Why?
1) Evaluate model for **bias** in predictions or limitation in application. Model may be changed (process descriptions added) to accommodate limitations.
2) Determine model accuracy. What **confidence** do we put in the predictions?
3) To **rank** the required inputs as to sensitivity. Which parameters are most important (in which conditions), which are poorly- or well-defined? Where to expend effort for model parameter, data, input.

Verification = does the code do what you want it to do?
Validation = evaluate the model’s accuracy and sensitivity
Calibration = parameter estimation, characterization, and system relation to empirical vs. process/physically-based models

**Validation**
1) Sensitivity Analysis
2) Error Analysis
3) Comparisons to independent data sets

Sensitivity is the rate of change in one factor with respect to change in another factor.
Consider the model:
\[ O = f[F_1, F_2, F_3, ..., F_n] \]
The change in O with respect to \( F_i \) is: (Series expansion)
\[
f\left[ F_i + \Delta F_i, F_{/i} \right] = O_0 + \frac{\partial O_0}{\partial F_i} \Delta F_i + \frac{1}{2!} \frac{\partial^2 O_0}{\partial F_i^2} \Delta F_i^2 + ...
\]
where \( O_0 \) is the value of O at some specified \( F_i \) level.

**IF** the non-linear terms are insignificant with respect to the linear terms
\[
f\left[ F_i + \Delta F_i, F_{/i} \right] \approx O_0 + \frac{\partial O_0}{\partial F_i} \Delta F_i
\]
Thus \( \Delta O = \frac{\partial O_0}{\partial F_i} \Delta F_i \)

**Linearized Sensitivity Equation:**
Linear Sensitivity Coefficient, \( S \)
\[
S = \frac{\partial O_0}{\partial F_i} = \lim_{\Delta F_i \to 0} \frac{f(F_i + \Delta F_i, F_{/i}) - f(F_i, F_2, ..., F_n)}{\Delta F_i}
\]
Computing \( S = \frac{\partial O_0}{\Delta F_i} \)
1) Analytically – limited by model complexity
2) Factor Perturbation – finite difference form

\[ S = \frac{\Delta O_o}{\Delta F_i} = \frac{f(F_i + \Delta F_i, F_{/i}) - f(F_i, F_{2},..., F_n)}{\Delta F_i} \]

Sensitivity may be performed on model parameters, P or inputs, I

\[ I \rightarrow \text{function w/parameters P} \rightarrow O \]

Model
(Respons Function)
Distribution Function

Parametric Sensitivity: \[ S_{pi} = \frac{\partial O}{\partial P_i} \]

Input Sensitivity: \[ S_{pi} = \frac{\partial O}{\partial I_i} \]

Sensitivity of Model is function of the State of the System

\[ S = \frac{\partial O_o}{\partial F} = \lim_{\Delta F_i \to 0} \frac{f(F_i + \Delta F_i, F_{/i}) - f(F_i, F_{2},..., F_n)}{\Delta F} \]

\[ S_i = f_i c(F_i, F_2,...F_n) \] the “baseline” values

*The sensitivity of a factor is a function of the baseline values chosen.

Example: from WEPP sensitivity paper
Table 10 for Gradient = 9%

<table>
<thead>
<tr>
<th>Slope length</th>
<th>Sensitivity to Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>0.66</td>
</tr>
<tr>
<td>50</td>
<td>0.44</td>
</tr>
<tr>
<td>200</td>
<td>0.18</td>
</tr>
</tbody>
</table>

More sensitive on short slopes

Why?

How to address this issue of dependence of S on baseline variables

1) Perform sensitivity on the application of interest. May want to look at S=f[temporal changes in site conditions]
2) If desired is sensitivity over many applications choose representative applications
   Conventional Corn Iowa
   Conservative Till Corn Iowa
   Cotton MS
Select parameter value levels. Evaluate sensitivity for random level combinations. May use surface exploration technique to find maximum or surface—or—plot the surface.

Setting Levels – 5 variable model

<table>
<thead>
<tr>
<th>INPUT</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>X₁</td>
<td>Min</td>
</tr>
<tr>
<td>X₂</td>
<td></td>
</tr>
<tr>
<td>X₃</td>
<td></td>
</tr>
<tr>
<td>X₄</td>
<td></td>
</tr>
<tr>
<td>X₅</td>
<td></td>
</tr>
</tbody>
</table>

Randomly select combinations of Xᵢ
Plot Rᵢ vs X₁, X₂, X₃, X₄, X₅
Regress Rᵢ vs X₁, X₂, X₃, X₄, X₅
Plot Distribution of Rᵢ and identify max, min case
How many combinations?
N=?
The more, the better!
How much better?

Brooks (1958)
Surface exploration technique for identifying maximum

\[ n = \frac{\log (1-S)}{\log (1-a)} \]

where \( n \) = number of trials required to obtain probability \( S \) of finding at least 1 treatment combination falling into subregion \( a \) of the factor space which maximizes the response.

*If you want to be 95% certain of your response values, ONE is within 5% of the maximum then you need

\[ n = \frac{\log (1-0.95)}{\log (1-0.05)} = 59 \]

does not matter the number of variables or levels!
If your response is R₁ (Relative sensitivity to X₁) then of 60 treatment combinations you’ll likely (95% or so) be within 5% of maximum R₁.
Criticism of the random/level method of sensitivity analysis is that certain combinations of variable values may not be realistic. Eg. if CLAY & KSAT are INPUTS they would be expected to be correlated (positively) LOW CLAY, May not be realistic HIGH KSAT

*A random selection method which takes into account input correlation would (could) address the issue.

Error Analysis

![Flowchart of Error Analysis]

Error Sources
Inputs: Identification from measured data (for Gaged Conditions – may act as INPUT values)
Estimation from input value prediction eqs. (for Ungaged Conditions)

Identification Errors
Measurement errors “Fitting” errors

Estimation Errors
Eq. Dev. error
Uncertainty in predictor variable

Example
Soil Rill Erodibility
Run erosion simulation experiments
* Measurement errors techniques simulator
spill the bucket
misread the scale
etc.

* Fitting Errors

\[ Dr = Kr(\tau - \tau_c)(1 - \frac{G}{T_c}) \]

Optimized procedure
Random variation – nature

“Measured” \( K_R \)
w/Identification Errors

Relate \( k_r \) to Soil Properties

\( \text{clay} \)
\( \text{silt} \)
\( \text{sand} \)
\( \text{OM} \)
\( \text{density} \)
\( \text{Na}^+ \)
\( \text{Mg}^{2+} \)
\( \text{agg stability} \)
\( \text{color} \)

↑ Errors in measuring

\( \text{Vs. } Kr \Rightarrow \text{Regression Equation} \)

\( R^2 = 0.80 \)

Lost 20% of predictability

A) Equation Development Error
B) Equation Application Error

Your Site (ungaged) ⇒ CLAY
⇒ OM
⇒ Na\(^+\)
⇒ Mg\(^{2+}\)
⇒ aggregate stability

Model Error
1) Aggregation Errors – Process lumping, rill erosion ⇒ scour but also wall sloughing, headcutting
   At some scale all models lump system properties. i.e., assume spatial and temporal homogeneity. They assume that the mean value of the system property distribution represents the property. If: linear response & symmetric freq. dist⇒ OK
2) Solution: Numerical errors of solution
3) Parametric Errors
   Natural variability of model parameters
Unaccounted dependence of parameters
Ex. \( G_c = (1 - e^{\beta g}) \) for ground cover effect

\( \beta \) may be a constant model parameter but truly dependent upon soil, crop type, etc. and naturally variable.

4) Structural errors
Mathematical formulation \( G_c = 1 - e^{\beta g} \)

Very difficult to address
“Easiest” and certainly most often looked at:
Error associated w/ input error

Variance of Input
\[ \downarrow \]
Variance of Output

Methods:
- Monte Carlo
- 1\(^{st}\) order
- Point estimate

**The Concept Of Parameter Uncertainty Analysis**

1. Estimate Distributions Of Values For Parameters \( X, Y, \) And \( Z \)

\[ f(x) \quad x \quad f(y) \quad y \quad f(z) \quad z \]

2. Input Distributions Into Model

\[ \text{SOIL LOSSES} = g(x,y,z) \]

3. Produce Distributions Of Model Predictions
Method:
1st-Order analysis

\[ \sigma_y^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} \left( \frac{\partial y}{\partial x_i} \right) \left( \frac{\partial y}{\partial x_j} \right) \sigma_{x_i} \sigma_{x_j} \theta_{x_i x_j} \]

For Uncorrelated Inputs ( \( \theta_{x_i x_j} = 0 \) )

\[ \sigma_y^2 = \sum_{i=1}^{m} \left( \frac{\partial y}{\partial x_i} \right)^2 \sigma_{x_i}^2 \]

S = “raw” sensitivity

\[ \sigma_y^2 = \sum_{i=1}^{m} S_i^2 \sigma_{x_i}^2 \]

Monte-Carlo
Validation
Data Comparison
Quantifying the error between predicted and measured response variables using an independent data set.

Issues and Difference:
1) Continuous vs. Event Model
2) Gaged vs. ungaged applications

Continuous: may compare annuals, event totals, event peaks, and distribution.
Event model: event totals, peaks, OR perhaps dynamic response.

Methods
1. Measured vs. Predicted or time-series Graphs Visual-Comparison
2. Statistical Analysis of error terms
3. Statistical comparison of frequency distributions.
Goodness-of-Fit Criteria

2*

Relative Error = \frac{y_{obs} - y_{pred}}{y_{obs}}

Squared Error

F^2 = \sum (y_{pred} - y_{obs})^2 \quad [\text{Residual variance of regression}]

Nash & Sutcliffe (1970)

Efficiency coefficient

F^2 = \sum (y_{pred} - y_{obs})^2 \quad (\text{residual variance})

F_o^2 = \sum (y_{obs} - \bar{y})^2 \quad (\text{initial variance})

R^2 = \frac{F^2 - F_o^2}{F_o^2} \quad \text{efficiency}

R^2 = 1 - \frac{F^2}{F_o^2}
• If \( F^2 = F_o^2 \) then \( R^2 = 0 \) Means that mean value does as well as the model as a predictor
• If \( F^2 = 0 \) then \( R^2 = 1 \) Model accounts for all variance
• If \( F^2 > F_o^2 \) then \( R^2 < 0 \) you’re better off using the mean value

**A problem:**

With relative error in any normalized statistic, equal weight is given to large and small values, i.e., for large and small events. However, large events contribute a large percent (sometimes disproportionately large percent) of loads. Small events are HARDER to predict.

? Does the inclusion of equal weight for events bias the error term?
? Would an absolute error term be more appropriate for NPS loads?

For comparing dynamic response e.g. (hydrograph, sedigraph, P vs. time, etc)

Sum of square residuals for a single event

\[
G = \sum_{i=1}^{n} (y_{obs} - y_{pred})^2
\]

Total sum of square residuals for multiple events

\[
TSSR = \sum_{j=1}^{m} \left( \sum_{i=1}^{n} (y_{obs} - y_{pred})^2 \right)_j
\]

\( n = \) number of data points within event \( j \)
\( m = \) number of events

Refs.
Perhaps Model doesn’t give good 1:1 event comparisons, but does replicate the long-term frequency distribution?

Follows approximately a $\chi^2$ distribution

Degree of Freedom $\nu = k - 1$

$P[D \geq \chi^2_{\nu, k-1}] = \alpha$

1) Computer $D_1$
2) Compare $D_1$ to $\chi^2_{\nu, k-1}$
3) $H_0$: The two distributions are the SAME
   if $D_1 \leq \chi^2 \Rightarrow$ Accept $H_0$
   if $D_1 > \chi^2 \Rightarrow$ Reject $H_0$
Kolmogorov-Smirnov Test

\[ F = \frac{i}{n} \]

<table>
<thead>
<tr>
<th>( n )</th>
<th>1/n</th>
<th>0.01</th>
<th>2/n</th>
<th>0.03</th>
<th>3/n</th>
<th>0.48</th>
<th>4/n</th>
<th>0.62</th>
<th>. . .</th>
<th>n</th>
<th>1</th>
<th>3.84</th>
</tr>
</thead>
</table>

\[ D_2 = \max_{i=1}^{n} | F \cdot (y_i) - F(y_i) | \]

\( H_0: \) The distributions are the same \( P(D \geq C) = \alpha \)

\( H_1: \) They are not the same

If \( D_2 \leq C \): Accept \( H_0 \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \alpha = 0.10 )</th>
<th>( \alpha = 0.05 )</th>
<th>( \alpha = 0.01 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.51</td>
<td>0.56</td>
<td>0.67</td>
</tr>
<tr>
<td>10</td>
<td>0.37</td>
<td>0.41</td>
<td>0.49</td>
</tr>
<tr>
<td>15</td>
<td>0.30</td>
<td>0.34</td>
<td>0.40</td>
</tr>
<tr>
<td>20</td>
<td>0.26</td>
<td>0.29</td>
<td>0.35</td>
</tr>
<tr>
<td>25</td>
<td>0.24</td>
<td>0.26</td>
<td>0.32</td>
</tr>
<tr>
<td>30</td>
<td>0.22</td>
<td>0.24</td>
<td>0.29</td>
</tr>
<tr>
<td>40</td>
<td>0.19</td>
<td>0.21</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Relationship between validation & parameter identification

Define objective function

e.q./Residual Variance

\[ F^2 = \sum (y_{pred} - y_{obs})^2 \]
Find model parameters of input values which minimize the object function:

1) Steepest Descent
2) Global Search

Problem with Parameter Identification
Insensitive to $X_3$

Interdependence between $X_1X_2$

Local Minima
Relative Sensitivity (Resilience Sensitivity)

\[ R_i = \frac{\partial 0}{\partial F_i} \bigg|_{F_i} = \frac{\partial 0}{\partial F_i} \left( \frac{F_i}{0} \right) \]

This relative sensitivity is not dependent on magnitude of values (i.e., units) and allows for relative comparison of sensitivity values among Factors.

Fraction change output

Fraction change input

Example:

I_1 = Precipitation in mm

\[ I_1 = .050 \text{ m} \quad \Delta I_1 = .002 \text{ m} \quad 0_o = 250 \text{ g/m}^2 \quad \Delta 0_o = 5 \text{ g/m}^2 \]

I_2 = Interill erodibility \( \frac{ky - s}{m^4} \)

\[ I_2 = 1,500,000 \quad \Delta I_2 = 60,000 \quad 0_o = 250 \text{ g/m}^2 \quad \Delta 0_o = 5 \text{ g/m}^2 \]

\[ S_1 = \frac{\Delta 0_o}{\Delta I_1} = \frac{25 \text{ g/m}^2}{.002 \text{ m}} = 2,500 \text{ g/m}^3 \]

\[ R_1 = \frac{\Delta 0_o}{\Delta I_1} \frac{I_1}{0_o} = 0.5 \]

\[ S_2 = \frac{\Delta 0_o}{\Delta I_2} = \frac{5 \text{ g/m}^2}{60,000 \text{ kg/s/m}^2/n^x} = 8.33 \times 10^{-5} \text{ g/m}^2/ky \text{s/m}^2 = 8.33 \times 10^{-5} \text{ m}^2/s \]

\[ R_2 = \frac{\Delta 0_o}{\Delta I_1} \frac{I_2}{0_o} = 0.5 \]

R = 0.5 \( \Rightarrow \) 4% change in input gives 2% change in output
FIG. 11.1 Scattergrams showing presence of systematic errors.