Incorporating Measurement Uncertainty in Water Quality Monitoring and Model Evaluation

Daren Harmel – USDA-ARS, Temple, TX

USDA

Agricultural Research Service

Objectives

- **Objective #1 - Determine Measurement Uncertainty in H/WQ Data**
  - **Briefly describe the Data Uncertainty Estimation Tool for Hydrology and Water Quality (DUET-H/WQ)**
- **Objective #2 - Incorporate Uncertainty in Model Evaluation**
  - **Briefly describe modified "goodness-of-fit" indicators that incorporate measurement uncertainty in model calibration and validation**

USDA

Agricultural Research Service

Determine Uncertainty – INTRODUCTION

"Should it not be required that every… (field and modeling study) …attempt to evaluate the uncertainty in the results?" - Beven (2006)

"The use of uncertainty estimation… (should be)…routine in hydrological and hydraulic science." - Pappenberger and Beven (2006)

- **All measurements are inherently uncertain; so why is uncertainty typically ignored?? Until recently…**
  - **An adequate understanding of measurement uncertainty had not been established**
  - **No complete uncertainty (error propagation) analysis had been conducted on measured H/WQ data**
  - **No easy-to-use tool was available to assist with uncertainty estimation**

Determine Uncertainty - METHODS

- In 2006, published uncertainty estimation framework (Harmel et al. 2006)
  - Developed for small watersheds
  - Focused on streamflow and water quality data (TSS, N, P)
  - Provided published uncertainty estimates for data collection steps within each of four procedural categories
    - discharge measurement, sample collection, sample preservation/storage, laboratory analysis
- Used to create Data Uncertainty Estimation Tool or "DUET-H/WQ"
  - Added another procedural category - data processing and management

USDA

Agricultural Research Service

DUET-H/WQ - LookUp Table for calculation of uncertainty in sample collection

Click to select sample type

- Baseflow
- Storm
- Other

Sample Collection Method - Constituent Type	Storm Sample Collection Uncertainty	Reference
Automated sampling (single intake) - dissolved (P)	Median +0% to +0%	Martin et al. (1992)
Automated sampling (single intake) - dissolved (P)	±1 to 38%	Ging (1999)
Automated sampling (single intake) - dissolved (PO4-P)	±0 to 52%	Ging (1999)
Automated sampling (single intake) - total (N)	Median +0% to +0%	Martin et al. (1992)
Automated sampling (single intake) - total (N)	±0 to 19%	Ging (1999)
Automated sampling (single intake) - total (P)	Median +0% to +17%	Martin et al. (1992)
Automated sampling (single intake) - total (P)	+-0 to 44%	Ging (1999)

± 15 %
(Click to change)

Sampling Frequency (interval)		
Time-interval <= 120 minutes (5 minutes discrete - field scale)	±0% to 11%	Miller et al.(2000)
Time-interval <= 120 minutes (5 minutes discrete - field scale)	±0% to 18%	Miller et al.(2007)
Time-interval <= 120 minutes (5-15 minutes discrete)	-1% to +1%	King and Harmel (2003)
Time-interval <= 120 minutes (10 minutes discrete - field scale)	±1% to 26%	Miller et al.(2000)
Time-interval <= 120 minutes (10 minutes discrete - field scale)	±0% to 40%	Miller et al.(2007)
Time-interval <= 120 minutes (30 minutes discrete - field scale)	±3% to 42%	Miller et al.(2000)
Time-interval <= 120 minutes (30 minutes discrete - field scale)	±0% to 72%	Miller et al.(2007)

± 10 %
(Click to change)

Discrete or Composite Sampling		
Discrete samples assumed to represent the true value	0%	N/A
Composite - 3 composite time-interval samples per bottle (5-360 minutes)	±1 to 33%	King and Harmel (2003)
Composite - 6 composite time-interval samples per bottle (5-360 minutes)	±4 to 50%	King and Harmel (2003)
Composite - 6 composite time-interval samples per bottle (60 minutes)	±0 to 19%	Miller et al. (2007)
Composite - 3 composite flow-interval samples per bottle (2.5-15mm)	±0 to 5%	King and Harmel (2003)
Composite - 6 composite flow-interval samples per bottle (2.5-15mm)	±0 to 8%	King and Harmel (2003)
Composite - up to 6 composite flow-interval samples per bottle (1.32, 2.64, and 5.28mm)	Median ±0.4% (avg 9.8%)	Harmel and King (2005)
OTHER -	---	N/A

± 0 %
(Click to change)

Minimum Storm Threshold (storm sampling only)		
"Low" threshold, disgregard flow and concentration outside storm	±1 to 5% (median = 3%)	Based on Harmel et al. (2002)
"High" threshold, disgregard flow and concentration outside storm	±9 to 81% (median = 35%)	Based on Harmel et al. (2002)
"Low" threshold, extrapolate flow and concentration outside storm	±2%	Based on Harmel et al. (2002)
"High" threshold, extrapolate flow and concentration outside storm	±20%	Based on Harmel et al. (2002)
OTHER -	---	N/A

± 5 %
(Click to change)

Cancel

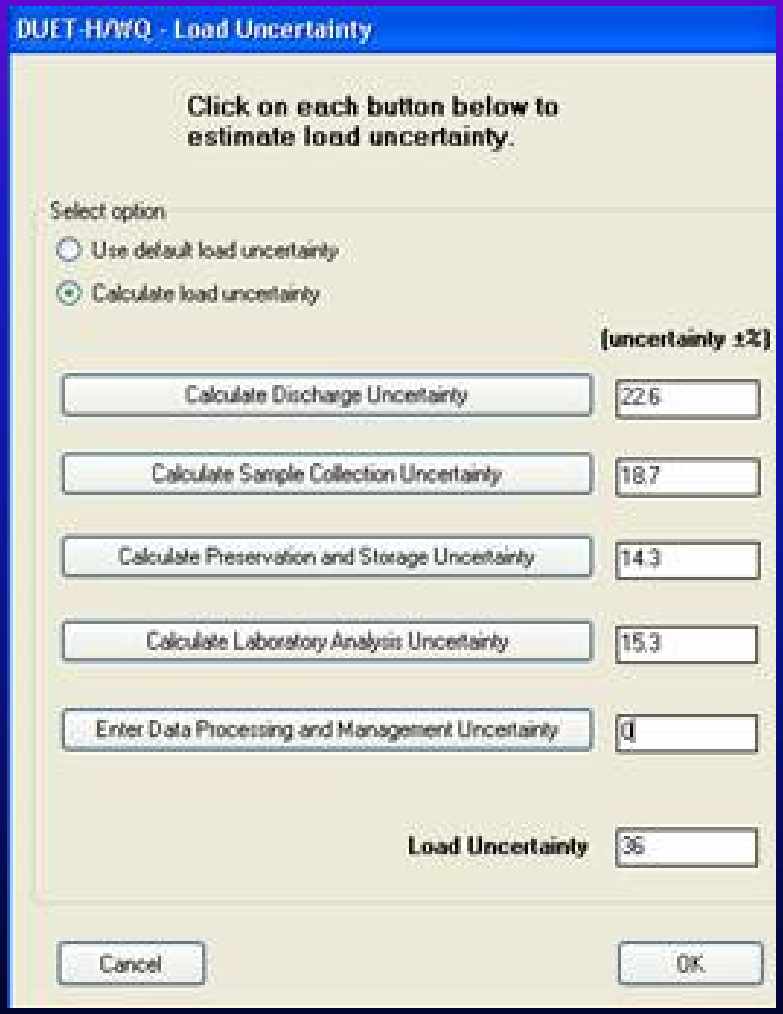
OK

Cumulative uncertainty in sample collection ± 18.7 %

Determine Uncertainty - METHODS

- Framework (and DUET-H/WQ) utilize the RMSE method to determine the "combined" uncertainty
  - within each procedural category
  - for individual measured values

$$EP = \sqrt{\sum \left(E_Q^{\ 2} + E_C^{\ 2} + E_{PS}^{\ 2} + E_A^{\ 2} + E_{DPM}^{\ 2} \right)}$$



Determine Uncertainty - METHODS

- Some have argued that the RMSE method is too “simple”

$$EP = \sqrt{\sum \left(E_Q^{\,2} + E_{\bullet}^{\,2} + E_{PS}^{\,2} + E_A^{\,2} + E_{DPM}^{\,2} \right)}$$

- Responses:
  - Certain value of simplicity is too often underappreciated in this age of rapidly advancing science and technology.
  - A simple method is beneficial (possibly required) for “data collectors” to commit added time and effort to estimate uncertainty for measured data.
  - RMSE method is certainly valid in the present application.

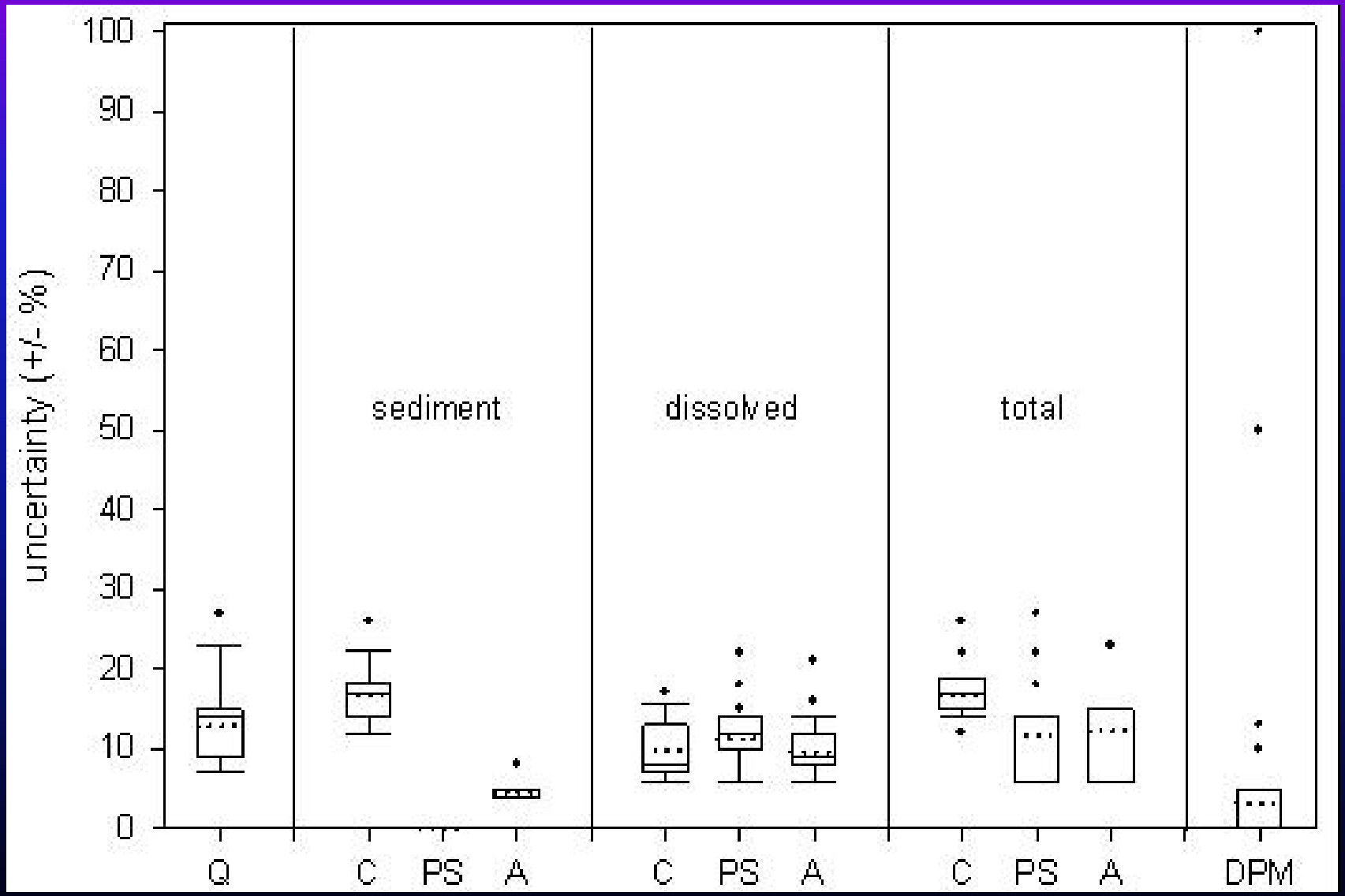
Determine Uncertainty - METHODS

- Applied DUET-H/WQ to estimate measurement uncertainty in real-world data
  - 131 storm events
  - Q, TSS, NO_3-N, PO_4-P, total N, and total P
  - Five monitoring projects
    - wide range of hydrologic settings, land uses, watershed sizes, and field and laboratory techniques

USDA

Agricultural Research Service

Determine Uncertainty - RESULTS



USDA

Agricultural Research Service

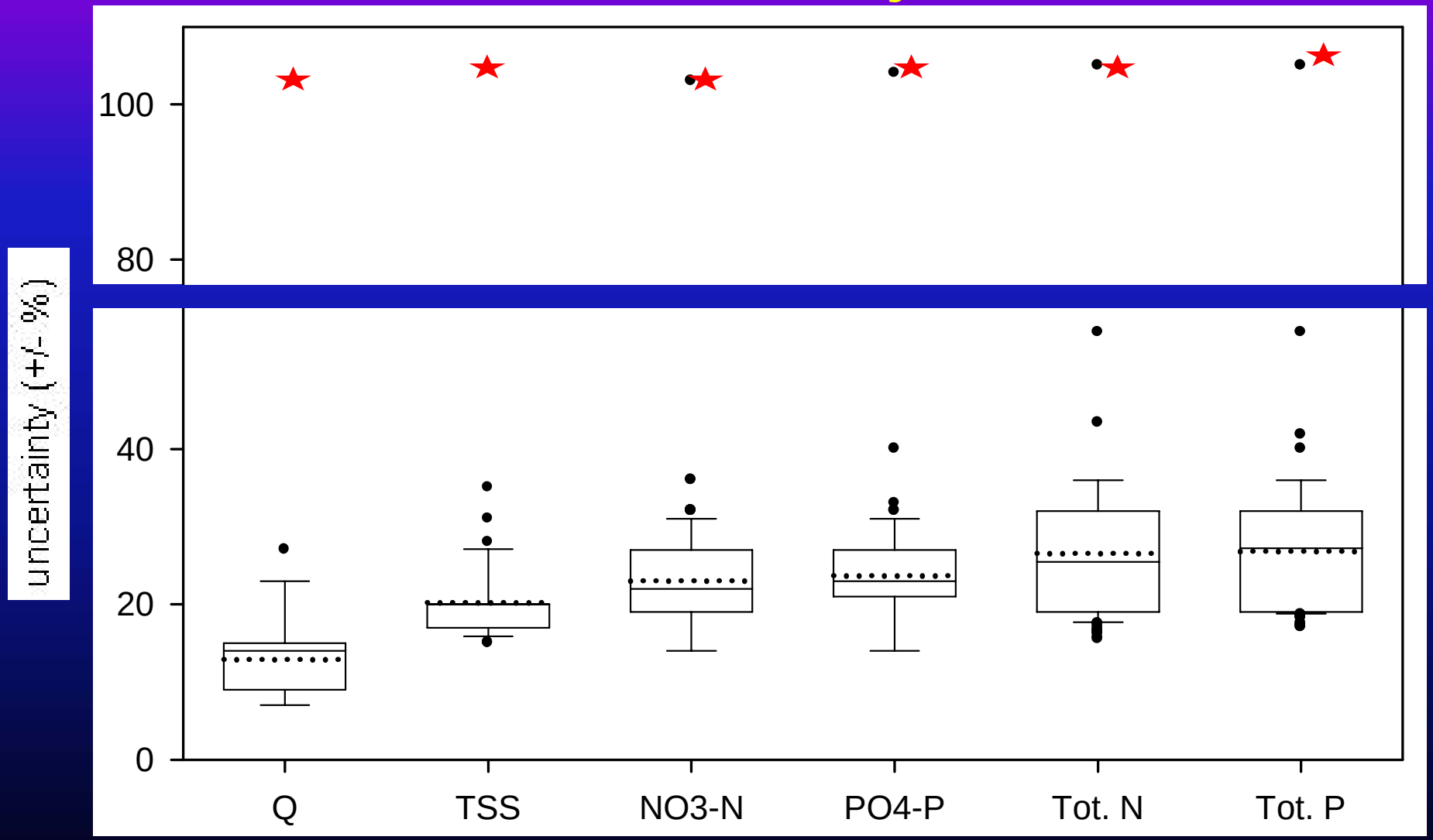
Determine Uncertainty - RESULTS

- Each procedural category can contribute substantial uncertainty; therefore, QA/QC protocols should address them all
  - discharge measurement - typically adequately addressed
  - sample collection - typically poorly addressed
  - sample preservation/storage - typically emphasized
  - laboratory analysis - typically emphasized

USDA

Agricultural Research Service

Determine Uncertainty - RESULTS



USDA

Agricultural Research Service

Determine Uncertainty - CONCLUSIONS

- All sources of measurement uncertainty deserve QA attention
- Unacceptable to ignore uncertainty in H/WQ data collection
- Measured data with corresponding uncertainty estimates provide many benefits
  - enhanced monitoring design, decision-making, model application and evaluation, and… scientific integrity
- What's next...???
  - Attempt to make uncertainty estimation a routine task in H/WQ data collection and reporting
  - Publicize DUET-H/WQ as a valid, user-friendly tool

Model Evaluation – INTRODUCTION

- **One important source of uncertainty in H/WQ modeling is measurement uncertainty; therefore, it should be considered in model evaluation.**
- **However, when "measurement uncertainty" is included in uncertainty analysis**
  - **focuses almost exclusively on model inputs or parameter estimation (e.g. hydraulic conductivity, curve number, fertilizer application)**
  - **does *not* address uncertainty in measured data, against which model outputs are compared (e.g. flow, water quality)**
- **This discussion focuses on uncertainty in measured data used to calibrate or validate H/WQ models.**

Model Evaluation – INTRODUCTION

- Our belief is that models should:
  - not be expected to simulate/reproduce uncertain data values
  - produce output within the uncertainty range of measured data
- To represent this mathematically, we modified several popular model goodness-of-fit indicators (E_{NS}, *d*, RMSE, MAE)
- Specifically, we modified the error or deviation term ($e_i = O_i - P_i$)

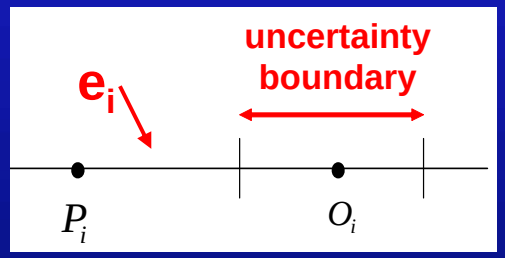
USDA

Agricultural Research Service

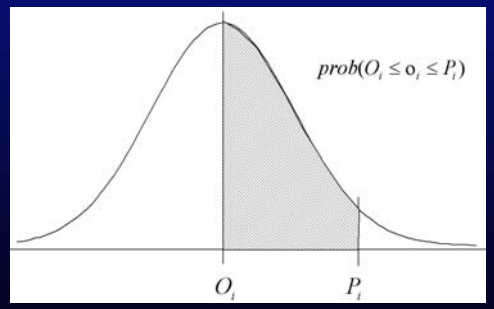
Model Evaluation – METHODS

- **Developed two error term modifications based on information available (and willingness to assume a distribution) of measurement uncertainty for individual measured values (O_i) - *NOT* for the entire population of measured data!!!**

- **Modification 1 is most appropriate if:**
  - **distribution not reasonably assumed**

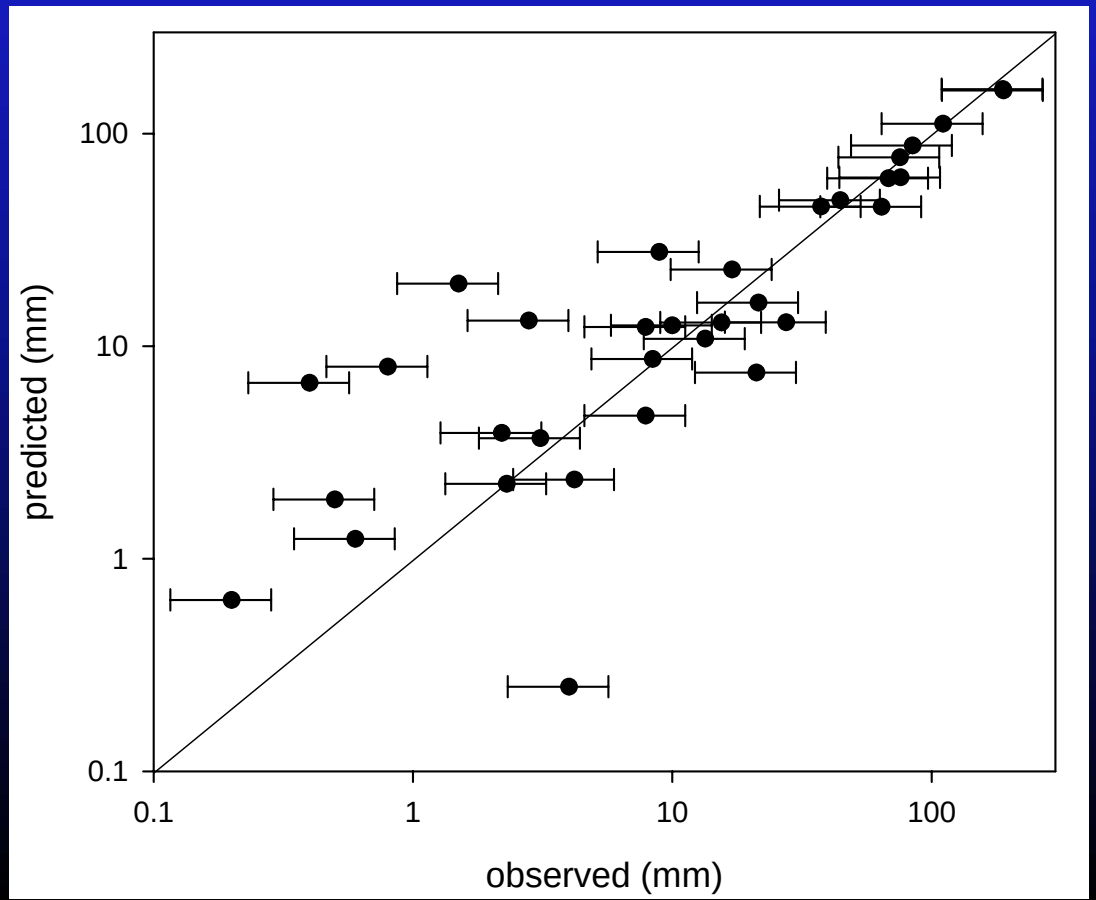


- **Modification 2 is most appropriate if:**
  - **uncertainty distribution known or assumed**



Model Evaluation – RESULTS

- **Modification 1 –** produces large improvements in indicator values (minimizes e_i); thus produces "generous" goodness-of-fit conclusions
  - Thus, best suited for use in visual (graphic) assessment

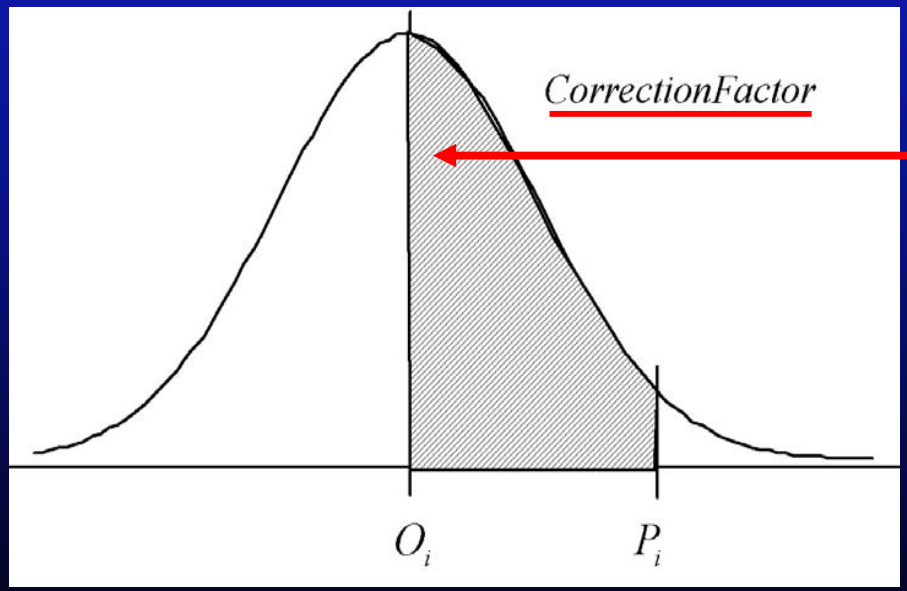


USDA

Agricultural Research Service

Model Evaluation – RESULTS

- **Modification 2** - provides more realistic goodness-of-fit conclusions
  - Minimal goodness-of-fit improvement for data with little uncertainty
  - Modest improvement when data with substantial uncertainty were compared with both poor and good model predictions **Important**



$$e_i = \frac{CF_i}{0.5} \times (O_i - P_i)$$

Model Evaluation – CONCLUSIONS

- Measurement uncertainty should be incorporated in H/WQ model goodness-of-fit evaluations
- Advantageous for modelers to quantify the "quality" of calibration and validation data

- What's next...???
  - Finalize method to consider both measurement and model uncertainty in model goodness-of-fit evaluation
  - Incorporate uncertainty estimates and modified goodness-of-fit indicators in SWAT, EPIC/APEX interface

Any Questions??

Daren Harmel, PhD
USDA-ARS
808 E. Blackland Rd.
Temple, TX 76502

(254) 770-6521
daren.harmel@ars.usda.gov

www.ars.usda.gov/spa/dharmel

USDA

Agricultural Research Service