NPS-PM-06-089

EXCERPT FROM THE PROCEEDINGS

OF THE

THIRD ANNUAL ACQUISITION RESEARCH SYMPOSIUM

MEASURING UNCERTAINTY IN EARNED VALUE

Published: 30 April 2006

by

Richard Suter

3rd Annual Acquisition Research Symposium of the Naval Postgraduate School:

Acquisition Research: Creating Synergy for Informed Change

May 17-18, 2006

Approved for public release, distribution unlimited.

Prepared for: Naval Postgraduate School, Monterey, California 93943

The research presented at the symposium was supported by the Acquisition Chair of the Graduate School of Business & Public Policy at the Naval Postgraduate School.

To request Defense Acquisition Research or to become a research sponsor, please contact:

NPS Acquisition Research Program
Attn: James B. Greene, RADM, USN, (Ret)
Acquisition Chair
Graduate School of Business and Public Policy
Naval Postgraduate School
555 Dyer Road, Room 332
Monterey, CA 93943-5103
Tel: (831) 656-2092
Fax: (831) 656-2253
E-mail: jbgreene@nps.edu

Copies of the Acquisition Sponsored Research Reports may be printed from our website www.acquisitionresearch.org

Conference Website:
www.researchsymposium.org

Proceedings of the Annual Acquisition Research Program

The following article is taken as an excerpt from the proceedings of the annual Acquisition Research Program. This annual event showcases the research projects funded through the Acquisition Research Program at the Graduate School of Business and Public Policy at the Naval Postgraduate School. Featuring keynote speakers, plenary panels, multiple panel sessions, a student research poster show and social events, the Annual Acquisition Research Symposium offers a candid environment where high-ranking Department of Defense (DoD) officials, industry officials, accomplished faculty and military students are encouraged to collaborate on finding applicable solutions to the challenges facing acquisition policies and processes within the DoD today. By jointly and publicly questioning the norms of industry and academia, the resulting research benefits from myriad perspectives and collaborations which can identify better solutions and practices in acquisition, contract, financial, logistics and program management.

For further information regarding the Acquisition Research Program, electronic copies of additional research, or to learn more about becoming a sponsor, please visit our program website at:

www.acquistionresearch.org

For further information on or to register for the next Acquisition Research Symposium during the third week of May, please visit our conference website at:

www.researchsymposium.org

THIS PAGE INTENTIONALLY LEFT BLANK

Measuring Uncertainty in Earned Value

Presenter: Richard Suter
Author Affiliation: RS Consulting, 240-350-7586, rsx@ieee.org

Abstract

The Department of Defense is transforming its logistics and business systems to become agile, global-in-reach, and readily adaptable to evolving threats—all with significantly reduced Total Ownership Costs. However, the scope and complexity of these systems pose significant technical and programmatic challenges, successful management of which requires accurate engineering, planning, and cost estimation data. Because these programs and systems are information intensive, the costs of data acquisition are governed by the efficiency of communication, coordination and control activities. Likewise, they govern the capability of tools such as Earned Value Management (EVM). Unfortunately, much of the information essential to formulating accurate Planned Value (PV) estimates is not available until after a program is well underway. The key to information/data accuracy lies in the rate and extent to which uncertainty surrounding estimates is eliminated.

The confidence that can be placed in estimates, such as Planned Value, depends on a range of factors—all dominated by the maturity and discipline of Project Management, Quality Assurance, Enterprise Architecture, and Systems Engineering. Unfortunately, measures of their effectiveness have traditionally proven to be hard to implement, hard to interpret, and lack a clear relationship to the accuracy of Planned Value calculations.

However, several observations from Information Theory can be applied to these estimation problems. These include: (1) directly measuring the often unknown and usually unobservable "true" Planned Value parameters; (2) measuring the indirect costs for coordination and control—which represent the vast majority of activity costs for information-intensive organizations and programs—which could pave the way for more efficient and more comprehensive Activity-based Costing.

The strategy employed in this paper is to develop measurement models based on estimation techniques borrowed from Adaptive Control Theory (i.e., for closed-loop systems with unidentified components). The models predict the extent and rate of change (reduction) of uncertainty with respect to the confidence intervals bounding Planned Value calculations. By implication, the reduction (convergence) rate also indirectly measures the efficiency of information utilization of an organization—and, thus, System Effectiveness.

The measurement models outlined in this paper incorporate metrics from standard program management "Dashboards," (a few of which are provided in the Glossary) along with measures of response delay and uncertainty that can be implemented as a discrete event simulation whose outputs can be compared against project data repositories—such as NASA's SEL (Software Engineering Laboratory). The benefits of this approach include providing Decision Makers with: (1) "on-demand" capability for assessing both confidence levels for EV estimates, their underlying Planned Value calculations, and other project management parameters; (2) the rate of improvement in those confidence levels; (3) heuristic insight into the dynamics and consequences of decisions for their projects under a range of uncertain conditions.

The proposed measurement models are part of the shift to performance and model-based acquisitions in which cost, performance, and schedule trade-offs are quantitatively integrated to enhance the decision support available to program managers throughout a program's lifecycle.

Key Terms: Earned Value, Planned Value, Information Theory, Enterprise Architecture, Systems Engineering, Adaptive Estimation, Control Theory, Project Management, Information Productivity Paradox

Introduction

Earned Value Management (EVM) enables managers to anticipate problems and to take pre-emptive action. But, EVM implicitly assumes a level of accuracy for Planned Value (PV) that may not be justified, especially at the onset of a project, even for organizations with demonstrated capabilities. This is due to the inherent complexity and scope of the large-scale COTS Acquisition/IT modernization initiatives, rapidly evolving environments environment, and the continual evolution of technology.

However, organizations with strong Program Management and Systems Engineering capabilities can rapidly improve their estimates of project control variables such as scope, risk, schedule, and cost. These capabilities determine the rate at which the uncertainty can be removed from the information employed by an organization. The processes governing these rates and associated uncertainty levels can be modeled using traditional state variable methods and several results from Information Theory. The models generate (indirect) measures of the gap between estimated and "true" (and unobservable) parameter values that quantify the level of non-specificity (uncertainty) of the information resident in PV and related estimates. This provides a basis for determining whether and when enough information is available to satisfy specific confidence levels for estimates. The steady growth of best practices, as advocated by the CMMI, 6-Sigma, OPM3, and the availability of project management tools, indicate that the methods discussed in this paper can be applied at reasonable cost to provide previously unavailable decision support capabilities.

The approach outlined in this paper also scales up to large-scale, COTS-based IT modernization projects, which have minimal software development requirements, but nonetheless a large number of unknowns. For example, a "typical" SAP business system implementation will have thousands of critical parameters, each of which may be associated with a range of interdependencies that generate (unrecognized) ripple effects. Compounding these effects is a range of Information Assurance, Inter-operability, mission and agency-related requirements, undocumented complexities associated with yet-to-be phased-out legacy systems, all in addition to the competing demands of the program's stakeholders. The outcome is substantial integration, cost, performance, and schedule risk that results in the high level of uncertainty that drives the "Information Productivity Paradox."

The Information Productivity Paradox

The *Information Productivity Paradox* results from technology investments that do not improve productivity, because these investments do not contribute to technical and programmatic integration. That is, the new technology does not eliminate traditional organizational "stove-pipes"). Invariably, this is a consequence of immature organizational processes (e.g., as defined by the CMMI, OPM3, 6-Sigma) that result in poor planning and oversight. The low level of integration exhibited by bureaucratic organizations drives the low level of rate-of-change and the absence of timely feedback, which limits the capability to “learn,” thus constraining the integration needed to improve productivity. The consequences of this adverse feedback loop include limited capability to control the prime cost driver of i*nformation intensive* organizations—the effort consumed in the coordination and control of information (“10 Myths,” construx.com).

Absent that control, uncertainty levels will be high, thus precluding the “agility” needed to achieve the pre-requisites for accurate EVM. The attributes of that agility include:

- Commonality of data and information processes
- Efficiencies of scale
- Integration across functions
- Availability of real-time information
- Processing efficiency

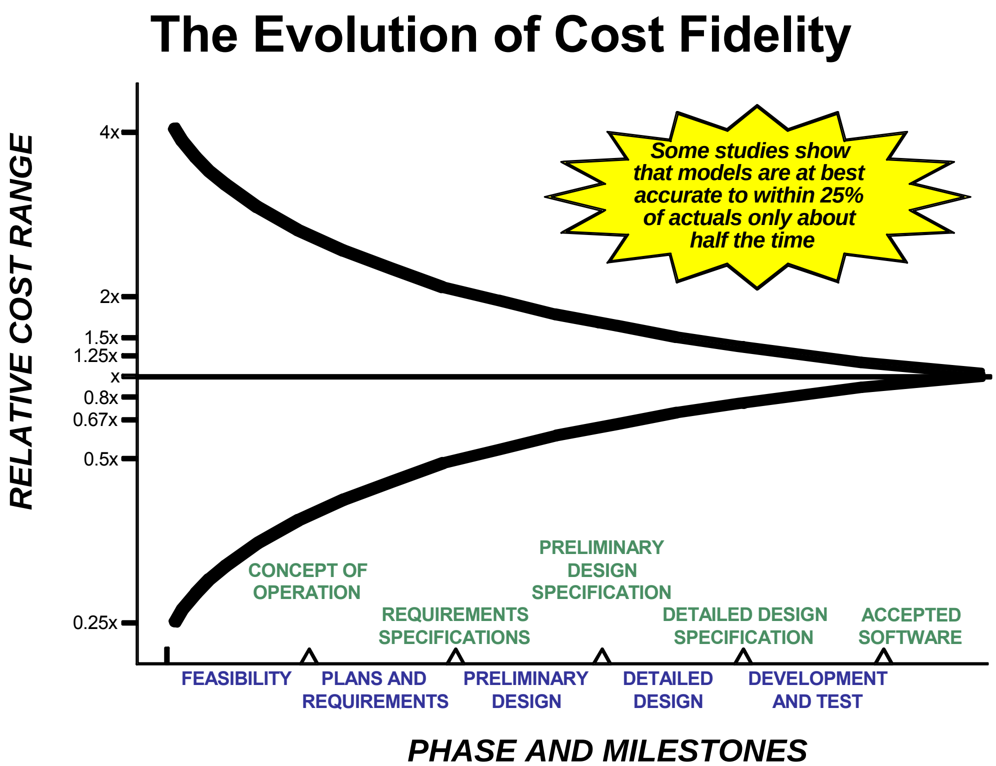
The relationships underlying these attributes can be expressed as a state variable system of organizational dynamics, using the methods pioneered by Jay Forrester (1999) that can establish the convergence rate of estimated and true Planned Values.

The strategy behind this approach is called “adaptive estimation” and has been applied to a wide range of processing and signal control applications in electrical utilities, manufacturing, and aerospace (Schweppe, 1973), which will be discussed after some basic EVM concepts are introduced.

How EVM Works—An Example

EV measures work—accomplished against a predefined schedule, thus enabling decision makers to systematically assess progress. As the elements of work are completed, their budgets are “earned,” thus quantifying the amount of work accomplished over time. This is Earned Value Management (EVM).

Figure 1. The Evolution of Cost Fidelity (McConnell, 1997, p. 31)



But, EVM assumes that Planned Value data is accurate, a condition that is rarely satisfied at the onset of a large scale IT infrastructure modernization, regardless of capability level. The variability of estimates is particularly pronounced at the onset of an IT modernization project, as illustrated in Figure 2, below.

EV is calculated as follows.

Schedule Variance (SV) is defined as:

[4.1] SV = EV – PV

Progress against project schedule can be measured by evaluating [4.1] over a sequence of points in time, noting at each such time point whether {SV < 0, or SV >= 0}. PV is the *a priori* estimate of the work to be accomplished, and EV is what we observe at the end of each reporting period. If EV < PV at the end of a reporting period, then SV < 0 for that period. This means that the project is slipping schedule, and value is not “earned” since work is *not* completed on schedule. But, if SV >=0 then work is completed on schedule; so, the dollar value of the budget is “earned.”

For example, if a widget worth $100.00 is to be delivered at the end of the month (this is the Planned Value), and the widget is completed by the end of the month, the Earned Value is $100.00, and the Schedule Variance is 0.

But, if the widget is only half completed at the end of the month, then the value earned is EV = \$50.00, resulting in SV = -\$50.00, which indicates that the project is slipping schedule, since only one-half the PV (Planned Value) was, in fact, delivered, or "earned."

Similarly, Cost Variance (CV) is defined as:

[4.2] CV = EV - AC

So, if the cost of producing one-half of a widget is \$200.00 – the Actual Cost (AC), then the CV, from Eqn [4.2], would be \$50.00 – \$200.00, or -\$150.00.

But, these calculations assume that PV accurately represents the "true" workload, which takes time to calculate accurately, even for highly capable organizations, as illustrated in Figure 1. Indeed, the NASA Software Engineering Laboratory (SEL) includes in its estimation process a 40% increase in the estimate of total workload made at project inception that will be needed to complete a project (Suter, 2005).

The variance equations [4.1], [4.2] can be treated as rate equations for closed loop State Variable systems, from which the amount of uncertainty associated with the convergence rates portrayed in Figure 1 can be estimated. To that end, we consider next the construction of a State Variable system.

Rate Equations, Organizational Dynamics and EVM

The complex mixture of organizational time-lagged response rates, transient and steady state conditions generate different time shapes due to delay modulation that varies as a consequence of differing levels of information availability. Rate (action) variables indicate how fast levels of funding, resources, quality, risk, rework, action items, products developed/integrated/delivered, are changing. They determine not present, but future value, as indicated by the rate change in level per unit of time.

The Cost and Schedule variance equations [4.1], [4.2] are rate equations defined by organizational policy, and can be derived using the following methodology:

- Define the goal (e.g., an objective function defining cost, schedule, quality and other to be optimized (i.e., maximized or minimized as appropriate)
- Observe the condition of the system (e.g., using methods such as periodic reviews of program progress, burn rates, requirement churn rates, quality, acceptance rates for tasks completed, etc.)
- Provide the means to express the discrepancy between goal and observed condition; e.g., between "true" Planned Value (PV) and the estimated Planned Value.
- Indicate what action is to occur, given the discrepancy observed

The rate equations [4.1] and [4.2] are instances of state variable systems which have the general form:

[5.2] $[dSV(t)/dt] = d\mathbf{x}(t)/dt = \mathbf{A}*\mathbf{x} + \mathbf{b}*\mathbf{u} + \mathbf{e_1}*\mathbf{v_1}$

[dCV(t)]/dt]

The composition of the state variable vector, **x,** is arbitrary and can consist of standard project variables such as schedule, cost, functionality, risk, quality, performance, each of which could be decomposed into ever finer levels of detail.

A, the coefficient matrix describes the processing efficiency of the organization in responding to events. It can represent system efficiency as a ratio of Input/Output coefficients as measured in terms of processing rates/hr, tasks completed/week, etc. (To maintain mathematical tractability, it is normally treated as a constant, but can be made function of time.)

u(t) is the control vector, and consists of adjustments to the state variables as defined by management-concerning factors such as resource and budget allocations, schedules. The effectiveness **u**(t) is constrained by situational awareness and organizational capability, and, hence, the quality and timeliness of the information available.

v(t) is vector of observational errors caused by factors such as incomplete, poor, or delayed information.

b, d are vector coefficients of the control variables **u**(t), **v**(t)

$\mathbf{e_1}$, $\mathbf{e_2}$ are vector coefficients of the control variables $\mathbf{v_1}$(t), $\mathbf{v_2}$(t)

"**t**" is time, and is the yardstick for measuring delay effects, task time, interrupt time, transient and steady state responses, etc. (It is implicit on the right-hand side of [5.2])

Model Based Measures of Uncertainty

Delay is inherent to organizations because information cannot be gathered, analyzed, or transmitted instantaneously. Thus, changes in the environment, slips in schedule may, or may not, be recognized when they occur. For example, decreases in data quality typically generate increased disruption in operation. As more resources are shifted to fixing and correcting data records, the rate at which information is processed decreases. The resulting inefficiency generates increased correction and rework rates, along with increased delays in task completion.

The net effect is a decrease in "situational awareness" that adversely impacts Planned Value calculations. The consequence is a "Nash Equilibrium" which is the point at which the cost of acquiring the information needed to identify a better solution exceeds the perceived benefit (In Game Theory, a Nash Equilibrium occurs when no player has any incentive to unilaterally change his action, since a change in strategy by any one of them would lead that player to "earn" less than if remaining with his current strategy). H. Simon termed this "satisficing" (March & Simon, 1967). The location of that equilibrium point can be inferred by measuring the uncertainty inherent in state variable estimates. Capable organizations (as defined by the CMMI, OPM3, 6-Sigma, etc.) will systematically shift the equilibrium point over the course of a project to one affording more accurate assessments of "true" PV (Suter, 2005).

That convergence is possible because organizational policy drives the level of organizational and technical integration that govern the timeliness and quality of information available to decision makers. Unlike resource flows, information flows are not conserved (i.e., the use information does not deplete it). But, the value of information decays overtime and does so more rapidly in environments characterized by high levels of Entropy. The impact is especially pronounced for organizations with limited capacity for information processing that reside in rapidly evolving environments.

The interrelationships among these variables constitute a closed-loop feedback system that impacts PV estimates in various ways such as:

- The rate of change (e.g., improvement in the accuracy of PV estimates), while influenced by many factors, can be considered as proportional to delays in decision-making. That is a surrogate indicator of system effectiveness, a key component of which is information processing capability.

- Fluctuations in the variability of (cost, schedule, quality, etc.) estimates are a consequence of (multiple) response lags arising from the interaction of factors such as: open-action items, unmanaged issues, delays in recognizing and adjusting to changes in requirements, scope, budget, market conditions.

(The glossary lists some measures of these factors that could be built from a standard collection of project management dashboard metrics.)

The “damp-out” rates for these fluctuations reflect different adjustment intervals that correspond to the level of organizational integration. Where the integration level is “low,” information transmission delays and distortion rates will result in sub-optimal policy decisions—the “Nash Equilibrium” effect. The consequences include an inability to control the continual stream of transient effects because of the greatly diminished timeliness and the value (quality) of the available information that precludes acquiring a true picture of the situation. Among the unfortunate results is a continual stream of “brush fires” that must be brought under control.

The following Figure illustrates a few of these (overly simplified) feedback dynamics.

Organizational Capability/Information Effectiveness-Feedback Loop

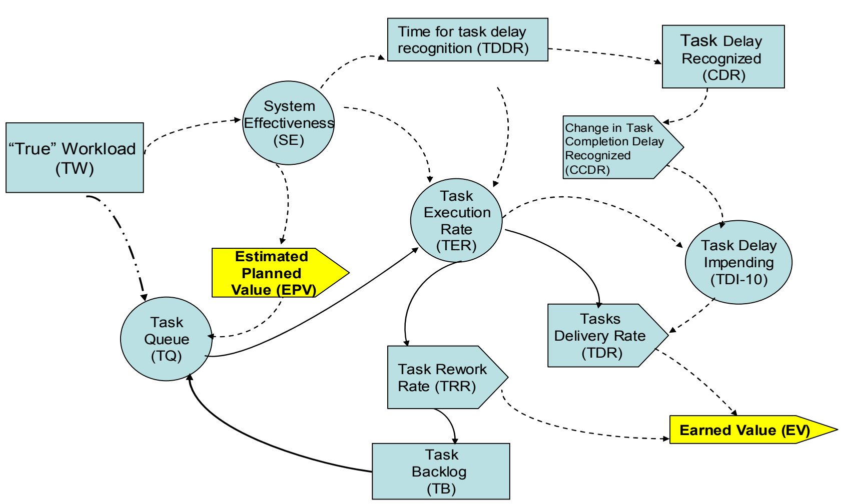


Figure 2. Organizational Capability/Information Effectiveness Feedback System

The key relationships in Figure 2 can be represented as a set of equations comprised of two basic entities:

(1) Levels - the amount of some quantity

(2) Rates - the measure of change in level per unit of time.

There are two types of rate:

(2.1) Controllable: denoted as rectangles pointing to the right (these are the decision variables available to Management)

(2.2) Not controllable: denoted as circles (which are functions of the controllable rates and their interactions with other rate parameters).

There are two types of delay impact rates:

(1) Task Execution (physical delay)

(2) Time to recognize changes (informational delay)

The impacts of the physical rates and levels on the flow of tasks are considered next, while those of information flows are considered in Section 8, below.

[6.1] $TB_{present} = TB_{previous} + \square *(TQ – TC)$

The current Task Backlog ($TB_{present}$) is the product of the reporting interval, □. and the backlog incurred during previous reporting which is defined as the difference between Tasks-in-Queue (TQ) and tasks completed (TC).

TC decomposes into Task Delivered (TD)—those accepted by the customer; and, TR, those not accepted which must be reworked. Thus, if TD > TQ, the present backlog is reduced; otherwise, it increases.

[6.2] TC = TD + TR

Tasks completed is the sum of Tasks Delivered (accepted by the customer) and those to be reworked (TR)

[6.3] TQ = EPV+ TR

Indirectly, TQ depends on EPV (which will vary inversely to accuracy of the resource and time requirement estimate) and the amount of Task Rework (TR)—due to defects, the failure to satisfy requirements, etc. “True Workload” (TW) is unknown because project scope typically is not well defined, requirements are not well understood and are subject to change. While TW is not directly observable, the gap between it and TQ is a function of the amount of (relevant) information available to decision makers—which is a function of overall System Effectiveness, a quantity that can be estimated, as explained in Section 8, below.

[6.4] ITD = TB/TDR

Impending Task Delays (ITD) can be expressed as the ratio of the Task Backlog (TB) (measured in units) to the Task Delivery Rate (TDR)—measured in units/month, which leaves ITD as an estimate of the time needed to complete backlogged tasks. This delay is based on physical capacity to handle the workload.

There is also a second type of delay based on Entropy/Uncertainty-driven time delays. The first of these is:

[6.5] RCTD = (1/TDDR)*(ITD – CDR)

Recognized Change in Task Delay (RCTD) is defined as proportional to the difference between ITD and the time required to recognize delay in task completion, labeled as “Completion Delay Recognized” (CDR). The product of this difference and the fraction of Time for Delivery Delay Recognition (TDDR) indicate how quickly an organization can adjust to the gap between ITD and CDR (i.e., to the difference between physically driven and informational delays)—and this is a function of the amount of new information becoming available to decision-makers in each reporting period.

[6.6] $CDR_{present} = CDR_{previous}$ + □ *RCTD

Completion Delay Recognized (CDR) equals the Completion Delay Recognized for the previous period plus the product of the reporting time period, □ and RCTD.

[6.7] $TER_{present} = TER_{previous}$ + □ * K_{SE} *SE

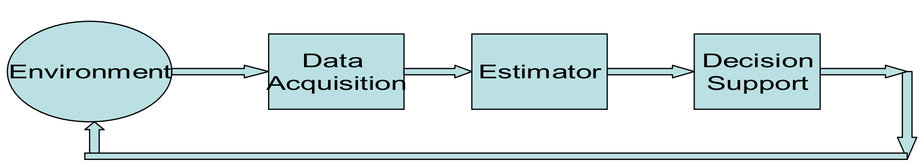
Current Task Execution Rate (TER) is defined as equal to TER for the previous time period plus an amount proportional to System Effectiveness.

The relationships of Figure 2, above, illustrate the role of information flows on system operations and on the capability to develop accurate PV estimates. Assessing that impact is the province of Information Theory.

Information Theory and Its Applications to EV Estimation

For many applications, measures on state variable system parameters are either distorted or are outright impossible. Consequently, the observation process itself must be modeled (in its simplest form the process is illustrated in Figure 3). The first step in modeling estimate accuracy is to distinguish between two general types of noise and their effects. First are those caused by imperfections in the measurement of the output variables; the second are those caused by excluding (simplifying) processes from state space models with the aim of simplifying them. The effects of both must be factored into the models.

Figure 3. The Estimation Environment—Signals, Measurement, Design



Both types of noise can be modeled using any (combination) of the four general estimation models found in the signal processing and statistical research literature. These are: the Fisher, Unknown-but-Bounded, Weighted Least Squares (WLS) and Bayesian. Of these, we shall only consider the first. By way of background, the WLS is limited to correlational analysis, and does not make any assumptions about underlying physical processes, which means that it is of limited value for the purposes of this paper.

While WLS imposes the fewest assumptions, Bayesian models impose the strongest assumptions; namely, all of the state variable parameters have known underlying probability distributions as do its [Bayesian's] corresponding error and observation models. Unknown-but-Bounded (U-b-B) methods assume that observations on **x** can be viewed either as means to find: (1) the center of some set, or (2) a point estimate; with the covariance matrix defining the size and shape of the set (often assumed in Signal Processing to be ellipsoid in shape). U-b-B models can be used to analyze systems such as [6.1] – [6.6], where both **x** and **v** are assumed to be unknown. The Fisher estimation model assumes no a priori knowledge of the vector of state variable vector, "**x**" (i.e., it assumes no underlying probability distribution, and is thus defined as "unknown"). Only the noise vector "**v**" is characterized as a random variable (i.e., it has an underlying probability distribution).

The questions of interest in this paper include:

- What do differences in response times indicate for the accuracy of PV estimates?
- What does the "time shape" (e.g., attributes such as lag, curvature, frequency, amplitude, variability) of a response indicate about the level of confidence that could be placed in estimate accuracy?
- When, and under what conditions, can the accuracy of PV estimates be considered acceptable?
- What effects do modeling errors have on the design and cost of decision support systems?
- How can measures of information uncertainty be used to establish confidence levels for various parameter estimates? (Klir, 2006—This text provides a comprehensive introduction to Information Theory.)

The first step in answering these questions is to develop a (static) linear estimation model of the observation process:

[7.1] **z** = **H*x** + **v**

Where:

z: Is the set of observations on **x** as filtered (e.g., "distorted") by **H** and **v**

H: Defines the coefficient matrix of structural relationships defining the observability of the (unknown) state variables, **x**, that impact the observations **z**. These relationships can be extracted from models such as those outlined in Figure 4, above.

v: In the Fisher model, it represents disturbances to observation of an uncertain nature, with:

[7.2] E {**v**} = 0

[7.3] E {**v*v'**} = **R**, which serves as a measure **v**

The objective is to find a "best" estimation model that minimizes the error:

[7.4] | $\mathbf{x}_{\text{true}}$ – **x^** |< □

Where:

□□: Is some arbitrarily small amount

$\mathbf{x}_{\text{true}}$: Is not known

x^: Is the estimated, and distorted, value of $\mathbf{x}_{\text{true,}}$ the "true" value of **x** is based on

$\mathbf{z}_{\text{actual}}$: The vector of recorded observation values

$\mathbf{v}_{\text{actual}}$: The actual value of uncertainty in the observation

Using these (redefined) variables, [7.1] becomes

[7.5] $\mathbf{z}_{\text{actual}} = \mathbf{H}^* \mathbf{x}_{\text{true}} + \mathbf{v}_{\text{actual}}$

$\mathbf{z}_{\text{actual}}$ and **H** are known, but $\mathbf{x}_{\text{true}}$ and $\mathbf{v}_{\text{actual}}$ are not known. So, **x^** is constrained to depend on the known terms and on the uncertainty models for the unknown terms. For example, one element of the state variable vector $\mathbf{x}_{\text{true}}$ is "True Workload" (TW), while EPV is an element of **x^**; $\mathbf{v}_{\text{actual}}$ consists of errors in recording observations (observed data values), and **H** is the structure of organizational relations that systematically filter/distort $\mathbf{z}_{\text{actual}}$

Using the known terms and candidate uncertainty models, the task is to:

(1) Develop a computational model that best minimizes the error (gap) in [7.4].

(2) Determine how close is to **x^** is to $\mathbf{x}_{\text{true}}$, which has the corollary problem of determining whether and how long it will take **x^** to converge to $\mathbf{x}_{\text{true}}$.

For Planned Value, [7.4] becomes

[7.4'] |TPV – EPV| < □

Which can be read as: the gap between "true" and estimated planned value is acceptably small.

For the Fisher model, the covariance matrix, □ can be pre-computed independently of **z** as follows:

[7.6] $□_{\text{Fisher}} = [\mathbf{H'}^*\mathbf{R}^{-1*}\mathbf{H}]^{-1}$

The error uncertainty, as measured in terms of the covariance matrix ($\square_{Fisher}$) can be used to determine whether and when the estimates, **x^**, will satisfy a pre-specified degree of accuracy, even though $\mathbf{x}_{true}$ itself is not observable (Harley, 1928—Harley developed a measure of uncertainty for finite sets, which Shannon adapted to Communications Theory.).

The estimate, $\mathbf{x\hat{}}_{Fisher}$ is:

[7.7] $\mathbf{x\hat{}}_{Fisher} = \square_{Fisher} * \mathbf{H'} * \mathbf{R}^{-1} * \mathbf{z}$

[7.8] □ (N|N) → □□□□**Q** > 0, **R** > 0

If [7.8] is satisfied, then the covariance matrix is unique, positive definite, and satisfies Controllability (**Q**>0) and Observability (**R**>0) conditions—discussion of which is beyond the scope of this paper, except to note that they determine when [7.6], [7.7] will provide satisfactory estimates of $\mathbf{x}_{true}$. Also, beyond the scope of this paper are the conditions under which ill-conditioned covariance matrices, biased, bounded, weighted, non-optimum estimators, and the conditions under which the “whiteness” of residuals can be used to define estimates that satisfy pre-defined confidence intervals.

Adaptive Estimation—Planned Value Estimation and Uncertainty

The organizational dynamics illustrated in Figure 2, above, constitute an incompletely specified closed loop state space system (i.e., one with unknown components). For these situations, estimates of state variables are updated as new information becomes available. This strategy is known as “adaptive estimation” and can be implemented using any of the standard estimation models, depending on the assumptions we make concerning the physical and information processes of interest.

Schematically, the estimation problem can be portrayed as:

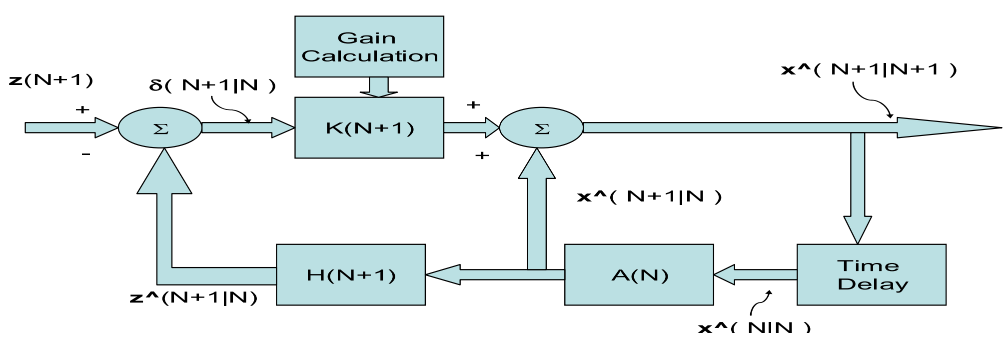


Figure 4. Estimation for Decision Support Systems

The first step in applying adaptive estimation to Planned Value estimates is to note that the state variable model, [5.2], can be represented in discrete time-case state variable model as:

[8.1] **x**(n+1) = **A**(n)**x**(n) + **G**(n)***w**(n)

Assuming **H, A, G** are known functions of time, [8.1] predicts the system state at time 'n+1', over a set of discrete points in time, n = 0, 1, 2…

In discrete time form, the estimation model, Eqn [7.5], becomes:

[8.2] $\mathbf{z}(n)_{actual} = \mathbf{H}(n)*\mathbf{x}(n)_{true} + \mathbf{v}(n)_{actual}$

$\mathbf{z}_{actual}$ and **H** are assumed to be known, but $\mathbf{x}_{true}$ and $\mathbf{v}_{actual}$ are not known

The expectation and covariance of **v** are:

[8.3] $E(\mathbf{v}) = 0$; $\mathbf{R} = [\mathbf{v}*\mathbf{v}']$

Applying [7.6], an estimator could be defined as:

[8.4] $\mathbf{W}_{Fisher} = \square_{Fisher} *\mathbf{H}'\ \mathbf{R}^{-1}$

Then

[8.5] $\mathbf{x}\hat{}\ (n)_{Fisher} = \mathbf{W}_{Fisher}*\mathbf{z}(n)$ n = 1, 2, …

The conditions that would make $\mathbf{W}_{Fisher}$ a "best" estimator are those of [7.8], above. And, they indicate when **x^** (e.g., EPV) is sufficiently close to "true" Planned Value ($\mathbf{x}_{true}$), as measured against a pre-defined confidence level.

The observations up to time n, **z**(1)…**z**(n) provide an estimate the state **x^**(n+1). The following table summarizes the key parameters of the state variable estimation problem, Eqn [8.1], [8.2].

Table 1. State Variable Parameters

Variable	Description
z(n)	Observations of **x**(n) filtered by **z** = **H*x** +**v.** Example, if z(1) is EPV, then the estimate of actual PV is x^(1) = W*z(1)
v(n)	Recording errors—observation uncertainty, which may be due to limited or incomplete data
w(n)	Uncertain inputs to organization processes—due to changes in project scope, environment
A(n)	Structural determinants of organizational dynamics
x^$(n_1\|n_2)$	Is the best estimate of x(n1) using observations z(1)… $z(n_2)$ (One element of this vector is Estimated PV)
$\mathbf{x}_{true}$	Actual system state (i.e., "true" PV, which accurately represents the "true workload)

G(n)	The (structured) relationships governing the handling of uncertain inputs
H(n)	Structural relationships governing observation (recording) accuracy
"n"	Discrete time points n = 0, 1, 2…k
x(0)	Vector of Initial conditions, which may be uncertain, such as a first estimate of Planned Value
R	Observability covariance error matrix with E{**v**}= 0 and with **R** = E{**v*v'**} for the model **z**= **H*x** + **v**
Q	**Q**= E{**w*w**'} is the Controllability covariance matrix of uncertain inputs to organizational processes, and E{**w**} =0

To determine the amount of new information that becomes available to a project, and moves a project closer to satisfying the conditions of [7.9], we define:

[8.6] **J(**N|N**)** = $\square^{-1}{}_{\text{Fisher}}$ (N|N**)**, N = 0, 1, 2…

J(N|N) is the inverse of the covariance matrix is the Fisher Information Matrix, and measures the ***amount*** of information contained in **x**^(N|N); that is, **z**(1)… **z**(N) about **x**(N). (A discussion of why this is so is beyond the scope of this paper but can be found in Klir (2006), and Schweppe (1973).

[8.7] **X**(N|N) = **J(**N|N**)*x^(** N|N**)**

Is the ***actual information*** in x^(N|N); i.e., is contained in z(1)…z(N)] about x(N),

where **x**(N |N) is read as the state vector **x** at time "N" given "N" observations (Schweppe, 1973, sec. 6.2).

Without going into detail, **J** and **X** can be used to measure how much information is *lost* due to the presence of uncertain inputs **w**(n) to the system. They also provide a means to construct an "Information Discount Rate (IDR)" against which the value and rate of investment in policies, tools aimed at reducing uncertainty, could be assessed. (IDR is a rate used to determine the present value of future information that can be constructed from estimates of the rate at which "uncertainty" is removed over a succession of estimates. This provides one mechanism to assess various projects estimates.)

System Effectiveness (SE) can be defined as proportional to the ratio of *amount* of change in *new* information acquired between the present time period (N+1) to that acquired in previous time period (N), as measured by [8.8], [8.9], both of which provide feedback to the organizational models of [6.1] **-** [6.7]. For example, TDDR of Eqn [6.4], above, is dependent on the amount of time required for sufficient information to be acquired for decision making, thus making it proportional to SE, where:

[8.8] SE = J(N+1)/J(N) – measures the percent gain in information between reporting periods N and (N+1). (Other measures are, of course, possible and may prove more useful)

[8.9] TDDR = K_3*SE (N+1) = K_4* J(N+1)/J(N)

Thus, various types of delay such as TDDR can be made explicit functions (changes in information available to decision makers) and can be used to explicitly model the information flows that govern organizational effectiveness as in Eqn [6.5].

Some of these effects, including their impact on the evolution of accuracy of Planned Value estimates, are illustrated in the following graph—using "synthetic" (and "smoothed-out") data.

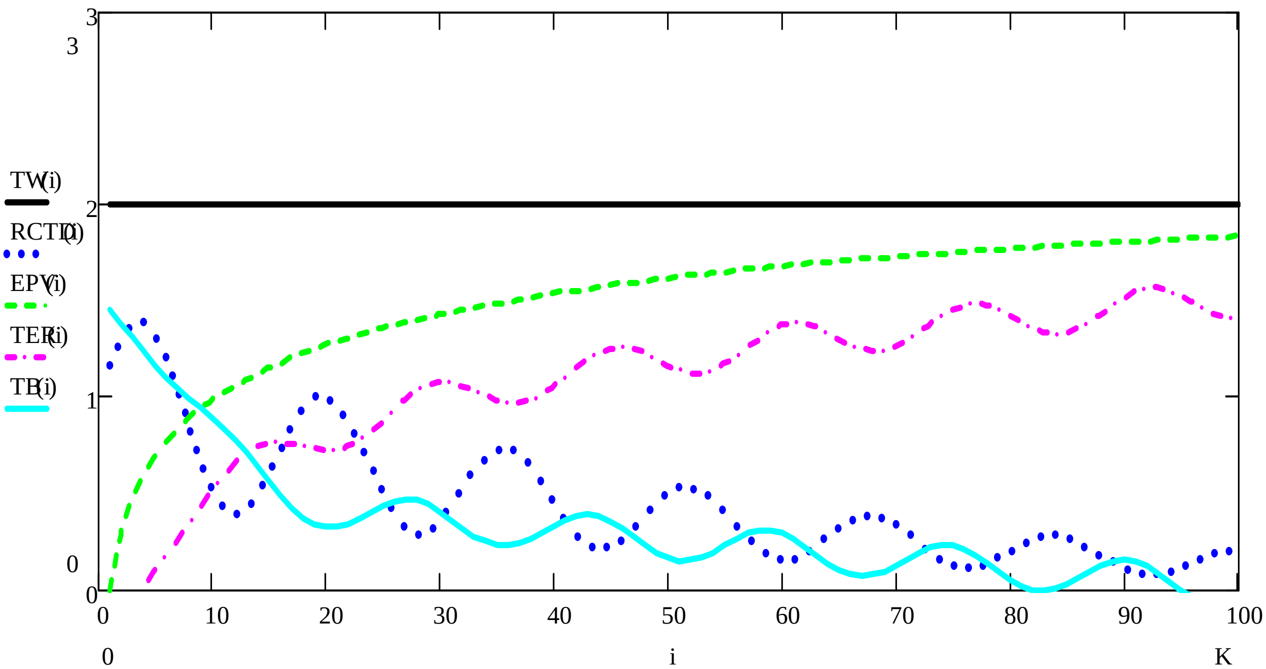


Figure 5. Information Uncertainty Impacts to Planned Value Estimates

The State Variables referenced in Figure 5 are:

TW = True Workload (assumed proportional to True Planned Value)

RCTD = Time to Recognize Delay in Task Completion

EPV = Estimated Planned Value

TER = Task Execution Rate

TB = Task Backlog

The Figure is a heuristic device to illustrate the fluctuations in the coupled feedback loop systems of Eqn [6.1] – [6.7], the interactions of which govern task flows and the associated delays (such as RCTD) in information flows and other system parameters such as task execution rates (TER), and Task Backlog (TB) levels. The interactions determine the rate at which uncertainty (non-specificity) is reduced and, thus, the degree of confidence that can be placed in EV.

Thus, in the Figure, EV steadily approaches “true” PV (with a slope that approximates the inverse of the uncertainty level (entropy), while RCTD and TER share a (coupled) time-lagged oscillation rate that declines over time as does the Task Backlog (TB) level.

Summary

EVM is a valuable tool for managing complex projects, but it rests upon assumptions that can be difficult to satisfy, especially at the onset of a project, and which may never be satisfied by projects with low management capability levels.

However, state variables methods, combined with results from Information Theory can be used to assess the accuracy of Planned Value estimates, the specificity of the underlying information, and, thus, the degree of confidence they merit. These are effects of information/system efficiency that can be inferred from measures such as the variability and time-lagged responses of rate parameters in response to perturbations and shifts in levels of uncertainty.

The next step is to complete and to refine the models, their associated measures, and then validate them against actual project data. Then, they can be implemented as software based tools for use with existing Project Planning tools.

The measurement models outlined in the paper provide the means to provide decision support in a cost-effective manner where they can be integrated with the automated data-acquisition tools; where improvements in organizational capabilities levels are present, the methods outlined in this paper can be implemented.

Glossary

Term	Definition
Activity-based Costing (ABC)	Is based on the assumption that products directly consume activities, not resources. Therefore, the cost of a product is the sum of all the costs of the activities performed to produce that product.
Actual Cost (AC)	The funds spent on work as of some specific date
Controllability	Is satisfied if an input to a system exists which takes the state of the system from any point to any other point in a specified time
Discount Rate	The interest rate used in determining the present value of future cash flows.
Cost Variance (CV)	CV = EV – AC
Information Discount Rate	The rate used to determine the present value of future information

(IDR)	flows that can be constructed from estimates of the rate at which “uncertainty” (i.e., the non-specificity) of the information is removed from estimates provided to decision makers.
Entropy	The uncertainty (non-specificity) resident in information. The formal theory and measurement of uncertainty was first developed by Harley (1928, p. 535-563). Shannon employed the measure to quantify uncertainty in communication systems, and the amount of information needed to reduce that uncertainty acceptable levels. It is a point-wise information metric that quantifies the association strength between 2 events by measuring (in probability terms) the amount of information that event 1 tells us about a second event.
Earned Value (EV)	The measure of work completed within a pre-determined time period. Thus, if the Planned Value of the work to be completed within a month is $100.00, if that amount of work is completed within that time period, then the budgeted amount for that work is “earned.”
Earned Value Management (EVM)	The set of methods, policies and procedures use to estimate EV
Observability	Is satisfied if it is possible to determine the state of a system from knowledge of the output, and input, without knowledge of initial conditions
Planned Value (PV)	The amount of work budgeted for completion within a specific period of time
SEL	Software Engineering Laboratory at NASA, Goddard
Schedule Variance (SV)	SV = EV – PV
Organizational Capability Measures include:	
BGCI: The Binary Group Index	Measures whether the vendor passes or fails a group of Binary Exit or Entry Criteria
DAI: Deliverable Acceptance Index	Measures the Quality performance standard for all acceptance-based deliverables. DAI = the number of times a developer submits the final version of a deliverable before approval by the customer
RA/RE: Results Achieved/Results Expected Index	Measures the percentage of expected results actually achieved for results-based deliverables. The type of results will differ by deliverable (e.g., training results for training deliverable, test results for a test deliverable) but the method to collect and assess the results (the RA/RE Index) will be consistent
SVI: Schedule Variance	SVI is the difference in the number of days between the expected

Index	and actual delivery date for a milestone or deliverable. It provides a schedule performance standard for all deliverables and milestones

References

10 myths of rapid development. Retrieved from http://www.construx.com

Forrester, J. (1999). *Industrial dynamics.* Pegasus.

Harley, R. (1928). Transmission of information. *The Bell System Technical Journal,* 7(3), 1928, 535-563,

Klir, G. (2006). *Uncertainty and information.* Wiley Inter-science.

March, J., & Simon, H. (1967). *Organizations.* Wiley.

McConnell, S. (1997). *Software project survival guide.* Microsoft Press, p.31.

Schweppe, F. (1973). *Uncertain dynamic systems.* Prentice-Hall, Sec. 6.2.

Suter, R. (2005). *A framework for calculating the indirect costs components of earned value for IT infrastructure modernization programs. Proceedings.* Second Annual Symposium on Acquisition Research. Naval Postgraduate School. Monterey, CA. May 2005. (NPS-AM-05-004, p. 261-275).

THIS PAGE INTENTIONALLY LEFT BLANK

2003 - 2006 Sponsored Acquisition Research Topics

Acquisition Management

- Software Requirements for OA
- Managing Services Supply Chain
- Acquiring Combat Capability via Public-Private Partnerships (PPPs)
- Knowledge Value Added (KVA) + Real Options (RO) Applied to Shipyard Planning Processes
- Portfolio Optimization via KVA + RO
- MOSA Contracting Implications
- Strategy for Defense Acquisition Research
- Spiral Development
- BCA: Contractor vs. Organic Growth

Contract Management

- USAF IT Commodity Council
- Contractors in 21st Century Combat Zone
- Joint Contingency Contracting
- Navy Contract Writing Guide
- Commodity Sourcing Strategies
- Past Performance in Source Selection
- USMC Contingency Contracting
- Transforming DoD Contract Closeout
- Model for Optimizing Contingency Contracting Planning and Execution

Financial Management

- PPPs and Government Financing
- Energy Saving Contracts/DoD Mobile Assets
- Capital Budgeting for DoD
- Financing DoD Budget via PPPs
- ROI of Information Warfare Systems
- Acquisitions via leasing: MPS case
- Special Termination Liability in MDAPs

Logistics Management

- R-TOC Aegis Microwave Power Tubes

- Privatization-NOSL/NAWCI
- Army LOG MOD
- PBL (4)
- Contractors Supporting Military Operations
- RFID (4)
- Strategic Sourcing
- ASDS Product Support Analysis
- Analysis of LAV Depot Maintenance
- Diffusion/Variability on Vendor Performance Evaluation
- Optimizing CIWS Life Cycle Support (LCS)

Program Management

- Building Collaborative Capacity
- Knowledge, Responsibilities and Decision Rights in MDAPs
- KVA Applied to Aegis and SSDS
- Business Process Reengineering (BPR) for LCS Mission Module Acquisition
- Terminating Your Own Program
- Collaborative IT Tools Leveraging Competence

A complete listing and electronic copies of published research within the Acquisition Research Program are available on our website: www.acquisitionresearch.org

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CALIFORNIA 93943

www.acquisitionresearch.org