TECOLOTE RESEARCH, INC.

Bridging Engineering and Economics

Since 1973

The Statistical Approach for USCM9 CER Development

24 - 27 June 2008

Dr. Shu-Ping Hu, Michael Pfeifer, Nick Lozzi

▪ Los Angeles ▪ Washington, D.C. ▪ Boston ▪ Chantilly ▪ Huntsville ▪ Dayton ▪ Santa Barbara
▪ Albuquerque ▪ Colorado Springs ▪ Columbus ▪ Ft. Meade ▪ Ft. Monmouth ▪ Montgomery ▪ Ogden ▪ Patuxent River ▪ Pensacola ▪ San Diego
▪ Charleston ▪ Cleveland ▪ Denver ▪ New Orleans ▪ Oklahoma City ▪ Silver Spring ▪ Warner Robins AFB ▪ Vandenberg AFB

TECOLOTE RESEARCH, INC.

Objectives

- **Finalize all statistical criteria for USCM9 publication**
  - Regression Technique
  - Functional Forms
    - Factor, Linear, Power, and Triad (same as before)
  - Statistical Measures for Evaluation
    - Two-tiered evaluations: one for fit measures and one for predictive measures
- **Discuss USCM9-related topics**

TECOLOTE RESEARCH, INC.

Outline

- **Objectives**
- **Regression Technique**
  - Multiplicative Error Models: Log-Error, **MUPE**, ZPB/MPE (**ZMPE**)
  - Properties of MUPE Method
- **USCM CER Selection Criteria**
  - Fit Measures vs. Predictive Measures
  - R^2 Issues
- **Prediction Intervals (for use in uncertainty analysis)**
- **Cost Improvement Curve (CIC) Analysis**
  - T1 vs. QAIV
  - Disjoint vs. Sequential
- **Conclusions**

TECOLOTE RESEARCH, INC.

Outline

- **Objectives**
- ***Regression Technique***
  - *Multiplicative Error Models: Log-Error,* ***MUPE****, ZPB/MPE (****ZMPE****)*
  - Properties of MUPE Method
- **USCM CER Selection Criteria**
  - Fit Measures vs. Predictive Measures
  - R^2 Issues
- **Prediction Intervals (for use in uncertainty analysis)**
- **Cost Improvement Curve (CIC) Analysis**
  - T1 vs. QAIV
  - Disjoint vs. Sequential
- **Conclusions**

TECOLOTE RESEARCH, INC.

Multiplicative Error Models – Log-Error, MUPE, & ZMPE

Definition of error term for $Y = f(x)*\varepsilon$

- **Log-Error: $\varepsilon \sim LN(0, \sigma^2)$ ⇨ Least squares in log space**
  - Error = Log (Y) - Log f(X)
  - Minimize the sum of squared errors; process done in log space
- **MUPE: $E(\varepsilon) = 1$, $V(\varepsilon) = \sigma^2$ ⇨ Least squares in weighted space**
  - Error = (Y-f(X))/f(X)
  - Minimize the sum of squared (percentage) errors iteratively

Note: $E((Y-f(X))/f(X)) = 0$

$V((Y-f(X))/f(X)) = \sigma^2$

- **ZMPE: $E(\varepsilon) = 1$, $V(\varepsilon) = \sigma^2$ ⇨ Least squares in weighted space**
  - Error = (Y-f(X))/f(X)
  - Minimize the sum of squared (percentage) errors with a constraint

$\Sigma_i (\text{Error}_i) = 0$

- **(MPE: Same as ZMPE but with no constraint)**

TECOLOTE RESEARCH, INC.

Regression Technique

- **MUPE regression method (to model multiplicative errors) is selected for USCM9 CER development**
  - MUPE is an Iteratively Reweighted Least Squares (IRLS) regression technique
  - MUPE is consistent with previous USCM publications such as USCM7 and USCM8
  - Multiplicative error assumption is appropriate when
    - Errors in the dependent variable are believed to be proportional to the value of the variable
    - Dependent random variable ranges over more than one order of magnitude

TECOLOTE RESEARCH, INC.

Regression Technique – Properties of MUPE Method

- **The MUPE method relies on the nonlinear regression technique to derive a solution, except for linear CERs**
- **The MUPE CER has zero proportional error for all points in the database (no sample bias) through the minimization process**
- **The MUPE CER does not need correction factor adjustments**
- **MUPE is a well-established technique and has been in use for decades. We can use the asymptotic goodness-of-fit measures to evaluate whether the fitted coefficients are significant.**
- **The MUPE method produces consistent estimates of the parameters and the mean function; it is the best linear unbiased estimate (BLUE) for linear models (see Reference 13)**
- **The estimated parameters using the MUPE method are also the maximum quasi-likelihood estimates (MQLE) of the parameters**
- **CO$TAT provides PIs for MUPE CERs for uncertainty analysis**
  - PIs for nonlinear CERs are already given in SAS, Statistica, CO$TAT, etc. and CO$TAT's results are consistent with the PI estimates from these commercial tools

TECOLOTE RESEARCH, INC.

Regression Technique – Summary Table for MUPE

- **MUPE regression method meets all USCM9 modeling requirements**

Features/Functions/Requirements	MUPE
Require Transformation?	No
Require CF Adjustment (for Mean)?	No
Consistent Estimates?	Yes
BLUE for Linear Model?	Yes
MQLE for Parameters?	Yes
Provide CI for Regression Coefficient?	Yes
Provide PI for Future Observation?	Yes
Require Nonlinear Regression?	Yes
Trapped in Local Minima?	Rarely
Nonlinear Equation Stable?	Mostly Yes

TECOLOTE RESEARCH, INC.

Cautions on Selecting a "Best" CER from Multiple Methods

- **<u>One</u> regression method will be used for USCM9 CER development to avoid any possible confusion**
- **A fitting method is selected based upon (1) model/error assumption and (2) certain specific criteria; statistical and/or <u>engineering rationale</u> should be provided for changing the fitting criteria when evaluating CERs**
  - For example, if one method produces a CER weight exponent greater than one while another method generates an exponent less than one, analysts (with prior info) may choose the latter instead of the former
- **How to select a "best" CER from multiple methods is not well-defined**
  - Based upon SPE, approximated t-stat, GRSQ, Adj.R^2, or a combo?
  - Different functional <u>forms</u> (with several candidate <u>drivers</u>) should be evaluated. (If MUPE generates a logical exponent for driver A but not for driver B while ZMPE does the opposite, which one should be chosen?)
  - It is subjective, tedious, and may cause confusion

TECOLOTE RESEARCH, INC.

Outline

- **Objectives**
- **Regression Technique**
  - Multiplicative Error Models: Log-Error, **MUPE**, ZPB/MPE (**ZMPE**)
  - Properties of MUPE Method
- ***USCM CER Selection Criteria***
  - Fit Measures vs. Predictive Measures
  - R^2 Issues
- **Prediction Intervals (for use in uncertainty analysis)**
- **Cost Improvement Curve (CIC) Analysis**
  - T1 vs. QAIV
  - Disjoint vs. Sequential
- **Conclusions**

TECOLOTE RESEARCH, INC.

Selection Criteria

- **The fit measures should be analyzed separately from the predictive measures when evaluating USCM9 CERs**
- **Fit measures are used for model/coefficient evaluation**
  - Use SPE (or SEE) to measure the model's overall error of estimation
  - Use the asymptotic confidence interval (CI) or approximated t-stat for each coefficient to determine whether the coefficient is significant
- **Predictive measures are for supplemental evaluations**
  - Adjusted R^2 for MUPE and Pearson's r^2 (i.e., GRSQ)
- **Based on users' feedback (too few stats listed in USCM8), we will report four statistical measures for each USCM9 CER (see below)**

See color code on page 23 for detailed selection criteria

Statistics	**For Review**	**For Report**
SPE (Multiplicative Error)	Yes	Yes
Approx T-Stats	Yes	Yes
Adjusted R^2 for MUPE	Yes	Yes
Pearson's r^2 (GRSQ)	Yes	Yes
MAD of % Errors	Yes	No
RMS of % Errors	Yes	No

The combined evaluations of the fit and predictive measures (not just one measure) will help us select a "best" CER

TECOLOTE RESEARCH, INC.

Outline

- **Objectives**
- **Regression Technique**
  - Multiplicative Error Models: Log-Error, **MUPE**, ZPB/MPE (**ZMPE**)
  - Properties of MUPE Method
- ***USCM CER Selection Criteria***
  - *Fit Measures vs. Predictive Measures*
  - R^2 Issues
- **Prediction Intervals (for use in uncertainty analysis)**
- **Cost Improvement Curve (CIC) Analysis**
  - T1 vs. QAIV
  - Disjoint vs. Sequential
- **Conclusions**

TECOLOTE RESEARCH, INC.

Definition – SPE

- **Standard Percent Error (SPE) or Multiplicative Error:**

$$\text{SPE} = \text{SEE} = \sqrt{\frac{1}{n-p}\sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i}{\hat{y}_i}\right)^2}$$

(n = sample size and p = total number of estimated coefficients)

Note: SPE is CER's standard error of estimate (SEE)

- **SPE is used to measure the model's overall error of estimation; it is the one-sigma spread of the MUPE CER**
- **SPE is based upon the objective function; the smaller the value of SPE, the tighter the fit becomes**

TECOLOTE RESEARCH, INC.

Interpretation of SPE

$$SPE = \sqrt{\sum_{i=1}^{n} ((Y_i - \hat{Y}_i)/\hat{Y}_i)^2 / (n-p)}$$

$Y = f(x)^*\varepsilon$, $\varepsilon \sim$ Distrn(1, σ^2)

- **SPE^2 is an estimate of σ^2**
- **SPE cannot be used to determine the significance of the regressed coefficients**
  - For example, if $f(X) = aX^b$, SPE is not used to test whether the regressed coefficient b is significant
  - Compare two regression models where the exponent of one equation is 10 times larger than the other:

    $Y = a_1 X_1^{0.5} \varepsilon_1$ vs. $Y = a_2 X_2^{0.05} \varepsilon_2$ ($\varepsilon_1 \sim$ Distrn(1,σ_1^2), $\varepsilon_2 \sim$ Distrn(1,σ_2^2))

    Assume also that σ_2 in the second equation is much tighter than σ_1 (i.e., $SPE_1 >> SPE_2$). However, the exponent coefficient in the first CER should be more significant than the second one.
- **Be wary of using SPE alone for selecting CERs**

TECOLOTE RESEARCH, INC.

Significance of Regression Coefficients (1/2)

- **Any of the three measures below can be used to determine whether the regression coefficient is significant**
  - Asymptotic confidence interval (CI) for the estimated coefficient
  - Approximated standard error of the coefficient
  - Approximated t-stat for each coefficient
- **Note: The asymptotic CI and standard error of each coefficient are listed in the nonlinear regression output for most statistical packages such as SAS, Systat, Statistica, CO$TAT, etc.**
- **We decided to report the approximated t-stats (instead of the CIs) in the USCM9 publication for readability**

TECOLOTE RESEARCH, INC.

Significance of Regression Coefficients (2/2)

- Asymptotic confidence interval (CI) or approximated t-stat for the estimated coefficient is derived under the normality assumption for the error term
- CIs, as well as approx t-stats, can be used as guidelines to determine whether the fitted coefficients are significant
- From our initial study, no significant changes were found in the approximated t-stats when using the log-normal distribution assumption

Equation	Method	SPE	T-Stat 1	T-Stat 2	T-Stat 3
$Y = 488.5 + 0.086 * X$	MUPE	41.4%	2.52	4.88	
$Y = 437.8 + 0.080 * X$	Log Error	46.3%	2.20	4.37	
$Y = 482.3 + 0.088 * X$	PING Factor	41.4%			
$T1 = 441.5*Wt^{0.4914}*1.130^{GEO\_Orbit}$	MUPE	17.6%	2.31	5.43	10.83
$T1 = 455.3*Wt^{0.4807}*1.147^{GEO\_Orbit}$	Log Error	18%	2.28	5.10	10.86
$T1 = 461.1*Wt^{0.4807}*1.147^{GEO\_Orbit}$	PING Factor	17.7%			
$T1 = 19.90 * Wt^{0.7472} * Mech^{0.50}$	MUPE	23.8%	2.49	12.13	3.18
$T1 = 19.55 * Wt^{0.7511} * Mech^{0.48}$	Log Error	23.9%	2.43	11.88	2.95
$T1 = 20.01 * Wt^{0.7511} * Mech^{0.48}$	PING Factor				

TECOLOTE RESEARCH, INC.

Definitions – Pearson's r and GRSQ

- **Pearson's correlation coefficient between two sets of numbers $\{x_i\}$ and $\{y_i\}$:**

$$r_{xy} = \frac{\sum_{i=1}^{n} (y_i - \overline{y})(x_i - \overline{x})}{\sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2} \sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2}}$$

- **Generalized R^2 (GRSQ): GRSQ is Pearson's r^2 between the actual $\{y_i\}$ and predicted $\{\hat{y}_i\}$ in unit space, i.e., GRSQ = $r^2(y, \hat{y})$**

$$r^2(y, \hat{y}) = \frac{\left(\sum_{i=1}^{n} (y_i - \overline{y})\left(\hat{y}_i - \overline{\hat{y}}\right)\right)^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2 \sum_{i=1}^{n} \left(\hat{y}_i - \overline{\hat{y}}\right)^2}$$

- **GRSQ was first listed in CO$TAT's predictive measures circa 1991**

TECOLOTE RESEARCH, INC.

Definition – Adj. R^2

- **Adjusted R^2 in unit space:**

$$Adj.\,R^2 = 1 - \frac{SSE/(n-p)}{SST/(n-1)} = 1 - \frac{\sum (y_i - \hat{y}_i)^2/(n-p)}{\sum (y_i - \bar{y})^2/(n-1)}$$

(n = sample size and p = total number of estimated coefficients)

Note: Adj. R^2 can be evaluated in both the fit and unit spaces

- ***Adjusted R^2 in unit space*** **translates SSE from the absolute scale to the relative scale by**
  - Comparing SSE to SST
  - Adjusting degrees of freedom for small samples
- **Note: We compare SSE to SST because $\bar{y}$ is the unbiased estimate for the univariate models (y = a + ε or y = a*ε)**

TECOLOTE RESEARCH, INC.

Interpretation of Adj. R^2

- **What is Adjusted R^2 in unit space?**

$$Adj.\,R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2/(n-p)}{\sum (y_i - \bar{y})^2/(n-1)} = \frac{MSE_{\bar{Y}} - MSE_f}{MSE_{\bar{Y}}}$$

for additive models

- **This statistic is well-defined and applicable. It measures the percent difference between the CER's estimated variance and the sample variance of Y**
  - For example, if a CER's estimated variance is 0.1 while the sample variance of y is 0.5, then the CER's variance is only 20% of the sample variance. This reduction of variance, 80%, is the Adjusted R^2.
  - The reduction of variance is considered to be an "improvement" when applying the CER
- **We can use *Adjusted R^2 in unit space* to compare a CER's performance to the starting point, i. e., MSE of an average CER (when the driver variables are not available)**

TECOLOTE RESEARCH, INC.

Modify Adj. R^2 for MUPE

- **Adjusted R^2 in unit space:**

$$Adj.\ R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2 / (n-p)}{\sum (y_i - \bar{y})^2 / (n-1)} = \frac{MSE_{\bar{Y}} - MSE_f}{MSE_{\bar{Y}}}$$

- **Adjusted R^2 for MUPE:**

$$Adj.\ R^2 (MUPE) = 1 - \frac{\sum ((y_i - \hat{y}_i)/\hat{y}_i)^2 / (n-p)}{\sum ((y_i - \bar{y})/\bar{y})^2 / (n-1)} = \frac{SPE_{\bar{Y}}^{\ 2} - SPE_f^{\ 2}}{SPE_{\bar{Y}}^{\ 2}}$$

- **This modified adjusted R^2 compares MUPE's SPE^2 to its baseline (i. e., SPE^2 of an average CER); it is more pertinent to the fitting methodology**
- ***Adjusted R^2 for MUPE* puts SPE in perspective**

TECOLOTE RESEARCH, INC.

Outline

- **Objectives**
- **Regression Technique**
  - Multiplicative Error Models: Log-Error, **MUPE**, ZPB/MPE (**ZMPE**)
  - Properties of MUPE Method
- ***USCM CER Selection Criteria***
  - Fit Measures vs. Predictive Measures
  - *R^2 Issues*
- **Prediction Intervals (for use in uncertainty analysis)**
- **Cost Improvement Curve (CIC) Analysis**
  - T1 vs. QAIV
  - Disjoint vs. Sequential
- **Conclusions**

TECOLOTE RESEARCH, INC.

Concerns About R^2 & Adj R^2

- ***R*²**, **as well as *Adjusted R*², has no value as a metric in cases other than OLS**
- **The formulas of *R*² and *Adjusted R*² are inapplicable**
- **Many good CERs may be dismissed when using *Adjusted R*² because they might have a negative *Adjusted R*²**
  - As noted in the USCM8 document, a negative Adj. R^2 is a warning flag
  - This warning flag has probably led to the rejection of a number of **good** USCM8 CERs

TECOLOTE RESEARCH, INC.

No Worries About Adj R^2

- **No <u>single</u> measure is relied on to select the best CER**
  - <u>Not</u> possible to reject "logical" CERs just because of negative Adj. R^2
- **Several fit as well as predictive measures will be examined for USCM9; they were also examined during USCM7 & USCM8**
  - SPE, Approx T-stats, Pearson's r, Adjusted R^2, MAD of % Errors, etc.

	Red	Yellow	Green
SPE (Multiplicative Error)	> 0.5	0.25 ~ 0.5	< 0.25
T-Stats	< 1.5	1.5 ~ 2.3	> 2.3
Pearson's r	< 0.6	0.6 ~ 0.8	> 0.8
Adj. R^2	< 0.45	0.45 ~ 0.65	> 0.65
MAD of % Errors	> 50%	25% ~ 50%	< 25%
RMS OF % Errors	> 60%	30% ~ 60%	< 30%

  - The first two evaluate the significance of the coefficients and model; the remainder are predictive measures
- **Neither R^2 nor Adj. R^2 was used to indicate the proportion of the variation explained by the MUPE CER**
- **Counter examples (good SPE with negative Adj. R^2) can be found**

TECOLOTE RESEARCH, INC.

Measures of Fit: T1 vs. Wt & GEO orbit dummy variable

MULTIPLICATIVE		N= 17		TT&C SUBSYSTEM RECURRING COST CER						
RUN	DEPENDENT VARIABLE	SE (MULT)	PEARSON'S CORR COEF	MAD of % ERROR	RMS OF % ERRORS	ADJ R^2	X1 COEF T SCORE***	X2 COEF T SCORE***	X3 COEF T SCORE***	CONST T SCORE***
FACTOR	T1K	0.2497	0.7130	18.0780	23.4464	-0.1158	13.8208	2.4570	0.0000	0.0000
	EQUATION T1K = 41.794 * WEIGHTLB + 1064.627 * ORBIT									
	REMARKS									
LINEAR	T1K	0.1832	0.7110	11.3870	16.6326	0.4201	5.1526	1.1167	0.0000	3.5334
	EQUATION T1K = 1737.572 + 24.408 * WEIGHTLB + 434.634 * ORBIT									
	REMARKS									
CURVE	T1K	0.1757	0.7480	10.6460	15.9338	0.4946	2.2980	5.2796	10.8654	0.0000
1	EQUATION **T1K = 441.546 * WEIGHTLB ^ 0.491 * 1.13 ^ ORBIT**									
	REMARKS									
TRIAD	T1K	0.1825	0.7460	10.6370	15.9661	0.4528	0.3270	1.1012	6.8263	0.0904
	EQUATION T1K = 292.978 + 339.597 * WEIGHTLB ^ 0.533 * 1.14 ^ ORBIT									
	REMARKS									

1. A significant intercept is omitted from the factor equation, which cannot be detected by examining the CER's SE (i.e., SPE) or Pearson's r^2
2. Neither SPE nor Pearson's r^2 can detect model flaws; **Adj R^2 may provide a clue**

TECOLOTE RESEARCH, INC.

Outline

- **Objectives**
- **Regression Technique**
  - Multiplicative Error Models: Log-Error, **MUPE**, ZPB/MPE (**ZMPE**)
  - Properties of MUPE Method
- **USCM CER Selection Criteria**
  - Fit Measures vs. Predictive Measures
  - R^2 Issues
- ***Prediction Intervals (for use in uncertainty analysis)***
- **Cost Improvement Curve (CIC) Analysis**
  - T1 vs. QAIV
  - Disjoint vs. Sequential
- **Conclusions**

TECOLOTE RESEARCH, INC.

Prediction Interval for Uncertainty Analysis - Analytic

- **For nonlinear regression, the confidence intervals (CI) and prediction intervals (PI) are part of the standard output for SAS, Statistica, CO$TAT, etc.**
  - CO$TAT can generate PIs for all MUPE CERs
- **Note: PIs are derived by the Taylor series expansion using first order derivatives**
  - $V(\hat{y}|_{x=x_o}) \cong z_o(Z'WZ)^{-1}z_o'\sigma^2$ for weighted least squares

    $Z = \partial f/\partial\theta$ (evaluated at $\hat{\theta}$),

    W = weighting matrix (W = **I** if not weighted)
  - $V(y - \hat{y}\ |_{x=x_o}) \cong (f(x_o)^2 + z_o(Z'WZ)^{-1}z_o')\sigma^2$ (for MUPE CERs)
- **If the data set is available, prediction intervals can be generated. We will evaluate the feasibility and impact of publishing PIs for the USCM9 CERs.**

TECOLOTE RESEARCH, INC.

Outline

- **Objectives**
- **Regression Technique**
  - Multiplicative Error Models: Log-Error, **MUPE**, ZPB/MPE (**ZMPE**)
  - Properties of MUPE Method
- **USCM CER Selection Criteria**
  - Fit Measures vs. Predictive Measures
  - R^2 Issues
- **Prediction Intervals (for use in uncertainty analysis)**
- ***Cost Improvement Curve (CIC) Analysis***
  - T1 vs. QAIV
  - Disjoint vs. Sequential
- **Conclusions**

TECOLOTE RESEARCH, INC.

CIC Analysis: T1 vs. QAIV (1/2)

T1 CER vs. QAIV CER (using CAC theory):

- T1: $T1 = a_1 *(X)^{c1} *(Z)^{d1}$
  - T1 is derived using a fixed quantity slope, e.g., 95% for USCM
- QAIV: $LAC = a_2 *(X)^{c2} *(Z)^{d2} *(LastUnit^{(b+1)} - (FirstUnit - 1)^{(b+1)})/LotQty$
- QAIV: $LAC = a_2 *(X)^{c2} *(Z)^{d2} *(LotQty)^{b}$ if FirstUnit = 1
  - Note that the exponent b is determined by the regression equation using the appropriate data set

The respective lot total costs are given by

- T1: $LTC = a_1 *(X)^{c1} *(Z)^{d1} *(LastUnit^{(-0.074+1)}-(FirstUnit - 1)^{(-0.074+1)})$
- QAIV: $LTC = a_2 *(X)^{c2} *(Z)^{d2} *(LastUnit^{(b+1)}-(FirstUnit - 1)^{(b+1)})$

Note: ln(0.95)/ln(2) = -0.074

TECOLOTE RESEARCH, INC.

CIC Analysis: T1 vs. QAIV (2/2)

Benefits of using QAIV:

- The first unit cost (T1) is a hypothetical cost, not a real dependent variable
- It is an objective way to let the data set determine the quantity slope rather than using a subjective assumption
- The potential of "understating" the error of the estimate with the T1 CERs (due to an incorrect quantity slope) can be avoided

Downsides of using QAIV:

- Not enough lot data, insignificant quantity slope, very significant learning slope but > 100%, etc.

QAIV for USCM9?

- We are still evaluating whether we should use the QAIV approach vs. the traditional T1 CERs
- We welcome any inputs regarding either of the methods

TECOLOTE RESEARCH, INC.

CIC Analysis: Disjoint vs. Sequential

- **Disjoint theory assumes <u>no</u> cost improvement to carry over from one contract to the next**
  - Every lot starts with <u>unit one</u>; there are no meaningful follow-on programs (sequential lots) to be considered
  - For the DSCS program, although units 8-14 are essentially identical to units 4-7, we would assume no learning to carry over to the 8-14 block under the disjoint theory. Therefore, unit 8 would be the first unit, unit 9 would be the second unit, etc.
  - Learning curves will become rate curves under the disjoint theory, treating the lot quantity variable as a surrogate for rate
- **Sequential theory assumes cost improvement carryover to the follow-on programs**
  - For the DSCS program, unit 8 would be unit 5, unit 9 would be unit 6, etc., when using the sequential theory
- **We should determine whether cost improvement occurs to the follow-on programs for USCM9**

TECOLOTE RESEARCH, INC.

Cost Improvement Carryover Excursion (from Ref 3)

- **Here is a QAIV analysis summary report (from Reference 3) using the USCM8 database, including an excursion assuming no cost improvement carryover between contracts. Note that the average slope is fairly close to 95%.**
- **The USCM8 CERs were based upon a database of 44 military, NASA, and commercial satellite programs; USCM9 includes 30+ new programs**

Subsystem or Suite	Implied Slope w carryover	Subsystems Only (w carryover)	**No Carryover Between Lots**
Electrical Power Subsystem	103%	103%	104%
Power Generation Suite (Primary CER)	97%		99%
Power Generation Suite (Secondary CER)	98%		99%
Power Storage Suite	93%		88%
Power Conditioning and Distribution Suite	112%		114%
Attitude Determination and Control Subsystem	100%	100%	94%
Attitude Determination Suite	91%		87%
Reaction Control System - Mechanical	101%		95%
Reaction Control System - Propellant	103%		102%
Structure/Thermal Subsystem (Primary CER)	91%	91%	89%
Structure/Thermal Subsystem (Secondary CER)	91%	91%	90%
TT&C Subsystem - Non GEO Satellites	89%	89%	89%
TT&C Subsystem - GEO Satellites	95%	95%	93%
Communications - Unique Programs	94%	94%	90%
Communications - Standard Buses	100%	100%	100%
Communications Antenna Suite	97%		98%
Minimum	89%	89%	87%
Maximum	112%	103%	114%
Average	97%	95%	96%
Standard deviation	6%	5%	7%

TECOLOTE RESEARCH, INC.

Conclusions

- **We will use the MUPE method to develop USCM9 CERs**
- **We plan to use the statistics below for selecting the USCM9 CERs**
  - Use relevant fit measures when evaluating CERs
    - Use SPE to measure the model's overall error of estimation
    - Use the asymptotic confidence interval or approximated t-stat for each coefficient to determine whether the coefficient is significant
  - Use predictive measures for supplemental evaluations
    - Adjusted R^2 for MUPE and Pearson's r^2 (i.e., GRSQ)
  - The combined evaluations between the fit and predictive measures will help us select a "best" equation
- **We will also evaluate reporting PIs for uncertainty analysis**
- **For cost improvement, we will explore the database first to determine whether we should (1) use disjoint or sequential theory to normalize data and (2) develop QAIV or T1 CERs**
  - Analysts can assess different approaches using total recurring costs

TECOLOTE RESEARCH, INC.

References

1. Hu, S. and Smith, A., “Why ZMPE When You Can MUPE,” 6th Joint Annual ISPA/SCEA International Conference, New Orleans, LA, 12-15 June 2007
2. Hu, S., “Prediction Interval Analysis for Nonlinear Equations,” 2006 Annual SCEA International Conference, Tysons Corner, VA, 13-16 June 2006
3. Hu, S., Fong, F., and Enser, B., “Cost Improvement Curve Analysis of the USCM8 Database Using Quantity As an Independent Variable (QAIV),” 2006 Annual SCEA International Conference, Tysons Corner, VA, 13-16 June 2006
4. Book, S. A. and Young, P. H., “The Trouble with R2,” International Society of Parametric Analysts, Journal of Parametrics, Vol. XXV, No 1 (Summer 2006), pages 87-112
5. Book, S. A., “Prediction Bounds for General-Error-Regression CERs,” 39th DoDCAS, VA, 14-17 Feb 2006
6. Hu, S., “The Impact of Using Log-Error CERs Outside the Data Range and PING Factor,” 5th Joint Annual ISPA/SCEA Conference, Broomfield, CO, 14-17 June 2005
7. Hu, S., “The Minimum-Unbiased-Percentage-Error (MUPE) Method in CER Development,” 3rd Joint Annual ISPA/SCEA International Conference, Vienna, VA, 12-15 June 2001
8. Goldberg, M.S. and Tuow, A., “Statistical Considerations in Estimating Learning Curves and Multiplicative CERs,” 33rd Annual DoDCAS, Williamsburg, VA, 1-4 February 2000
9. Book, S. A. and Lao, N. Y., “Minimum-Percentage-Error Regression under Zero-Bias Constraints,” Proceedings of the 4th Annual U.S. Army Conference on Applied Statistics, 21-23 Oct 1998, U.S. Army Research Laboratory, Report No. ARL-SR-84, November 1999, pages 47-56
10. Tjur, Tue, “Nonlinear Regression, Quasi Likelihood, and Overdispersion in Generalized Linear Models,” The American Statistician, Vol. 52 (1998), pages 222-227
11. Nguyen, P., et al., “Unmanned Spacecraft Cost Model, Seventh Edition,” U.S. Air Force Space and Missile Systems Center (SMC/FMC), Los Angeles AFB, CA, August 1994
12. Young, P. H., “Generalized Coefficient of Determination,” 26th Annual DoD Cost Analysis Symposium, Leesburg, VA, September 1992
13. Draper, N.R., and H. Smith, “Applied Regression Analysis, 2nd Edition,” NY: John Wiley & Sons, 1981
14. Wedderburn, R.W.M., “Quasi-likelihood Functions, Generalized Linear Models, and the Gauss-Newton Method,” Biometrika, Vol. 61, Number 3 (1974), pages 439-447

TECOLOTE RESEARCH, INC.
Bridging Engineering and Economics
Since 1973

Backup Slides

TECOLOTE RESEARCH, INC.

Regression Technique – Pros and Cons of ZMPE Method

- **The ZMPE method is a constrained minimization process, which requires Solver to derive a solution**
- **Solver problems encountered when developing ZMPE CERs**
  - The ZMPE CER appears to be trapped in local minima more often than the corresponding MUPE CER (Solver is sensitive to the starting points)
  - The ZMPE equations are found be less stable than the MUPE ones, especially for small samples or complicated nonlinear equations such as Triad (see Reference 1)
- **The statistical properties of ZMPE CERs are not readily available** (see Reference 1 for details)
  - Difficult to interpret ZMPE CERs (Mean, Median, or Mode) & their SPEs
  - Difficult to examine whether the fitted coefficients are significant or not
    - SPE and Pearson's r^2 are not used to decide whether the CER is significant; they are insufficient to detect model flaws either
- **Difficulties noted when generating PIs for ZMPE CER** (see Ref 1)
- **The ZMPE CER has zero proportional error for all sample data points under the given constraint (but it's not necessarily unbiased)**
- **This method requires no transformations or CF adjustments**

TECOLOTE RESEARCH, INC.

Regression Technique – Pros & Cons of Log-Error Method

- **This method involves a two-step process:**
  - Perform the curve fitting in log space
    - It requires nonlinear regression for linear and triad CERs
  - Transform the results back to unit space
- **We need to apply a correction factor (Goldberger's Factor or the PING Factor) to the unit space CER to adjust for the mean because the resultant equation will be biased low when transforming back to unit space**
  - Smearing factor ($\Sigma(y_i/\hat{y}_i)/n$) is a non-parametric CF for bias
- **Because of the use of the correction factor (CF), the predictive measures should be adjusted accordingly**
- **If the hypothesized equation is log-linear, e.g., $y = ax^b\varepsilon$, the regression can be done in log space linearly under the logarithmic transformation**

TECOLOTE RESEARCH, INC.

Regression Technique – Comparison Table

Features/Functions/Requirements	MUPE	ZMPE	Log-Error
Require Transformation?	No	No	Yes
Require CF Adjustment (for Mean)?	No	No	Yes
Require Pred Measures Updates?	No	No	Yes
Consistent Estimates?	Yes	No*	No*
BLUE for Linear Model?	Yes	No*	No*
MQLE for Parameters?	Yes	No*	No*
Provide CI for Regr Coefficient?	Yes	No	Yes
Provide PI for Future Observation?	Yes	No	Yes
Require Nonlinear Regression?	Yes	Yes	Yes
Trapped in Local Minima?	Rarely	Often♦	Occasionally•
Nonlinear Equation Stable?	Mostly Yes	Mostly No	See note@

- * Proof is required
- ♦ Solver seems to be fairly sensitive to the starting points
- • It is occasionally trapped in local minima for triad
- @ It appears to be less stable than MUPE when fitting triad equations

CIC Analysis: T1 vs. QAIV (3/3)

QAIV Analysis Summary for USCM8 Database (see Ref 3):

- QAIV model implied quantity slopes cluster in the 90 – 100% range for USCM8 (average around 95%)
- No clear case for Baseline or QAIV as the better method for USCM8 data based on goodness-of-fit and predictive measures
- QAIV CERs tend to more closely represent underlying data (estimate trend lines versus actuals in 12 of 16 cases)
- No clear estimating bias in QAIV or Baseline CERs (QAIV CERs estimate higher than Baseline CERs in 9 of 16 cases)
- Cost improvement carryover assumption does not have a significant impact on USCM8 QAIV models