Facilitating Predictive Cost Analytics via Modelling V&V

John Swaren, PRICE Systems

April 10, 2015

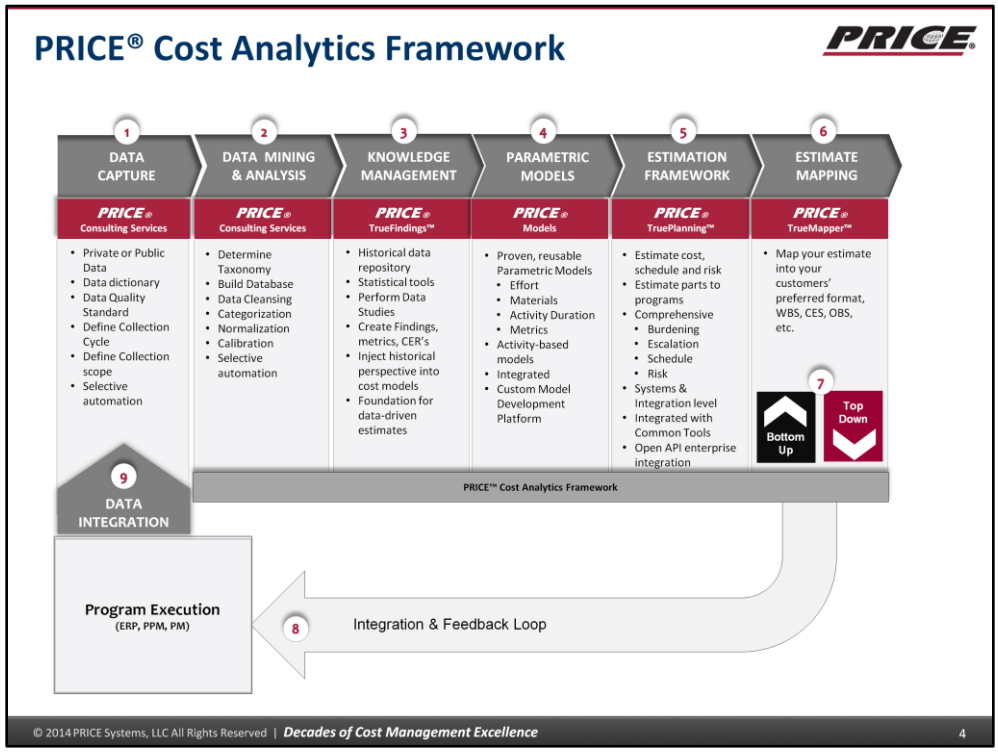
Why Verify/ Validate?

- Analogies often are not based on rigor and do not represent organizational experience/productivity, nor do they capture the complexity of project deliveries driven by requirements
- TruePlanning allows for calibration of the latter. TrueFindings allows for aggregation, filtering and analysis of these calibrations
- How can this knowledgebase of calibrated metrics best be utilized?
- **The answer is aligning knowledge of corresponding requirements that drove the costs that informed the calibrations**
- In this way, we predict an appropriate set of model input-drivers based on the organization and its project/program histories!
- **Utilize Predictive Analytics by applying "calibrated" knowledge to estimate programs/products based on relevant cost drivers**

Overview

- The PRICE® Cost Analytics Framework
  - Provides parametric models based on calibration of industry-wide data
  - Also provides for integration of a feedback loop from program execution through data capture/mining/analysis to customer-specific knowledge management
  - Customers typically wish to leverage their knowledge to fine-tune PRICE® models to their (or their contractors') specific histories for specific programs/products
- Key PRICE® Tools
  - TruePlanning™: Calibration of key parameters and Hosting of parametric models
  - TrueFindings ™: Knowledge capture and Development of custom relationships
  - Combining both -> Customers can leverage PRICE®'s knowledgebase as well as derive their own "tuned" estimating models, specific to their histories of process, experience, material, technology, producibility/productivity, complexity, etc.
- **PRICE® Cost Analytics is the scientific process of transforming data into insight for making better decisions**

 Decades of Cost Management Excellence



By using techniques such as mathematical modeling to analyze complex circumstances, predictive analytics gives executives the power to make more effective decisions and build more productive systems based on:

- More complete data
- More comprehensive analysis of this data
- Consideration of all available options
- Careful predictions of cost outcomes and estimates of risk

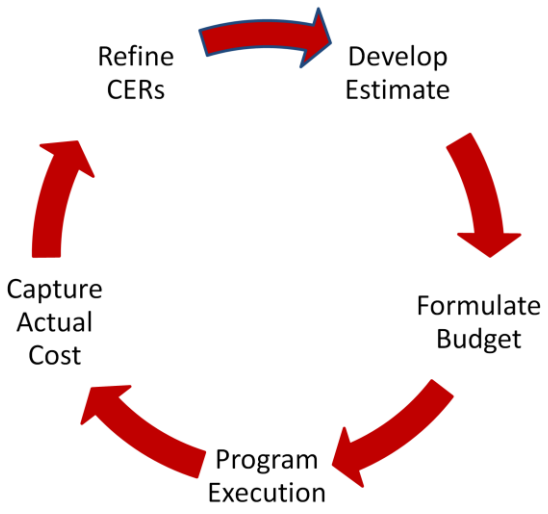
Why Calibrate?

- Typical parametric models use inputs with default values that represent industry standards, averages, and practices
- Calibration is the process of tuning these factors to your specific organization, based on data from previous projects
- The calibration tool can be used….
  - To apply information captured from similar projects to your new estimates (refining the estimating accuracy)
  - To identify the adjustments that are necessary to achieve project goals (e.g., target costs)
  - To benchmark an organization's actual experience
    - *With the industry standard*
    - *Over time (trend analysis)*
    - *To the competition*

Parametric model inputs are typically preset to default, industry average values. It's a great starting point, but we should fine tune our estimate to reflect a particular organization'. Calibration helps us do that fine tuning as an ESSENTIAL V&V tool.

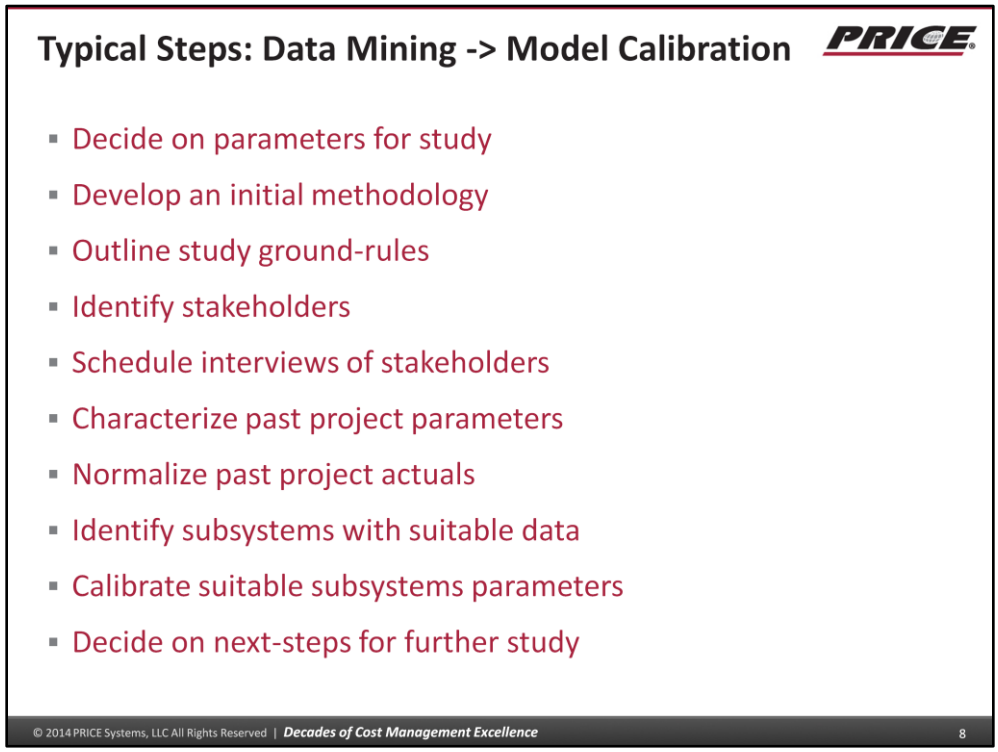
Calibration refines our estimate accuracy, we leverage existing information on similar projects.
Calibration also benchmarks the organization's actual experience, for comparing to the industry average, performing trend analysis, or "ghosting" the competition.

Cost Estimating is Cyclical Process



Types of Calibration

- Product
  - Calibrates technology or productivity level
- Organization
  - Adjusts TruePlanning® output from "industry average" to a particular firm
  - Adjusts TruePlanning® output to the accounting system of a particular firm
- Schedule
  - Use actual, historic schedule
- Cross Object
  - Adjusts TruePlanning® for a single complexity value based on the actual cost/effort for several lower Cost Objects
- External via EXCEL Companion Solution
  - Use TruePlanning® EXCEL Solution to quickly calibrate large data sets



Upcoming Webinar: **9 July: "Leveraging Historical Data in Affordability Analysis"**

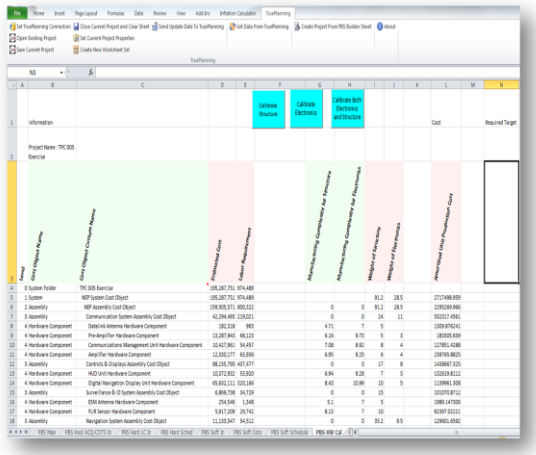
How to Use Calibration Results

- Create library of “like” historical projects
- Perform calibration on key cost drivers for each project
- Perform cross object calibration on project library
- Utilize calibrated parameters on future projects when more accurate information is not yet available
- Utilize range of calibrated values in library as risk inputs

 Decades of Cost Management Excellence

Calibrating with Excel Solver Companion Application

- Alternative to quickly calibrate multiple inputs in Excel
- Ability to send calibrated values to TruePlanning® and update estimate
- Ability to choose cost drivers for hardware or software
  - Manufacturing Complexities
  - Organizational Productivity
  - Functional Complexity
  - Code Size
  - Development Team Complexity



 Decades of Cost Management Excellence

TrueFindings Introduction

- Example Software Development data in Excel
  - Calibrated Complexity and Productivity Drivers
  - SW Code Descriptors and Hours/Costs
  - Key Performance Parameters

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Name	Space Sensor/Detector Software Component (Text)	Functional Complexity (Number)	Organizational Productivity (Number)	New Code Size - SLOC (Number)	Language (Text)	Normalized Dev. Time - Hours (Number)	Normalized Dev. Cost - 2013 $USD (Number)	KPP#1 Operational Effectiveness (Number)	KPP#2 Differentiation Effectiveness (Number)	KPP#3 Operational Availability (Number)	KPP#4 Spectral Resolution (Number)	KPP#5 Electromagnetic Intensity (Number)
2	Payload 1	Spectrometer SW	5.94	1.1	4,459	JavaScript	1,681	$ 77,871	95	96	97	95	94
3	Payload 1	Electro-Static Analyzer SW	6.13	1.08	9,875	C++	8,172	$ 254,287	95	96	95	94	95
4	Payload 1	Gamma Sensor SW	7.34	0.99	14,129	Java	12,048	$ 357,427	96	97	96	95	96
5	Payload 1	Neutron Sensor SW	7.54	0.94	14,750	Java	12,257	$ 367,696	97	97	98	98	99
6	Payload 1	Radiometer SW	8.32	0.89	153,824	Ada95	22,642	$ 746,347	98	98	99	99	99
7	Payload 2	Spectrometer SW	6.43	1.03	6,237	JavaScript	3,252	$ 133,449	95	96	95	94	95
8	Payload 2	Electro-Static Analyzer SW	5.82	1.19	8,032	C++	4,139	$ 185,747	95	95	96	96	97
9	Payload 2	Gamma Sensor SW	7.01	1.02	12,056	Java	10,214	$ 320,078	96	96	95	94	95
29	Payload 7	Gamma Sensor SW	6.12	1.08	7,495	JavaScript	3,967	$ 145,581	95	96	97	97	96
30	Payload 7	Neutron Sensor SW	6.22	1.05	12,050	C++	9,413	$ 306,413	95	96	97	95	94
31	Payload 8	Radiometer SW	5.84	1.18	6,132	Python	3,071	$ 121,318	95	95	94	93	94
32	Payload 8	Spectrometer SW	6.17	1.06	9,860	C#	7,679	$ 254,287	95	96	95	98	93
33	Payload 9	Electro-Static Analyzer SW	6.18	1.06	8,176	JavaScript	4,327	$ 185,747	95	96	97	97	97
34	Payload 9	Gamma Sensor SW	6.28	1.05	13,146	C++	11,943	$ 353,178	95	96	95	98	96
35	Payload 10	Neutron Sensor SW	5.79	1.24	5,451	Python	2,437	$ 104,469	95	95	94	97	93
36	Payload 10	Radiometer SW	6.14	1.08	8,764	C#	4,668	$ 189,816	95	96	97	95	95

- Perform Predictive Analytics to determine method for fine-tuning software cost drivers as a function of KPPs

Here's an example for the following notional Software Development data in Excel.

In blue are the project names and SW components delivered.
In green are the two key TruePlanning for Software ("True-S") inputs, derived by calibration of the gray actuals.
In yellow are component Key Performance Parameter scores.

{Note: these KPPs could have been KSAs or any other measure of requirements for past/future deliveries.}

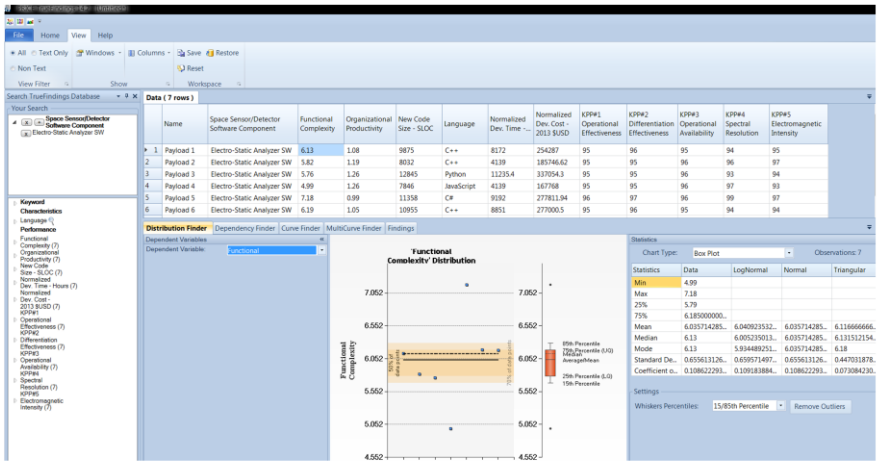
Example Scenario

- Space Sensor Software
  - Estimator wishes to fully leverage in-house knowledge
  - Initially uses TruePlanning Space Catalog values for Sensor Hardware
  - The TruePlanning Software Catalog can estimate sufficiently using PRICE's knowledgebase and estimating models
  - HOWEVER, the estimator also has relevant program data histories to leverage
  - The challenge (typical at NASA Phase-A) is the estimator ONLY knows mission requirements in the form of Key Performance Parameters
  - These KPPs represent mission-specific percentage (0%-100%) ratings for:
    - *Operational Effectiveness, Differentiation Effectiveness, Operational Availability, Spectral Resolution, Electromagnetic Intensity*

➢ First perform Calibration to fine-tune parametric models. Take advantage of relevant calibrated comparables (i.e., knowledge) to produce defendable estimates tied directly to parametric modelling drivers, in this example early-stage KPP requirements

TrueFindings Visualization

- Example Software Development data: Distribution Finder
  - Spreadsheet rows become knowledgebase fields for search/filtering
  - 1st tab-function shows descriptive statistics for all or (below) selected data



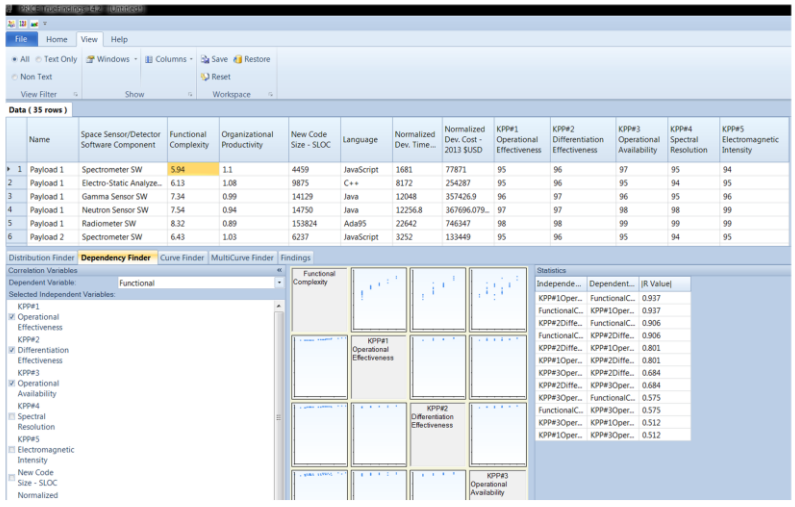
Again, the goal is to perform Predictive Analytics to determine a method for best justifying cost driver-inputs as a function of KPPs {or KSAs}.

Certainly, we can estimate sufficiently using PRICE's knowledgebase and True-S estimating models.

HOWEVER, we also have the above relevant program histories to leverage. The challenge (typical at early stage, e.g., NASA Phase-A) is the estimator ONLY knows mission requirements without yet an adequate description of product/functional requirements and detailed architecture.

TrueFindings Visualization

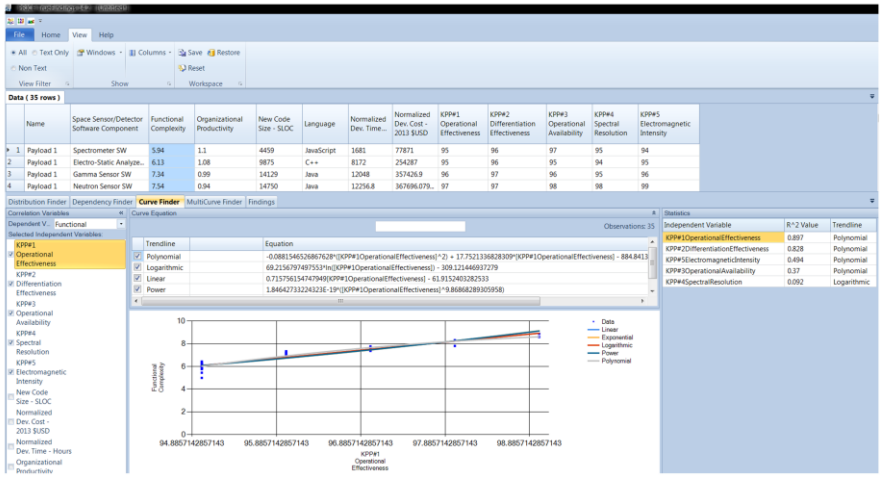
- Example Software Development data: Dependency Finder
  - 2nd tab-function shows bivariate correlations coefficients



Our job is to fine-tune the use of True-S models by best applying knowledge from KPPs to predict key drivers, in this case above Functional Complexity.

TrueFindings Visualization

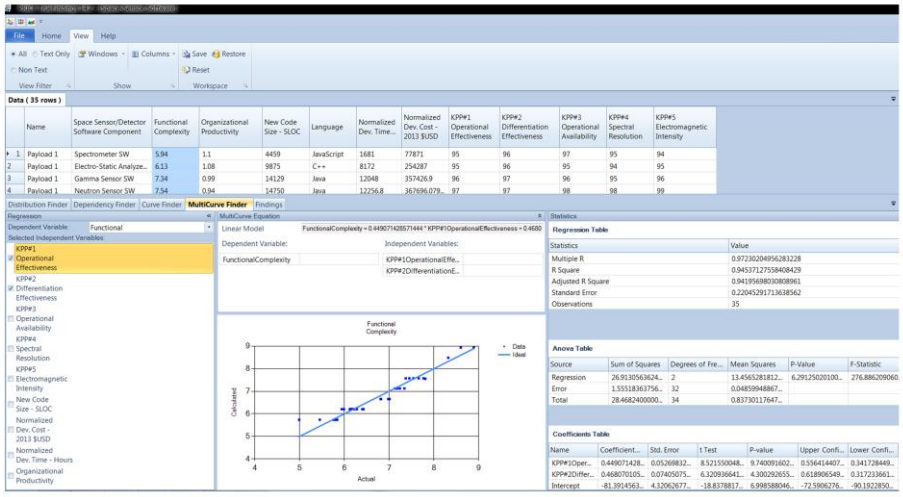
- Example Software Development data: Curve Finder
  - 3rd tab-function shows simple (one predictor) regression fits w/ R^2s



Assuming we know the future program's scores (and again, the only quantified understanding of new requirements), the solution is to perform regression analysis of the calibrated values of functional complexity versus their *associated* KPP scores as candidate predictors. Certainly, a simple bivariate regression is feasible.

TrueFindings Visualization

- Example Software Development data: Multicurve Finder
  - 4th tab-function shows multiple (two or more predictors) regression fits w/ R^2s



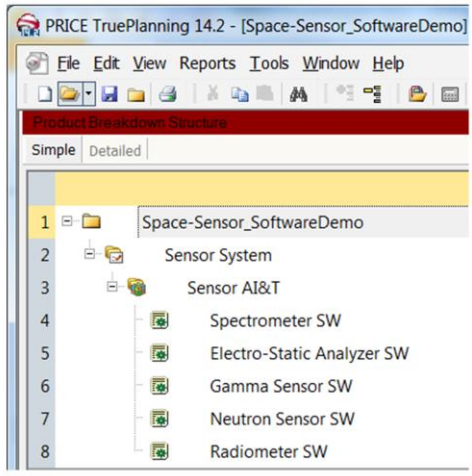
But why stop at linear regression? Use the new TrueFindings application to produce a most appropriate CER using multiple regression.
The result will look like this equation, with associated residual plot and statistical-significance metrics—

Functional Complexity = 0.449 * KPP#1 + 0.468 * KPP#2 - 81.39

We now have a linear model that predicts Functional Complexity based on KPPs. Specifically, it suggests that this key True-S input driver is best predicted by KPP#1 and KPP#2. In fact, any additional KPP predictors are not statistically-significant, per inspection of the ANOVA p-values.

TruePlanning Model PBS

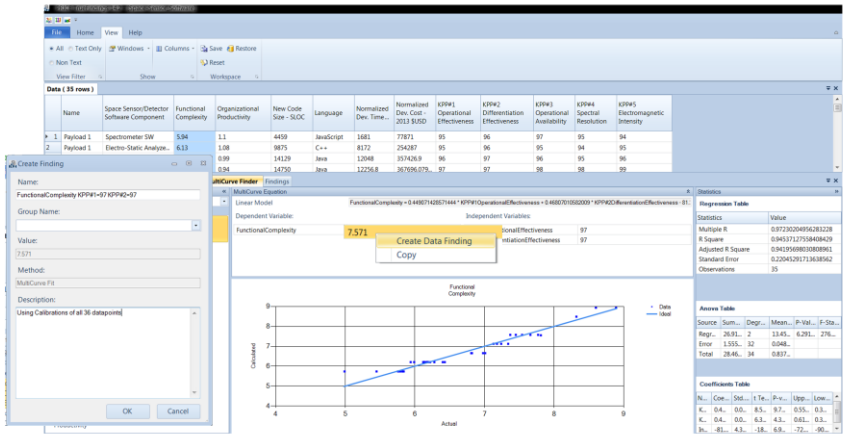
- Example using Input Drivers predicted via TrueFindings CER



TrueFindings -> TruePlanning

PRICE

- Find Functional Complexity using custom equation via TrueFindings



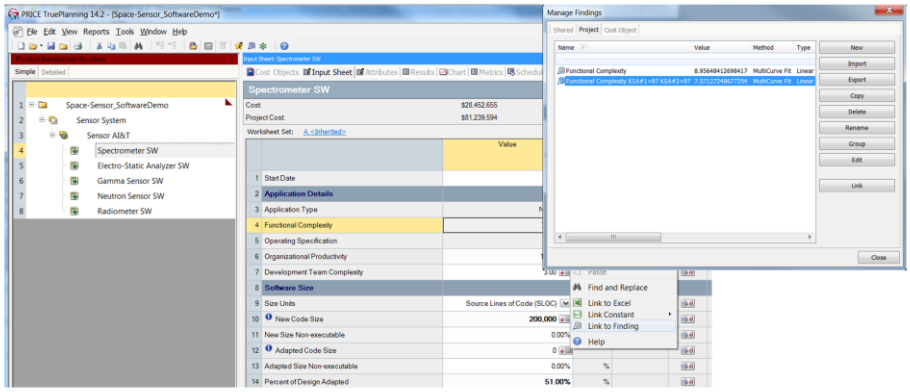
- CER produces Complexity of 7.571 for KPP#1=97 and KPP#2=97 based on best-fit regression analysis thru **calibrated** data

Appreciate that we can indeed take advantage of relevant calibrated comparables (i.e., knowledge) to produce defendable estimates tied directly to parametric modelling drivers that are based (only!) on early-stage requirements.

Today, the latter were expressed as KPP (or KSA) scores. For you, the pool of candidate predictors can be any metric. Essentially, you can not only develop a calibration knowledgebase; you can now also leverage the latter to best fine-tune TruePlanning reflecting your requirement metrics.

TrueFindings -> TruePlanning

- Apply Functional Complexity from saved Finding

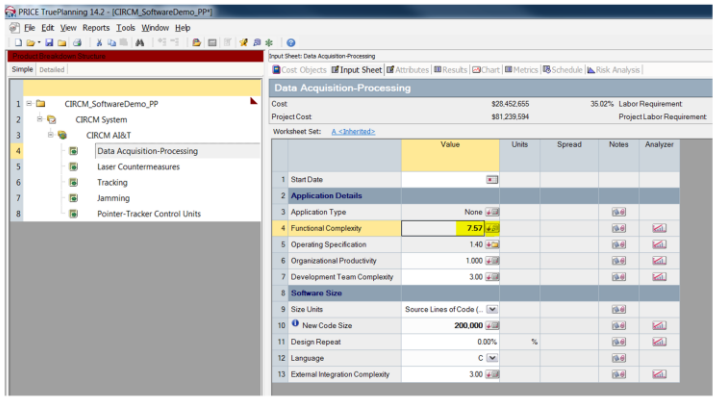


- CER produces Complexity of 7.571 for KSA#1=97 and KSA#2=97 based on best-fit regression analysis thru **calibrated** data

What's important to realize is the predicted values from an equation "Finding" are now immediately available within TruePlanning. By deriving equations (hopefully based on calibrated data), you are essentially creating your own custom-calculator, as a pre-processor to our knowledgebase algorithms!

TrueFindings -> TruePlanning

- Updated Project based on Input Driver linked to Findings



- Improved validity / defensibility of estimate, with modeling input directly linked to Analysis based on relevant calibrated data

This repeatable process allows us to perform Predictive Analytics to develop a method (i.e., a data-driven CER) that fine tunes TruePlanning for Software input driver(s) as a function of requirements, in this example KPPs. {Of course, the same process would apply for analyzing Hardware inputs.}

In either case, you can leverage a calibrated knowledgebase based on your data to best fine-tune TruePlanning reflective of your past and future requirement metrics.

Calibrating past knowledge and leveraging those calibrations is the essence of model V&V. Questions?