THE ASSESSMENT OF MODEL ADEQUACY

Luis Tedeschi, Texas A&M University
NANP-NRSP-9 Modeling Committee

Available online at www.sciencedirect.com

SCIENCE DIRECT®

AGRICULTURAL SYSTEMS

Agricultural Systems 89 (2006) 225–247

www.elsevier.com/locate/agsy

Review

Assessment of the adequacy of mathematical models

Luis Orlindo Tedeschi *

Department of Animal Science, Texas A&M University, 230 Kleberg Center, 2471 TAMU, College Station, TX 77843-2371, United States

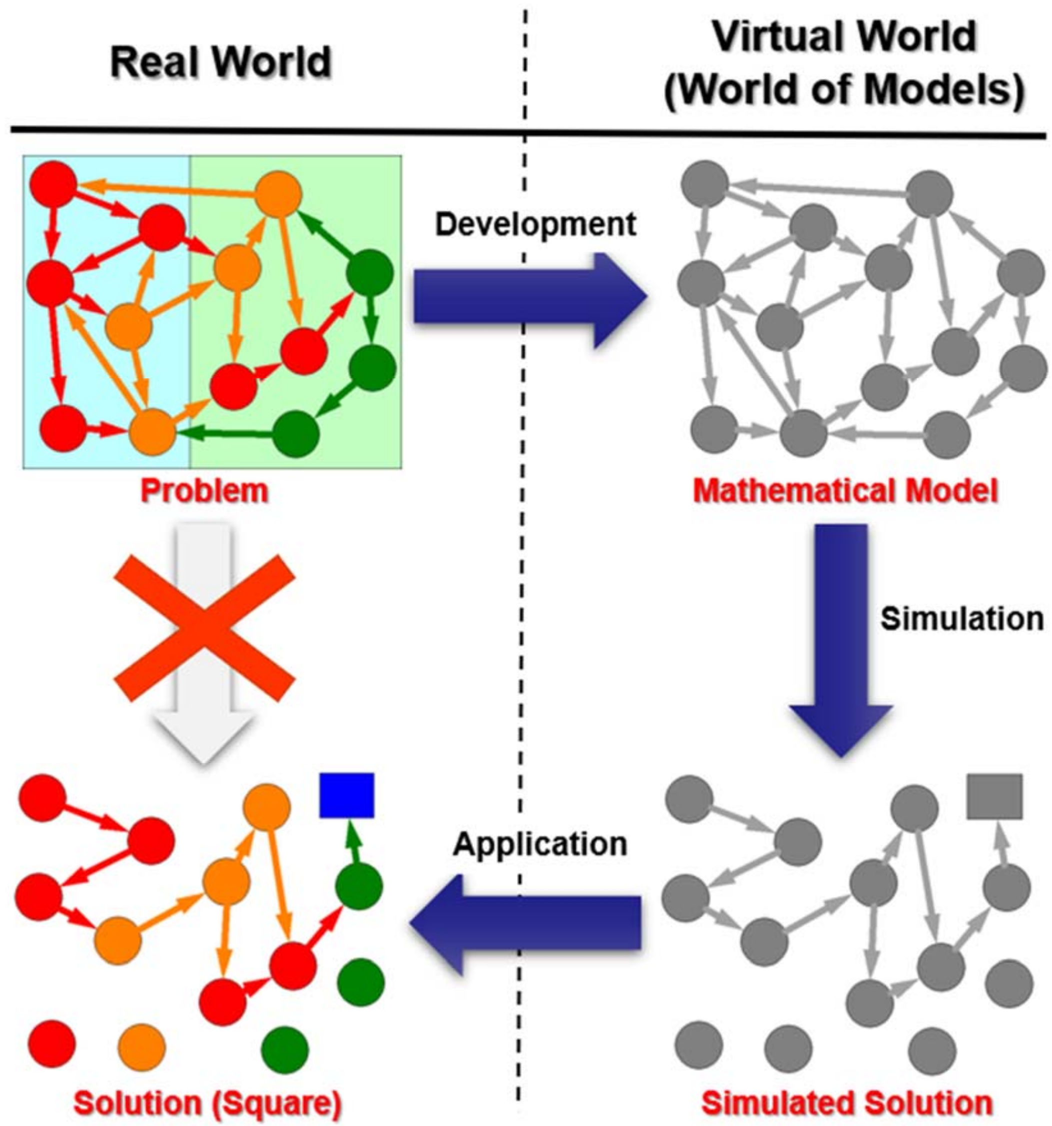
http://www.nutritionmodels.com/mes.html

What are models?

"Mathematical models (**MM**) are mental conceptualizations, enclosed in a virtual domain, whose purpose is to translate real-life situations into mathematical formulations (symbolically or numerically) to describe existing patterns or forecast future behaviors in the real-life situations"

-- Tedeschi (2019, 97:1921 J. Anim. Sci.)

Why Models?



Tedeschi (2019, 97:1921 J. Anim. Sci.)

"All models are wrong (false), but some are useful"

-- Box (1979)

- Understand and acceptance:
  - To strengthen the modeling process
  - To be more resilient to pitfalls during development and evaluation

- Improvement of the current model

- Understand the complex behavior of a phenomena by identifying small patterns in the process

Tedeschi (2006: 89:225 Ag. Syst.)

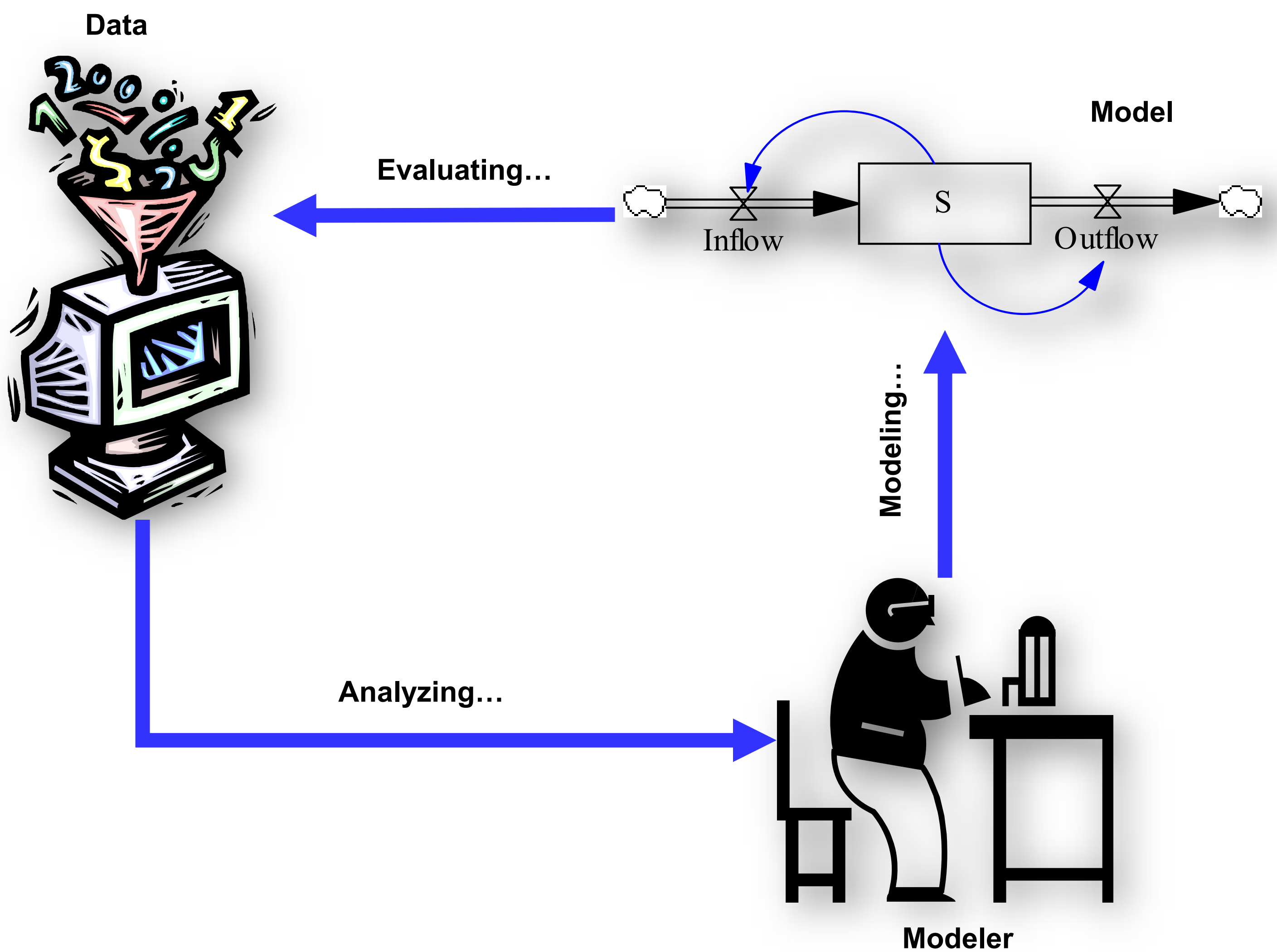
John Sterman

"In systems thinking, the understanding that models are wrong and humility about the limitations of our knowledge is essential in creating an environment [model] in which we can learn about the complexity of systems in which we are embedded"

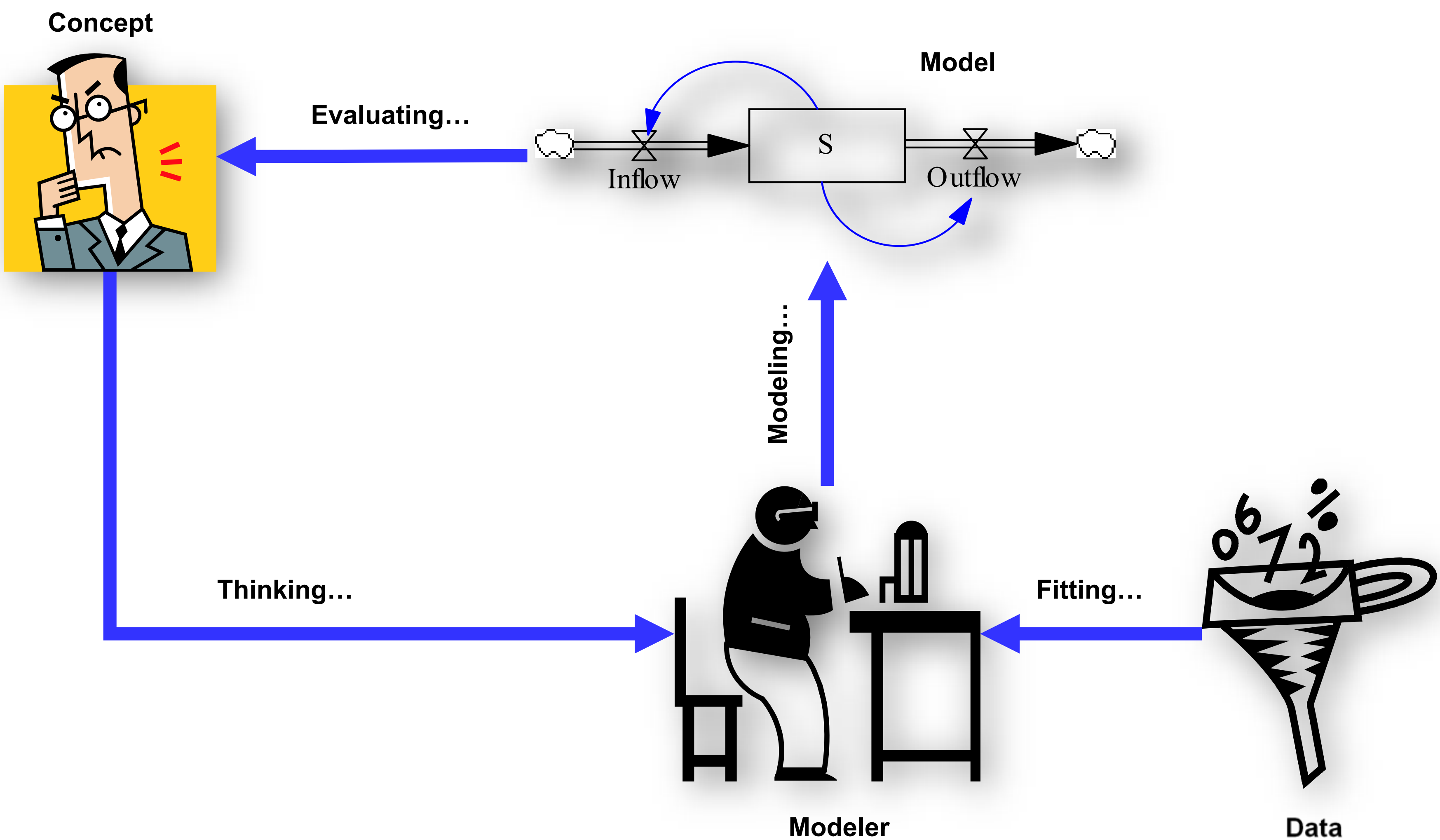
-- Sterman (2000)

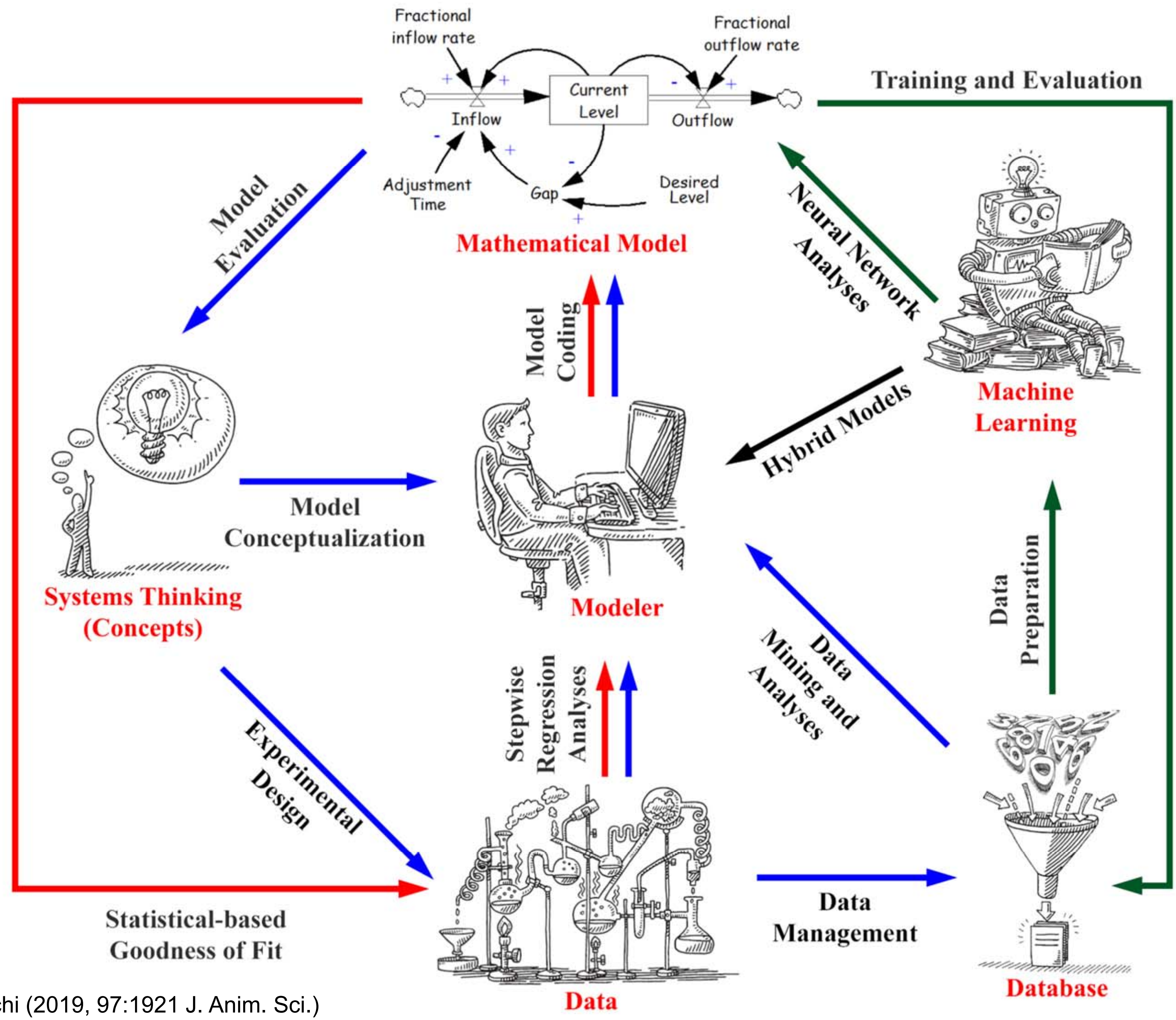
Processes for Model Development using Systems Thinking

Empirical or Relational Models



Conceptual or Theoretical Models





Tedeschi (2019, 97:1921 J. Anim. Sci.)

Model Evaluation

“Model testing is often designed to demonstrate the rightness of a model and the tests are typically presented as evidences to promote its acceptance and usability”

-- Sterman (2002)

John Sterman

¿Can models be 'validated'?

- Models cannot be validated
  - It is impossible to prove that all components of models or real systems are true or correct
- Models can never mimic the reality since they are representation of it
  - Some models can be programmed to predict quantities that cannot be measured in real systems

Tedeschi (2006: 89:225 Ag. Syst.)

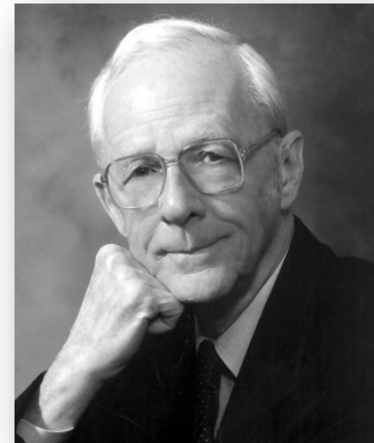
Models can be evaluated !

- Models can be evaluated or tested, but never validated
  - Validation means "*having a conclusion correctly derived from premises*"
  - Verification means "*establishment of the truth, accuracy, or reality of* "
- Calibration means model fine tuning or fitting; it is the estimation of values or parameters or unmeasured variables

Tedeschi (2006: 89:225 Ag. Syst.)

“Validity of a mathematical model has to be judged by its sustainability for a particular purpose; that means, it is a valid and sound model if it accomplishes what is expected of it”

-- Forrester (1961)



Jay Forrester

DEFINITIONS OF EVALUATION

Shaeffer (1980)

- Model examination
- Algorithm examination
- Data evaluation
- Sensitivity analysis
- Validation studies
- Code comparison studies

Hamilton (1991)

- Verification
  - Design, programming, and checking processes of the program
- Sensitivity Analysis
  - Behavior of each component of the model
- Evaluation
  - Comparison of model outcomes with real data

Evaluating Errors

TWO-WAY DECISION PROCESS

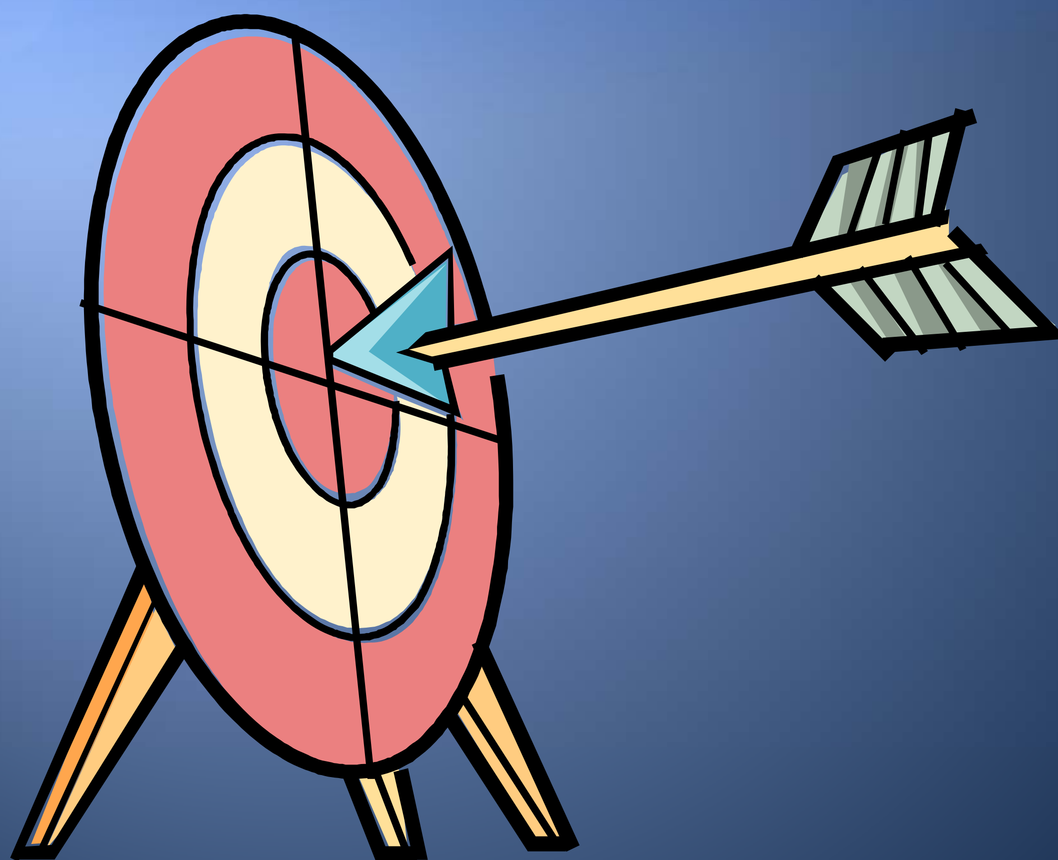
	Model Predictions	
Decision	Correct	Wrong
Reject	**Type I Error (α)**	Correct (1 - β)
Accept	Correct (1 - α)	**Type II Error (β)**

How does it happen?

- Type I Error (α): Rejecting an appropriate model
  - Biased or incorrect observations are chosen to evaluate a model
- Type II Error (β): Accepting a wrong model
  - Biased or incorrect observations are used to develop and evaluate a model
  - Conceptual model cannot be tested because lack of data

Tedeschi (2006: 89:225 Ag. Syst.)

Accuracy x Precision



Definitions

- Accuracy
  - It measures how closely model-predicted values are to the true values
  - Ability to predict the right values
- Precision
  - It measures how closely individual model-predicted values are within each other
  - Ability to predict similar values consistently

Tedeschi (2006: 89:225 Ag. Syst.)

More Definitions

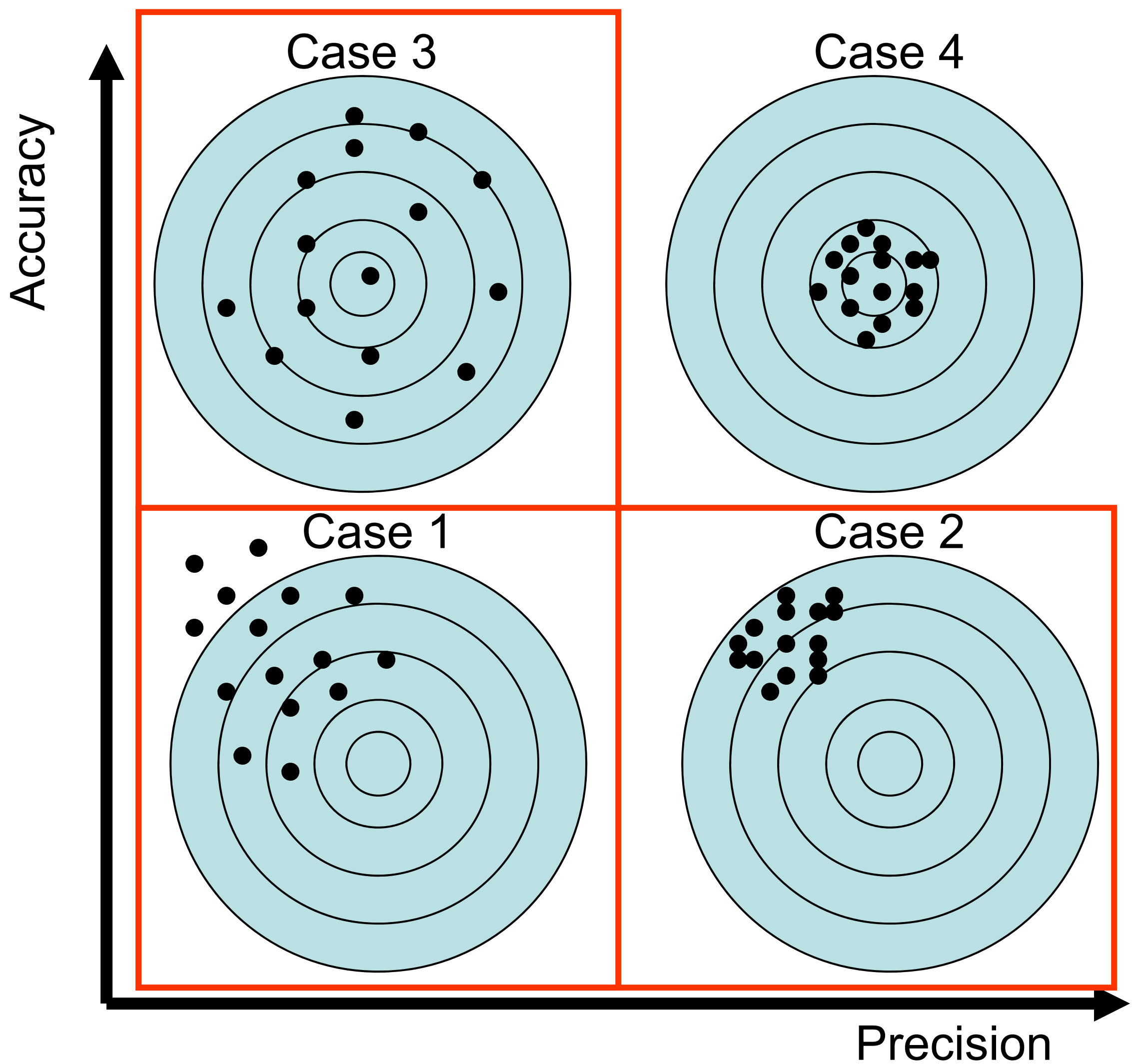
- Inaccuracy or bias
  – Systematic deviation from the truth

- Imprecision or uncertainty
  – Magnitude of the scatter about the average mean

Tedeschi (2006: 89:225 Ag. Syst.)

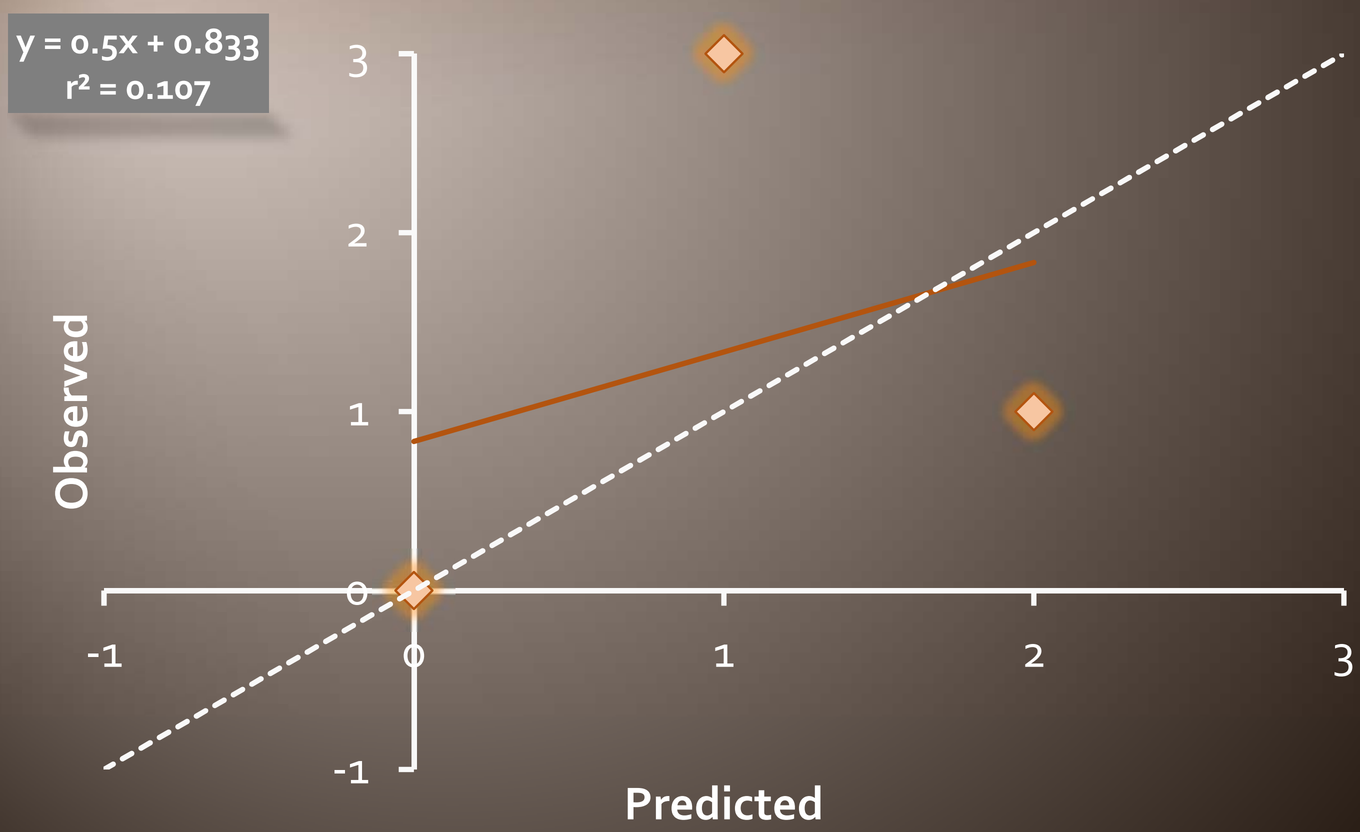
Imprecise or Uncertain

Inaccurate or biased

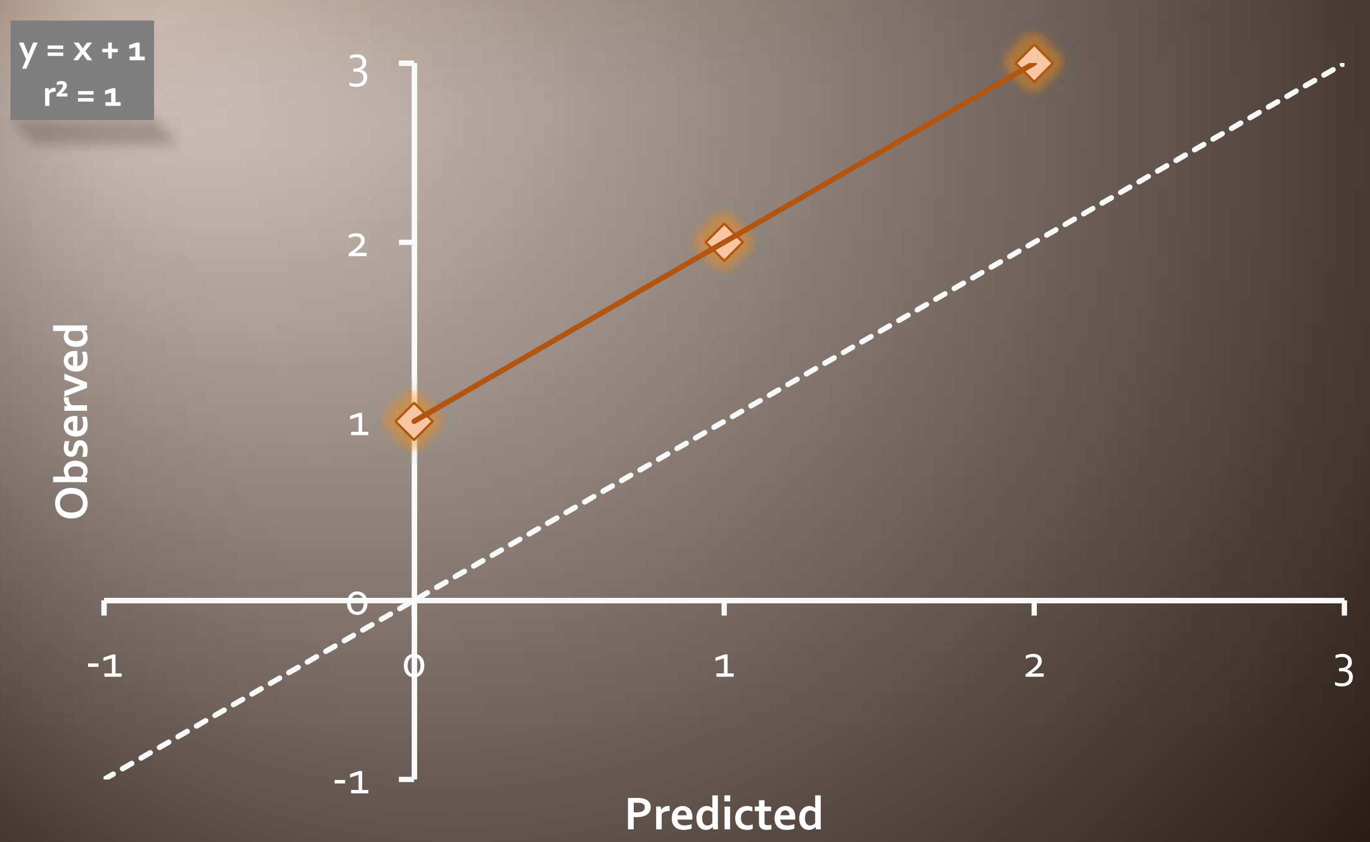


Tedeschi (2006: 89:225 Ag. Syst.)

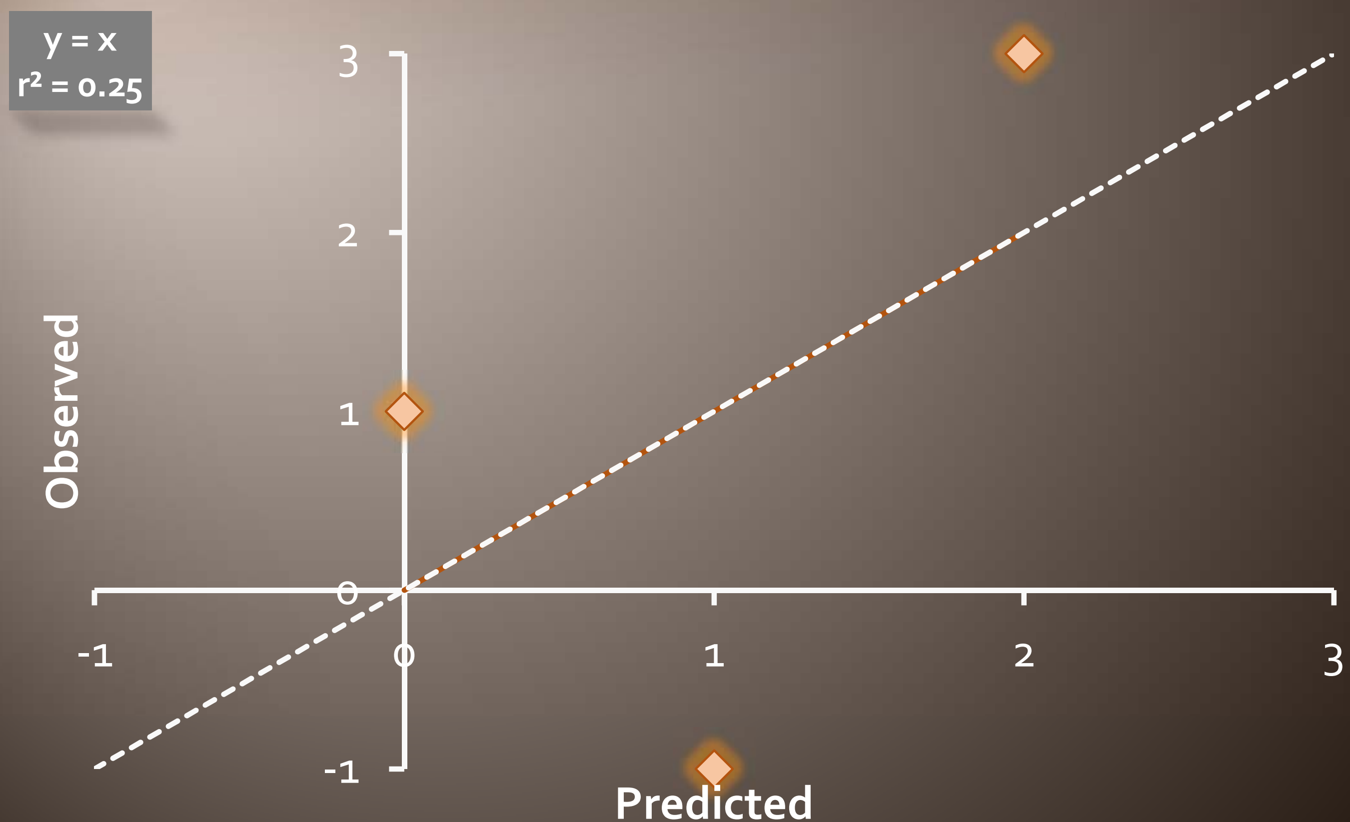
Case 1 - ↓ Precision ↓ Accuracy



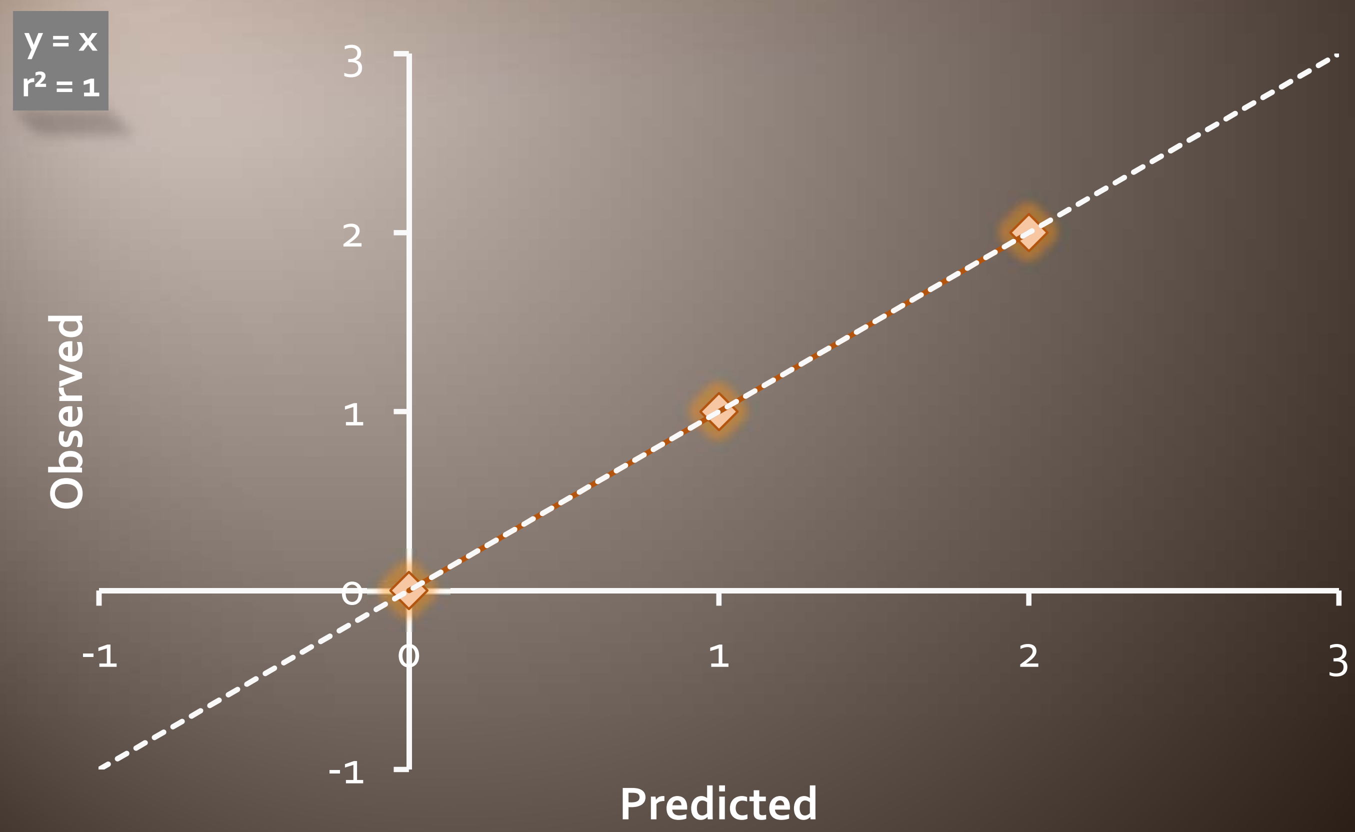
Case 2 - ↑ Precision ↓ Accuracy



Case 3 - ↓Precision ↑Accuracy



Case 4 - ↑ Precision ↑ Accuracy



Which one is better?

- Accuracy and Precision are independent
  - ↑ Accuracy does not imply ↑ Precision and vice-versa
- Imprecise model can get the right value using large number of data points (e.g. case 3)
- True mean is irrelevant for model comparison if the model is consistent (e.g. case 2)

Techniques for Model Evaluation: Regression Analysis

¿ Y-axis or X-axis ?

We regress the observed data (Y-axis) on the model-predicted (X-axis)

When using least-squares technique the vertical difference is minimized to estimate the parameters

Observed data has the random error, not the model-predicted values assuming deterministic model

Even stochastic models can be re-run several times, decreasing the error

Tedeschi (2006: 89:225 Ag. Syst.)

Why linear regression?

- Hypothesis is that when regression Y (Obs) on $f(X_{1,\ldots,}X_p)_i$ (Model-Pred), a perfect prediction would have intercept = 0 and slope = 1
- Little interest since the predicted value (by the linear regression) is useless in evaluating the mathematical model
- r^2 is irrelevant since one does not intend to make predictions using the fitted line!
  – May use it to adjust for model imprecision!

Tedeschi (2006: 89:225 Ag. Syst.)

Assumptions for LR

The X-axis values are known without errors (deterministic)

The Y-axis values MUST be independent, random, and homoscedastic

Residuals are independent and identically distributed ~ $N(0,\sigma^2)$

Tedeschi (2006: 89:225 Ag. Syst.)

Caution about r^2

A high coefficient of correlation (r) does not indicate that useful predictions can be made by a given mathematical model since it measures precision not accuracy

A high r does not imply the estimated line is a good fit (curvilinear)

An r near zero does not indicate that observed and model-predicted are not correlated since they may have a curvilinear shape

Tedeschi (2006: 89:225 Ag. Syst.)

MEAN SQUARE ERROR (MSE)

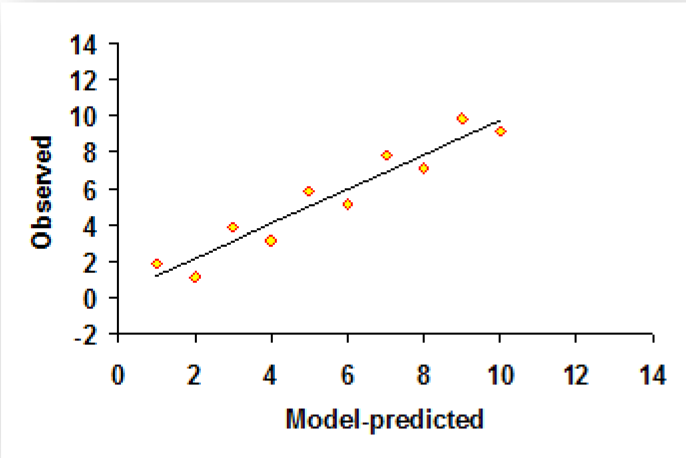
- Also known as residual mean square or standard error of the estimate
- This statistic may be used to compare model 'validity' when comparing models

$$MSE = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n-2}$$

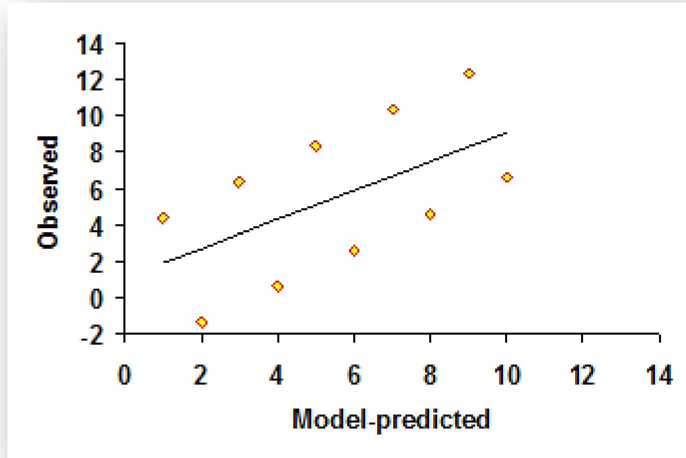
$$MSE = \frac{s_Y^2 \times (n-1) \times (1-r^2)}{n-2}$$

Tedeschi (2006: 89:225 Ag. Syst.)

Comparison of Model Prediction



- $Y_1 = a + b \times X \pm N(0,1)_{\alpha=0.2}$
- a = 0.28 ± 0.63
- b = 0.95 ± 0.10
- *P* (a=0) = 0.67
- *P* (b=1) = 0.63
- *P* (a=0 & b=1) = 0.90
- **r^2 = 0.92**
- **MSE = 0.89**



- $Y_2 = a + b \times X \pm N(0,4)_{\alpha=0.2}$
- a = 1.12 ± 2.53
- b = 0.80 ± 0.41
- *P* (a=0) = 0.67
- *P* (b=1) = 0.63
- *P* (a=0 & b=1) = 0.90
- **r^2 = 0.32**
- **MSE = 13.7**

Concerns about LR

Assumptions of normality and homoscedasticity are rarely satisfied

Ambiguous results depending on the scatter of the data

Regression lacks sensitivity to distinguish between random clouds and data points

Stochastic models require different technique to derive the parameters

Tedeschi (2006: 89:225 Ag. Syst.)

Is r^2 a good indicator of adequacy?

- r^2 measures how far (close) the observations (Y values) deviate from the best-fit regression
- The best-fit regression IS NOT the model-predicted values
- r^2 does not distinguish between:
  - observed and model-predicted values strongly agree
  - a strong linear relationship exists but the measurements do not agree

Tedeschi (2006: 89:225 Ag. Syst.)

Fitting Errors:
Analysis of Deviation

Analysis of Deviation

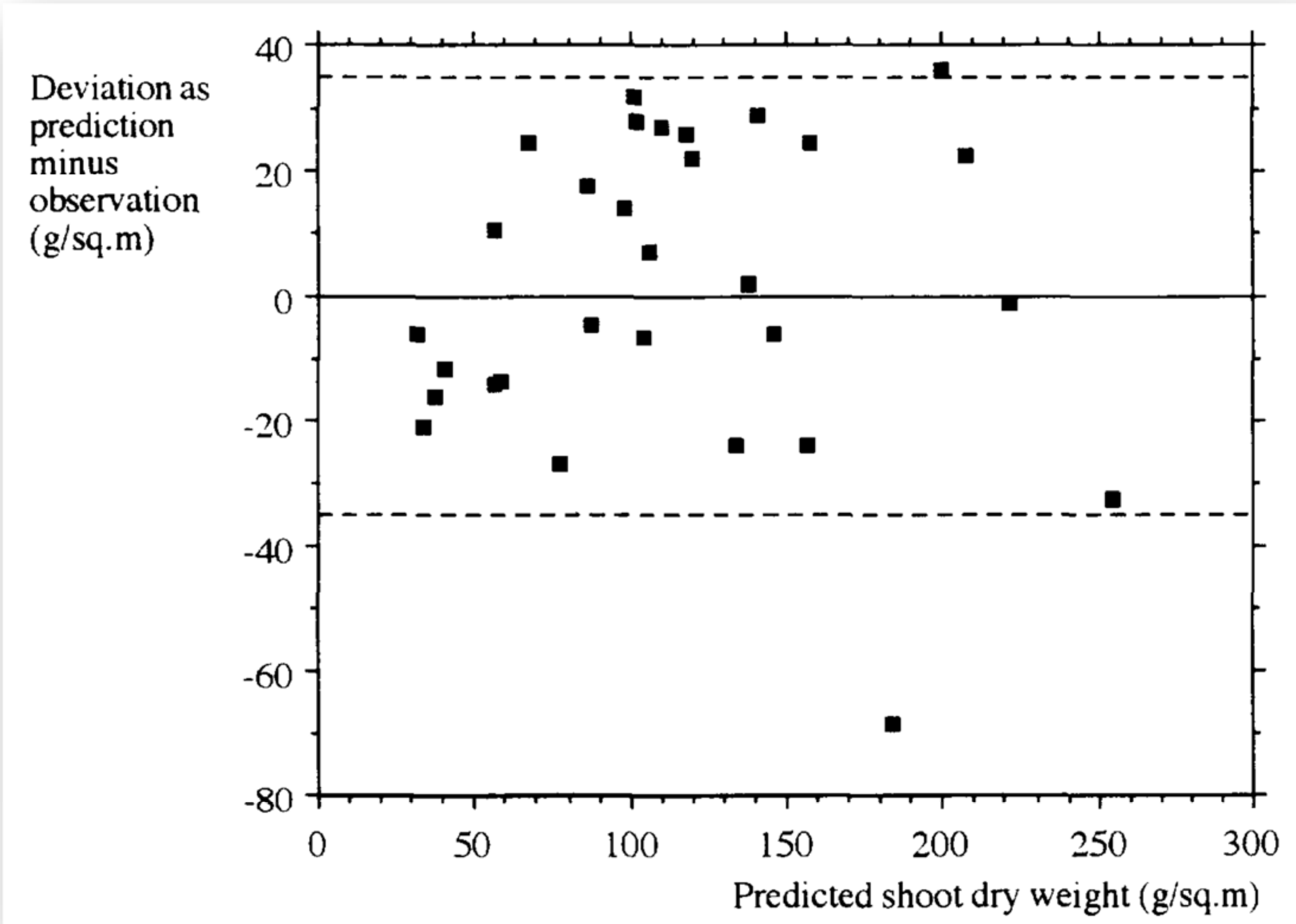
Empirical but powerful analysis

Deviation is the difference between **model-predicted** minus **observed** values

Usually, an acceptable range is used to accept or not the model performance

Tedeschi (2006: 89:225 Ag. Syst.)

Deviation Plot Analysis



Mitchell and Sheehy (1997)

Fitting Errors:
Extreme and Influential Points

FITTING ERRORS

- Extreme Points
  - Leverage
  - Studentized residue
  - PRESS$_p$
- Influential Points
  - DFFITS
  - Cook's distance

$$Semi\text{-}studentized = \frac{y_i - \hat{y}_i}{\sqrt{MSE}}$$

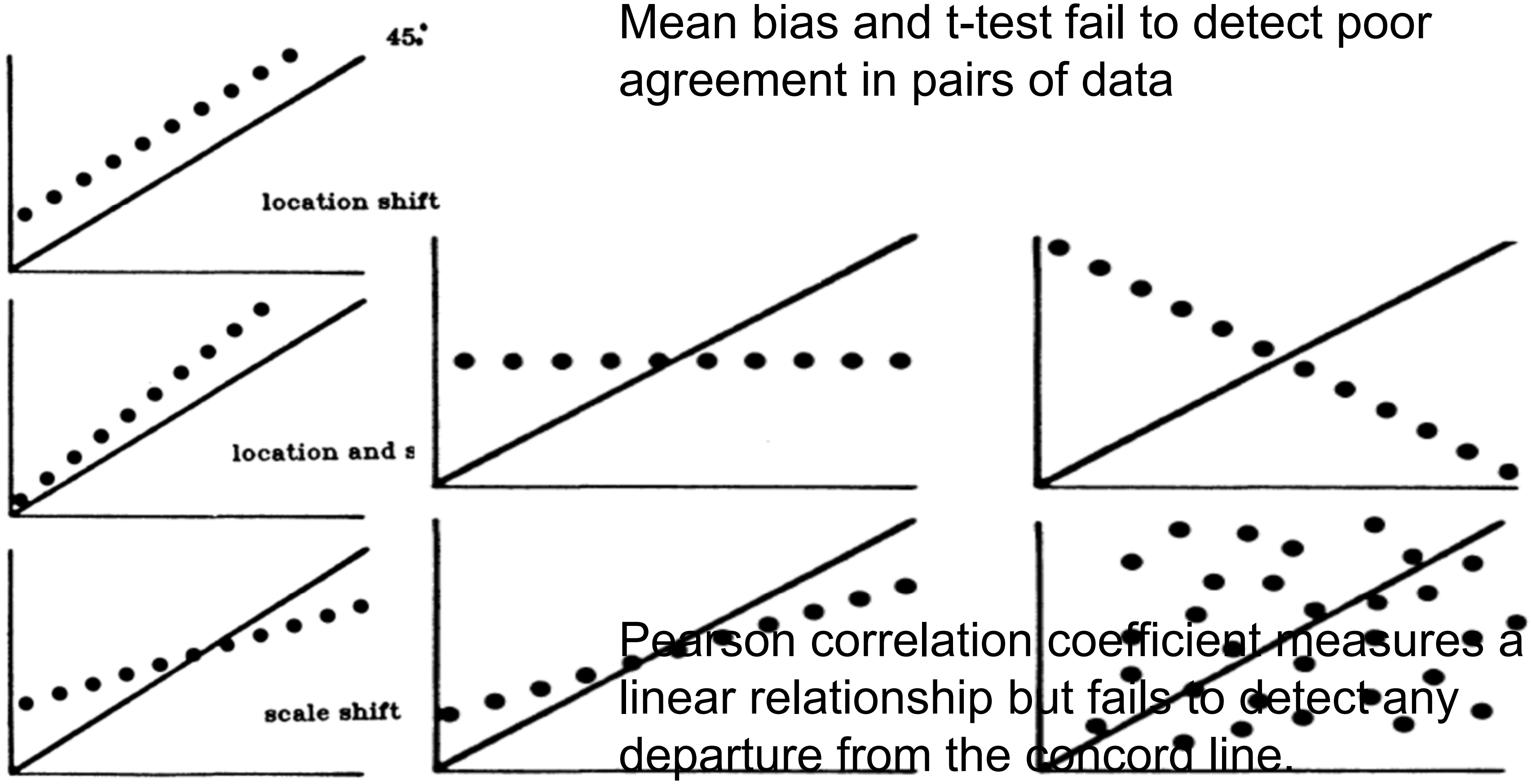
$$Cooks'D = \frac{\sum_{j=1}^{n}\left(\hat{y}_j - \hat{y}_{j(i)}\right)^2}{2 \times MSE} = \frac{\left(y_i - \hat{y}_i\right)^2}{2 \times MSE} \times \sqrt{\frac{h_{ii}}{\left(1 - h_{ii}\right)^2}}$$

Tedeschi (2006: 89:225 Ag. Syst.)

CONCORDANCE CORRELATION COEFFICIENT (CCC)

Failure of Agreement Measures

Mean bias and t-test fail to detect poor agreement in pairs of data



Pearson correlation coefficient measures a linear relationship but fails to detect any departure from the concord line.

Lin (1989, 45:255 Biometrics)

¿What is CCC?

CCC aka reproducibility index

Are the model-predicted values precise and accurate at the same time across a range and are tightly amalgamated along the unity line through the origin?

CCC accounts for precision and accuracy at the same time

Proposed initially by Krippendorff (1970) and modified by Lin (1989)

Tedeschi (2006: 89:225 Ag. Syst.)

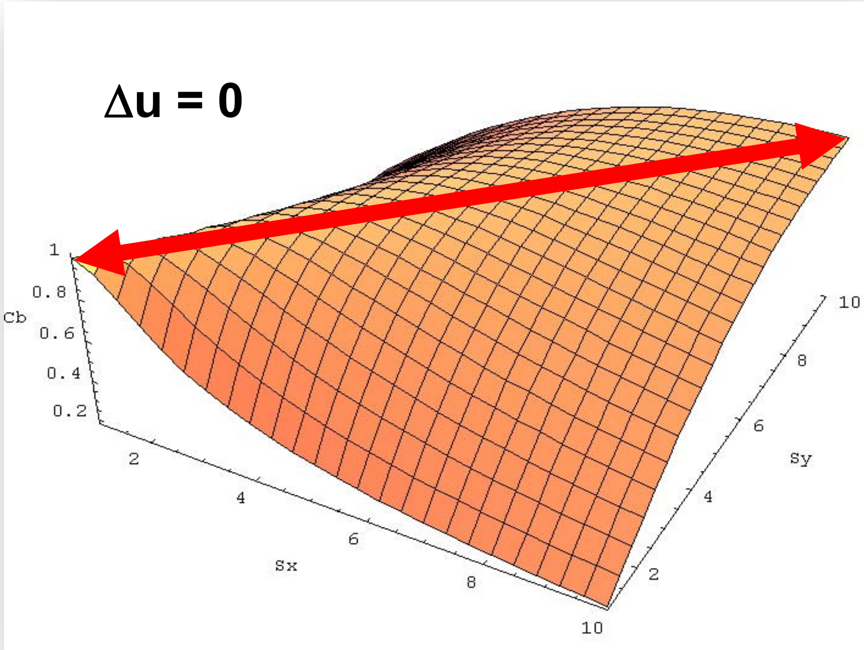
HOW IS CCC COMPUTED?

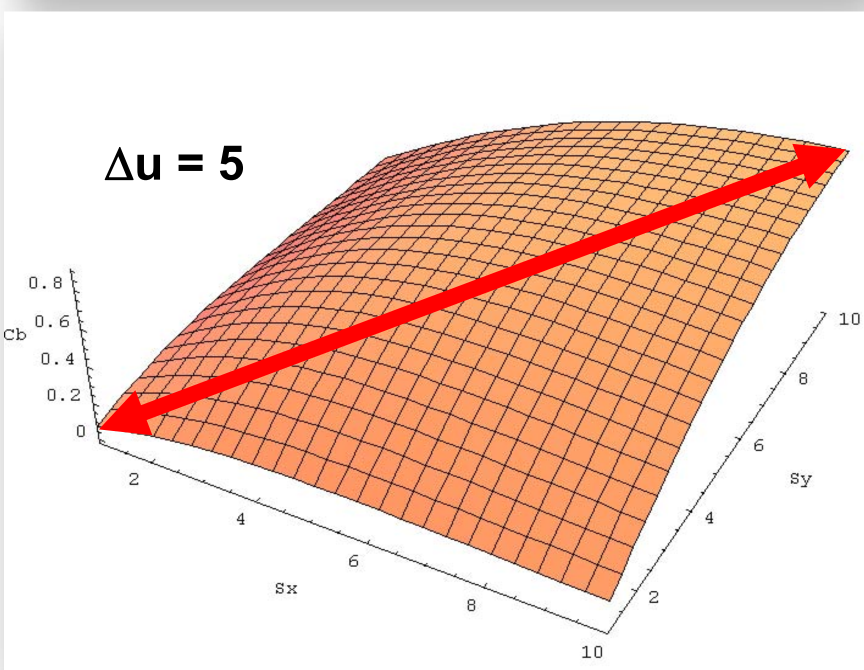
$$\hat{\rho}_c = \frac{2 \times s_{f(X_1,\ldots,X_p)Y}}{s_Y^2 + s_{f(X_1,\ldots,X_p)}^2 + (\bar{Y} - \bar{f}(X_1,\ldots,X_p))^2}$$

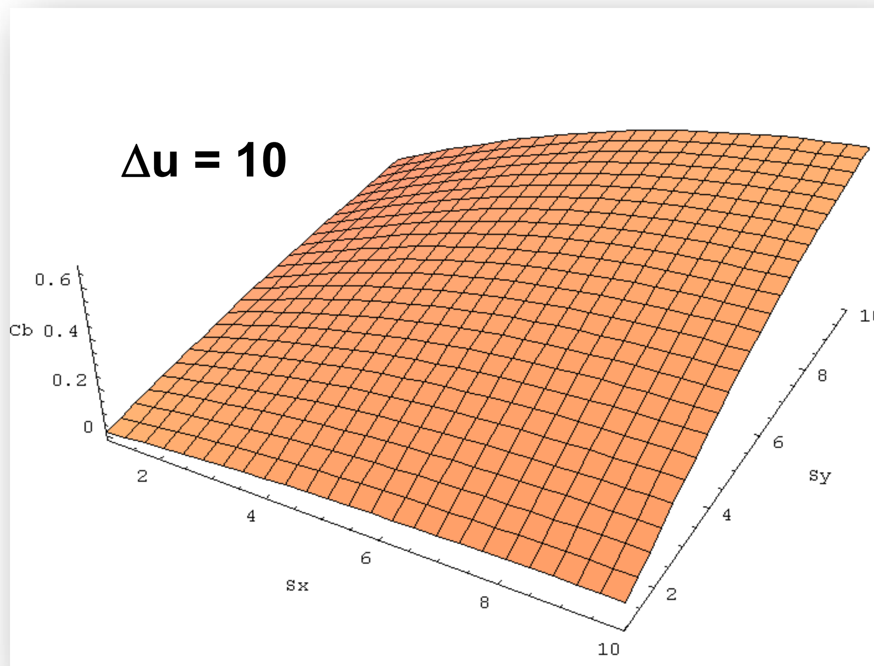
$$\hat{\rho}_c = \hat{\rho} \times C_b \quad \text{and} \quad C_b = \frac{2}{\left[v + \frac{1}{v} + \mu^2 \right]}$$

$$v = \begin{cases} = \dfrac{\sigma_1}{\sigma_2} & \text{for population} \\ = \dfrac{s_Y}{s_{f(X_1,\ldots,X_p)}} & \text{for sample} \end{cases} \quad \text{and} \quad \mu = \begin{cases} = \dfrac{\mu_1 - \mu_2}{\sqrt{\sigma_1 \sigma_2}} & \text{for population} \\ = \dfrac{\bar{Y} - \bar{f}(X_1,\ldots,X_p)}{\sqrt{s_Y s_{f(X_1,\ldots,X_p)}}} & \text{for sample} \end{cases}$$

Effects of Δu and Δv on Accuracy (Cb)





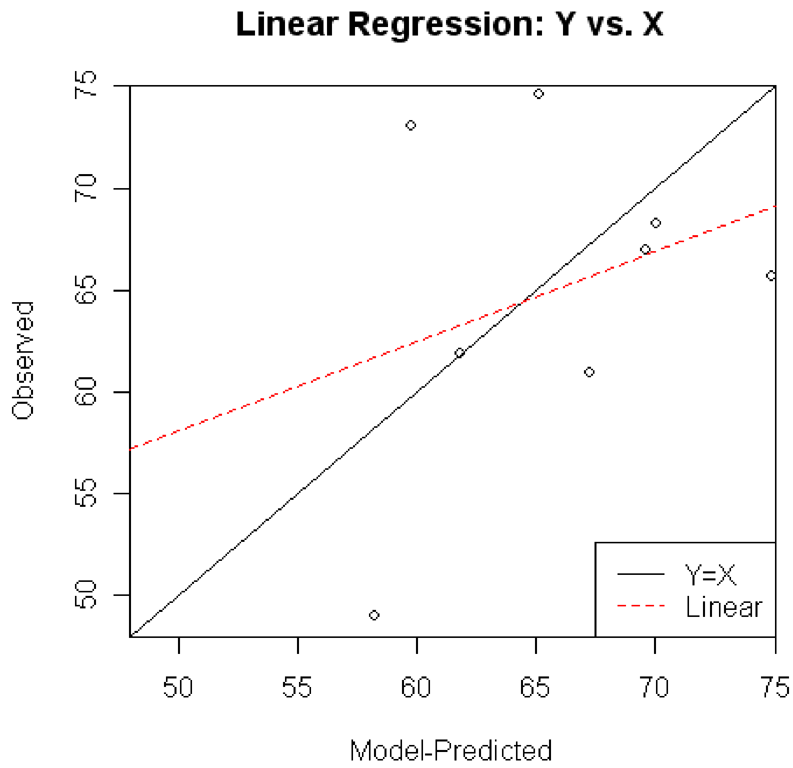


Example – Moment Statistics

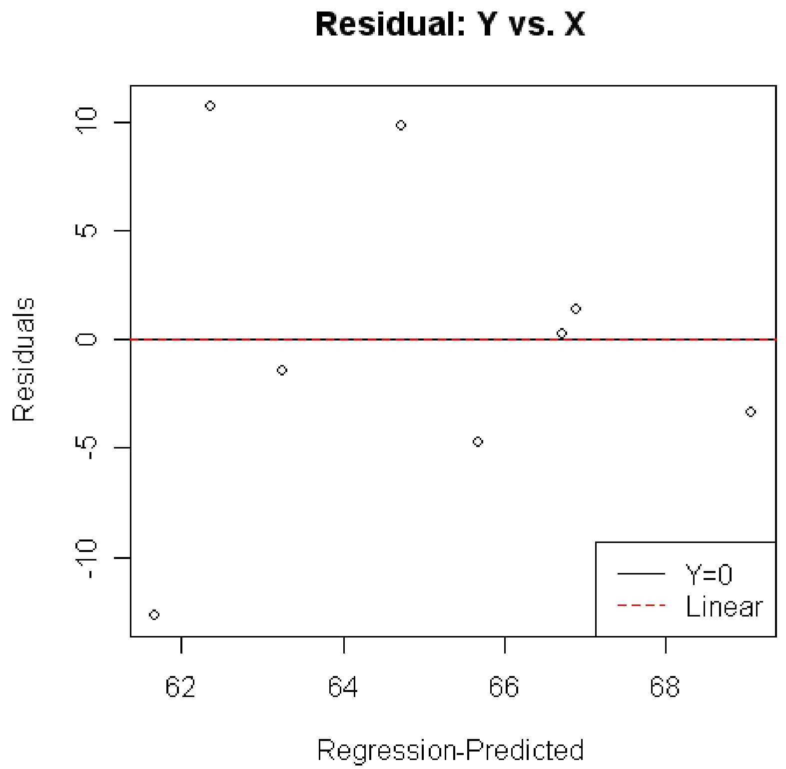
X Abs Mean	4.59500	Y Abs Mean	5.83031
X Min	58.24000	Y Min	49.00000
X Max	74.90000	Y Max	74.56000
X Mean	65.84750	Y Mean	65.04375
X Median	66.20000	Y Median	66.32000
X Variance	32.34248	Y Variance	64.80948
X Std. Dev.	5.68704	Y Std. Dev.	8.05043
X Skewness	0.09757	Y Skewness	-0.67381
X Kurtosis	1.45303	Y Kurtosis	2.35381
X - Y Mean	0.80375	X - Y Var	68.47100
Covariance	12.54792		

- **Cb = 0.94**

Example

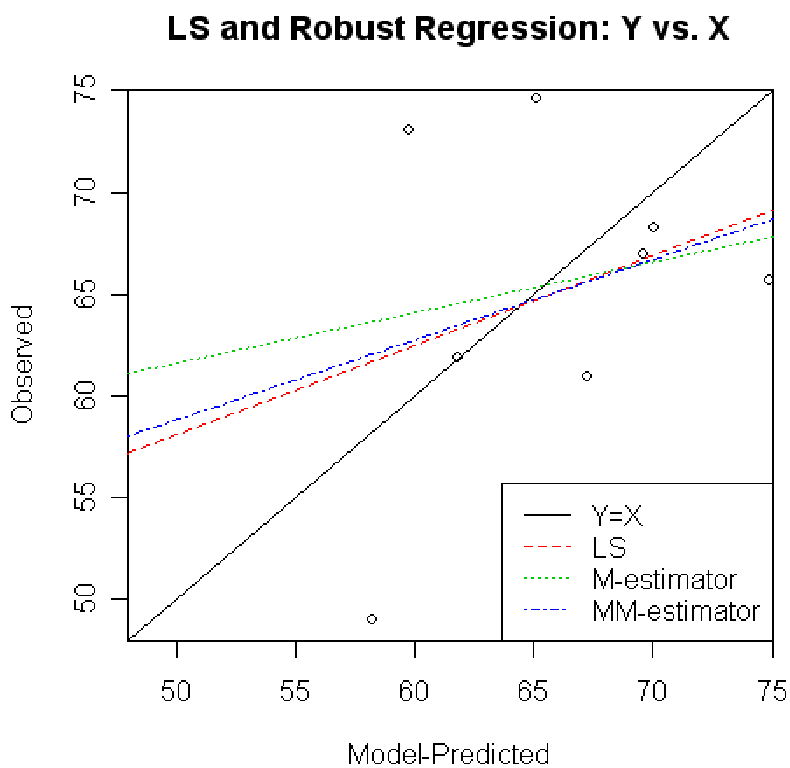


Linear Regression: Y vs. X
Observed
Model-Predicted
50
55
60
65
70
75
Y=X
Linear



Residual: Y vs. X
Residuals
Regression-Predicted
62
64
66
68
10
5
0
-5
-10
Y=0
Linear

Example – Robust regression



Limitations of CCC

Assumes that each pair of data point are interchangeable, that means, the order of the data point does not matter; there is no covariance

Nickerson (1997) suggested an adaptation to the CCC

Tedeschi (2006: 89:225 Ag. Syst.)

An improved CCC estimate

- CCC uses squared perpendicular distance $(Y_1 - Y_2)^2$ of any paired data point to the unity line
- Unfortunately, it measures only how close the data point is to the unity line and not which direction it goes

Tedeschi (2006: 89:225 Ag. Syst.)

An improved CCC estimate

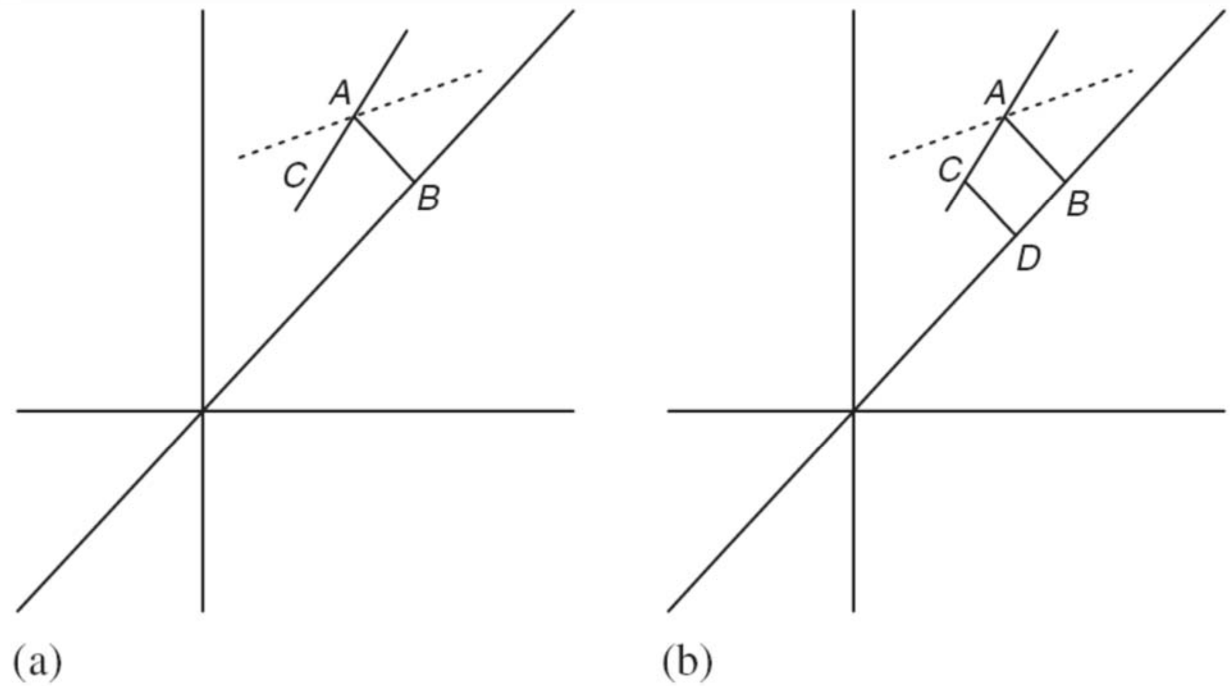


Figure 1. Comparison of the two criteria: (a) Lin's criteria; (b) new criterion.

Liao (2003)

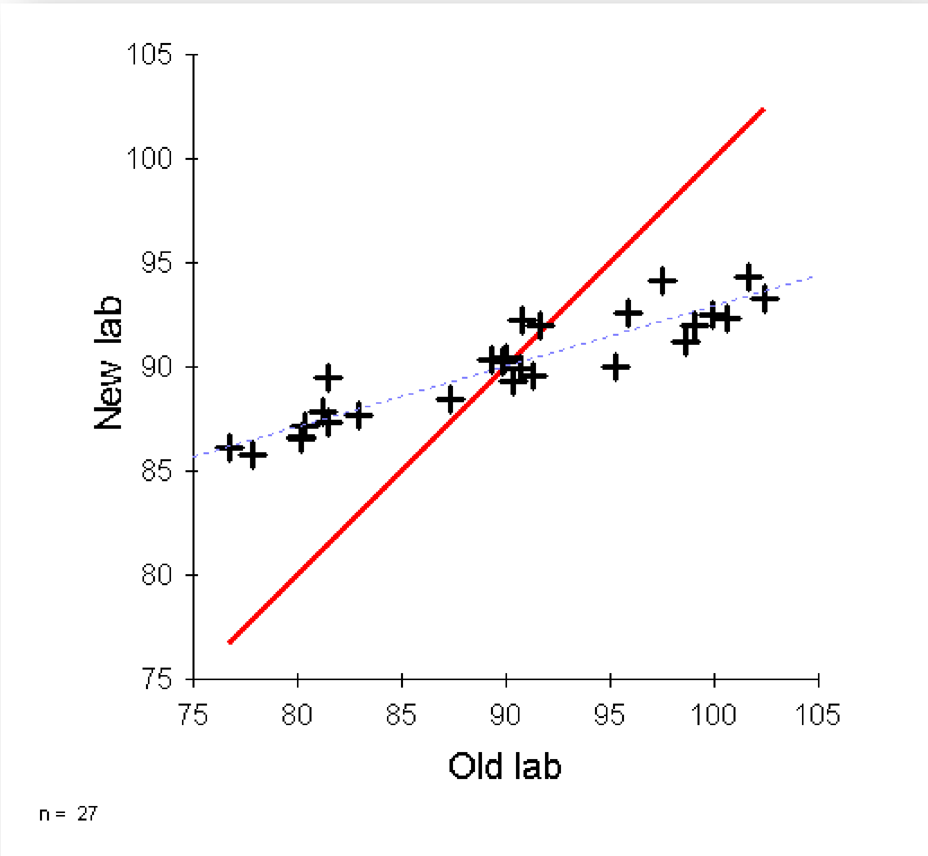
AN IMPROVED CCC ESTIMATE

- It is a quadratic area function of ρ whereas in Lin's it is quadratic distance function of ρ
- Accuracy (A_ρ) includes ρ whereas in Lin's (C_b) it does not

$$A_\rho = \frac{4\times\left(\frac{s_{f(X_1,\ldots,X_p)}}{s_Y}\right) - \hat{\rho}\times\left[1+\left(\frac{s_{f(X_1,\ldots,X_p)}}{s_Y}\right)^2\right]}{\left(2-\hat{\rho}\right)\times\left[1+\left(\frac{s_{f(X_1,\ldots,X_p)}}{s_Y}\right)^2\right]\times\left(\frac{\overline{Y}-\overline{f}(X_1,\ldots,X_p)}{s_Y}\right)^2}$$

$$\hat{\gamma}_\rho = \hat{\rho}\times A_\rho$$

Comparison Lin's x Liao's CCC



- Lin's CCC
  - $C_b = 0.571$
  - $r_c = 0.527$
- Liao's CCC
  - $A_r\ (C_b) = 0.205$
  - $G_r\ (r_c) = 0.189$
- Chinchilli's CCC
  - $GCCC_w = 0.179$

Liao (2003)

DIVERSE EVALUATION MEASUREMENTS

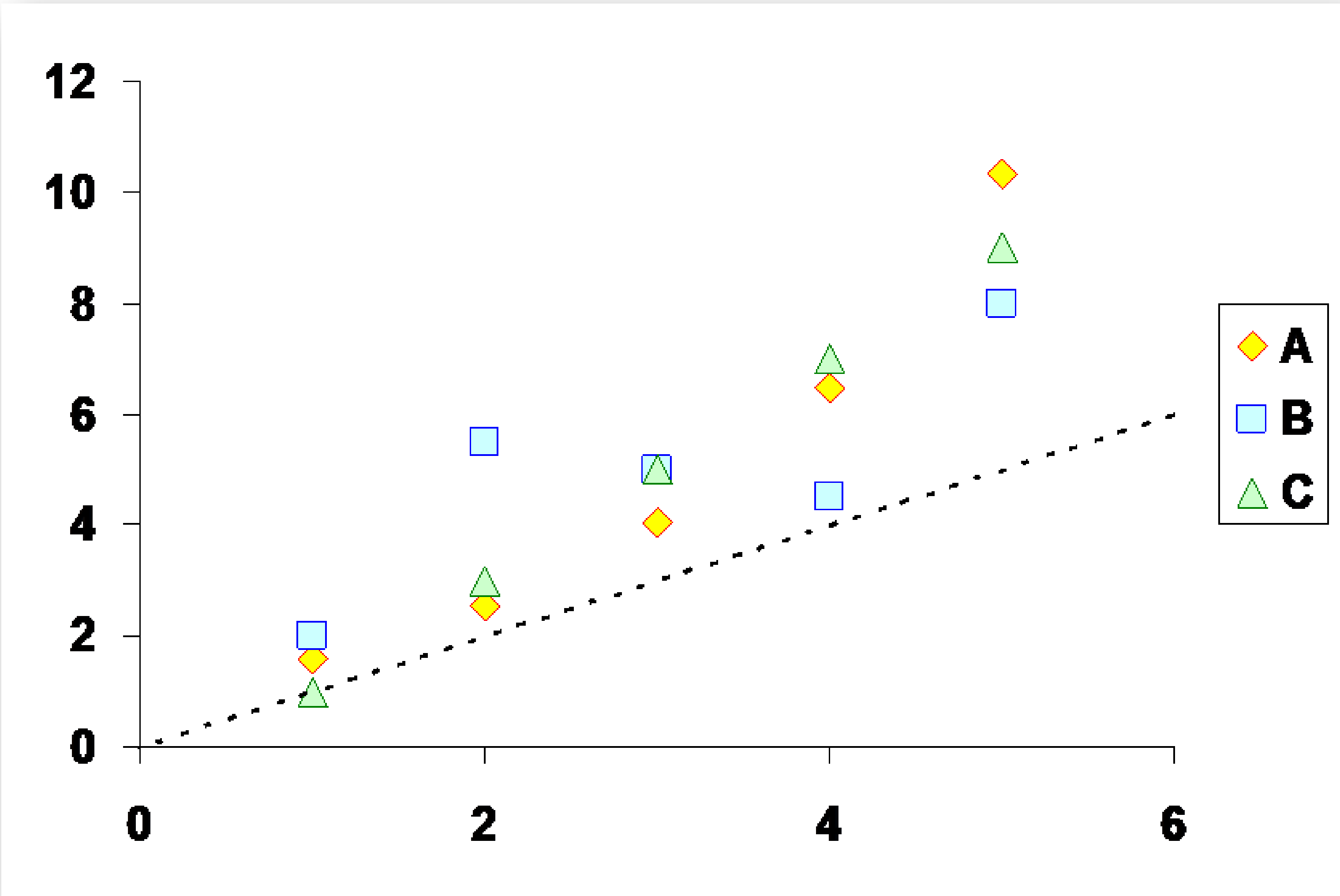
MEAN BIAS (MB)

- Likely to be the oldest and most used statistic to assess model accuracy
- Likely to be the weakest one too !

$$MB = \frac{\sum_{i=1}^{n} (Y_i - f(X_1, \ldots, X_p)_i)}{n}$$

$$t_{MB} = \frac{MB}{\sqrt{\frac{\sum_{i=1}^{n} \left(\left(Y_i - f(X_1, \ldots, X_p)_i \right) - MB \right)^2}{n \times (n-1)}}}$$

Which model has the lowest MB?



- All models (A, B, and C) have the same MB = 2
- *t*-test for Model A (exponential)
  - Assuming $\sigma_1 = \sigma_2$: $P = 0.29$
  - Assuming $\sigma_1 \neq \sigma_2$: $P = 0.28$
  - Assuming covariance: $P = 0.09$
- *t*-test for Model B
  - Assuming $\sigma_1 = \sigma_2$: $P = 0.14$
  - Assuming $\sigma_1 \neq \sigma_2$: $P = 0.13$
  - Assuming covariance: $P = 0.02$
- *t*-test for Model C (linear)
  - Assuming $\sigma_1 = \sigma_2$: $P = 0.25$
  - Assuming $\sigma_1 \neq \sigma_2$: $P = 0.24$
  - Assuming covariance: $P = 0.05$

Mean bias

- Must be adjusted for covariance!
- Rejection rates of the H_0 hypothesis increases as correlated errors increase
- Cannot be used as the main statistics for model evaluation

Tedeschi (2006: 89:225 Ag. Syst.)

RESISTANT R^2

- Resistant means it is insensible to outliers or extreme points
- Uses the median instead of mean

$$r_r^2 = 1 - \left(\frac{\underset{i=1}{\overset{n}{\mathrm{M}}}\left(\left|Y_i - \hat{Y}_i\right|\right)}{\underset{i=1}{\overset{n}{\mathrm{M}}}\left(\left|Y_i - \overline{Y}\right|\right)} \right)^2$$

MODELING EFFICIENCY (MEF)

- Proportion of variation explained by the line Y = f(X1,…,Xp); range [-∞ to 1]; MEF = 1 is better

$$MEF = \frac{\left(\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2 - \sum_{i=1}^{n}\left(Y_i - f(X_1,\ldots,X_p)_i\right)^2\right)}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2} = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - f(X_1,\ldots,X_p)_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2}$$

$$r = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2} = \frac{s_{f(X_1,\ldots,X_p)Y}}{s_Y \times s_{f(X_1,\ldots,X_p)}}$$

NASH-SUTCLIFFE MODEL EFFICIENCY COEFFICIENT (= MEF)

- Nash–Sutcliffe efficiency can range from −∞ to 1
  - *NSE* = 1 ➔ corresponds to a perfect match of modeled values to the observed data
  - *NSE* = 0 ➔ indicates that the model predictions are as accurate as the mean of the observed data
  - *NSE* < 0 ➔ observed mean is a better predictor than the model or, in other words, when the residual variance (described by the numerator in the expression above), is larger than the data variance (described by the denominator).

$$NSE = 1 - \frac{\sum_{t=1}^{T} \left(Q_m^t - Q_o^t\right)^2}{\sum_{t=1}^{T} \left(Q_o^t - \overline{Q_o}\right)^2}$$

Nash and Sutcliffe (1970)

COEFFICIENT OF DETERMINATION (CD)

- Ratio of total variance of observed data to the squared of the difference between model-predicted and mean of observed
- It is the proportion of the total variance of the observed values explained by the predicted data
- CD = 1 is better

$$CD = \frac{\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}{\sum_{i=1}^{n}\left(f(X_1,\ldots,X_p)_i - \bar{Y}\right)^2}$$

Mean Square Error of Prediction (MSEP)

MSE x MSEP

MSE assesses the precision of the fitted linear regression using the difference between observed and regression-predicted values

MSEP consists the difference between observed and model-predicted values

Tedeschi (2006: 89:225 Ag. Syst.)

MSEP X MSE

$$MSE = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n-2}$$

$$MSEP = \frac{\sum_{i=1}^{n}\left(Y_i - f(X_1,\ldots,X_p)_i\right)^2}{n}$$

Limitations of MSEP

- Removes the negative sign
- Weights the deviation by their squares, thus giving more influence to larger data points
- Does not provide information about model precision

Tedeschi (2006: 89:225 Ag. Syst.)

DECOMPOSITION OF MSEP

- Expanded MSEP equation and solved for known linear measures of linear regression (Theil, 1961)

$$MSEP = \frac{\sum_{i=1}^{n}\left(Y_i - f(X_1,\ldots,X_p)_i\right)^2}{n}$$

$$MSEP = \frac{\sum_{i=1}^{n}\left[\left(\overline{f}(X_1,\ldots,X_p)-\overline{Y}\right)+\left(f(X_1,\ldots,X_p)_i-\overline{f}(X_1,\ldots,X_p)\right)-\left(Y_i-\overline{Y}\right)\right]^2}{n}$$

$$MSEP = \left(\overline{f}(X_1,\ldots,X_p)-\overline{Y}\right)^2 + s^2_{f(X_1,\ldots,X_p)} + s^2_Y - 2\cdot r\cdot s_{f(X_1,\ldots,X_p)}\cdot s_Y$$

Understanding MSEP

$$MSEP_1 = \left(\bar{f}(X_1,...,X_p) - \bar{Y}\right)^2 + \left(s_{f(X_1,...,X_p)} - s_Y\right)^2 + 2\cdot(1-r)\cdot s_{f(X_1,...,X_p)}\cdot s_Y$$

$$MSEP_2 = \left(\bar{f}(X_1,...,X_p) - \bar{Y}\right)^2 + \left(s_{f(X_1,...,X_p)} - r\cdot s_Y\right)^2 + (1-r^2)\cdot s_Y^2$$

$$MSEP_3 = \left(\bar{f}(X_1,...,X_p) - \bar{Y}\right)^2 + s^2_{f(X_1,...,X_p)}\cdot(1-b)^2 + (1-r^2)\cdot s_Y^2$$

Inequality Proportions	Equations	Descriptions
U^M	$\left(\bar{f}(X_1,...,X_p) - \bar{Y}\right)^2 / MSEP$	Mean bias
U^S	$\left(s_{f(X_1,...,X_p)} - s_Y\right)^2 / MSEP$	Unequal variances
U^C	$2\times(1-r)\times s_{f(X_1,...,X_p)}\times s_Y / MSEP$	Incomplete (co)variation
U^R	$s^2_{f(X_1,...,X_p)}\times(1-b)^2 / MSEP$	Systematic or slope bias
U^D	$\left(1-r^2\right)\times s_Y^2 / MSEP$	Random errors

[a] Note that $U^M + U^S + U^C = U^M + U^R + U^D = 1$

Tedeschi (2006: 89:225 Ag. Syst.)

Understanding MSEP

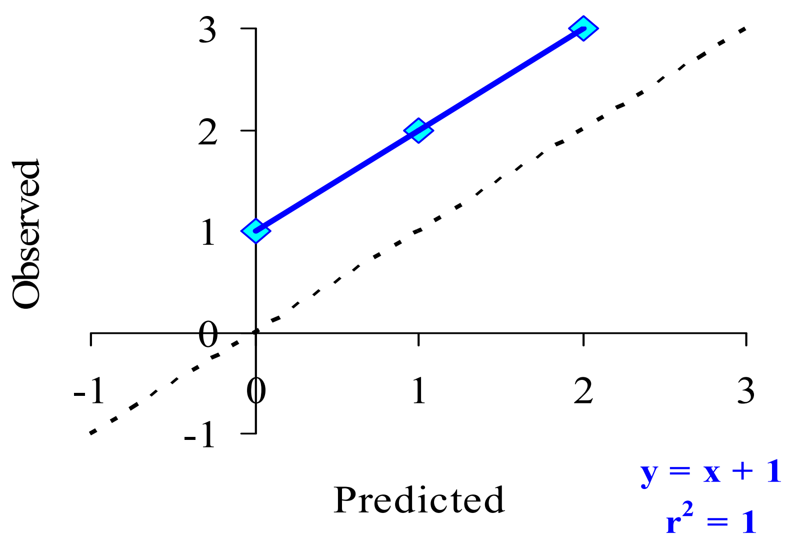
Mean bias indicate the error in central tendency

Systematic bias indicate how much the regression deviates from Y = X line, that means, errors due to regression

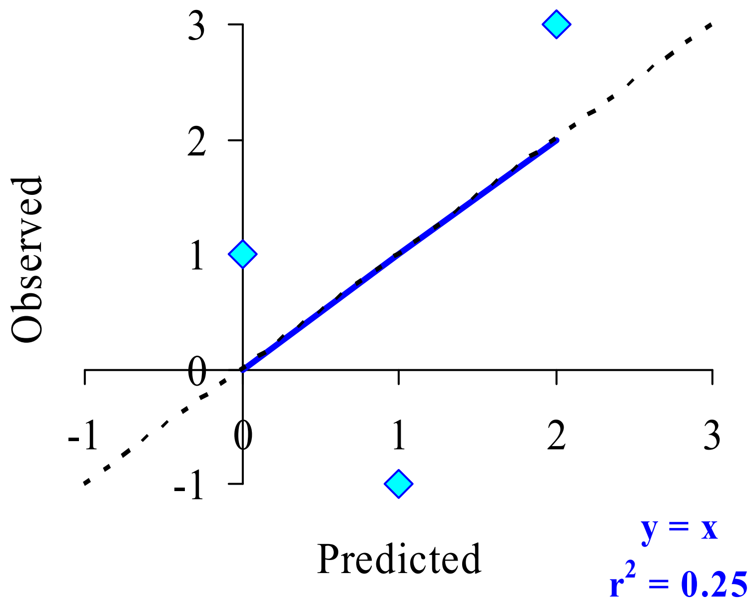
Random errors indicate the unexplained variation that cannot be accounted for by the relationship

Case 2 - ↑ Precision ↓ Accuracy

- MSEP = 1
- Mean Bias = 100%



Case 3 - ↓ Precision ↑ Accuracy

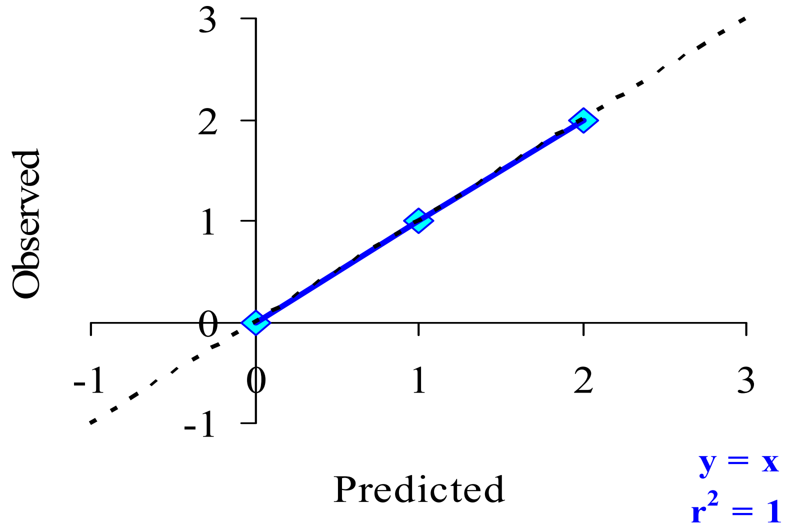


- MSEP = 2
- Mean bias = 0
- Slope bias = 0
- Random = 100%

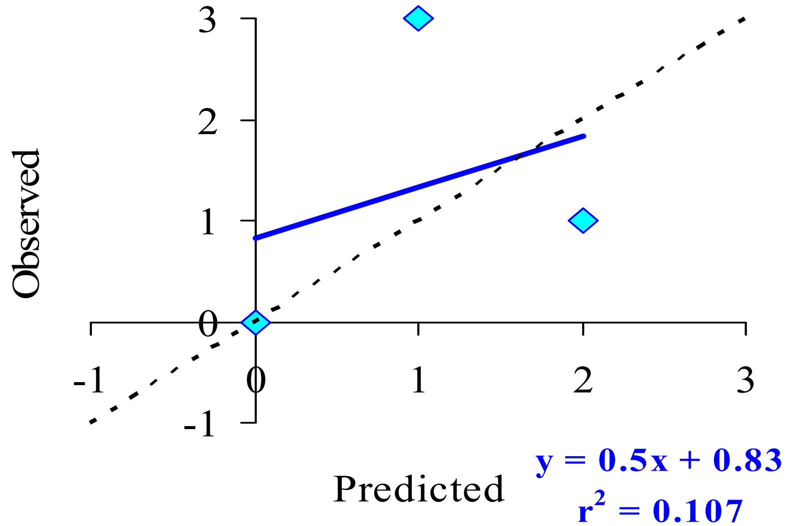
- Unequal variance = 33.3%
- Incomplete (co)variation = 66.7%

Case 4 - ↑ Precision ↑ Accuracy

- MSEP = 0



Case 1 - ↓ Precision ↓ Accuracy



- MSEP = 1.67
- Mean bias = 6.7%
- Slope bias = 10%
- Random = 83.3%

- Unequal variances = 11.1%
- Incomplete (co)variation = 82.2%

Nonparametric Analysis

Why nonparametric?

- One might be interested in the comparison of the ranking of real-observed values versus those predicted by models
  - Bull's EPD for efficiency
- More resilient to abnormalities of the data
  - Outliers and influential points

NONPARAMETRIC TESTS

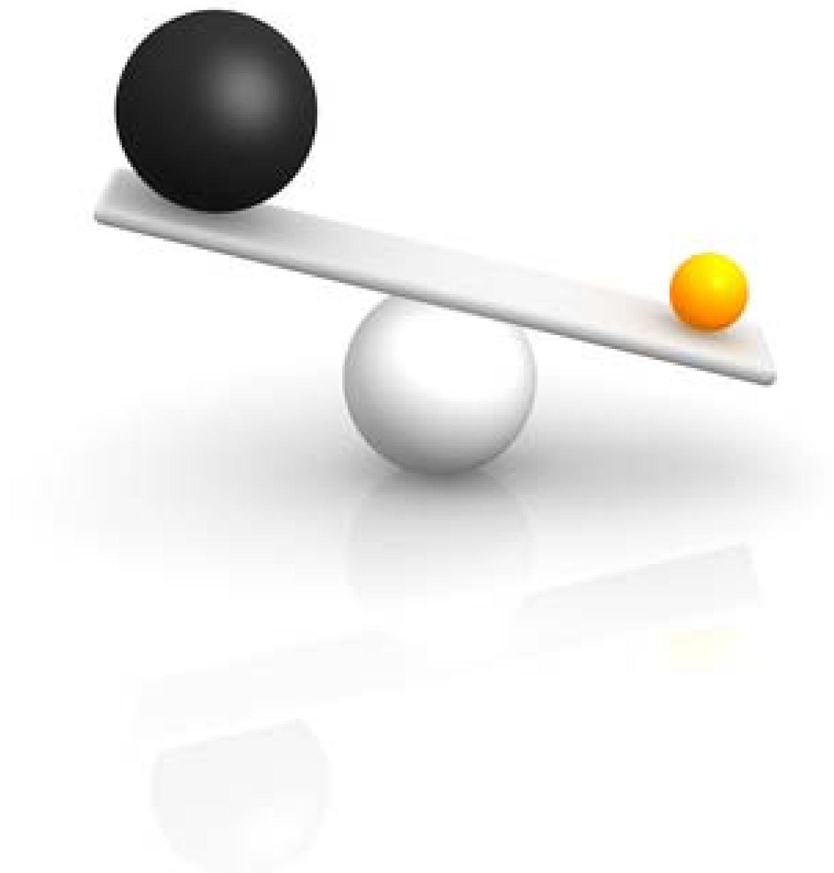
- Spearman's r_s is the linear correlation coefficient of the ranks

$$r_S = \frac{\sum_{i=1}^{n}\left(R_i - \overline{R}\right)\left(S_i - \overline{S}\right)}{\sqrt{\sum_{i=1}^{n}\left(R_i - \overline{R}\right)^2}\sqrt{\sum_{i=1}^{n}\left(S_i - \overline{S}\right)^2}}$$

- Kendall's τ measures the ordinal concordance of ½·n· (n-1) data points where a data point cannot be paired with itself

$$\tau = \frac{Concordant - Discordant}{\sqrt{Concordant + Discordant - ExtraY} \times \sqrt{Concordant + Discordant - ExtraX}}$$

Balance Analysis



Tedeschi (2006: 89:225 Ag. Syst.)

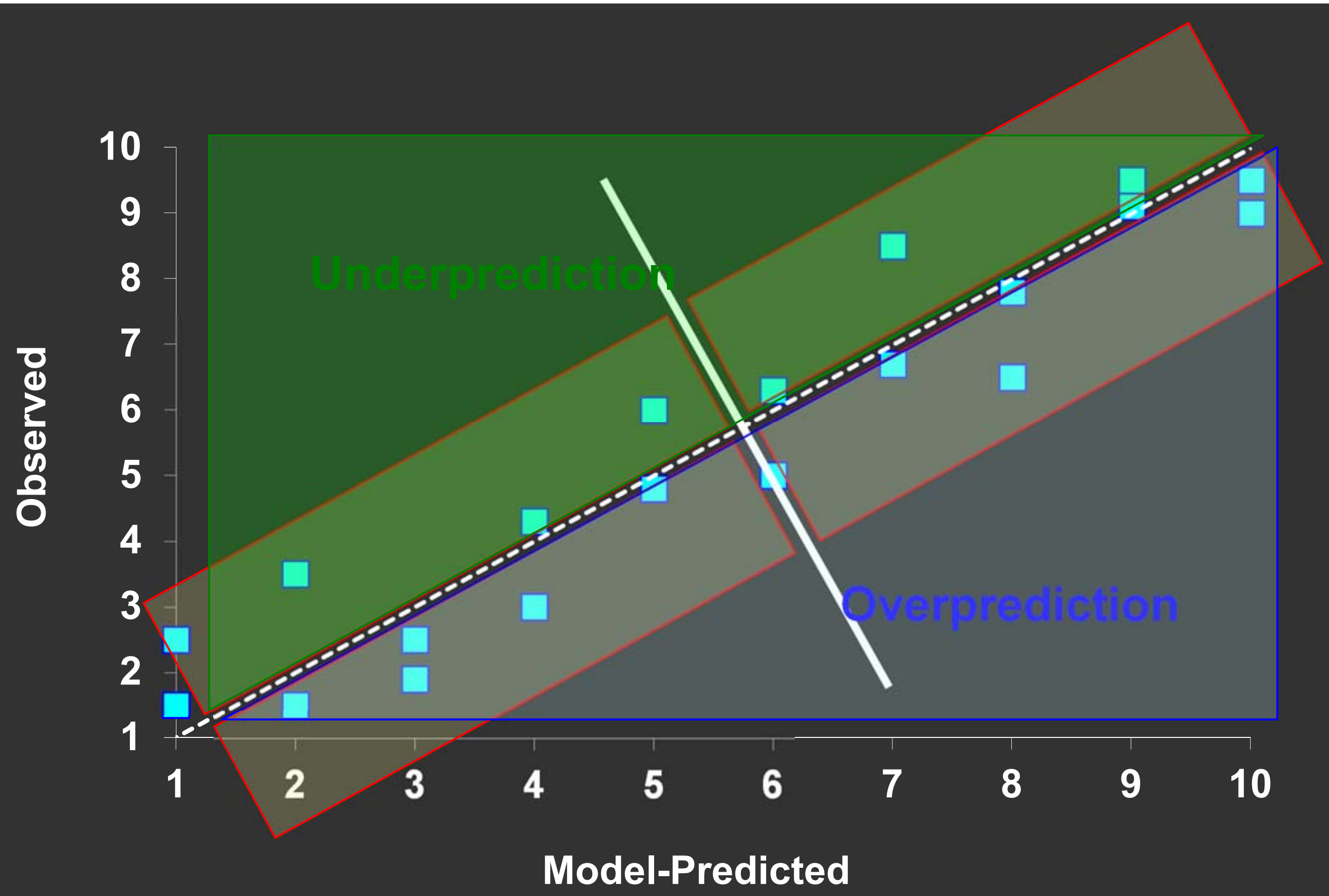
BALANCE ANALYSIS

- Evaluates the balance of number of data points under- and overpredicted by the model above and below the observed and model-predicted mean, orthogonal or not to the regression line.

Model prediction	Observed or Model-Predicted Mean	
	Below	Above
Overpredicted	n_{11}	n_{12}
Underpredicted	n_{21}	n_{22}

Tedeschi (2006: 89:225 Ag. Syst.)

BALANCE ANALYSIS



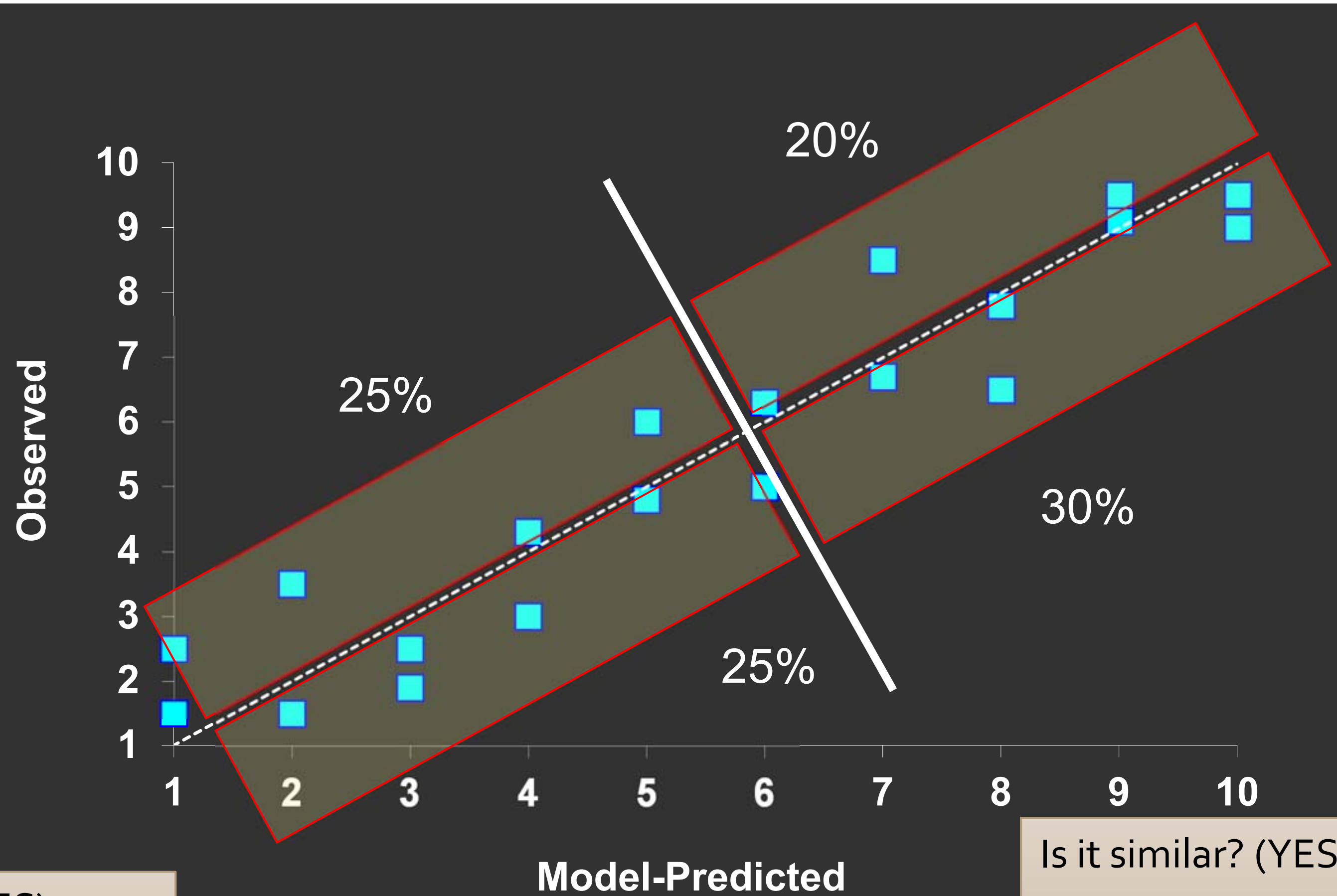
Tedeschi (2006: 89:225 Ag. Syst.)

What do we check in the Balance Analysis?

- Is the trend of under- or over-prediction similar?
- Is it similar below and above the mean?
- Use χ^2 analysis to check if the number of points is not different
  - Check if they are 25%
  - Check if the distribution is similar

Tedeschi (2006: 89:225 Ag. Syst.)

BALANCED ANALYSIS



Is it 25%? (YES)

$\chi^2 = 0.4$ (P = 0.53)

Tedeschi (2006: 89:225 Ag. Syst.)

Is it similar? (YES)

$\chi^2 = 0.2$ (P = 0.65)

Odds ratio = 0.69 (P = 0.67)

Data Distribution

Data Distribution

Stochastic models

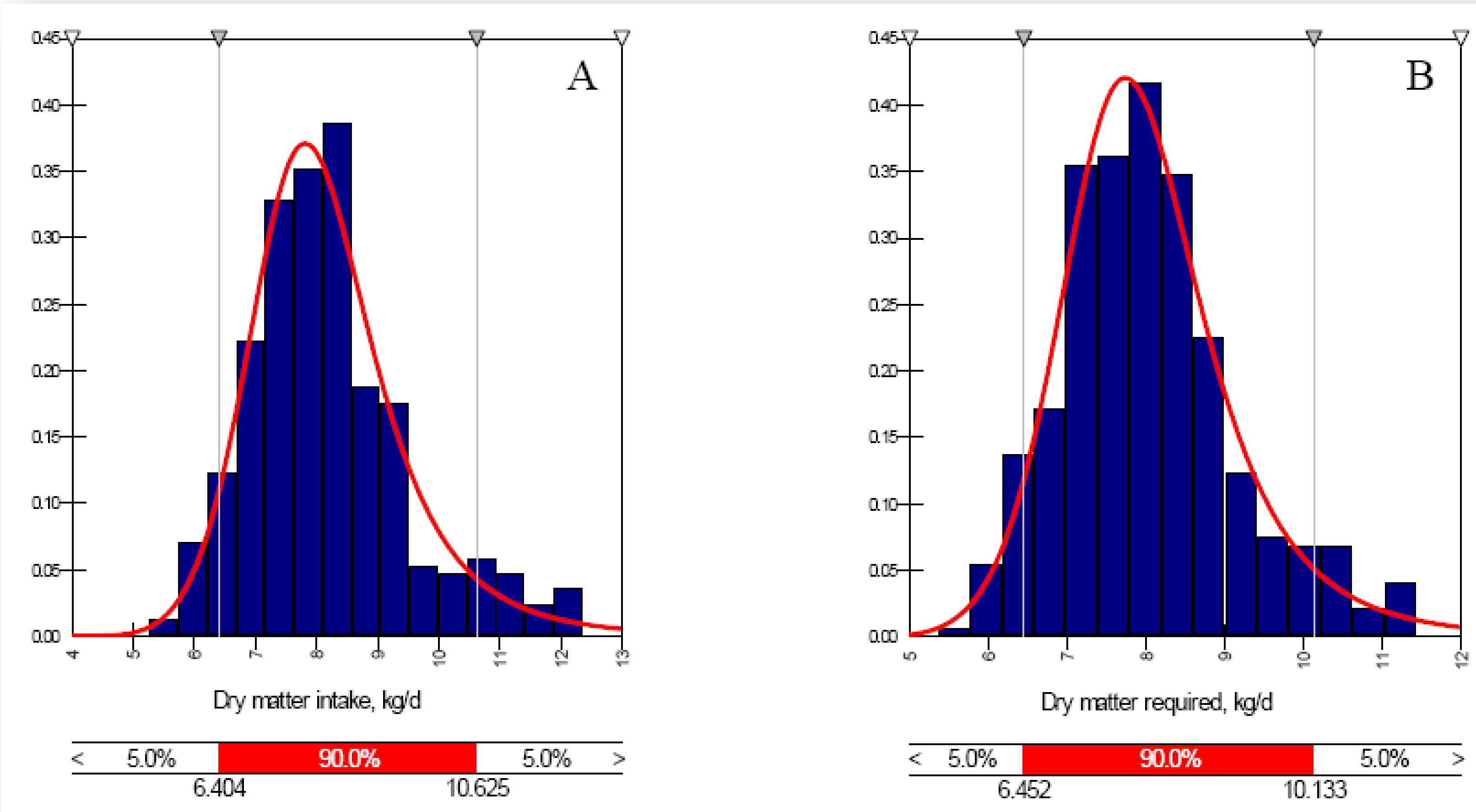
-- Reynolds and Deaton (1982)

Deterministic models

-- Dent and Blackie (1979)

The idea is to check if observed and predicted values come from the same distribution

Data Distribution



Summary (1/2)

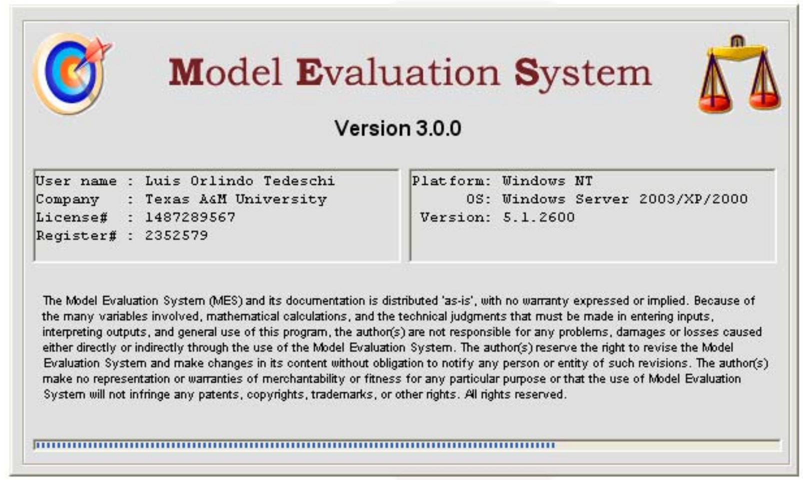
- Acceptance of model wrongness is important to ensure that more reliable and accurate models are developed
- The usefulness of a model depends on the purpose it was developed for

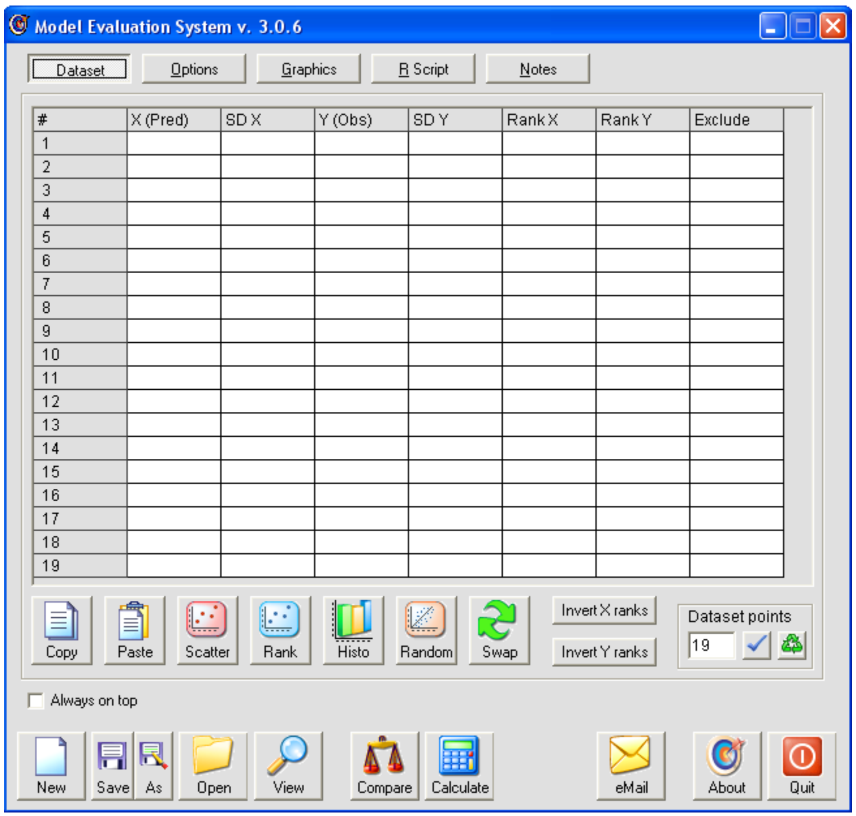
Summary (2/2)

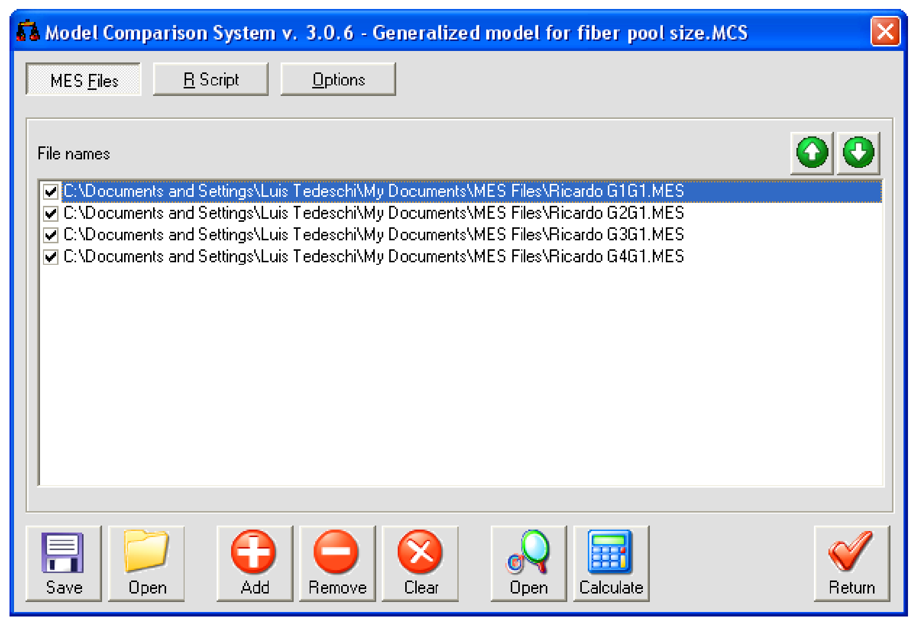
- High accuracy and high precision of a model for a given database implies NOTHING regarding future predictions of the model

- Model evaluation MUST be assessed using several statistical techniques; each technique measures different characteristics of the model

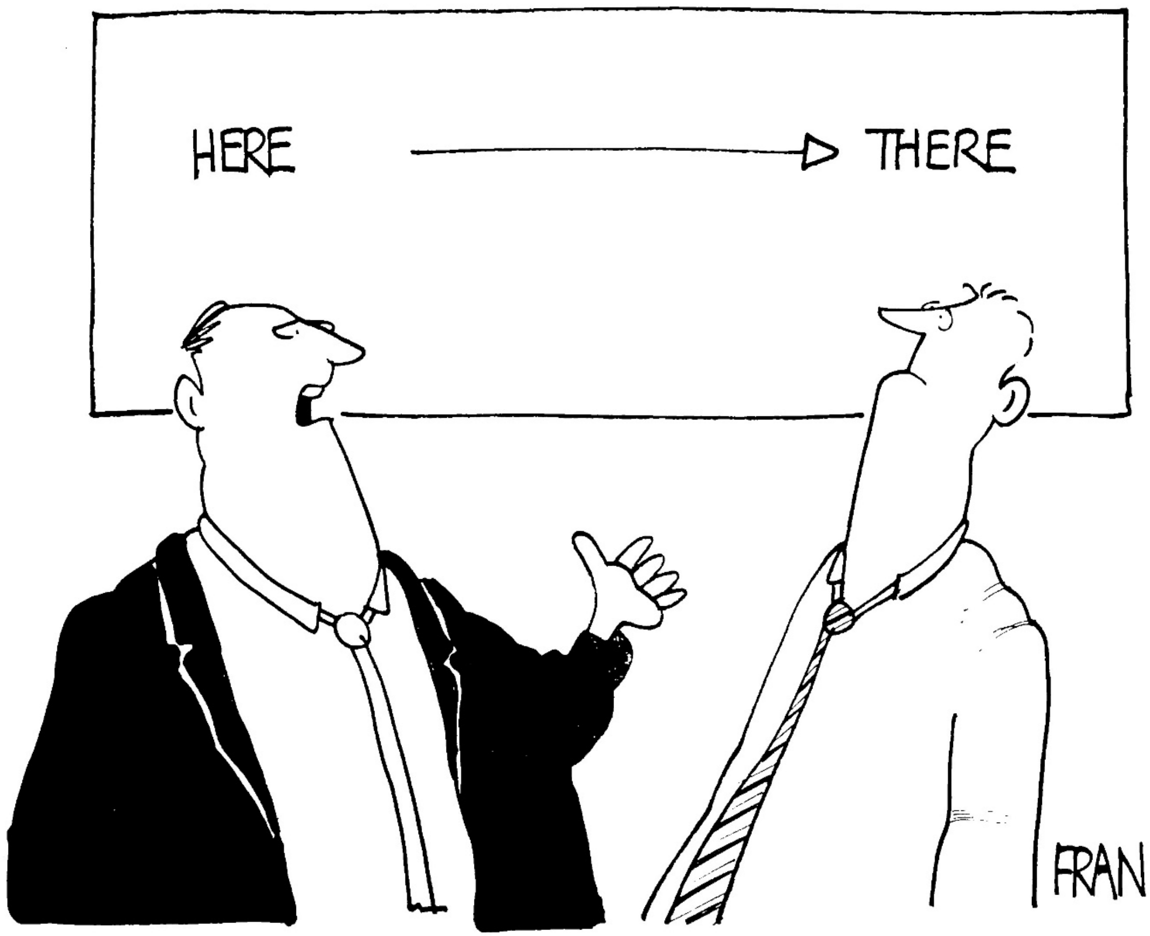
A handy assistant …







http://nutritionmodels.tamu.edu OR http://www.nutritionmodels.com/mes.html



"It's a simple model... but it works for me..."

TEXAS A&M UNIVERSITY

TEXAS A&M AGRILIFE RESEARCH

National Animal Nutrition Program
Leveraging Resources, Linking Researchers