TECOLOTE RESEARCH, INC.

Bridging Engineering and Economics

Since 1973

# Why ZMPE When You Can MUPE?

## 6th Joint ISPA/SCEA Conference

Dr. Shu-Ping Hu and Alfred Smith

12 - 15 June 2007



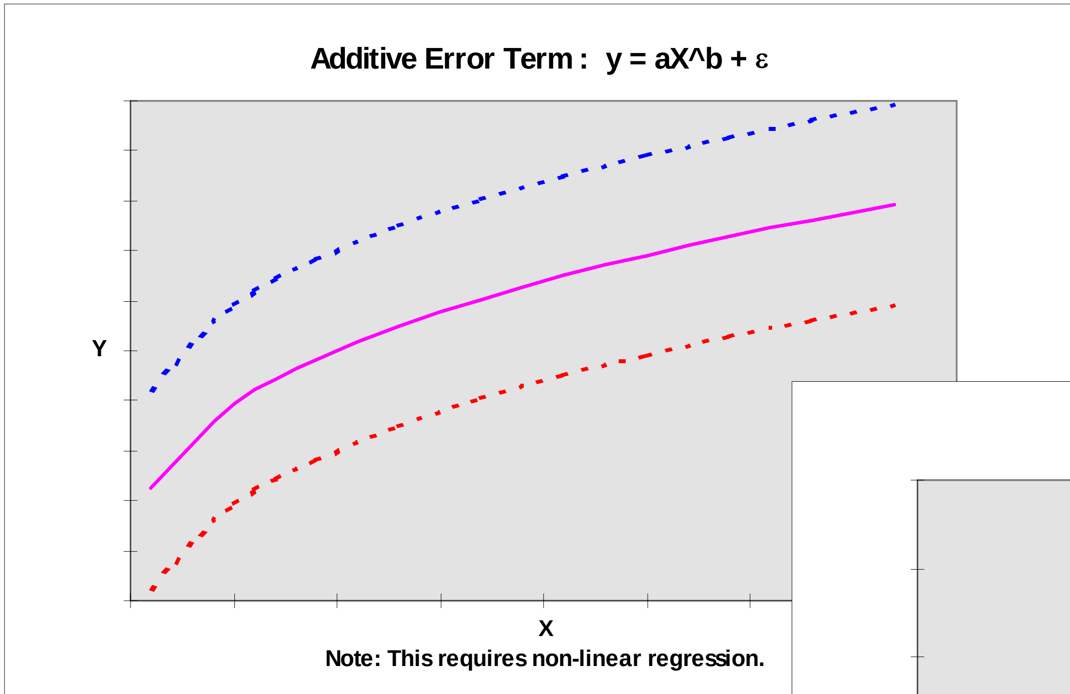
# Mission

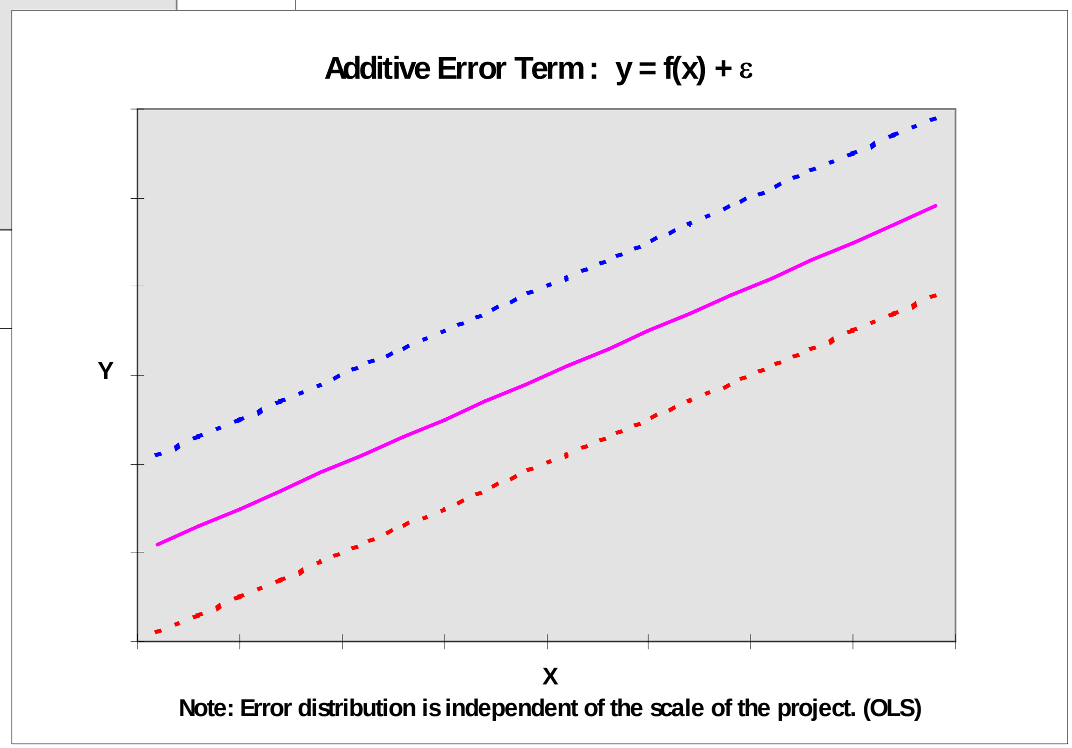
- **Find a regression method to fit a multiplicative CER in unit space directly (avoid transformations) and without bias**
- **Examine the goodness-of-fit measures to determine if the regression equation and coefficients are significant**
- **Interpret the CER result and use a validated and already implemented method to compute prediction intervals (PI) for cost uncertainty analysis**
  - Regardless of the CER methodologies, it is very important that the user knows (1) the CER result meaning and (2) how the error should be modeled.

# Outline

- **Mission**
- **Introduction/Background**
  - Error Term Assumption (Additive vs. Multiplicative)
  - Multiplicative Error Models: Log-Error, **MUPE**, ZPB/MPE (**ZMPE**)
- **Properties of MUPE and ZMPE CERs**
- **Common pros and cons of using MUPE and ZMPE**
- **Bad news about ZMPE CERs**
- **A Ground Antenna Example**
- **Good news about MUPE CERs**
- **Conclusions**

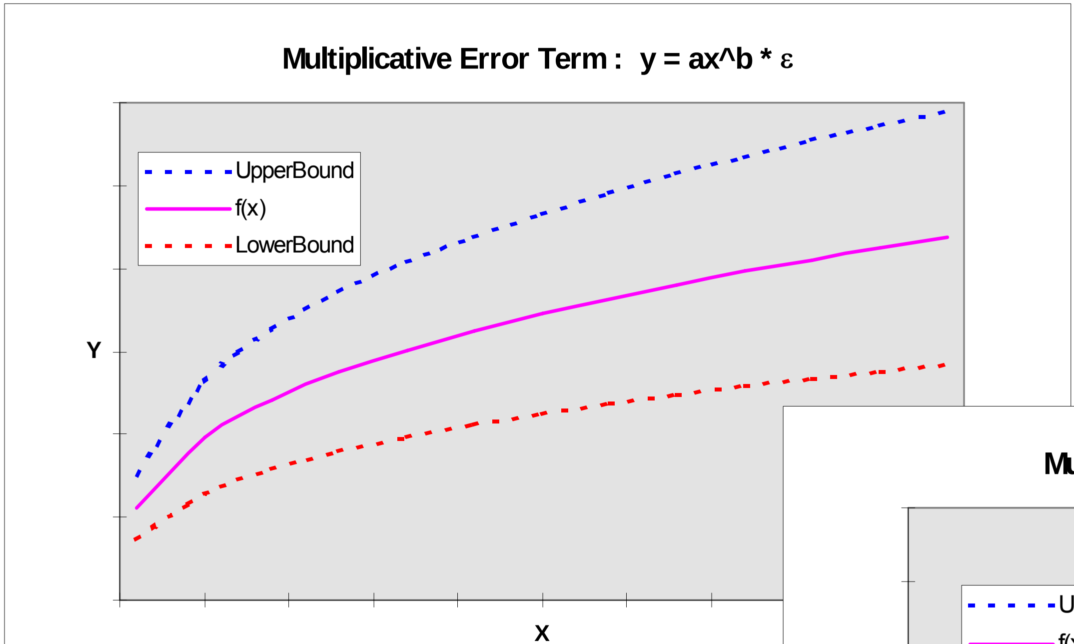
# Additive Error Term



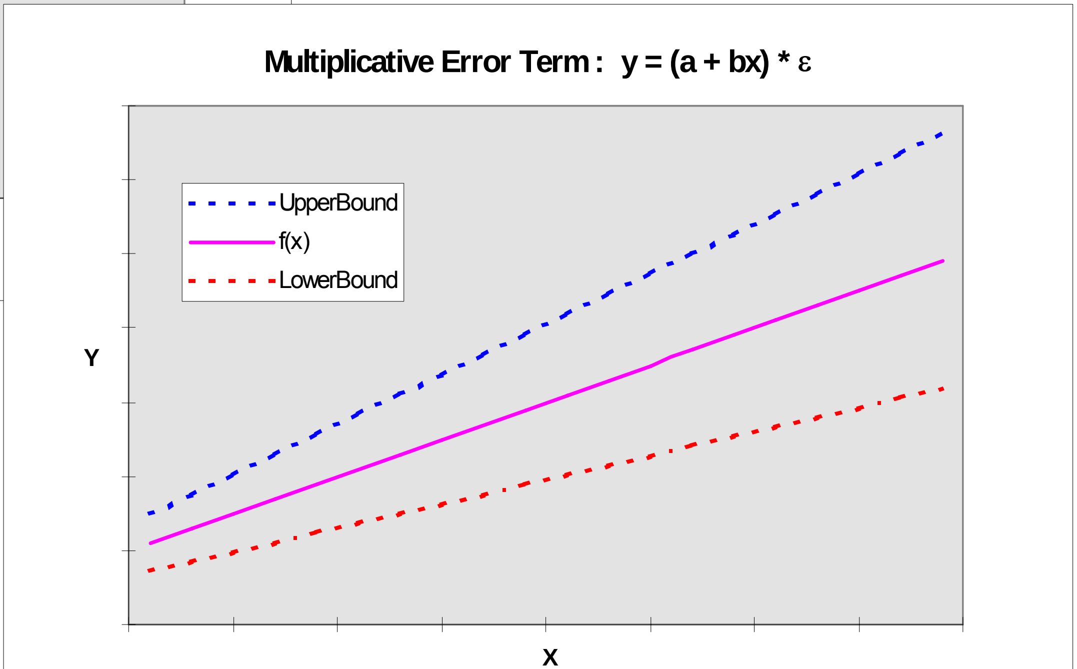




**Cost variation is independent of the scale of the project**

# Multiplicative Error Term




Note: This equation is linear in log space.




Note: This requires non-linear regression.

**Cost variation is proportional to the scale of the project**

# 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

# MUPE and ZMPE Methods

- **Two methods to perform the optimization for the weighted least squares using the <u>predicted</u> value, not the actual, as the basis. Sample percentage bias removed in both methods**
  - **MUPE** ⇨ *bias eliminated by fixing the denominator in IRLS*

$$Minimize \quad \sum_{i=1}^{n} \left( \frac{y_i - f(x_i)}{f_{\boldsymbol{k-1}}(x_i)} \right)^2$$

where $k$ is the iteration number

  - **ZMPE** ⇨ *sample bias eliminated through a constrained minimization process*

$$Minimize \quad \sum_{i=1}^{n} \left( \frac{y_i - f(x_i)}{f(x_i)} \right)^2$$

$$Subject\ \ to \sum_{i=1}^{n} \left( \frac{y_i - f(x_i)}{f(x_i)} \right) = 0$$

# Properties of MUPE and ZMPE Methods (1/3)

- **Both methods have zero percentage bias (ZPB) for the sample data points:**

$$\frac{1}{n}\sum_{i=1}^{n}\frac{y_i - \hat{y}_i}{\hat{y}_i} = 0$$

  - For MUPE, this condition is derived through the minimization process, which can be proved mathematically
  - For ZMPE, ZPB is obtained by using a constraint

- **If a CER is unbiased, then E($\hat{Y}$) = E(Y) = f(X,$\beta$)**

  i.e., the mean of the predicted CER is the hypothetical equation

- **Does “ZPB” imply that the CER is unbiased?**
  - The answer is **NO**
  - The ZPB constraint can be applied to any proposed methodologies (i.e., objective functions), but this is no guarantee that the CER result will be unbiased, i.e., this condition “E($\hat{Y}$) = f(X,$\beta$)” may not be true

# Properties of MUPE and ZMPE Methods (2/3)

- **The ZMPE method has a smaller standard percent error (SPE), i.e., multiplicative error, when compared to MUPE**
  - $SPE_{(ZMPE)} \leq SPE_{(MUPE)}$

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

- **Is ZMPE's SPE an unbiased estimator of $\sigma^2$?** ($\sigma^2 = V(\varepsilon)$)
  - Is a smaller SPE *better*?
    - If so, then we should develop MPE CERs
  - Do we know the statistical meaning of ZMPE's SPE?
  - Do we know the statistical interpretation of the ZMPE CER? Is it mean, median, mode, or what?

# Properties of MUPE and ZMPE Methods (3/3)

- **For linear CERs, such as Y = (a + bX + cZ)*ε, the MUPE method produces unbiased estimates of the function mean, i.e., $E(\hat{Y}) = E(Y) = a + bX + cZ$; the ZMPE method may not**
  - MUPE produces the best linear unbiased estimator (**BLUE**) of the parameters, a, b, and c and, consequently, BLUE of the function mean (under **no** distribution assumptions). See Ref 7 for details
  - ZMPE's estimated parameters are different from MUPE; they are certainly **not** BLUE
- **The MUPE CER produces consistent estimates of the parameters and the mean of the equation**
  - For non-linear CERs, unbiased estimates of the CER mean in general cannot be derived; the best to be found is consistency
- **MUPE's parameter estimators are also the maximum likelihood estimators (MLE) of the parameters**
- **We do not know the statistical properties of the ZMPE CER, but we know that the statement "ZMPE is unbiased" has NOT been proven, even for a linear CER**

# Common Pros and Cons for MUPE and ZMPE

- Both MUPE and ZMPE CERs have "zero percentage bias" for all points in the data set (no sample bias)
- Both methods require no transformation and no correction factor adjustment to the CER result
- Both rely on nonlinear regression technique to derive a solution
- Both methods do not always converge, especially when regressing learning curves

# Bad News About ZMPE for CER Development

**The ZMPE method is a constrained minimization process**

- **The ZMPE CER appears to get trapped in local minima more often than the corresponding MUPE CER**
  - Optimizers (i.e. Solver) are **sensitive** to the starting points for ZMPE, especially when regressing complicated non-linear equations
- **The ZMPE equations are found be less stable than the MUPE ones, especially for small samples**
- **Difficult to examine whether the calibrated coefficients are significant or not**
  - Optimizers generally do **not** provide any goodness-of-fit measures to the fitted regression equation other than SPE, which is simply based upon the objective function
  - SPE and Pearson's $r^2$ are insufficient
    - to determine if a ZMPE, a MUPE or a non-linear CER is significant, or
    - for detecting CER model flaws (see chart 14)

TECOLOTE RESEARCH, INC.

# Bad News About ZMPE for Uncertainty Analysis

- **No objective interpretation of the ZMPE CER**
  - Difficult to interpret ZMPE CERs - Mean, Median, Mode, or What?
- **Prediction intervals (PI) not readily available**
- **Although the Bootstrap method was suggested to construct PI bounds on the CERs, several shortcomings were reported in Reference 1, for instance:**
  - The actual implementation of the Bootstrap method to develop PI bounds could be tedious. For non-linear CERs, such as Triad ($y=a+bx^c+dz^e$), the process involves fitting hundreds of non-linear equations. Some of them may not converge or may be trapped in local minima, especially when the sample size is small
  - The Bootstrap sampling may **not** provide representative samples of the error distribution for **small** samples (see References 1 and 3)
  - The Bootstrap-based PI bounds were observed to be narrower than expected and not centered on the CER result for factor and linear equations. See Reference 1 for examples.

# A Ground Antenna Example from Ref 2 (1/5)

- **Data Set (see reference below):**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Y: Cost (FY99$K) | 3.595 | 1.900 | 3.300 | 10.900 | 15.434 | 16.074 | 17.274 |
| X: Diameter (feet) | 7.900 | 8.200 | 9.800 | 11.500 | 16.400 | 19.700 | 23.600 |

- **CERs and Stats:**

| Method | CER | SPE | Pearson's $r^2$ | Note |
|---|---|---|---|---|
| MUPE: | $(-28.45) + 13.49 * X^{0.404}$ | 40.5% | 89.5% | y < 0 if x < 6.34 |
| ZMPE #1: | $(-236.11) + 212.42 * X^{0.06}$ | 39.5% | 91.4% | y < 0 if x < 5.82 |
| ZMPE #2: | $75.661 + (-111.258) * X^{-0.2047}$ | 39.4% | 92.3% | y < 0 if x < 6.57 |

- **Not sufficient using SPE and Pearson's $r^2$**
  - It can be risky when selecting a CER solely based upon SPE and Pearson's $r^2$, especially when predicting outside the data range
  - We cannot interpret ZMPE CER and its SPE; this measure does not help identify the flaws in the CERs. (ZMPE CER #2 is totally absurd)
- **Use approximated std errors of the coefs to evaluate CERs**

Ref: Book, S., "IRLS/MUPE CERs Are Not MPE-ZPB CERs," 2006 Annual ISPA International Conference, Seattle, WA, 23-26 May 2006

# A Ground Antenna Example CO$TAT Output (2/5)

TECOLOTE RESEARCH, INC.

**I. Equation Form & Error Term**

| | |
|---|---|
| **Model Form:** | Weighted Non-Linear Model |
| **Non-Linear Equation:** | Y = (-28.45) + 13.49 * X ^ 0.404 |
| **Error Term:** | MUPE (Minimum-Unbiased-Percentage Error) |
| **Minimization Method:** | Downhill Simplex |

**II. Fit Measures**

**Coefficient Statistics Summary**

| Variable/Term | Coefficient Estimate | Approximate Std Error | Approximate Lower 95% Confidence | Approximate Upper 95% Confidence |
|---|---|---|---|---|
| **Fixed_Cost** | -28.4544 | 143.9821 | -428.4365 | 371.5278 |
| **a** | 13.4852 | 106.4782 | -282.3113 | 309.2817 |
| **b** | 0.4040 | 1.5859 | -4.0016 | 4.8096 |

**Least Squares Minimization Summary Statistics**

| Source | DF | Sum of Squares (SS) | Mean SQ = SS/DF |
|---|---|---|---|
| **Residual (Error)** | 4 | 0.6551 | 0.1638 |
| **Total (Corrected)** | 6 | 3.0113 | |

**Goodness-of-Fit Statistics**

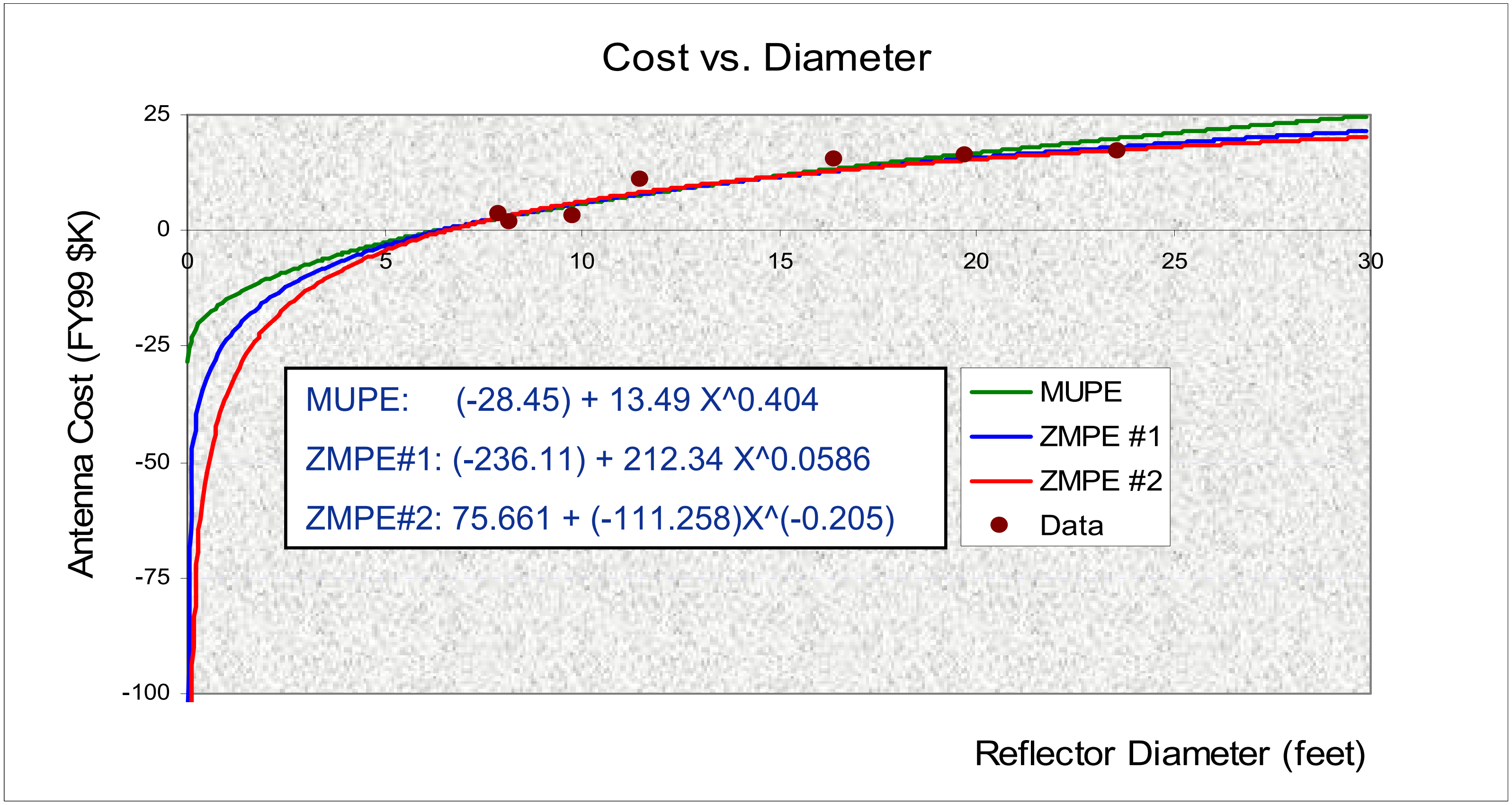
| Std. Error (SE) | Approx. R-Squared | Approx. R-Squared (Adj) |
|---|---|---|
| 0.4047 | 78.24% | 67.37% |

**Approximate Correlation Matrix of the Coefficients**

| Correlation | Coef 1 | Coef 2 | Coef 3 |
|---|---|---|---|
| **Coef 1** | 1.0000 | -0.9997 | 0.9982 |
| **Coef 2** | -0.9997 | 1.0000 | -0.9994 |
| **Coef 3** | 0.9982 | -0.9994 | 1.0000 |

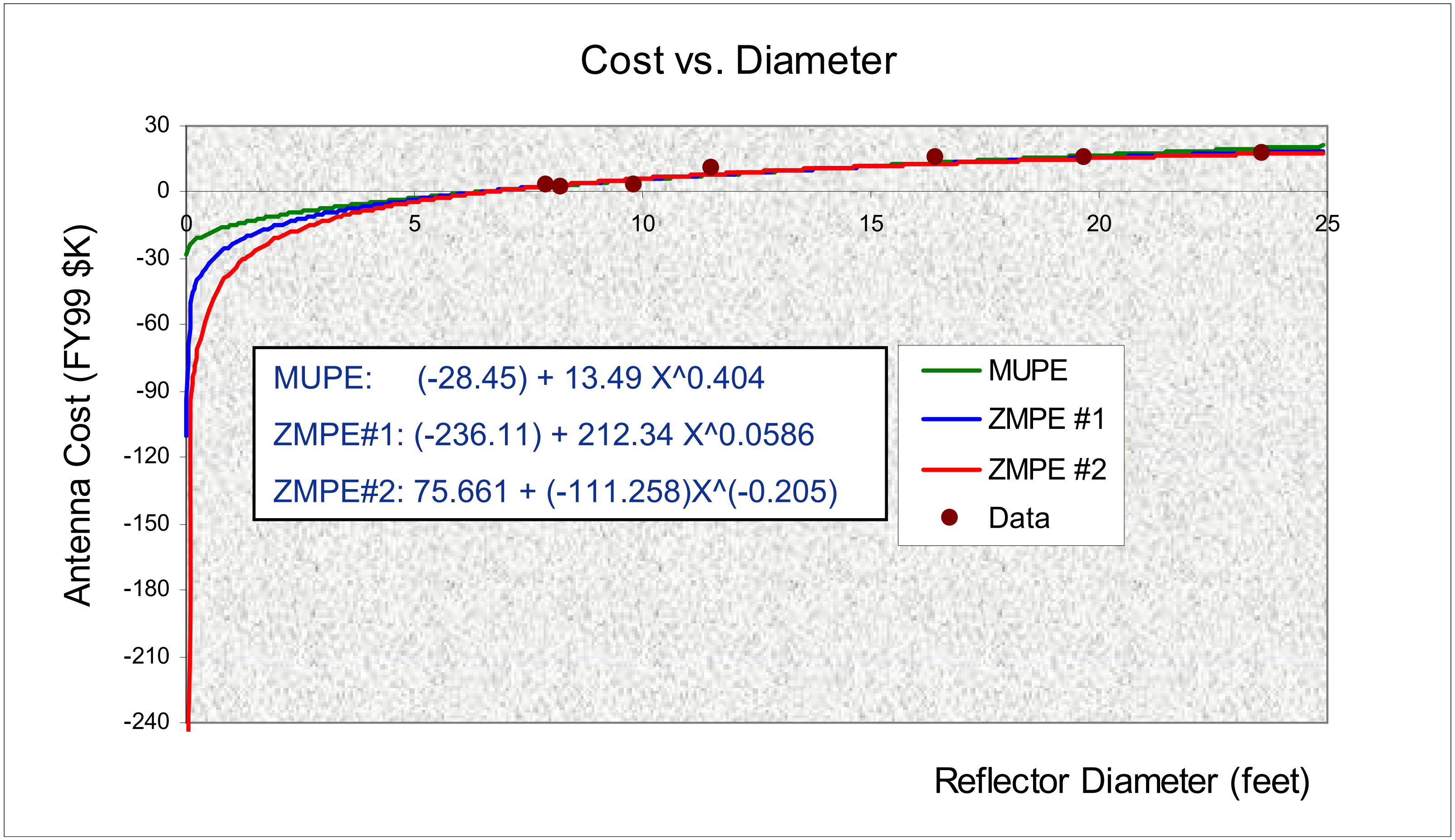
1. Systat produces very similar results
2. This CER should be rejected based upon the estimated std error of the coef

# A Ground Antenna Example Scatter Plot (3/5)




1. Over the data range, all three CERs are similar and appear to be normal
2. ZMPE CER #2 closely follows ZMPE CER#1 even beyond the data range
3. Note that when using 2 significant digits rather than four for the coefficients, ZMPE CER #1 is biased high by **21**%

# A Ground Antenna Example Scatter Plot (4/5)




1. This graph plots the CERs and the data points from “0” to 25 feet diameter
2. It shows the peculiarities of the triad equation

# A Ground Antenna Example from Ref 2 - Summary (5/5)

TECOLOTE RESEARCH, INC.

**ZMPE: Cost = -236.11 + 212.42 $X^{0.0586}$** **MUPE: Cost = -28.45 + 13.49 $X^{0.404}$**

- It makes no sense to have a negative set-up cost in the equation
- The ZMPE CER will generate a negative cost when the reflector has a diameter less than 5.82 feet (for MUPE, the threshold is 6.34 feet)
- The coefficients generated by the MUPE method are insignificant based upon the approximated standard errors of the coefficients (see the coefficient table in CO$TAT output)
- For the Triad MUPE CER, the fixed cost term (-28.45) is competing against the scale parameter (13.49) of the equation; the fixed cost term is also almost perfectly correlated with the exponent. (See the correlation matrix of the coefficients)
- When using two significant digits for the exponent, as suggested by Figure 10 of Ref 2, the ZMPE equation is biased high by **21**%. This also indicates that the ZMPE equation is not stable
- Neither Pearson's $r^2$ nor SPE can detect the flaws in this triad model

# What About Large Samples?

- **No significant differences found between MUPE and ZMPE CERs for large data sets**
  - ZMPE and MUPE CERs are quite close to each other for large samples generated by simulation runs
- **Goodness-of-fit measures still not available for ZMPE CERs**
  - SPE and Pearson's $r^2$ are not sufficient; it can be risky to evaluate ZMPE CERs based upon their SPE and Pearson's $r^2$
- **It takes extra effort to generate and validate the Bootstrap-based PIs for cost uncertainty analysis**
  - The Bootstrap sampling should provide representative samples of the error distribution for **large** samples. However, the Bootstrap-based PI bounds (i.e., Bootstrap Bound ± SEE) should be validated (Ref 1, 3 & 4)
  - There may be a **consistency** issue when users are specifying PIs at different probability levels, e.g., ($15^{th}$,$85^{th}$) or ($20^{th}$,$80^{th}$) vs. ($10^{th}$,$90^{th}$)

# Good News About MUPE

- **Approximated standard errors of the coefficients can be applied to judge the quality of the regression coefficients under the normality assumption**
- **The MUPE CER produces consistent estimates of the parameters and the mean of the equation**
  - MUPE CERs estimate the mean of the function for linear CERs
  - For non-linear CERs, the best to be found is consistency
- **The MUPE estimated parameters are the maximum likelihood estimators (MLE) of the parameters (see References 5, 6, 8)**
- **The PIs for MUPE CERs (as well as non-linear CERs) can be found in several statistical packages, including SAS, Statistica, and CO$TAT**

TECOLOTE RESEARCH, INC.

# Conclusions

- **Besides the common pros and cons, MUPE (not ZMPE) offers informative and useful statistics**
  - It provides consistent estimates of the parameters and the CER mean; it is BLUE for linear models
  - It provides (asymptotic) goodness-of-fit measures for evaluating CER coefficients
  - Its PI is readily available for cost uncertainty analysis (see SAS, Statistica, and CO$TAT)
- **Shortcomings found for ZMPE if sample size is small** (<15)
  - The ZMPE CER appears to be trapped in local minima more often than the corresponding MUPE CER
  - The ZMPE equations are found to be less stable than the MUPE ones and more sensitive to the starting points
- **While no significant differences found between MUPE and ZMPE CERs for large data sets, MUPE is the better one!**
  - ZMPE does not offer CER meaning, goodness-of-fit measures, or PIs
  - SPE and Pearson's $r^2$ cannot help detect the model flaws

# References


1. Hu, S., “Prediction Interval Analysis for Nonlinear Equations,” 2006 Annual SCEA National Conference, Tysons Corner, VA, 13-16 June 2006
2. Book, S., “IRLS/MUPE CERs are Not MPE-ZPB CERs,” 2006 Annual ISPA International Conference, Seattle, WA, 23-26 May 2006
3. Book, S., “Prediction Bounds for General-Error-Regression CERs,” 39th DoDCAS, Williamsburg VA, 14-17 February 2006
4. Book, S., “Prediction Intervals for CER-Based Estimates (With Cost Driver Values Outside the Data Range),” 37th DoDCAS, Williamsburg, VA, 10-13 February 2004
5. Seber, G. A. F. and Wild, C. J., “Nonlinear Regression,” John Wiley & Sons, Inc., 1989
6. Jørgensen, B., “Maximum Likelihood Estimation and Lare-Sample Inference for Generalized Linear and Nonlinear Regression Models,” Biometrika, 70, pages 19-28, 1983
7. Draper, N. R. and Smith, H., “Applied Regression Analysis,” 2nd Edition, John Wiley & Sons, 1981
8. Jennrich, R. I. and Moore, R. H., “Maximum Likelihood Estimation by Means of Nonlinear Least Squares,” American Statistical Assoc., Proc. Statistical Computing Section, pages 57-65, 1975

TECOLOTE RESEARCH, INC.

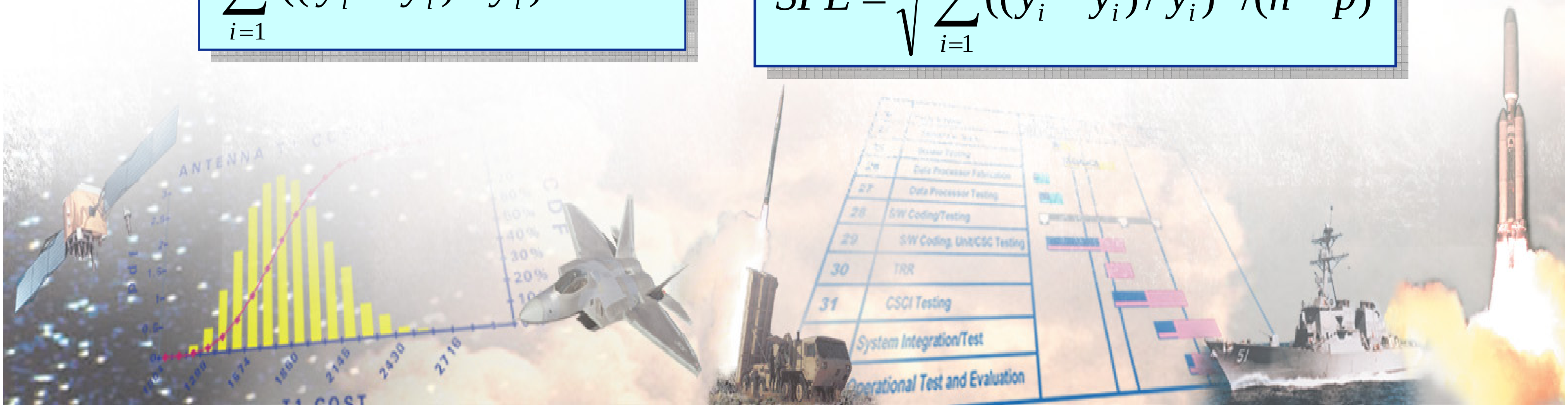
Bridging Engineering and Economics

Since 1973

# Backup Slides

$$\sum_{i=1}^{n}((y_i - \hat{y}_i)/\hat{y}_i)/n = 0$$

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



# Comparison between MPE and ZMPE

- **For most equations (i.e., $Y = a\ X^bZ^c$, $Y = a + bX + cZ$, etc.)**
  - Sensitivity coefficients (associated with the driver variables) are the same between MPE & ZMPE equations
  - Only leading term or level of function adjusted
  - Findings also proven by mathematical derivations
  - See reference below for details

- **For triad equations (i.e., $Y = a + b\ X^cZ^d$)**
  - All coefficients changed

Ref: 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.