Uncertainty Analysis in a Field-scale P Loss Model

Carl H. Bolster
USDA-ARS, Bowling Green, KY

Peter A. Vadas
USDA-ARS, Madison, WI.

JEQ 42:1109-1118 (2013)
Sources of Model Uncertainty

• Model structure error
  • All models are approximations
  • “All models are wrong, some are useful”

• Model parameter error (Generally obtained through calibration)
  • Incorrect optimization targets
  • Inaccurate, incomplete, or unrepresentative calibration data

• Model input error (variables such as rainfall, soil test P)
  • Measurement errors
  • Unrepresentative values
Sources of Model Uncertainty

• Model structure error
  • All models are approximations
  • “All models are wrong, some are useful”

• Model parameter error (Generally obtained through calibration)
  • Incorrect optimization targets
  • Inaccurate, incomplete, or unrepresentative calibration data

• Model input error (variables such as rainfall, soil test P)
  • Measurement errors
  • Unrepresentative values
Annual Phosphorus Loss Estimator

APLE is a spreadsheet model that simulates dissolved and sediment bound phosphorus loss in surface runoff.

Click here to download spreadsheet for:
- Data Entry
- Output Graphs
- Calculations

Click here to view APLE User’s Manual Version 2.0, Spring 2011

Click here to view APLE Theoretical Documentation Version 2.0, Spring 2011

For user questions or to report potential errors in calculations, contact the APLE creator:
Peter Vadas, Research Soil Scientist
U.S. Dairy Forage Research Center
USDA-Agricultural Research Service
Madison, Wisconsin, USA
Email: Peter.Vadas@ars.usda.gov

Last Modified: 05/03/2011
Objective:

Evaluate model predictions when uncertainties in both model output and measurements are included

How large are the errors?

Do uncertainties help in model evaluation?
### APLE input model variables

#### P Loss Pathway

<table>
<thead>
<tr>
<th></th>
<th>$D_{P_{man}}$</th>
<th>$D_{P_{fert}}$</th>
<th>$D_{P_{soil}}$</th>
<th>$P_{sed}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runoff/Precip</td>
<td>Runoff/Precip</td>
<td>Runoff</td>
<td>Soil Loss</td>
<td></td>
</tr>
<tr>
<td>Total manure applied</td>
<td>Total P applied</td>
<td>Labile P</td>
<td>Labile P</td>
<td></td>
</tr>
<tr>
<td>Percent manure solids</td>
<td>Percent fertilizer incorporation</td>
<td></td>
<td>Soil clay content</td>
<td></td>
</tr>
<tr>
<td>Manure TP content</td>
<td></td>
<td></td>
<td>Soil organic matter content</td>
<td></td>
</tr>
<tr>
<td>Water extractable P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent manure incorporation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mineralization rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Assumed Errors in Model Input Variables

<table>
<thead>
<tr>
<th>Model Variable</th>
<th>Small Uncertainty</th>
<th>Large Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runoff (weir)</td>
<td>± 5%</td>
<td>± 10%</td>
</tr>
<tr>
<td>Runoff (direct)</td>
<td>± 1%</td>
<td>± 3%</td>
</tr>
<tr>
<td>Erosion</td>
<td>± 2.5%</td>
<td>± 5%</td>
</tr>
<tr>
<td>Manure mineralization</td>
<td>± 2.5%</td>
<td>± 5%</td>
</tr>
<tr>
<td>P incorporation rates</td>
<td>± 5% (constant)</td>
<td>± 10% (constant)</td>
</tr>
<tr>
<td>All other variables</td>
<td>± 5%</td>
<td>± 15%</td>
</tr>
</tbody>
</table>

Based on Harmel et al. 2006
### Assumed Errors in Measured P Loads

<table>
<thead>
<tr>
<th>Model Variable</th>
<th>Small Uncertainty</th>
<th>Large Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample collection</td>
<td>± 5%</td>
<td>± 15%</td>
</tr>
<tr>
<td>Sample preservation</td>
<td>± 5%</td>
<td>± 15%</td>
</tr>
<tr>
<td>Laboratory analysis</td>
<td>± 5%</td>
<td>± 20%</td>
</tr>
<tr>
<td>Contributing area</td>
<td>± 2%</td>
<td>± 5%</td>
</tr>
<tr>
<td>Total Error in P Loads</td>
<td>± 9%</td>
<td>± 29%</td>
</tr>
</tbody>
</table>

*Based on Harmel et al. 2006*
Correlations between predicted and observed P loss

$$E = 0.71$$
Small Uncertainty

- Observed/Predicted P loss (kg/ha)
  - Observed
  - Predicted

CI range
- ± 3.5 x10^{-5} to 2.8 kg/ha
- ± 2 to 32%

# overlapping CIs
- 66 out of 255
Large Uncertainty

Observed/Predicted P loss (kg/ha)

CI range
± 9.4 x10⁻⁵ to 6.4 kg/ha

CI range
± 6 to 64 %

# overlapping CIs
156 out of 255
Distributions of absolute and relative errors for model predictions.
## Goodness-of-fit statistics

<table>
<thead>
<tr>
<th>Statistic</th>
<th>No uncertainty</th>
<th>Small uncertainty</th>
<th>Large uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$</td>
<td>0.71</td>
<td>0.71</td>
<td>0.72</td>
</tr>
<tr>
<td>RMSE, kg P ha(^{-1})</td>
<td>1.84</td>
<td>1.84</td>
<td>1.82</td>
</tr>
<tr>
<td>MAE, kg P ha(^{-1})</td>
<td>0.88</td>
<td>0.87</td>
<td>0.83</td>
</tr>
<tr>
<td>MAPE, %</td>
<td>59.1</td>
<td>58.8</td>
<td>57.0</td>
</tr>
<tr>
<td># of overlapping CIs</td>
<td>65</td>
<td></td>
<td>155</td>
</tr>
</tbody>
</table>
Conclusions

- Uncertainties in model predictions are a fact of life
- Ignoring them may do more harm than good
- Uncertainties in model predictions can help us better evaluate our models
- As modelers it is our responsibility to faithfully present the limitations with our model predictions to our audience
FOA Method

- Calculates variance in model output as product of input variable variance and sensitivity of model output to changes in that variable

\[ \sigma^2(\theta) = \sum_{i=1}^{k} \left( \frac{\partial \theta}{\partial I_i} \right)^2 \cdot \sigma^2(I_i) \]

\[ \frac{\partial \theta}{\partial I} \approx \frac{\theta_{I+\Delta I} - \theta_{I-\Delta I}}{2\Delta I} \]

\( \theta \) is model output

\( I \) is model input variable
MCS Method

- Model input variables are selected randomly from a pre-defined distribution (triangular)
- The model is run, and the output is stored.
- The process is repeated numerous times
- Statistical distributions of the output ensemble are used to assess uncertainty in model output.
95% confidence intervals for MCS Model Output Values

Cumulative Probability

Model Output Values
Methods

- Simulated P loss from two different field conditions:
  - DRP from soil, particulate P loss, DRP from fertilizer
  - DRP from soil, particulate P loss, DRP from manure
- Four error ranges:
  - ± 5%, ± 15%, ± 25%, ± 50%
- Triangular distribution
Triangular vs Normal distribution

![Graph comparing Triangular and Gaussian distributions for Labile P (mg kg\(^{-1}\)).]
Objective 1:

Compare First-Order Approximation (FOA) method with Monte Carlo Simulation (MCS) Method using APLE
% Differences in CIs between MCS and FOA: P loss from STP, erosion, and fertilizer P

![Graphs showing % Differences at 5%, 15%, 25%, and 50% error levels for Mean, Lower CI, and Upper CI.](image-url)
% Differences in CIs between MCS and FOA: P loss from STP, erosion, and manure P

<table>
<thead>
<tr>
<th>% Diff</th>
<th>Mean</th>
<th>Lower CI</th>
<th>Upper CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>5% error</td>
<td>-10</td>
<td>-5</td>
<td>0</td>
</tr>
<tr>
<td>15% error</td>
<td>5</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>25% error</td>
<td>0</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>50% error</td>
<td>10</td>
<td>20</td>
<td>30</td>
</tr>
</tbody>
</table>
Correlations between predicted and measured P loss

Small Uncertainty

Large Uncertainty

$E = 0.71$
Sensitivity Coefficients

- For Manure incorporation (%):
  - $S_{DP_{man}}$

- For Manure solid content (%):
  - $S_{DP_{man}}$

- For Erosion rate (kg ha$^{-1}$):
  - $S_{P_{sed}}$

- For Runoff/Precipitation:
  - $S_{DP_{fert}}$
Objectives

1. Conduct sensitivity analysis with all APLE input variables
   • Help identify which input variables require most accurate measurements

2. Evaluate model predictions when uncertainties in both model output and measurements are included
Objective 1:

Sensitivity Analysis of APLE Input Variables
Relative Sensitivity Coefficient

- Dimensionless parameter that measures how sensitive model output ($\theta$) is to a given change in model input ($I$)

$$S_r = \frac{\partial \theta}{\partial I} \frac{I}{\theta}$$

$$\frac{\partial \theta}{\partial I} \approx \frac{\theta_{I+\Delta I} - \theta_{I-\Delta I}}{2\Delta I}$$

$\theta$ is model output

$I$ is model input variable
<table>
<thead>
<tr>
<th>Model Variable</th>
<th>Variable Range</th>
<th>$S_r$ mean and range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>$DP_{soil}$</strong></td>
</tr>
<tr>
<td>Runoff, mm</td>
<td>10 – 500</td>
<td>1.0</td>
</tr>
<tr>
<td>Labile soil P, mg kg$^{-1}$</td>
<td>10 – 400</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>$DP_{man}$</strong></td>
</tr>
<tr>
<td>RO/PT, %</td>
<td>1 – 35</td>
<td>1.2</td>
</tr>
<tr>
<td>Manure application rate, kg ha$^{-1}$</td>
<td>$2.2 \times 10^2$ – $5.6 \times 10^4$</td>
<td>1.0</td>
</tr>
<tr>
<td>Solid content of manure, %</td>
<td>5 – 75</td>
<td>1.2 (1.0, 9.3)$^\dagger$</td>
</tr>
<tr>
<td>P content of manure, %</td>
<td>0.1 – 5</td>
<td>1.0</td>
</tr>
<tr>
<td>WEP content of manure, %</td>
<td>25 – 75</td>
<td>0.76 (0.51, 0.93)</td>
</tr>
<tr>
<td>Manure incorporation rate, %</td>
<td>10 – 90</td>
<td>-1.5 (-9.0, -0.11)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>$DP_{fert}$</strong></td>
</tr>
<tr>
<td>RO/PT, %</td>
<td>1 – 35</td>
<td>1.8 (1.3, 2.2)</td>
</tr>
<tr>
<td>Fertilizer application rate, kg ha$^{-1}$</td>
<td>20 – 300</td>
<td>1.0</td>
</tr>
<tr>
<td>Fertilizer incorporation rate, %</td>
<td>10 – 90</td>
<td>-1.5 (-9.0, -0.11)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>$P_{sed}$</strong></td>
</tr>
<tr>
<td>Erosion rate, kg ha$^{-1}$</td>
<td>$22 – 5.6 \times 10^4$</td>
<td>0.93 (0.75, 1.0)</td>
</tr>
<tr>
<td>Labile P, mg kg$^{-1}$</td>
<td>10 – 400</td>
<td>0.19 (-0.30, 0.80)</td>
</tr>
<tr>
<td>Clay content, %</td>
<td>0.1 – 50</td>
<td>0.19 (0.10, 0.59)</td>
</tr>
<tr>
<td>Soil organic matter, %</td>
<td>0.1 – 10</td>
<td>0.46 (0.02, 2.2)</td>
</tr>
</tbody>
</table>
Objectives

1. Test whether the first-order approximation method can provide accurate estimates of confidence intervals for APLE predictions
   • Compare results with Monte Carlo simulations

2. Conduct sensitivity analysis with all APLE input variables
   • Help identify which input variables require most accurate measurements

3. Evaluate model predictions when uncertainties in both model output and measurements are included

JEQ 42:1109-1118 (2013)