimator with guidance as well as assurance that if the base cost is overrun, they will not be held responsible for the project's failure to perform or the impact of risks and so on. The statement also communicates a clear objective that estimate validation can assure has been achieved.

BUSINESS STRATEGY AND TOTAL COST ("THE NUMBER")

The cost strategy in the basis of estimate is intended to guide the estimating function in its base estimating practice. Validation assures that strategy is achieved in the base. The cost strategy reflects an overall business objective established by the business case. As discussed in the AACE **TCM Framework**, Chapter 6.1 on *Asset Performance Assessment*, profitability of the capital investment is usually a main objective, and Net Present Value (NPV) and other Return on Investment (ROI) methods are the most common means of measuring profitability [27]. In the non-profit world, minimizing capital spending will still be an objective even if revenue is not the measured benefit. Having validated the base estimate, there remains the step of validating or, as most would refer to it at this high level, *benchmarking* the total cost.

Even prior to developing a cost estimate, most businesses will have a general idea of the limit of capital spending for an investment that will result in a positive NPV for a given revenue projection (or what will be a successful tender). Experienced estimators are familiar with the tyranny of "*the number*" [22]; i.e., a total cost announced by the business but of indeterminate basis. The number often reflects a strong bias; usually an optimism bias.

AACE recommended practice for cost engineering professionals is not to accept numbers with an indeterminate basis. The total cost should be a formally estimated base plus the risk costs; i.e., contingency, management reserves and escalation. AACE RP 40R-08 establishes the principles for quantifying the risk including providing "probabilistic estimating results in a way the supports effective decision making and risk management" [7]. Therefore, unlike the base estimate, the total cost should not be expressed as a *number*, but a probabilistic distribution. Further, in accordance with 40R-08, a recommended quantitative risk analysis (QRA) method employs empiricism; i.e., it will be based on actual practices and results. As such, quantitative risk analysis is inherently a form of validation, albeit reflecting internal data. Based on probabilistic QRA, management decides on a *number* that is in accordance with their explicit bias, otherwise known as risk tolerance expressed as a probability or confidence level of underrun (e.g., "fund at p50").

If the total cost is based on probabilistic QRA, then what remains for estimate validation is to compare the decided upon number, expressed as a metric, to external metrics or benchmarks. As will be discussed later, these overall metrics are usually gross unit costs (cost/quantity such as $/m2) or cost-capacity ratios (cost/capacity such as $/barrel per day throughput). It is difficult to obtain robust external data at a more detailed level except through industry benchmarking consortia or consultants [20].

RELATIONSHIP OF ESTIMATE VALIDATION TO QUANTITATIVE RISK ANALYSIS (QRA)

RP 42R-08 on parametric risk analysis of systemic risks states that base estimate bias is one of the systemic risks being quantified in risk analysis [8]. It also states that "estimate validation (to detect bias among other objectives) is always a recommended practice in conjunction with risk analysis." Estimate validation provides an objective measure of whether and to what degree the base estimate is aggressive or conservative relative to the comparison metrics. Bias measurement is critical to realistic risk analysis and contingency determination because the cost risk or contingency is quantified relative to the base estimate. If the base estimate is conservative there will be less need for contingency and vice-versa for aggressiveness. Aggressiveness taken to extremes and/or not supported by excellent practices and appropriate contingency will add risks by stressing the project system (particularly for mega-projects). In any case, estimate validation and bias measurement is not only an estimating process step, but a QRA step as well.

While RP 42R-08 addresses rating base estimate bias, for the total cost, if the risk funding is based on probabilistic QRA methods, then management will be presented with a cost distribution that they can use to fund or tender the project at their desired level of confidence in underrunning. In short, estimate validation and QRA are closely tied processes across the board.

Estimate Validation Process and Metrics

ESTIMATE VALIDATION PROCESS

This article has defined the essential estimate validation method as "a form of benchmarking that compares relevant estimate cost, time and/or resource measures (e.g., metric ratios) to those of a selected basis of comparison." The practice relies on the concept of Cost Estimating Relationships (CERs). CERs per RP 10S-90 "show some resource (e.g., cost, quantity, or time) as a function of one or more parameters that quantify scope, execution strategies, or other defining elements" [1]. For example, if one knows the equipment costs for a process plant, one can estimate the rest of the cost of the plant using a CER ratio of total cost/equipment (e.g., a Lang factor). Such ratios are used in conceptual estimating and may be used in estimate validation as well.

The basic estimate validation process is fairly straight-forward as shown in Figure 2:

- Plan the estimate validation considering the cost strategy
- Calculate the metrics for the project estimate
- Obtain comparison metrics normalized to the estimate metrics basis
- Compare the estimate metrics to the comparison metrics and make assessment
- Determine if the cost strategy has been achieved (and other quality findings)
- Recommend estimate improvement actions and/or report bias for risk quantification

Figure 2 also shows how estimate validation relates to other processes in the TCM Framework or other RPs (connector symbols). In particular, database management is critical as a source of validation metrics. In lieu of an RP on databases, there are excellent papers available on the topic [12, 14, 17]. This article does not define these other processes. However, to understand validation, the topic of *normalization* (a sub-topic of database management) is addressed here at a summary level.

ESTIMATE VALIDATION METRICS OR RATIOS

Validation relies on the concept of CERs which for validation purposes are expressed mathematically as ratios (metrics) of resource measures found in a cost estimate. The resources are cost, hours and quantities. (For schedule validation, one would add "time" to this list. i.e., quantity/time is a production rate.) The nine possible ratios or metrics for these three resources are listed in Table 2. Using concrete as an example, the relationships that one could examine for variations from benchmarks might include the cost of concrete relative to the cost of steel (cost/cost), the unit cost of the concrete (cost/quantity) or the unit hours for concrete (hours/quantity). Each comparison and any deviations from target may tell a different story. However, before calculating the ratio metrics, one must normalize the historical data to the current basis of comparison.

cost/cost
hours/cost
cost/hours
cost/quantity
hours/hours
hours/quantity
quantity/quantity
quantity/cost
quantity/hours

TABLE 2 General Cost Estimate Validation Ratio or Metric Types

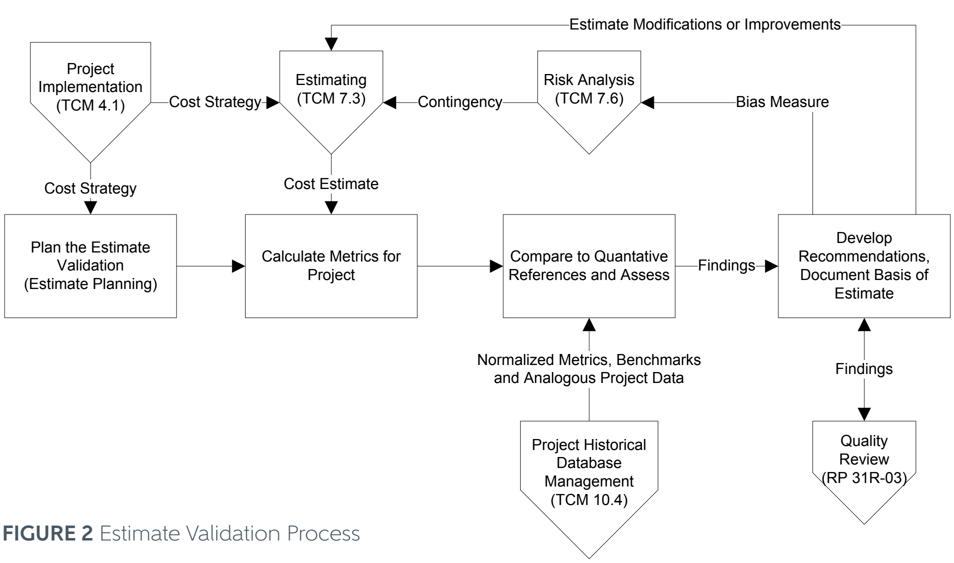

FIGURE 2 Estimate Validation Process

NORMALIZATION

A principle of estimate validation is that metrics comparisons should be on an apples-to-apples basis. For example, one would not compare the concrete unit cost for a nuclear containment structure to the concrete unit cost for a parking garage structure. However, there are some project differences that are reasonably explainable and can be adjusted for *a priori* to make maximum use of the available data. The adjustment process is called *normalization*. Per RP 10S-90, normalization is "a process used to modify data so that it conforms to a standard or norm (e.g., conform to a common basis in time, currency, location, etc.)" [1]. The following more fully describes the main drivers of metric deviations that can be corrected for:

- Escalation: changing economic and market conditions over time
- Currency: exchange rate and its change over time
- Location: adds regional productivity, labor rate and material cost differences to above

Escalation

Per AACE RP 10S-90, escalation is "a provision in costs or prices for uncertain changes in technical, economic, and market conditions over time. Inflation (or deflation) is a component of escalation" [1]. AACE RP 58R-10 further explains how to use historical price indices from a reliable source to normalize historic project costs to the basis of comparison reference date [9]. For example, if the database has the cost for concrete in 2009 (the year of expenditure), and the goal is to use this to compare to the estimated cost of concrete in 2019, the database cost item would be adjusted as follows:

- Concrete cost in 2019 = Concrete cost in 2009 x (2019 price index/2009 price index).

All of the database costs would be adjusted to the basis of comparison time period prior to calculating the comparison metrics. One must be aware that the adjustment, while reasonably reliable, adds some uncertainty to the validity of the comparison metrics. This is particularly true if volatile economic times occurred when the historical cost were expended or between the historical date and the current time.

Currency (and Escalation)

Historical costs are adjusted or normalized to a historical cost in one currency to the cost in a different currency using exchange rates. For example, if the cost of item in the database is one euro, and the basis of comparison is US dollars (and the historical data is for the same time period) the database item cost would be multiplied by the dollar/euro exchange rate. However, the historical item is usually from a prior time period. So, both escalation and currency need to be adjusted in the same normalization process step.

AACE RP 58R-10 describes a method to normalize an historical cost for currency and escalation together. It recommends adjusting for escalation to the basis of comparison time period first, using a price index for the historical location, then adjusting for currency using the exchange rate of the basis of comparison time period. An example is as follows for adjusting a $100 item in the database in 2009 Canadian dollars to a basis of comparison in 2019 US dollars:

Given:

- Canadian price indices from of 1.10 in 2009 and 1.32 in 2019
- Exchange rate in 2019 is 0.90 $US per $CAN

Then adjustments are:

- Escalation: ($100 in 2009 $CAN) x (1.32/1.10) = $100 x 1.20 = $120 in 2019 $CAN
- Exchange: ($120 in 2019 $CAN) x (0.90 $US/$CAN in 2019) = $108 in 2019 $US

All of the database costs would be adjusted to the basis of comparison time period and currency prior to calculating the comparison metrics. There are arguments for doing currency adjustment first, then escalation; it may give a different result. Just be aware that this adjustment is adding uncertainty to the validation that should be allowed for in assessment (i.e., significant concern over minor metric variations is not justified).

Location (Productivity, Labor Rates and Local Materials)

Location normalization is adjustment to both the hours (productivity) and cost resources (labor rates and material costs). This is the most uncertain adjustment because productivity is driven by both regional labor market characteristics and project-specific labor performance issues (i.e., risk drivers). Also, labor rates vary depending on crew make-up and construction methods. Some would argue that regional labor difference is like scope difference; for example, it might not be advisable to compare cost data from China to the US. Another uncertainty is that labor productivity changes over time. For example, since the 1970s, labor productivity in China has improved dramatically while labor productivity in the US Gulf Coast regions has arguably declined.

AACE RP 28R-03 describes a method to adjust costs for location including not only labor cost but for material costs [2]. It can be a fairly complex calculation. It starts with breaking a cost down into elements (i.e., labor hours, labor rates, materials, etc.), adjusting each element to the new location basis using what it calls *location factors*, then compiling the cost back into the new location basis. Escalation and exchange rate adjustment would then be made to bring the compiled cost to the new comparison basis in time and currency. While complex, once a dataset of location factors has been developed for various cost elements, they can be reused and periodically updated. The method is too complex for this paper; readers should refer to the RP.

Having normalized the historical database resource data to the basis of comparison in time, currency and location, the ratio metrics using this adjusted data can be calculated for use in estimate validation comparisons.

Estimate Validation Methods

BASIC METHOD: SUCCESSIVE DETAIL

The basic method in the "Compare" step of Figure 2 for quality review purposes is to calculate the +/- percentage difference between the estimate metric and a normalized comparison metric (e.g., an average derived or obtained from the database or other reference), and then assess causes for the percentage deviations. Pre-determined thresholds of acceptability may be established when planning for validation (e.g., within +/-10% or some other acceptable range considering the uncertainty of these metrics). Any delta outside the threshold is further assessed to ascertain cause. If the deviation is explainable, then the validation comparison step moves on to the next metric.

As better data is developed, the comparison becomes more statistically robust. The term *analytics* is increasingly used to describe this level of data-driven analysis; a precursor to machine learning and artificial intelligence [11]. For example, instead of comparing to an average or typical (rule-of-thumb) metric value, it may be possible to compare to a range of metrics from a comparison dataset drawn from the database (sometimes referred to in benchmarking parlance as the *compset*). For example, compare one's metric to low, average, and high values from the compset. Optimally, the compset is robust enough to support generating a full distribution or histogram such that it is possible to conclude the estimate metrics is say "p30 of the compset" (i.e., a competitive bias). This method has been used to calibrate a base cost estimating database as well [25].

The following provides example comparisons to single values (rules-of-thumb, targets, benchmarks, parametric model output, etc.), ranges, and distributions. The distribution method is optimal for setting and assuring cost strategy (e.g., ability to target a p-value).

- Single Value: Estimated 12 hours/tonne is 20% greater than rule-of-thumb benchmark of 10 hours/tonne. This is outside the +/-10% (or some other) desired threshold. Commonly used with historical norms cost strategy.
- Range: Estimated 12 hours/tonne is within the low-high range of 7 to 14 hours/tonne, but 20% greater than the average of 10 hours/tonne. Similar to single value but with added information as to extreme thresholds. Implies estimate is near the max in this case. May use median as the reference.
- Distribution: Estimated 12 hours/tonne is the p80 value of the historical compset distribution as shown in Table 3 and Figure 3. This is common with target cost strategy and high-level benchmarking; i.e., target may be expressed as p30 of compset as opposed to a rule of thumb or average value.

P-value	hours/tonne
10%	7.5
20%	8.0
30%	8.5
40%	9.0
50%	9.5
60%	10
70%	11
80%	12
90%	13

TABLE 3 Example of a Tabular Comparison Metric Distribution

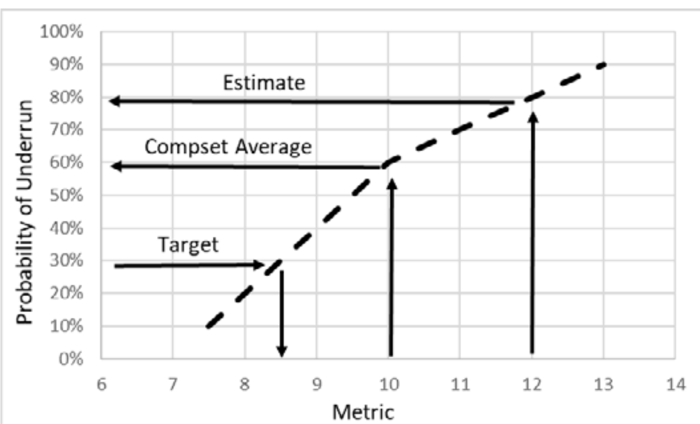

FIGURE 3 Example of a Graphical Comparison Metric Distribution

Examples of estimate validation tabular and graphical reports for the *Range* method are provided in Appendix A.

Note that in the single value and range examples, the comparison is typically an average or historical norm represented in the case as 10 hrs/tonne. However, for the example Table 3 or Figure 3 make it clear that 6 times out of 10 the company did better than 10 hrs/tonne; i.e., setting the base at 10 hrs/tonne is not competitive. Setting the goal at 8.5 hours per tonne would be aggressive but reasonably achievable and in effect adjusts the data to remove the impact of nominal risk events that should be funded by contingency, not in the base.

At a project level, the metric comparisons should be made in a structured way for the entire base estimate to assure that the project cost strategy is being achieved across the board. The typical structure is to examine metrics from the top down, level-by-level in the work breakdown structure (WBS) and cost code of accounts (disciplines) until estimate quality is assured. To do this it is necessary to have a standard code of accounts; a good example is the International Construction Measurement Standards (ICMS) structure [19].

At the top level, it is possible to start with benchmarking using a gross unit cost or cost-capacity measure where the cost includes the base as well as risk costs (contingency, management reserve and escalation). At this level, benchmarking is to assure the business case objective is met (e.g., NPV or ROI). Some example metrics at a business level include;

- Building project: total cost/square meter (cost/gross quantity)
- Infrastructure project: total/passenger kilometer (cost/capacity)
- Process plant: total cost/barrels per day (cost/capacity)

The validation would be against internal metrics if available (there may be a few greenfield projects in the database), but also external metrics from other sources such as benchmarking firms, consultants or publications. Some have taken to calling this *reference class forecasting* (RCF) [15], but in AACE terminology, and for benchmarking firms, it is simply estimate validation of the overall project using external benchmarks. Capturing and sharing industry cost-capacity data for estimating and benchmarking has been an interest of AACE since near its founding [16].

At a high level, but considering only the base costs, it is common to use cost/cost metrics. An example in the process industry is the Lang factor which is a ratio of the total cost/cost of equipment [13]. There are variations on this metric such as including only direct costs in the numerator, including only major process equipment in the denominator, and so on; documenting the specific metric being used is suggested.

Any deviation from strategy at the top level would be noted, but do not stop there. To further search for the source of a deviation, or to assure lower level deviations were not cancelling each other out, comparisons would be made for similar metrics at successively lower levels of detail in the WBS until satisfied or until the database would not support more detailed examination. For cost accounts or disciplines, it is good practice to examine metrics for each major account and then drill down. This could be done for the overall project and then for lower levels of the WBS for large projects. For example:

- Construction: total construction costs/total cost
- Field indirect costs: total field indirect costs/total construction costs
- Construction equipment: construction equipment costs/total field indirect costs

In addition to pure WBS/code of account views, examining metrics that represent various areas of responsibility is also possible. For example:

- Physical design: quantity/quantity of select commodities
- Procurement and design: costs/quantity of select commodities
- Engineering: hours/quantity of each discipline
- Construction: hours/quantity of each discipline
- Management: office hours/field hours
- Engineering/Construction: hours design/hours construction by discipline

Note the importance of *key quantities* to these metrics. Quantity data, which addresses scope, is essential to complete estimate validation; cost and hours data is not enough. For example, while examining the cost of engineering/cost of construction ratio, it might appear to be a competitive metric even though the cost in both the numerator and denominator are not competitive; i.e., both might be elevated due to overall poor project practices. However, examining the cost of engineering or construction as a ratio to quantity (by discipline) will highlight poor performance correctly because quantity is not as affected by performance.

As part of the metrics analysis, care should be taken to assure there is no perverse trading between management and/or engineering and construction costs. For example, low, but inadequate project management (PM) costs may result in high construction costs (due to quality or poor performance in the field). Management may perceive a low PM cost/construction costs ratio as success when in fact all they did was increase the construction costs through weak management (lower the metric by increasing the denominator). Measures of engineering and home office in relation to key quantities are typically better metrics than cost/cost ratios.

BASIC METHOD ENHANCEMENT: TRIANGULATION

In addition to seeking the source of strategy deviations by successive detail, one can leverage the fact that there are three *resources* to consider for each *account*: hours, cost, and quantity to help one search for evidence. Looking at one account using multiple ratios is called *triangulation* (even if there are more or less than three views). This may help identify whether the source of any deviation is in the hours, cost or quantity. Using the concrete as an example, there are multiple metrics that can be generated for the estimate and the compset; examples include:

- cost of concrete/m3 of concrete (cost/quantity)
- hours of concrete/cost concrete (hours/cost)
- cost of concrete/cost steel (cost/cost)

If the cost/quantity was high and the hours/cost was also high, this might suggest the cause of the concrete metric deviation is in the labor, not the material cost. Moving on, if the cost of concrete/steel was also high, it is possible to hypothesize that there was significant elevated concrete structure, which is more costly than foundation concrete on a unit basis. They hypothesis of cause could be confirmed by examining the design.

The validation assessment outcomes should be reported in the basis of estimate (BoE) as appropriate. Key sections in the BoE as regards bias are the description of allowances (additions) and exclusions (subtractions) from the base estimate or schedule. These should be carefully considered as to their appropriateness, with a particular eye out for hiding of contingency above the line and/or overly aggressive results of "cost savings" initiatives. An independent validator can help assure management is made aware of bias when it is found; this is a challenge when management is causing bias but not willing to document it as their strategy.

RATIO-TO-DRIVER METHOD

Ratio-to-driver is a stepped approach, but rather than just marching through a list of accounts in code-of-account order, this approach is based on the fact that one cost is usually driven by another cost or resource; hence the term *ratio-to-driver*. For example, the need for concrete is driven by the need to support steel and equipment, so a metric of concrete volume/steel weight should give an indication if the concrete volume is in line with historical norms. The following list shows the general sequence or order of metrics to look at starting with the quantities being designed and installed:

1. Quantity/Quantity (indicates the efficiency of the design)
2. Bulk Material Cost/Quantity
3. Direct Field Labor Cost/Bulk Materials Cost
4. Field Indirects Cost/Field Directs Costs
5. Engineering Costs/Quantities
6. Engineering Hours/Direct Field Hours
7. (PM and Owners Costs)/(Field and Engineering Labor Costs)

This sequence is seeking the root cause of deviations. Quantity deviations are the root (odd ratios may indicate design idiosyncrasies-check for them first). Bulk material costs are then driven by quantities. Labor is driven by the bulks being installed. Field indirect costs are driven by direct labor needs, and so on. This minimizes mistaken-cause hypotheses that sometimes result from lists of metrics or even triangulation. An example of a ratio-to-driver estimate validation tabular report is provided in Appendix A.

A practice to be avoided is using metrics that include a given cost in both the numerator and denominator (e.g., engineering costs/total project costs). These types of metrics obscure the cause and effect and variations in the numerator are buffered or hidden. For example, assume a base project had $20 engineering, $50 direct field costs and $100 total cost. If engineering costs increased from $20 to $40, total cost would increase from $100 to $120. An engineering/total cost ratio would be 0.20 (20/100) for the base and 0.33 (40/120) for the increased case; a 65% increase in the metric. If one had measured engineering/field cost, the metric would go from 0.40 (20/50) to 0.8 (40/50); a 100% increase that is more likely to be noticed.

Another consideration is to be sure the purpose of each metric can be appropriately described (i.e., one gets what one measures). For example, if the concern is the cost effectiveness of engineering, the ratio of engineering costs/total project costs is a poor metric to use because this metric tends to be lowest for those projects with the most field rework resulting from poor design. Also, engineering *costs* hide whether a problem is in the hours or the rates. A better metric would be something like mechanical engineering hours/tonne of piping because it is in accordance with the ratio-to-driver principle.

Validation and Quantitative Risk Analysis (QRA)

RP 42R-08 documents the parametric method of quantifying systemic risks. As mentioned, bias of the base is a strong determinant of the need for contingency (e.g., if there is above the line contingency, then little additional may be needed). When creating a parametric model, the input parameter for bias will likely be a rating as opposed to a direct percentage entry (validation is indicative estimate, not an exact accounting) [18]. The following is an example rating of bias on a 1 to 5 scale.

1 - Very Conservative (e.g., >p80)
2 - Somewhat Conservative (e.g., p60-p80)
3 - Average (i.e., Historical Norm or p40-p60)
4 - Somewhat Aggressive (p20-p40)
5 - Very Aggressive (e.g., <p20)

The p-values reflect the project's level of confidence in a distribution of compset metrics. Using the example from Figure 3, the 12 hours/tonne was the p80 of the compset. Hence the parametric input rating would be 2 (but close to 1) which would result in less contingency in the model algorithm. In practice the rating would reflect a weighting of the entire estimate, not just one item metric.

Conclusion

This article described a cost estimate validation process and methods, including an approach called *ratio-to-driver*. It highlighted that validation should start with establishing an explicit *cost strategy* (planning for bias) in the basis of estimate and estimating requirements documents. Validation's purpose is to assure this strategy was achieved. The article also discussed how validation provides a bias measure for parametric models of systemic risks. Finally, the article is offered as the basis for a potential AACE Recommended Practice (RP) for estimate validation. If accepted as such, updates to related RPs would be required (in particular; 10S-90, 31R-03, 34R-05, 35R-09, CE-81 and 42R-08.) In addition, the author looks forward to an RP on project historical database management which is a necessary practice to support estimate validation.

References

1. AACE International, Recommended Practice No. 10S-90, Cost Engineering Terminology, Morgantown, WV: AACE International, Latest revision.
2. AACE International, Recommended Practice No. 28R-03, Developing Location Factors by Factoring, Morgantown, WV: AACE International, Latest revision.
3. AACE International, Recommended Practice No. 31R-03, Reviewing, Validating, and Documenting the Estimate, Morgantown, WV: AACE International, Latest revision.
4. AACE International, Recommended Practice No. 34R-05, Basis of Estimate, Morgantown, WV: AACE International, Latest revision.
5. AACE International, Recommended Practices No. 35R-09, Development of Cost Estimate Plans – As Applied for the Building and General Construction Industries, Morgantown, WV: AACE International, Latest revisions.
6. AACE International, Recommended Practices No. 36R-08, Development of Cost Estimate Plans – As Applied in Engineering, Procurement, and Construction for the Process Industries, Morgantown, WV: AACE International, Latest revisions.
7. AACE International, Recommended Practice No. 40R-08, Contingency Estimating-General Principles, Morgantown, WV: AACE International, Latest revision.
8. AACE International, Recommended Practice No. 42R-08, Risk Analysis and Contingency Determination Using Parametric Estimating, Morgantown, WV: AACE International, Latest revision.
9. AACE International, Recommended Practice No. 58R-10, Escalation Estimating Principles and Methods Using Indices, Morgantown, WV: AACE International, Latest revision.
10. AACE International, DRAFT Recommended Practice CE-81, Owner's Estimate Requirements – As Applied in Engineering, Procurement, and Construction for the Process Industries, Morgantown, WV: AACE International, Latest revision.
11. Arrow, James E. et.al., Project Controls and Data Analytics in the era of Industry 4.0, AACE International Transactions, Morgantown WV, 2018.
12. Bomba, Susan W. and Lamis El Didi, Benchmarking and Predictive Analytics to Improve Estimates, Forecasts, and Performance Measurement, AACE International Transactions, Morgantown WV, 2018.
13. Dysert, Larry R., Chapter 9: Cost Estimating, Skills and Knowledge of Cost Engineering-Sixth Edition, AACE International, Morgantown WV, pp. 99-100, 2015.
14. Figueiredo, F. Cristina and Ray S. Philipenko, Taking a Project Knowledge Management System (PKMS) to the Next Level, AACE International Transactions, Morgantown, WV, 2011.
15. Flyvbjerg, Bent, Curbing optimism bias and strategic misrepresentation in planning: Reference class forecasting in practice. European Planning Studies, 16 (1). pp. 3-21, 2008.
16. Hand, W.E., Cost-Capacity Curves, AACE Bulletin, Vol 3, No. 2, AACE International, Morgantown, WV, 1961.
17. Hollmann, John K., Project History – Closing the Loop, AACE International Transactions, Morgantown, WV, 1995.
18. Hollmann, John K., "Project Risk Quantification", Probabilistic Publishing, Sugarland, TX, 2016. p105.
19. International Construction Measurement Standards Coalition, International Construction Measurement Standards (ICMS): Global Consistency in Presenting Construction Costs, 1st edition, July 2017.
20. IPA, Inc., Summary Cost Metrics and Detailed Cost Metrics, www.ipaglobal.com/summary-cost-metrics and detailed-unit-cost-metrics, 2018.
21. Kulkarni, Phyllis, Why Are So Many Site & Sustaining Capital Projects Schedule-Driven?, IPA Insites, IPA Inc. (www.ipaglobal.com), 31 March 2015.
22. McDonald, Donald F., "The Number" or "The Result"; Reliability, Accuracy, Precision, Confidence, or What?, Cost Engineering, AACE International, Morgantown WV, Vol 37/No.1, January 1995.
23. Pickett, Todd W., Developing Key Quantities, Unpublished presentation, Conquest Consulting Group (http://www.ccg-estimating.com/pdfs/WP-DevelopingKeyQuantities), 2015.
24. Pickett, Todd W., and Bruce G. Elliott, Transforming Historical Project Data into
25. Useful Information, AACE International Transactions, Morgantown WV, 2007.
26. Srivastava, Prashant K., et.al., Path to Predictability: The Evolution of An Owner Integrated Cost Engineering System, AACE International Transactions, Morgantown WV, 2018.
27. Stephenson, H.L., Ed., Total Cost Management Framework (Section 4.1.2.3; Establish Objectives and Targets), 2nd ed., Morgantown, WV: AACE International, Latest revision.
28. Stephenson, H.L., Ed., Total Cost Management Framework (Chapter 6.1 Asset Performance Assessment), 2nd ed., Morgantown, WV: AACE International, Latest revision.
29. Stephenson, H.L., Ed., Total Cost Management Framework (Chapter 7.3 Cost Estimating and Budgeting), 2nd ed., Morgantown, WV: AACE International, Latest revision.

Appendix A: Estimate Validation Range Method Examples

The following are example *range* method estimate validation reports. The range method usually reports the estimate metric and compares it to the compset average and low/high range. Some make the comparison to the compset median (p50) in addition to or instead of the mean.

Figure A1 is a tabular estimate validation report from a company "project historical retrieval and analysis system (PHRAS)" that was first reported on in 1995 [17, 25]. Note that in addition to the variance of the estimate metric from the average, this system also allowed comparison to a third-party benchmarking company's metrics.

Figure A2 shows an estimate validation report (range method) graphic screen from a company's "project knowledge management system (PKMS)" [17]. The authors state that "The system captures the information of the estimate to be compared, and then retrieves the data as per search criteria...Once the data set is selected, the system develops a table with graphical capabilities that compares the metrics for the current project estimate against the metrics for each project selected and includes respective ranges (min/max) and average values." This company has since converted its in-house database to a commercial software package. Several such packages with full capabilities including normalization have come on the market.

COST RELATIONSHIP	--- COMPANY HISTORY --- LOW	AVG	HIGH	EST	VAR	INDUSTRY DATA AVG	VAR
TOTAL PROJECT $ / EQUIPMENT $	2.38	4.97	6.98	**4.77**	-4%	4.86	-2%
TFC $ / EQUIPMENT $	1.65	3.45	4.79	**3.20**	-7%	3.59	-11%
TOC $ / EQUIPMENT $	0.51	1.21	1.72	**1.15**	-5%	1.18	-3%
TFL $ / EQUIPMENT $	1.09	1.99	2.41	**1.80**	-10%	1.89	-5%
PROJ MGMT $ / TOTAL PROJ $	0.022	0.043	0.065	**0.037**	-14%	0.046	-20%
PROJ MGMT $ / TFC $	0.009	0.063	0.099	**0.054**	-14%	0.064	-16%
PROJ MGMT $ / TOC $	0.152	0.233	0.320	**0.210**	-10%	0.212	-1%
PROJ MGMT $ / EQUIPMENT $	0.190	0.295	0.452	**0.260**	-12%	0.241	8%
ENGR+DESIGN $ / TOTAL PROJ $	0.122	0.168	0.272	**0.198**	18%	0.189	5%
ENGR+DESIGN $ / TFC $	0.141	0.243	0.342	**0.291**	22%	0.262	11%
ENGR+DESIGN $ / TOC $	0.420	0.690	0.880	**0.872**	26%	0.779	12%
ENGR+DESIGN $ / EQUIPMENT $	0.810	1.020	2.000	**1.173**	15%	1.020	15%
TOC $ / TOTAL PROJ $	0.162	0.231	0.292	**0.245**	6%	0.229	7%
TOC $ / TFC $	0.190	0.310	0.550	**0.341**	10%	0.319	7%
TFC $ / TOTAL PROJ $	0.560	0.710	0.888	**0.682**	-4%	0.783	-13%
TFL $ / TOTAL PROJ $	0.272	0.362	0.523	**0.341**	-6%	0.362	-6%
TFM $ / TOTAL PROJ $	0.302	0.481	0.691	**0.419**	-13%	0.419	0%
TFL $ / TFC $	0.282	0.485	0.590	**0.457**	-6%	0.466	-2%
TFM $ / TFC $	0.291	0.550	0.817	**0.598**	9%	0.534	12%
TFL $ / TFM $	0.510	1.125	1.721	**0.902**	-20%	0.980	-8%
EQUIPMENT $ / TOTAL PROJ $	0.096	0.223	0.572	**0.241**	8%	0.251	-4%
EQUIPMENT $ / TFC $	0.125	0.304	0.632	**0.334**	10%	0.304	10%
START-UP $ / TOTAL PROJ $	0.000	0.034	0.158	**0.041**	21%	0.033	24%
START-UP $ / TFC $	0.000	0.051	0.111	**0.066**	29%	0.047	40%
START-UP $ / EQUIPMENT $	0.000	0.171	0.362	**0.235**	37%	0.169	39%
CONSTR IFC $ / TOTAL PROJ $	0.059	0.085	0.110	**0.079**	-7%	0.088	-10%
OWNER COST $ / TOTAL PROJ $	0.105	0.175	0.219	**0.155**	-11%	0.186	-17%
ENGR+DESIGN HRS / EQUIP COUNT	841	990	1287	**1120**	13%	1039	8%
TOT FIELD HRS / EQUIP COUNT	3087	4117	4734	**3884**	-6%	4089	-5%
CONCRETE CY / EQUIP COUNT	16	20	33	**16**	-20%	22	-27%
PIPING LF / EQUIP COUNT	392	474	687	**484**	2%	470	3%
CONCRETE HRS / CY CONCRETE	12.1	13.4	19.2	**12.4**	-7%	13.8	-10%
PIPING HRS / LF PIPE	2.7	3.2	3.4	**3.3**	3%	3.2	3%
CONCRETE MATL $ / CY CONCRETE	219	243	306	**224**	-8%	255	-12%
PIPING MATL $ / LF PIPE	88	103	113	**104**	1%	102	2%

FIGURE A1 Example Tabular Report from a Database System [24] (Used With Permission)

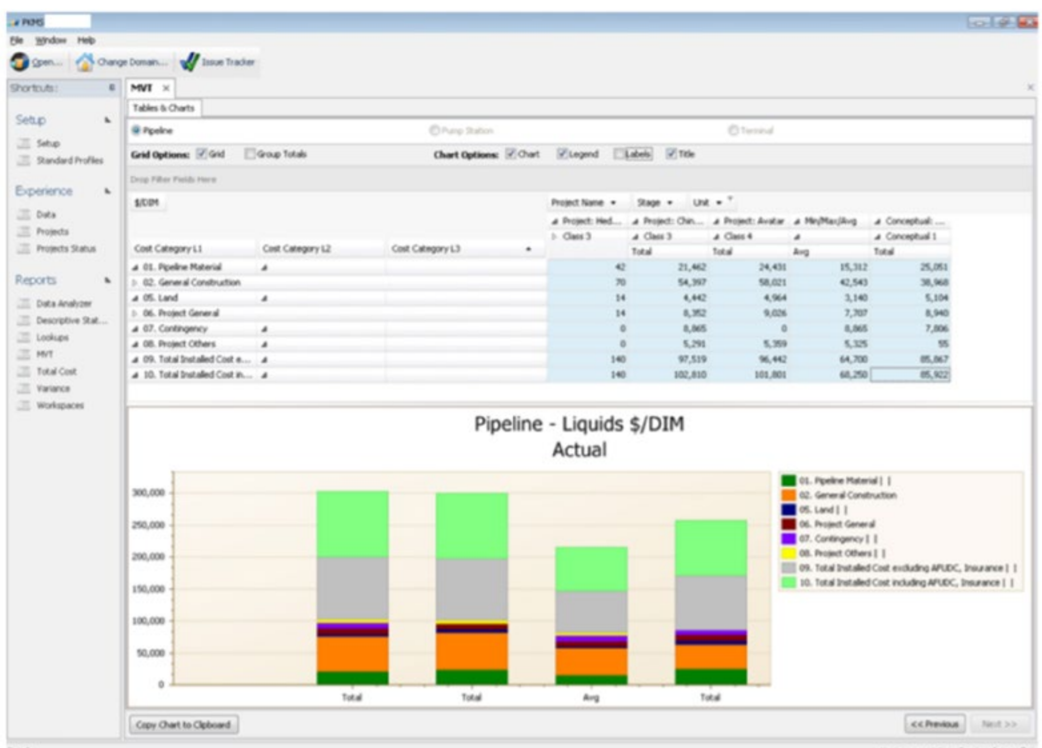

FIGURE A2 Example Graphic Report from a Database System [14] (Used With Permission)

Description	Quantity		Total Hour	Hours per Unit		Variance to Benchmark Average	Variance to Benchmark Median	Average		Median		Low		High	
								Hours per Unit		Hours per Unit		Hours per Unit		Hours per Unit	
Structural Steel	9,365	MT	368787	39.38	MT	-17.9%	3.3%	47.96	MT	38.14	MT	25.46	MT	69.47	MT
Piping	42,888	LM	527094	12.29	LM	6.0%	2.3%	11.59	LM	12.01	LM	7.51	LM	16.43	LM
Electrical	321,488	LM	247546	0.77	LM	18.5%	16.7%	0.65	LM	0.66	LM	0.52	LM	0.78	LM

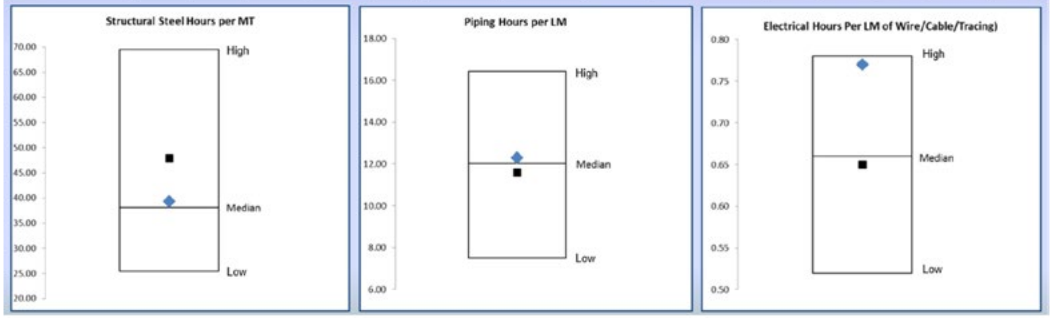

FIGURE A3 (at left) Example Tabular and Box-Plot Report [23] (Used With Permission)

TABLE A1 (bottom left) Example Tabular Report with Ratio-to-Driver Sequencing

METRICS	Units	Estimate	Low	Average	High	% Variance
Overall; LANG Factor						
DFC/Equipment (mech + E&I eqpt)	$/$					
Quantity /Quantity						
Earthwork/Concrete	m3/m3					
Concrete/Steel	m3/tonne					
Concrete/Equip	m3/each					
Steel/Equip	tonne/Each					
Piping/Equip	m/each					
Elec Wire&Cable/Equipment count	m/each					
Instr/Equip	each/each					
Direct Unit Costs/Quantity						
Earthwork	$/m3					
Concrete-CIP	$/m3					
Steel	$/tonne					
Buildings	$/m2					
Process Equipment	$/each					
Piping	$/m					
Electrical	$/m cable					
Instruments	$/each					
Bulk Materials/Quantity						
Concrete-CIP	$/m3					
Steel	$/tonne					
Piping	$/m					
Electrical	$/m cable					
Instruments	$/instr					
Direct Labor/Bulks Material						
Concrete-CIP	$/$					
Steel	$/$					
Piping	$/$					
Electrical	$/$					
Instruments	$/$					
Direct Hours/Quantity						
Earthwork	hr/m3					
Concrete-CIP	hr/m3					
Steel	hr/tonne					
Process Equipment	hr/each					
Piping	hr/m or t					
Electrical	hr/m cable					
Instruments	hr/each					
Construction Labor Rates						
Labor Rate (wage + benefits & burdens)	$/hr					
Engineering (hrs/quantity)						
Civil/concrete	hr/m3					
Structural/steel	hr/tonne					
Mech/equipment	hr/each					
Piping/piping	hr/m					
Electrical/electrical	hr/hr					
Control/Instruments	hr/each					
Indirect Metrics						
Eng/DFC	$/$					
Eng hours/direct field hours	hr/hr					
CM/DFC	$/$					
CM hours/direct field hours	hr/hr					
Eng $/Hr	$/hr					
CM $/Hr	$/hr					
Camp & Catering /(direct & indirect hours)	$/hr					
Total Indirects/DFC	$/$					
Owner and Risk						
Owner's/DFC	$/$					
Owner's/(DFC+Indirects)	$/$					
Contingency/(DFC+Indirect+Owner's)	$/$					
Escalation/(DFC+Indirects+Owner+Cont)	$/$					
DFC = Direct Field Costs, m=meters						

Figures A3 shows an estimate validation report (range method) at a discipline level with tabular and box-plot information and including both the median and the average.

Table A1 is an example report for a *Ratio-to-Driver* approach to the range method of estimate validation. This differs from the other Appendix A examples in the sequence of metrics starting with quantities and progressing to indirect costs. The exact metrics used would vary by the project types and code of account used.

ABOUT THE AUTHOR

John K. Hollmann, PE CCP CEP DRMP FAACE Hon.Life, is with Validation Estimating, LLC. He can be contacted by sending email to: **jhollmann@validest.com**