New Developments for Parameterizing Models for Forecasting and Uncertainty

Christine Shoemaker
Civil and Environmental Engineering & Operations Research and Information Engr.
Cornell University.

Funding Acknowledged from NSF and USDA
New Developments for Parameterizing Models for Forecasting and Uncertainty

Christine Shoemaker
Cornell University.

Results drawn from recent PhD Theses by:

Josh Woodbury on Watershed Analysis
(now Postdoc at Cornell)

Taimoor Akhtar on Multi Objective Analysis
Part I

Importance of Algorithm Choice for Calibration and Management Optimization
Points to Make

• Calibration of watershed models and management analysis are greatly assisted by optimization.

• Many/most watershed problems involving water quality as well as flow have multiple local minima and hence require global optimization algorithms; and calibration is much more difficult.

• The choice of algorithm makes a huge difference in the accuracy of the optimization solution obtained for hydrologic problems.
Simulation - Optimization

- In hydrology and environmental science/engineering, we are often coupling the optimization to a simulation model.
- The computational expense is primarily the simulations (\textit{=function evaluations}) unless the simulation runs very quickly (seconds).
- The following slide compares a large number of global optimization algorithms on the same (groundwater) problem.
Results (Umatilla): This plot shows that Stochastic RBF is most efficient of the tested algorithms (**lowest curve**) (From PhD thesis of my student Amandeep Singh, 2011)

Convergence Plot (note log scale) **Lowest curve is best.**
Impact of Algorithm Choice

• All results are averaged over 10 trials for each algorithm.
• Note that the vertical scale is log scale.
• The genetic algorithm, simulated annealing, shuffle complex evolution solution are 100 times higher (worse) than the best (our response surface approach “Stochastic RBF”) after 400 evaluations.
• Hence, the choice of algorithm makes a big difference in optimization (including for global optimization).
Genetic Algorithms are not very efficient for most hydrology problems

- GA are good for picking the best binary numbers in a long string (1001101000001).
- Hydrology calibration and management optimization problems tend to have continuous decisions (e.g. parameter values) for which GA is not very efficient.
- DDS (Tolson & Shoemaker, 2007) is a heuristic that works better than GA/SCE for watershed models, especially for higher dimensional problems.
- If your objective function is not expensive and you can afford to evaluate it thousands of times, then you don’t need to worry about using an efficient algorithm now.
Single Objective Optimization

• Some of the methods we have developed to improve optimization include the use of “response surfaces” so that in each step of an optimization search we build an approximation of the function to be optimized using all the objective function simulations done so far.

• This can be shown to greatly reduce the number of evaluations necessary to get a good solution.
Multi-Objective Optimization (MOO) Analysis

• Often we want to optimize over multiple objectives. Two examples:
  – 1. Goodness of fit for flow and 2. fit for water quality
  – 1. Cost of management and 2. environmental effectiveness (e.g. SUSTAIN)

• Multi objective optimization is typically much more computationally expensive (e.g. requires many more simulations) to obtain an accurate answer.
NSGAII

- NSGAII was developed years ago and continues to be most widely used method for MOO. NSGAII uses genetic algorithms.
- NSGAII is used in SUSTAIN
- There are newer methods that are much more efficient including our “GOMORS” (see following slides)
Convergence Time Comparison: GOMORS - NSGA2
Convergence Time Comparison:
GOMORS (414 Evals) better than NSGA2 (1200 Evals)
Convergence Time Comparison:  GOMORS (414 Evals) =
NSGAII (25,000)!!!
Example of Importance of Efficient Algorithms

• Estimate of difference in computation time for a simulation model that takes 1 minute:
  – With our GOMORS algorithm 414 evaluations takes 414 minutes or about 8 hours (serial).
  – With NSGAII 25,000 evaluations takes about 3 years

Hence it is not feasible to get accurate answers to this multi objective problem with NSGAII. The answer with 2 hours of NSGAII is not accurate.

Since people don’t run these genetic algorithms for years, the typical result is to get inaccurate answers, e.g. 1200 eval NSGAII takes 24 hours.
Convergence Time Comparison: GOMORS - NSGA2
Parallel Algorithms

• Parallel calculations are important in algorithms.
• We are implementing new algorithms, designed for parallel analysis (funded by NSF CISE)
• This is not discussed here.
Major Point

• Algorithm Efficiency is important unless the time to simulate the hydrology model is extremely fast (e.g. seconds)
• There are better algorithms available (not just ours) now.
• Contact me if you want to know more about our algorithms.
Part II

Using Models and Calibration to Investigate Competing Hydrologic Processes and Management

This includes the use of weighted *mixtures* of watershed models
Quantifying Relative Importance of Competing Dynamic Hydrological Processes

• Hortonian Flow and Variable Source Area Hydrology are competing hydrological processes because at a given time and location one process will dominate the other.

• However, it is reasonable to assume that this dominance is varies over time (i.e. is dynamic).

• This quantification is difficult to do over the scale of a large watershed.
Illustration of VSA hydrology

Agricultural field conditions in August (Dry) * Agricultural field conditions in March (Wet) *

* Soil and Water Lab, BEE, Cornell University,
http://soilandwater.bee.cornell.edu/research/VSA/extension.html
Application to Cannonsville Watershed and Reservoir

- Approximately 1200 square kilometers in the Catskill Region of Upstate New York
- The watershed is primarily forest and agricultural land, approximately 0.5% is urban

- Currently provides NYC with unfiltered drinking water
- Phosphorous is a problem
How Do We Protect This Water From Pollution?

(New York City water supply)

Problem is that phosphorous from surrounding Watershed can result in need for an $8 Billion Water Plant for NYC!
We have used a SWAT2005 Model of Cannonsville Watershed (1200 km$^2$)

Using a spatially distributed, process model helps us evaluate management options.

43 subbasins
758 HRUs
Avg HRU = 1.6 km$^2$
Quantifying Dynamic Dominance of Competing Hydrologic Processes

• Our approach is to use a mixture of two models
  – one which describes Hortonian flow (called SWAT TS or shortened to SWAT) and
  – one of which describes Variable Source Area Hydrology (called SWAT VSA or shortened to VSA)

• Each of the model has been calibrated individually to the observed data.

• The mixture model is based on weights that are dependent on time (here on calendar month)
Weighted Averaging Method

• The Weighted Averaging Method assigns weights to different outputs as follows:

\[ SSE_i = \sum_{t=1}^{T} \left[ \sum_{k=1}^{K} w_{k,i} f_k - Measured_t \right]^2 \]  

(2)

– Each output has a specific weighting value, \( w_k \), between 0 and 1.
– The sum of all weights equals unity.
– This method only outputs an expected value.
Finding the Weights

• DDS (and linear method) are used to find the mixtures weights as by minimizing the following equation:

\[ SSE_i = \sum_{t=1}^{T} \left[ \sum_{k=1}^{K} w_{k,i} f_k - Measured_t \right]^2 \]  

– Where: \( T = \) length of time, \( w_k \) is the weight for model \( k \), \( f_k \) is the output from model \( k \). \( i \) refers to the period of time. Index \( i \) goes over all data.
Weighting Schemes

• Scheme W1: Different weights for each Month
  – This method determines 72 different weights for each of the 72 months in the calibration period. This creates a better fit, but can not be used for future predictions.

• Scheme W2: Different weights for each calendar month
  – This method assigns 12 different weights to each of the twelve months in a year. This method can be extended into the future and gives us some insight into when one model outperforms the others.

• (Also analyzed weighting schemes W3) by season and W4) by recent precipitation
Model Studies

• Study 1: Model averaging for flow, sediment, TDP and PP with the SWAT-VSA and SWAT-TS models
  – For this study, the results for the previously calibrated models, SWAT-TS and SWAT-VSA are used. No new calibrations are performed.
  – Most prior work (by others) with mixture models has been done looking at flow only, but nutrients are included in this study
  – For each output, flow, sediment, TDP and PP, different weights are determined.
Does this improve Fit to calibration Data?

Yes!
Calibration Period Results: Mixture Model Outperforms Individual Models

Normalized SSE values for All Months

Key:
- W1: BMA weights
- W2: DDS weights
- W3: Linear weights
- W4: BMA Flow weights
- W5: DDS flow weights
- W6: Linear flow weights

- TDP
- PP

Individual models

Mixture Models
## Calibration Results for Study 1

- **Scheme 1:** Different weights for each cal.

<table>
<thead>
<tr>
<th>Flow</th>
<th>TS</th>
<th>VSA</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.85</td>
<td>0.84</td>
<td>0.86</td>
<td>0.89</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>2.29E+06</td>
<td>3.01E+06</td>
<td>2.35E+06</td>
<td>1.93E+06</td>
<td>1.93E+06</td>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>TDP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.68</td>
<td>0.72</td>
<td>0.71</td>
<td>0.76</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>3.63E+07</td>
<td>3.11E+07</td>
<td>3.21E+07</td>
<td>2.82E+07</td>
<td>2.82E+07</td>
<td>3.</td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.69</td>
<td>0.53</td>
<td>0.58</td>
<td>0.77</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>2.53E+08</td>
<td>3.80E+08</td>
<td>3.40E+08</td>
<td>1.84E+08</td>
<td>1.84E+08</td>
<td>2.</td>
<td></td>
</tr>
<tr>
<td>Sed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.79</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>7.48E+07</td>
<td>8.30E+07</td>
<td>7.79E+07</td>
<td>5.91E+07</td>
<td>5.91E+07</td>
<td>7.</td>
<td></td>
</tr>
</tbody>
</table>
Calibration Results for Study 1

- **Scheme 1: Different weights for each cal.**

<table>
<thead>
<tr>
<th>All_Months - individual constituent weights</th>
<th>Flow</th>
<th>TS</th>
<th>VSA</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.85</td>
<td>0.84</td>
<td>0.86</td>
<td>0.89</td>
<td>0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.29E+06</td>
<td>3.01E+06</td>
<td>2.35E+06</td>
<td>1.93E+06</td>
<td>1.93E+06</td>
<td></td>
<td>2.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TDP</td>
<td>0.68</td>
<td>0.72</td>
<td>0.71</td>
<td>0.76</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.63E+07</td>
<td>3.11E+07</td>
<td>3.21E+07</td>
<td>2.82E+07</td>
<td>2.82E+07</td>
<td></td>
<td>3.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP</td>
<td>0.69</td>
<td>0.53</td>
<td>0.58</td>
<td>0.77</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.53E+08</td>
<td>3.80E+08</td>
<td>3.40E+08</td>
<td>1.84E+08</td>
<td>1.84E+08</td>
<td></td>
<td>2.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sed</td>
<td>0.73</td>
<td>0.73</td>
<td>0.73</td>
<td>0.79</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.48E+07</td>
<td>8.30E+07</td>
<td>7.79E+07</td>
<td>5.91E+07</td>
<td>5.91E+07</td>
<td></td>
<td>7.</td>
</tr>
</tbody>
</table>
Optimized Weights make hydrologic sense. They weight VSA less in summer and indicate both processes important in Mar-May and Sept-Nov.
Note also that the weights differ for flow and two types of phosphorous (particulate PP and dissolved TDP).
Quantification of competing Process

• These results then help quantify the impact these two competing process on a very large scale (1200 km$^2$).
• Can complement with smaller field scale empirical study.
• Hard to see how else one can get this kind of large scale quantification of relative importance of competing processes.
The sum is better than the parts

• The mixture model gives a substantially better fit than either of the individual models (VSA or Hortonian Flow).

• This is not surprising since we would expect that under different weather conditions difference process would dominate.

• We are not aware that anyone has previously attempted to quantify the impact of two processes in this way.
Choice of Weighting Method

- The previous results are based on a simple weighting scheme.
- “Bayesian Averaging” is also a popular algorithm that we applied.
- The simple weighting scheme gave much better results than the Bayesian Averaging Method.
- The weights obtained by Bayesian averaged did not display any pattern that could be used to infer the relative importance of the competing hydrologic processes.
- Hence the Bayesian Averaging Analysis was a failure.
Using an alternative “Bayesian” weighting Scheme gives poor results

• Scheme 2: Different weights for each month

- It makes physical sense that the SWAT-VSA would be better in the Spring Months as with simpler weights

- BMA weights don’t really have any pattern as to when one model is better than the other.
Calibration Effort Required

• A major and time consuming effort in the preceding study was to calibrate both the VSA and conventional SWAT model to the historical data.

• We do this with a combination of optimization methods and “manual” calibration to get the best results.

• When trying to understand what is going on over large areas, good calibration of model parameters (and good algorithms) is essential because it is infeasible to get enough data to be able to select the best parameter values without calibration.
The mixture model is then used to do evaluations of Alternative **Management** Schemes for Reducing Phosphorous Loading
Conclusion on Algorithms

• It is worthwhile to find and use an efficient algorithm for your problem if your simulation model takes some time (e.g. minute or longer)

• Both single and multi objective optimization can be very useful.

• Algorithm efficiency is an important aspect of watershed analysis since there are many cases for calibration, management and uncertainty quantification where it is infeasible to do the analysis without highly efficient algorithms.
Conclusions on Mixture Models

• The process used here is able to give a quantitative estimate of the relative importance of two competing hydrologic processes.
• The same process could be used for two different model types (e.g. SWAT versus HSPF) and certainly to look at impacts of different processes
• The major computational effort is to calibrate the individual models and not to determine the weights.
Conclusions on Mixture Modeling

• We found the results for the simple weighting method were far preferable to Bayesian Averaging.

• The mixture model fit better on validation data, implying it is more accurate for prediction than either individual model.

• We could use the mixture model to get an improved estimate of the impact of management options.

• Hence the benefits are helpful both in terms of science (hydrologic processes) and management (BMP options)
Questions?
Conclusions

• The weighted average method tends to outperform the BMA method in terms of the expected value prediction. The weighted average method is limited because it does not provide a measure of uncertainty.

• The BMA method does often perform at least as well as the best individual model, but it seems like a lot of work to just get at least as good.
Conclusions

• The weights from the weighted average method tend to reflect the model performance better than BMA. These weights also seem to make sense in terms of when VSA hydrology would be an improvement.

• Weights derived for flow can be used for other model outputs.
Description

• Multi-model approaches simply combine the outputs of two or more models to achieve a better overall match to the measured data.
• The methods used in this study use weights for each model output in such a way that all weights sum to unity and are between 0 and 1.
  – For example, averaging the model outputs essentially assigns equal weights to all model outputs.
Weighted Averaging Method

• The Weighted Averaging Method assigns weights to different outputs as follows:

\[ \text{Weighted Output} = \sum_{k=1}^{K} w_k f_k \]

– Each output has a specific weighting value, \( w_k \), between 0 and 1.
– The sum of all weights equals unity.
– This method only outputs an expected value.
DDS and the constrained least squares technique

• Both DDS and the constrained least squares technique are used to find the weights as by minimizing the following equation:

$$SSE_i = \sum_{t=1}^{T} \left[ \sum_{k=1}^{K} w_{k,i} f_k - Measured_t \right]^2$$  \hspace{1cm} (2)

  — Where: $T =$ length of time, $w_k$ is the weight for model $k$, $f_k$ is the output from model $k$. $i$ refers to the period of time.

• The matlab function for the least squares technique is used here.
Weighting Schemes

• Scheme 1: Different weights for each Cal. Month
  – This method determines different weights for each of the 72 months in the calibration period. This creates a better fit, but can not be used for future predictions.

• Scheme 2: Different weights for each month
  – This method assigns different weights to each of the twelve months in a year. This method can be extended into the future and gives us some insight into when one model outperforms the others.

• Scheme 3: Different weights for each season
  – This method assigns different weights to each of the four seasons.
Calibration Results for Study 1

• Scheme 1: Different weights for each cal.

Normalized SSE values for All Months

Key:
- W1: BMA weights
- W2: DDS weights
- W3: Linear weights
- W4: BMA Flow weights
- W5: DDS flow weights
- W6: Linear flow weights

Legend:
- Flow
- TDP
- PP
Most scenarios predict a relatively large increase in Phosphorus export.

Only scenario close to no increase is eliminating P in manure
  - This is not a feasible option for the area
Outline

• The Cannonsville Watershed/Reservoir
• The SWAT 2005 Model
• SWAT vs SWAT-VSA
  • Calibration and Validation
• Model Weighting
• Conclusion
• Questions ?
Calibration

• Both of the models are calibrated first for flow, then sediment and finally phosphorous
• The calibration period is from Jan. 1994 to Dec. 1999
• Auto-calibration (DDS) and manual calibration techniques are used to get the best fit
## Results – Calibration Period

### Calibration Period, 1994 - 1999

<table>
<thead>
<tr>
<th></th>
<th>Flow</th>
<th>Sediment</th>
<th>TDP</th>
<th>PP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SWAT (Monthly)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R - Squared</td>
<td>0.85</td>
<td>.73</td>
<td>.7</td>
<td>.72</td>
</tr>
<tr>
<td>% Diff.</td>
<td>0.34</td>
<td>-3.3</td>
<td>-4.18</td>
<td>-6.22</td>
</tr>
<tr>
<td><strong>VSA (Monthly)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R - Squared</td>
<td>0.84</td>
<td>.73</td>
<td>.72</td>
<td>.53</td>
</tr>
<tr>
<td>% Diff.</td>
<td>0.44</td>
<td>-1.9</td>
<td>-4.77</td>
<td>-0.85</td>
</tr>
</tbody>
</table>

* Calibration plots are included at the end of the presentation if interested*
### Results – Validation Period

- **Validation Period, 1991 - 1993**

<table>
<thead>
<tr>
<th></th>
<th>Flow</th>
<th>TDP</th>
<th>PP</th>
<th>Sediment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SWAT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Monthly)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.97</td>
<td>0.90</td>
<td>0.696</td>
<td>0.84</td>
</tr>
<tr>
<td>% diff</td>
<td>5.33</td>
<td>-36.9</td>
<td>0.057</td>
<td>0.014</td>
</tr>
<tr>
<td><strong>VSA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Monthly)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.93</td>
<td>0.88</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>% diff</td>
<td>7.49</td>
<td>-40.6</td>
<td>-9.73</td>
<td>-2.84</td>
</tr>
</tbody>
</table>
Outline

• The Cannonsville Watershed/Reservoir
• The SWAT 2005 Model
• SWAT vs SWAT-VSA
• Calibration and Validation
• Model Weighting
• Conclusion
• Questions ?
Model Weighting

- Multi-model approaches are used to combine the outputs of two or more models to improve the overall prediction.
- Since there are uncertainties and weaknesses in all models, combining the outputs from different models can help eliminate the inadequacies and combine the strengths of different models.
- Different methods have been used to combine different models, such as simple averaging, weighted averaging, and more complex methods such as Bayesian model averaging.
What are the goals?

• Comparing the weighted average method with Bayesian Model Averaging (BMA).
  – The weighted average method is applied using DDS and the constrained least squares method.

• Determine when either SWAT-VSA or SWAT-TS performs better.
  – Yield an insight into when the addition of VSA is helpful.
What are the goals?

- Other studies have been limited to flow only.
- This study looks at using multi-model methods on outputs other than flow, i.e. Phosphorus and sediment.
  - Should the flow weights be used, or should new weights be determined?
- The study also looks at different weighting schemes for hydrologic models
Bayesian Model Averaging

• Bayesian model averaging (BMA) is a statistical scheme used to produce a probabilistic prediction.
• The intention is that this prediction will posses more skill than any of the individual outputs.
• Unlike the other two methods, BMA includes a measure of uncertainty in the prediction. This uncertainty is not used in this study.
Bayesian Model Averaging

• If we consider a quantity \( y \) to be the predicted variable, \( D \) is the training data with length \( T \), and \( f = [f_1, \ldots, f_k] \) is the ensemble of model predictions, then the \( p_k(y | f_k, D) \) is the conditional pdf of \( y \) given model prediction \( f_k \) and observational data set \( D \). Now, the BMA prediction pdf is as follows:

\[
\sum_{k=1}^{K} w_k = 1
\]

– Where \( w_k \) is the likelihood of model prediction \( f_k \) being correct, which is the weight assigned to the \( k^{th} \) model output and \( K \) is the number of models.
Bayesian Model Averaging

• The weights, $w_k$, and the variances, $\sigma_k$, for the BMA method are determined by maximizing the likelihood function (eq. 1).

• The likelihood function is maximized using the Expectation-Maximization Algorithm.
  – This is an iterative local search algorithm which starts with an initial guess and converges to an answer.
## Calibration Period Results

- **Scheme 1:** Different weights for each cal. month

### Flow

<table>
<thead>
<tr>
<th></th>
<th>TS</th>
<th>VSA</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2$</td>
<td>0.85</td>
<td>0.84</td>
<td><strong>0.86</strong></td>
<td><strong>0.89</strong></td>
<td><strong>0.89</strong></td>
</tr>
<tr>
<td>SSE</td>
<td>2.29E+06</td>
<td>3.01E+06</td>
<td><strong>2.35E+06</strong></td>
<td><strong>1.93E+06</strong></td>
<td><strong>1.93E+06</strong></td>
</tr>
</tbody>
</table>

### TDP

<table>
<thead>
<tr>
<th></th>
<th>TS</th>
<th>VSA</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
<th>W4</th>
<th>W5</th>
<th>W6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2$</td>
<td>0.68</td>
<td>0.72</td>
<td><strong>0.72</strong></td>
<td><strong>0.76</strong></td>
<td><strong>0.76</strong></td>
<td>0.72</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>SSE</td>
<td>3.63E+07</td>
<td>3.11E+07</td>
<td><strong>3.21E+07</strong></td>
<td><strong>2.82E+07</strong></td>
<td><strong>2.82E+07</strong></td>
<td>3.11E+07</td>
<td>2.85E+07</td>
<td>2.85E+07</td>
</tr>
</tbody>
</table>

### PP

<table>
<thead>
<tr>
<th></th>
<th>TS</th>
<th>VSA</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
<th>W4</th>
<th>W5</th>
<th>W6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2$</td>
<td>0.69</td>
<td>0.53</td>
<td><strong>0.70</strong></td>
<td><strong>0.77</strong></td>
<td><strong>0.77</strong></td>
<td>0.65</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>SSE</td>
<td>2.53E+08</td>
<td>3.80E+08</td>
<td><strong>3.40E+08</strong></td>
<td><strong>1.84E+08</strong></td>
<td><strong>1.84E+08</strong></td>
<td>2.84E+08</td>
<td>2.00E+08</td>
<td>2.00E+08</td>
</tr>
</tbody>
</table>

### Sediment

<table>
<thead>
<tr>
<th></th>
<th>TS</th>
<th>VSA</th>
<th>W1</th>
<th>W2</th>
<th>W3</th>
<th>W4</th>
<th>W5</th>
<th>W6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^2$</td>
<td>0.73</td>
<td>0.73</td>
<td><strong>0.74</strong></td>
<td><strong>0.79</strong></td>
<td><strong>0.79</strong></td>
<td>0.73</td>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>SSE</td>
<td>7.48E+07</td>
<td>8.30E+07</td>
<td><strong>7.79E+07</strong></td>
<td><strong>5.91E+07</strong></td>
<td><strong>5.91E+07</strong></td>
<td>7.58E+07</td>
<td>6.78E+07</td>
<td>6.78E+07</td>
</tr>
</tbody>
</table>
Calibration Period Results

- Scheme 2: Different weights for each month

![Normalized SSE values for Months graph](image)
Outline

• The Cannonsville Watershed/Reservoir
• The SWAT 2005 Model
• SWAT vs SWAT-VSA
• Calibration and Validation
• Model Weighting
• Conclusion
• Questions ?
The SWAT 2005 Model

• Soil and Water Assessment Tool
  – There have been several versions, but we are currently using the 2005 version

• The model is a continuous time, physically based watershed model developed by the USDA

• The model requires significant amounts of data and a large number of parameters
  – Climate data, topography data, soils data and land use information
Outline

• The Cannonsville Watershed/Reservoir
• The SWAT 2005 Model
• SWAT vs SWAT-VSA
• Calibration and Validation
• Model Weighting
• Conclusion
• Questions ?
The current SWAT-TS version is a replication of the SWAT 2000 model for the Cannonsville published by Tolson & Shoemaker (Jn. of Hydrology, 2007).

The SWAT-VSA model incorporates the model and file changes in the SWAT-TS model, as well as Variable Source Area (VSA) hydrology.

The main difference between the two models is the way in which the flow is distributed:

- Infiltration excess vs. Saturation excess.
SWAT-VSA

- In SWAT-VSA, VSA hydrology is modeled using wetness classes.
  - Wetness classes are used to determine how likely it is that an area will become saturated.
  - Our model uses ten wetness classes, with 1 being the driest, least likely to saturate and 10 being the wettest, most likely to saturate.
SWAT-VSA

• These wetness classes are then used to create HRUs
  – In the SWAT-VSA model, HRUs are unique combinations of wetness class and land use
  – Each wetness class has the necessary soil parameter values to run the model

• This has interesting implications for nutrient runoff
SWAT-VSA

• In SWAT-VSA, VSA hydrology is modeled using wetness classes
  – Wetness classes are used to determine how likely it is that an area will become saturated
  – Our model uses ten wetness classes, with 1 being the driest, least likely to saturate and 10 being the wettest, most likely to saturate
SWAT-VSA

• These wetness classes are then used to create HRUs
  – In the SWAT-VSA model, HRUs are unique combinations of wetness class and land use
  – Each wetness class has the necessary soil parameter values to run the model

• This has interesting implications for nutrient runoff
Validation Period Results (based on data to which parameters are not calibrated)

- **Scheme 2: Different weights for each month**

<table>
<thead>
<tr>
<th>Flow</th>
<th>RSQ</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.96</td>
<td>3.51E+05</td>
</tr>
<tr>
<td>VSA</td>
<td>0.95</td>
<td>9.49E+05</td>
</tr>
<tr>
<td>BMA</td>
<td>0.96</td>
<td>5.28E+05</td>
</tr>
<tr>
<td>DDS</td>
<td>0.97</td>
<td>3.50E+05</td>
</tr>
<tr>
<td>Linear</td>
<td>0.97</td>
<td>3.50E+05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TDP</th>
<th>RSQ</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.88</td>
<td>3.04E+07</td>
</tr>
<tr>
<td>VSA</td>
<td>0.91</td>
<td>3.51E+07</td>
</tr>
<tr>
<td>BMA</td>
<td>0.90</td>
<td>3.17E+07</td>
</tr>
<tr>
<td>DDS</td>
<td>0.92</td>
<td>3.00E+07</td>
</tr>
<tr>
<td>Linear</td>
<td>0.92</td>
<td>3.00E+07</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PP</th>
<th>RSQ</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.70</td>
<td>6.29E+07</td>
</tr>
<tr>
<td>VSA</td>
<td>0.73</td>
<td>5.54E+07</td>
</tr>
<tr>
<td>BMA</td>
<td>0.72</td>
<td>5.63E+07</td>
</tr>
<tr>
<td>DDS</td>
<td>0.74</td>
<td>5.30E+07</td>
</tr>
<tr>
<td>Linear</td>
<td>0.74</td>
<td>5.30E+07</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sed</th>
<th>RSQ</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.80</td>
<td>1.80E+07</td>
</tr>
<tr>
<td>VSA</td>
<td>0.73</td>
<td>2.62E+07</td>
</tr>
<tr>
<td>BMA</td>
<td>0.75</td>
<td>2.33E+07</td>
</tr>
<tr>
<td>DDS</td>
<td>0.81</td>
<td>1.76E+07</td>
</tr>
<tr>
<td>Linear</td>
<td>0.81</td>
<td>1.76E+07</td>
</tr>
</tbody>
</table>
Validation Period Results (based on data to which parameters are not calibrated)

- Scheme 2: Different weights for each month

<table>
<thead>
<tr>
<th>Flow</th>
<th>RSQ</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.96</td>
<td>3.51E+05</td>
</tr>
<tr>
<td>VSA</td>
<td>0.95</td>
<td>9.49E+05</td>
</tr>
<tr>
<td>BMA</td>
<td>0.96</td>
<td>5.28E+05</td>
</tr>
<tr>
<td>DDS</td>
<td>0.97</td>
<td>3.50E+05</td>
</tr>
<tr>
<td>Linear</td>
<td>0.97</td>
<td>3.50E+05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TDP</th>
<th>RSQ</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.88</td>
<td>3.04E+07</td>
</tr>
<tr>
<td>VSA</td>
<td>0.91</td>
<td>3.51E+07</td>
</tr>
<tr>
<td>BMA</td>
<td>0.90</td>
<td>3.17E+07</td>
</tr>
<tr>
<td>DDS</td>
<td>0.92</td>
<td>3.00E+07</td>
</tr>
<tr>
<td>Linear</td>
<td>0.92</td>
<td>3.00E+07</td>
</tr>
</tbody>
</table>
Validation Period Results (based on data to which parameters are not calibrated)

- Scheme 2: Particulate Phosphorous and Sediment each

<table>
<thead>
<tr>
<th>PP</th>
<th>RSQ</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.70</td>
<td>6.29E+07</td>
</tr>
<tr>
<td>VSA</td>
<td>0.73</td>
<td>5.54E+07</td>
</tr>
<tr>
<td>BMA</td>
<td>0.72</td>
<td>5.63E+07</td>
</tr>
<tr>
<td>DDS</td>
<td>0.74</td>
<td>5.30E+07</td>
</tr>
<tr>
<td>Linear</td>
<td>0.74</td>
<td>5.30E+07</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sed</th>
<th>RSQ</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>0.80</td>
<td>1.80E+07</td>
</tr>
<tr>
<td>VSA</td>
<td>0.73</td>
<td>2.62E+07</td>
</tr>
<tr>
<td>BMA</td>
<td>0.75</td>
<td>2.33E+07</td>
</tr>
<tr>
<td>DDS</td>
<td>0.81</td>
<td>1.76E+07</td>
</tr>
<tr>
<td>Linear</td>
<td>0.81</td>
<td>1.76E+07</td>
</tr>
</tbody>
</table>
Does this improve prediction of Validation Data?

• Yes!
Equations for mixture model

The month-specific weights allow incorporation of the statistical distribution of weather data in each month.
The Cannonsville Watershed/Reservoir

• Approximately 1200 square kilometers in the Catskill Region of Upstate New York

• The watershed is primarily forest and agricultural land, approximately 0.5% is urban
  – Dairy cow farming is the dominant form of agriculture

■ Currently provides NYC with unfiltered drinking water
** This analysis cannot be duplicated by small scale studies.

- It is highly worthwhile to do small scale field studies with intensive data collection designed to assess the impact of weather on these competing processes.
- However, to assess what is going on in a large watershed (1200km\(^2\) in this case) it is too expensive to collect intensive data over the whole watershed.
- Hence this mixture model approach is an important tool (in conjunction with small scale field studies) to understand when variable source area is significant over a large watershed.
Illustration of VSA hydrology (and its impact of flow and contaminant transport)

Agricultural field conditions in August (Dry)  *  Agricultural field conditions in March (Wet)  *

* Soil and Water Lab, BEE, Cornell University, http://soilandwater.bee.cornell.edu/research/VSA/extension.html
What is VSA hydrology?

- VSA hydrology is the concept that run-off generating areas will vary in size and location overtime.
  - Runoff generating areas can change depending on weather, season, topography and vegetation.

- Saturation Excess Overland Flow is the main mechanism behind VSA hydrology
  - Saturation excess overland flow occurs when the soil becomes saturated and any additional precipitation results in run-off
VSA Hydrology

1. Precipitations falls onto the surface and infiltrates the soil

2. As water continues to infiltrate the soil, the watertable rises close to the surface

3. With increased precipitation, the watertable rises above the surface. Any additional rain will create runoff

4. More precipitation continues to saturate the soil, creating a larger runoff area

5. Water table recedes when precipitation ends, decreasing the size of the saturation area

6. Saturated areas are eliminated once the watertable falls below the surface