Kentucky Nutrient Model Calibration Tutorial

Developed by the
Kentucky Water Resources Research Institute

Prepared for the
Kentucky Department of Environmental Protection (KYDEP)
Division of Water

September 30, 2014

Contents

List of Figures

List of Tables

1.0 Calibration Tutorial Introduction.

This tutorial has been developed to illustrate how to calibrate the KYNM using two different example watersheds: 1) the Upper Chenoweth Run watershed and 2) the Middle Chenoweth Run watershed. Both models are provided with the KYNM Users Manual. The Upper Chenoweth Run does not contain any significant point sources, while the Middle Chenoweth Run watershed contains several point sources.

Calibrating a model of a dynamic system is often not a simple linear process but an iterative one. The following discussion outlines "steps" in calibrating a model, but it must be understood that the phenomena being modeled are interconnected and affect one another, hence it will be likely that these steps may need to be revisited and possibly adjusted, until the desired calibration quality is achieved.

1.1 Hydrologic Calibration

Calibration of the hydrologic model parameters will typically occur in four basic steps: baseflow calibration, non-point source flow calibration, rainfall calibration, and point source calibration. Before the calibration process is begun, the data for each of these menus must first be entered into the program.

1.1.1 Hydrologic Data

Before the hydrologic calibration process is begun, the user must first input the basic hydrologic data along with the assumed values of the basic model parameters.

The basic hydrologic input data include:

1) Observed daily streamflow data.

2) Observed daily rainfall data.

3) Observed or assumed discharges from all permitted facilities entered either as:

a) daily observed values.

b) monthly observed values.

c) annual observed or assumed values.

4) The percentage of hydrologic soil types (i.e. A, B, C or D) associated with each landuse.

5) The areas of each of the major landuses for both MS4 and non-MS4 areas and the assumed normal, maximum, and minimum hydrologic curve numbers for each landuse. Prior to calibration, the user should simply use the provided literature values for the hydrologic curve numbers. These can be adjusted later if needed during model calibration.

1.1.2 Hydrologic Parameters

The user should make sure the basic calibration parameters are set equal to their default values before beginning model calibration (see Table 1.1).

Table 1.1 Default Values for Hydrologic Parameters

Symbol	Groundwater Parameter	Default Value	Spreadsheet Cell
α	Baseflow Filter Coefficient	.9	G30
BFI_{max}	BFI for perennial streams with porous aquifers	.8	G33
BFI_{max}	BFI for ephemeral streams with porous aquifers	.5	G33
BFI_{max}	BFI for perennial streams with rock aquifers	.25	G33
β_f	Rising Limb Percent	.3	H30
T_t	T_L Lag Time	1	H33
K_g	K_g Linear Reservoir Coefficient	.2	I30
GAF_G	GAF_G Global Adjustment Factor: Groundwater	1	I33
	Non Point Source Parameters		
K_{NPS}	Linear Reservoir Coefficient	.5	S26
GAF_{NPS}	Global Adjustment Factor	1	T26
	Point Source Parameters		
GAF_{PS}	Global Adjustment Factor	1	X32

1.1.3 Baseflow Calibration

Two different methods have been provided for developing a baseflow or groundwater time series: 1) baseflow separation or 2) baseflow generation. The baseflow generation method is reserved for those applications where no observed or synthesized streamflow time series is available. In such a case, no baseflow calibration is possible since there is no time series to compare against. As a result, in most cases the baseflow separation method will be used.

When providing an observed (or synthesized) streamflow time series as an input (i.e. COLUMN E), the user will always use the baseflow separation method. Two different options are available: 1) the filter method or 2) the heuristic method. For most instances, we will normally use the Filter Method.

Before beginning the calibration, the user should first select a baseflow calibration option by entering either a "1" or "2" in CELL J30, for the filter method or the heuristic method, respectively.

Once the point source values have been entered (i.e. COLUMN X), and the daily runoff from the various landuses have been calculated (i.e. COLUMN AW) the user should then adjust the baseflow parameters (i.e. G30 and G33 when using the filter option or H30 and H33 when using the heuristic method) until the volume of predicted surface runoff (i.e. CELL AW 36) matches that of the separated runoff (i.e. CELL AV 36).

1.1.4 Guidance for the Filter Method

There are two parameters that control the output from this method: the filter parameter α and the baseflow index *BFI max*. Both parameters α and BFI_{max} jointly determine what fraction of the observed streamflow is attributed to baseflow. With a fixed BFI_{max}, lower values of α will attribute a larger proportion of the observed streamflow to baseflow; similarly with a fixed α, higher values of BFI_{max} will attribute a larger proportion of the observed streamflow to baseflow. However, BFI_{max} is the stronger of the two at assigning all of the observed streamflow to baseflow: if BFI_{max} is set equal to 1, the computed baseflow

will exactly equal the observed streamflow regardless of the value of α (α and BFI max cannot both be set to 1 or we get division by zero). If α is set to 0, then the computed baseflow will equal the observed streamflow multiplied by BFI_{max}, leaving BFI_{max} to again assign the final proportion. The α parameter is the stronger of the two at the other extreme of the parameter values. If α is set to 1, then the computed baseflow is set equal to the previous time step baseflow, regardless of the value of BFI_{max}; because the minimum computed baseflow cannot be greater than the observed streamflow, this means that the computed baseflow will be numerically equal to the lowest previous observed streamflow. If BFI_{max} is set to 0, then the computed baseflow is equal to the previous time step baseflow multiplied by α, leaving α to assign the final proportion of the previous baseflow.

During the process of model calibration, the parameter α (i.e. CELL G30) will typically vary between 0.8 to 0.98 while the parameter BFI_{max} (CELL G33) will typically vary between 0.2 and 0.9.

For a given set of filter coefficients, the model will generate a time series of groundwater values (COLUMN J). By subtracting the sum of these values (CELL J36) along with the sum of the point source values (CELL X36) from the sum of the flows from the observed streamflow time series (CELL E36), an estimate of the total volume of runoff from non-point sources can be calculated (i.e. CELL AV36). Ideally, this volume should match the total runoff volume from non-point sources as predicted by the model (i.e. CELL AW36). If the separated volume (i.e. CELL AV36) is greater than the predicted volume (CELL AW36), then we have diverted more of the streamflow to runoff than is being reasonably estimated by landuse and rainfall data (or conversely, we have under predicted the amount of streamflow that should be diverted to baseflow). This can be corrected by lowering the filter coefficient α and/or raising the BFI. This concept is illustrated in Figures 1.1 and 1.2 and summarized in Table 1.2.

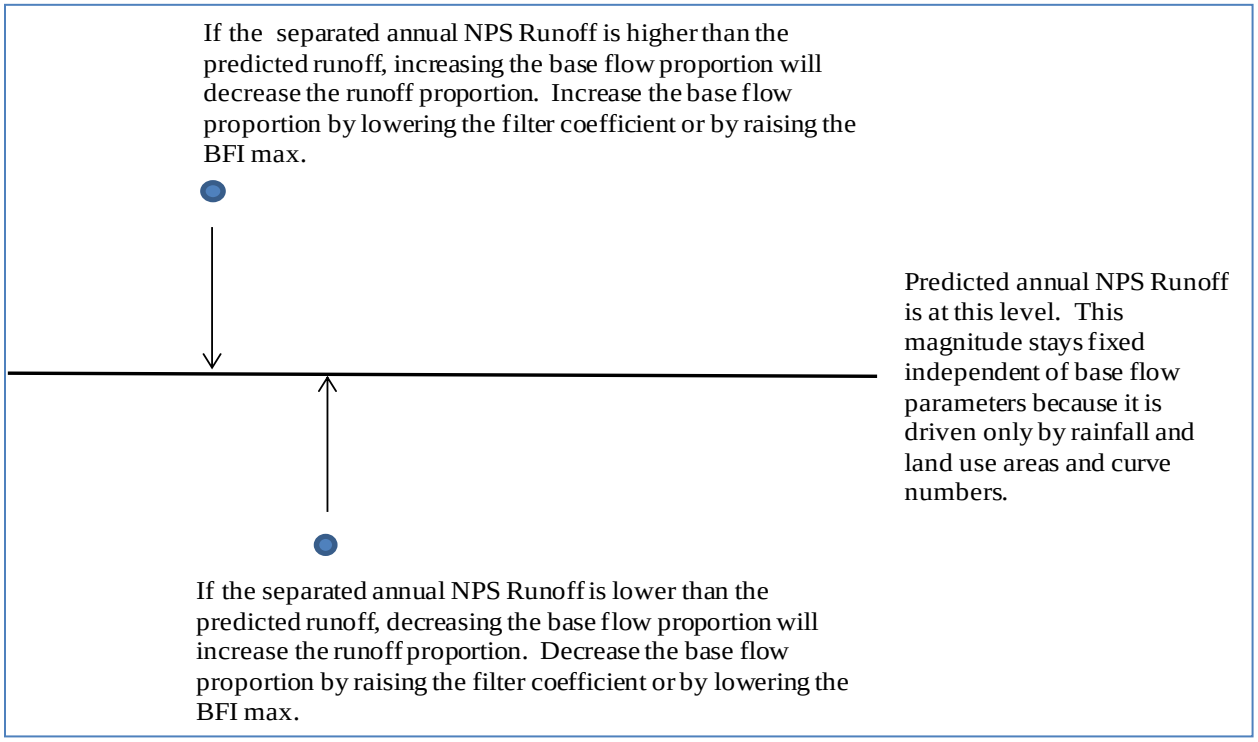

Figure 1.1 Illustration of the Baseflow Calibration Strategies

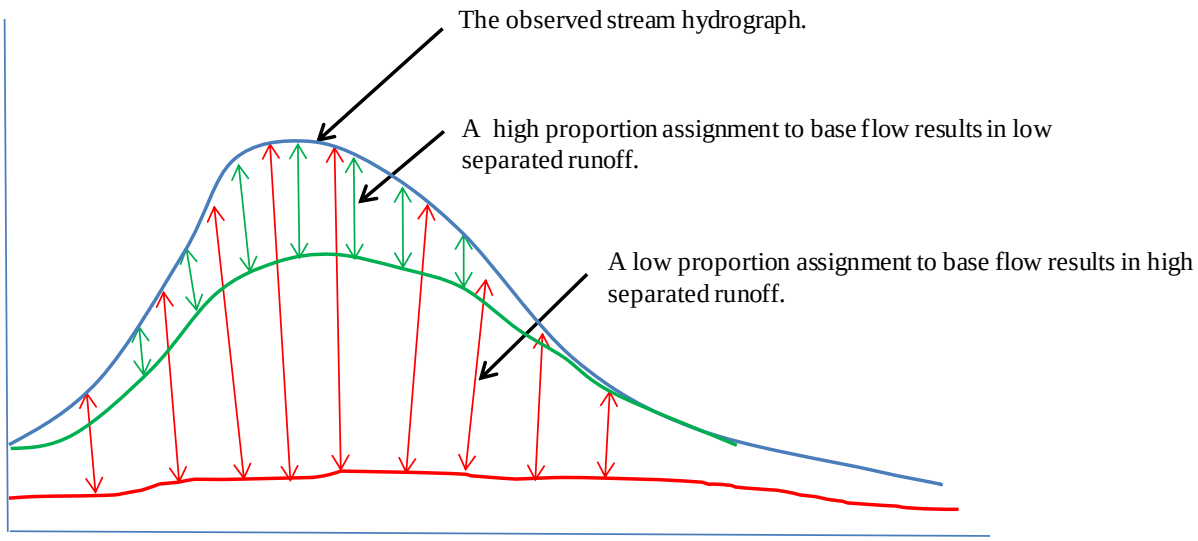

Figure 1.2 Illustration of the Impacts of Baseflow Calibration Strategies

Table 1.2 General Guidelines for Baseflow Calibration using the Filter Option

Condition	Evidence	Action
Too little flow is being diverted to baseflow	CELL AV36 > CELL AW 36	Decrease α Increase BFI max
Too much flow is being diverted to baseflow	CELL AV36 < CELL AW 36	Increase α Decrease BFI max

1.1.5 Non Point Source Calibration

Once the baseflow separation has been performed, we will normally assume that the volume of the predicted non-point source flows are essentially correct. However, if we plot the observed and predicted hydrographs (i.e. COLUMN M vs COLUMN N) it is likely that the shapes of the hydrographs and the associated peaks will not exactly match (see Figure 1.3). These characteristics can be adjusted using two additional calibration parameters: 1) the linear storage coefficient K (i.e. CELL S26) , and 2) the global runoff adjustment factor (i.e. CELL T26). General guidance on each of the parameters is provided in Table 1.3. The user should always try to correct the problem using K before trying to make any adjustments with GAF_{NPS}. Remember, if one adjusts GAF_{NPS}, this will affect the mass balance between the total observed and predicted runoff from the non-point sources, and thus could either increase or decrease that associated error.

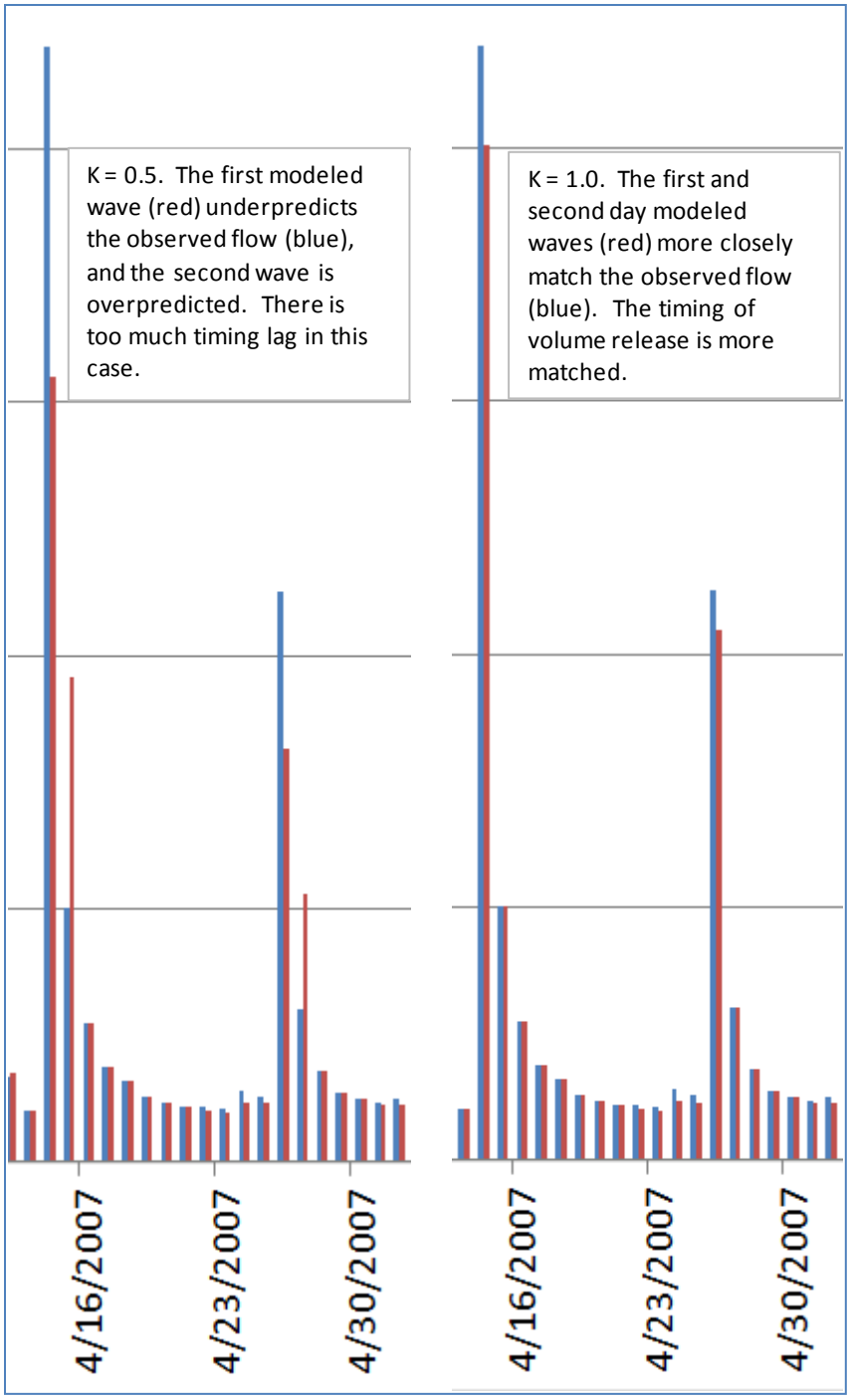

Figure 1.3 Illustration of Timing Differences between Observed and Predicted Time Series

Table 1.3 General Guidelines for Non-Point Source Hydrologic Parameter Calibration

Condition	Evidence	Action	Limits
Predicted peaks are greater than the observed peaks	COLUMN N > COLUMN M	Decrease K Decrease GAFNPS	K = 1, or $0 < K \leq 0.5$ $0.9 < GAFNPS < 1.1$
Observed peaks are greater than the predicted peaks	COLUMN M > COLUMN N	Increase K Increase GAFNPS	K = 1, or $0 < K \leq 0.5$ $0.9 < GAFNPS < 1.1$

While one can iteratively change the values of K and GAFNPS and then look at a plot of the predicted and the observed streamflow time series, this can be especially tedious when working with a long time series (e.g. one year). To help the user in the process, five calibration statistics have been provided in the KYNM to help guide the user in the calibration process. These five statistics are defined below. The primary two statistics with associated quantitative targets in the KYNM are the Nash Sutcliffe Coefficient and the percent bias. General guidance on desired ranges for these two statistics are provided in Table 1.4 below.

Table 1.4 Hydrologic Calibration Statistics

Statistic	Model CELL	Acceptable Range	Very Good
Nash Sutcliffe Coefficient	N32	>.50	> .75
Percent Bias	N31	< 25%	< 10%

Armed with these statistics, the user can then systematically make changes in K (or GAF_{NPS}) until the model values of the model parameters approach or satisfy these ranges. Once this process is completed, the user should still examine the actual time series to observe the general performance of the model prediction. The general goal will be for the predicted and observed time series to generally match. Individual daily deviations can sometimes be due not to the accuracy of the runoff prediction, but due to errors in the rainfall data or undocumented point source discharges such as a combined sewer overflow or a sanitary sewer overflow. Potential adjustments for these situations can be addressed either by rainfall calibration or point source calibration.

Equation 1.1 Nash Sutcliffe Efficiency

The Nash-Sutcliffe efficiency (NSE) is a normalized statistic that determines the relative magnitude of the residual variance ("noise") compared to the measured data variance ("information") (Nash and Sutcliffe, 1970). NSE indicates how well the plot of observed versus simulated data fits the 1:1 line. NSE ranges from negative infinity to 1.0. A NSE of 0 indicates that the mean of the observed data is as good a predictor of the data as the model. A negative NSE predicts worse than the mean of the observed data. A perfect prediction results in a NSE of 1.0.

$$E = 1.0 - \frac{\sum_{i=1}^{N} \left(Q_{obs,i} - Q_{sim,i}\right)^2}{\sum_{i=1}^{N} \left(Q_{obs,i} - \bar{Q}_{obs}\right)^2}.$$

Where $Q_{obs,i}$ is the *i*th observation for the constituent being evaluated, $Q_{sim,i}$ is the *i*th simulation for the constituent, and N is the total number of observations.

Equation 1.2 Percent Bias

Percent bias (PBIAS) measures the average tendency of the simulated data to be larger or smaller than their observed counterparts (Gupta et al., 1999). The optimal value of PBIAS is 0.0, with low-magnitude values indicating accurate model simulation. Positive values indicate model underestimation bias, and negative values indicate model overestimation bias (Gupta et al., 1999).

$$\text{PBIAS} = \left[\frac{\sum_{i=1}^{n} \left(Y_i^{obs} - Y_i^{sim}\right) * (100)}{\sum_{i=1}^{n} \left(Y_i^{obs}\right)} \right]$$

Where Y_i^{obs} is the *i*th observation for the constituent being evaluated, Y_i^{sim} is the *i*th simulated value for the constituent being evaluated, and n is the total number of observations.

Equation 1.3 Root Mean Squared Error (RMSE) and the RMSE Observations Standard Deviation Ratio (RSR)

RSR incorporates the benefits of error index statistics and includes a scaling/normalization factor, so that the resulting statistic and reported values can apply to various constituents. RSR varies from the optimal value of 0, which indicates zero RMSE or residual variation and therefore perfect model simulation, to a large positive value. The lower RSR, the lower the RMSE, and the better the model simulation performance.

$$\text{RSR} = \frac{\text{RMSE}}{\text{STDEV}_{obs}} = \frac{\left[\sqrt{\sum_{i=1}^{n} \left(Y_i^{obs} - Y_i^{sim}\right)^2}\right]}{\left[\sqrt{\sum_{i=1}^{n} \left(Y_i^{obs} - Y^{mean}\right)^2}\right]}$$

where Yi^{obs} is the *i*th observation for the constituent being evaluated, Yi^{sim} is the *i*th simulated value for the constituent being evaluated, Y^{mean}

Equation 1.4 Coefficient of Determination

In statistics, the coefficient of determination, denoted R^2, indicates how well data fit a statistical model. It is a statistic used in the context of models whose main purpose is either the prediction of future outcomes or the testing of hypotheses, on the basis of other related information. It provides a measure of how well observed outcomes are replicated by the model, as the proportion of total variation of outcomes explained by the model.

$$R^2 = \left\{ \frac{\sum_{i=1}^{N} \left(Q_{obs,i} - \bar{Q}_{obs}\right)\left(Q_{sim,i} - \bar{Q}_{sim}\right)}{\left[\sum_{i=1}^{N} \left(Q_{obs,i} - \bar{Q}_{obs}\right)^2\right]^{0.5} \left[\sum_{i=1}^{N} \left(Q_{sim,i} - \bar{Q}_{sim}\right)^2\right]^{0.5}} \right\}^2$$

Where $Q_{obs,i}$ is the *i*th observation for the constituent being evaluated, $Q_{sim,i}$ is the *i*th simulation for the constituent, $\bar{Q}_{obs}$ is the mean of the observed data, $\bar{Q}_{sim}$ is the mean of the simulated data, and N is the total number of observations.
is the mean of observed data for the constituent being evaluated, and *n* is the total number of observations.

1.1.6 Rainfall Calibration

One should recognize that the daily rainfall values input in the model in COLUMN C may not exactly represent the average of the rainfall across the entire basin, since such values are typically synthesized from either one rainfall gage or a weighted average of several gages. As a result, it is possible that "observed rainfall" input in COLUMN C may either underestimate or overestimate the actual rainfall that caused the observed streamflow in COLUMN E. This will result in a predicted streamflow (i.e. COLUMN N) that either underestimates or overestimates the observed streamflow (See Figure 1.4). Both cases are summarized in Table 1.5.

Table 1.5 General Guidelines for Rainfall Calibration

Situation	Cause	Calibration Steps
Observed streamflow much greater than predicted streamflow	Underestimated observed rainfall	Increase the daily rainfall amount in COLUMN C for the day of interest
Predicted streamflow much greater than observed streamflow	Overestimated observed rainfall	Decrease the daily rainfall amount in COLUMN C for the day of interest

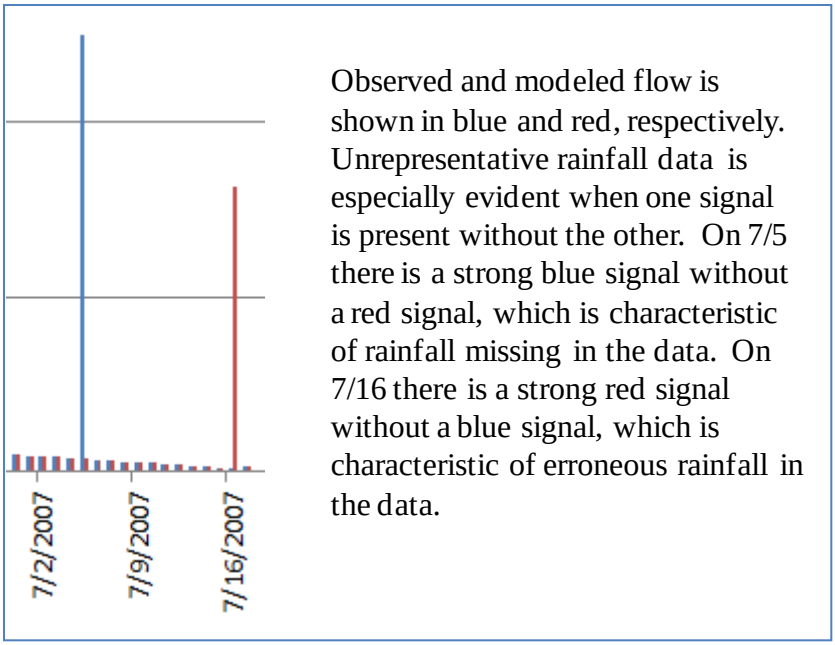

Figure 1.4 Illustration of Rainfall Data Screening

1.1.7 Point Source Discharge Calibration

Once the baseflow, surface runoff parameters, and rainfall data have been calibrated, the user should next inspect the observed and predicted time series to see if there any individual days in which the observed streamflow is significantly larger than the predicted daily streamflow values. Normally, this will be due to a rainfall error, which can be corrected as discussed above. However, in some cases, it is possible the deviation may be due to an undocumented or poorly documented SSO or CSO event. For example, we may know that a SSO or CSO event occurred on that date, but we may not know the exact volume of that event. In such cases, the user may try entering a SSO flow for that day in the general point source menu (i.e. COLUMN AG). The user may try adjusting this value until either the calibration statistics improve or the observed and predicted daily streamflows now match.

Alternatively, the user may make a small tweak to the point source flows by using the Point Source Global Adjustment Factor GAF_{PS}. Such an adjustment may be necessary to account for any unaccounted for point source discharges. Alternatively, by entering in a number less than 1, the user may account for any water withdrawals from the system.

1.1.8 Model Validation

Once the model has been calibrated to an acceptable level of performance (i.e. Table 1.4), the model and the associated calibrated parameters should be validated using another set of rainfall and streamflow data. For example, if we have used rainfall, streamflow, and point source data from 2007 in calibrating the model, we may then decide to use rainfall, streamflow, and point source data from 2008. If using the same calibrated model parameters, we are able to achieve relatively good model performance using the 2008 data - that is we still end up with good calibration statistics and the observed and predicted hydrographs seem to correlate well together, then we can have some confidence that are model parameters will be good not just for 2007, but also for other data sets. If our validation results are poor,

then we may want to consider adjusting the model parameters so that the net performance of the model for both years is relatively good. In order to do this, we might have to iterate back and forth between the two models (e.g. the one using the 2007 data and the one using the 2008 data). In some cases you may be able to calibrate both models separately and then use an average of the associated model parameters. Regardless of the approach employed, you will always want to validate your model before using it to make management decisions.

2.0 Hydrology Calibration Application to Upper Chenoweth Run Watershed (Jeffersontown)

The daily rainfall values for the Upper Chenoweth Run were obtained from the Louisville International Airport, NOAA weather station NCDC 93821. These values have been entered in COLUMN C in the model. The daily streamflow values for Upper Chenoweth Run (which drains Jeffersontown) were obtained from USGS gaging station 03298135. These values have been entered in COLUMN E in the model. The assumed values for the non-point data for the Upper Chenoweth Run are provided in Table 2.1.

Table 2.1 Non Point Source Parameter Values for Upper Chenoweth Run Watershed - Jeffersontown

Landuse	Area	Curve Number	Tot. Nitrogen	Tot. Phosphorus
Urban	(acres)		EMC (mg/L)	EMC (mg/L)
Barren Land	0.0	N/A	N/A	N/A
Residential	1608.3	80	3.76	0.555
Commercial	878.7	93	6.08	0.435
Industrial	325.5	91	2.9	0.305
Recreational				
Parks	0.0	N/A	N/A	N/A
Golf Course	0.0	N/A	N/A	N/A
Natural				
Forest	494.6	74	0.51	0.015
Grassland	5.1	77	2.8	0.08
Agriculture				
Pasture	113.0	77	5.09	0.61
Row Crops	39.7	84	13.89	1.025
Silvicuture	0.0	N/A	N/A	N/A
Hydraulic				
Open Water	0.0	N/A	N/A	N/A
Wetlands	0.0	N/A	N/A	N/A
Septic Systems			**lb/acre/day**	**lb/acre/day**
F. Septic Sys.	18.8	80	0.18	0.02
	Total Area (ac)	Average CN	Tot Load (lbs)	Tot Load (lbs)
	3484	83	1420	128

2.1 Baseflow Calibration

Using an initial value of α = .9 and BFI = .8, the **separated** annual runoff (CELL AV36) is 3,030 acre feet. Under the given rainfall data, land use data, and literature curve numbers, the **predicted** annual runoff (CELL AW36) is 2,431 acre feet. In accordance with the prior guidance, the **separated** runoff under the currently attempted base flow separation is too high and so therefore should be lowered, which means the base flow should be increased. The base flow can be increased by raising the BFI_{max} from 0.8 to 0.9. Under the new parameters, the observed annual runoff becomes 2,099 acre feet, which has dropped a fair amount below the 2,431 bench line. Now the runoff has become too low due to apportioning too much base flow. The base flow can now be lowered by either lowering the BFI_{max} back down or by raising the filter parameter. Taking the latter approach and raising the filter parameter from 0.9 to 0.925 the annual observed runoff becomes 2,404 acre feet. These annual runoff amounts are a good match so we leave base flow calibration for now, although it could be revisited.

Another point to note is that in this process of adjusting the base flow to match runoff volumes, the total annual flow volumes become matched as well: observed flow of 7,323 acft (i.e. CELL M36) vs predicted flow of 7,357 acft (i.e. CELL N36).

2.2 Non Point Source Calibration

After inputting rainfall and land use areas and setting the best known curve number values, the main overall hydrology calibration parameters are the storage coefficient, K, (CELL S26) and the Global Adjustment Factor for the Non Point Sources, GAF_{NPS}, (CELL T26). The storage coefficient can assume a value between 0 and 0.5, or 1. If no time lag is desired, the storage coefficient should be set equal to 1. To begin, the storage coefficient is assumed to be 0.5 and the global adjustment factor GAF_{NPS} is set equal to 1. The main calibration statistics to be used in adjusting the timing of the hydrograph are the coefficient of determination, R^2, and the Nash Sutcliffe model efficiency coefficient, NS, which are initially at 0.77 and 0.71 respectively (CELLS N28 and N32). These values are both satisfactory values; however, it may be possible to still achieve a better calibration. Since this is more of a small headwaters type basin, a storage coefficient of 1 may be more appropriate than the 0.5 value. Setting K = 1 increases the R^2 to 0.86 and the NS to 0.75. Next we observe that lowering the GAF_{NPS} to 0.9, 0.8, and 0.5 leads to improvements in the NS of 0.81, 0.86, and 0.91 respectively. However, since the annual volumes are so well matched, lowering the GAF does not seem like the best approach to achieve better calibration, but it does point toward some phenomenon that should be investigated. In addition to improving quantitative calibration measures, it is equally important to visually inspect plots of predicted and observed data to determine the quality of the modeling. Figure 2.1 below shows a portion of a modeled versus observed plot of the total flows.

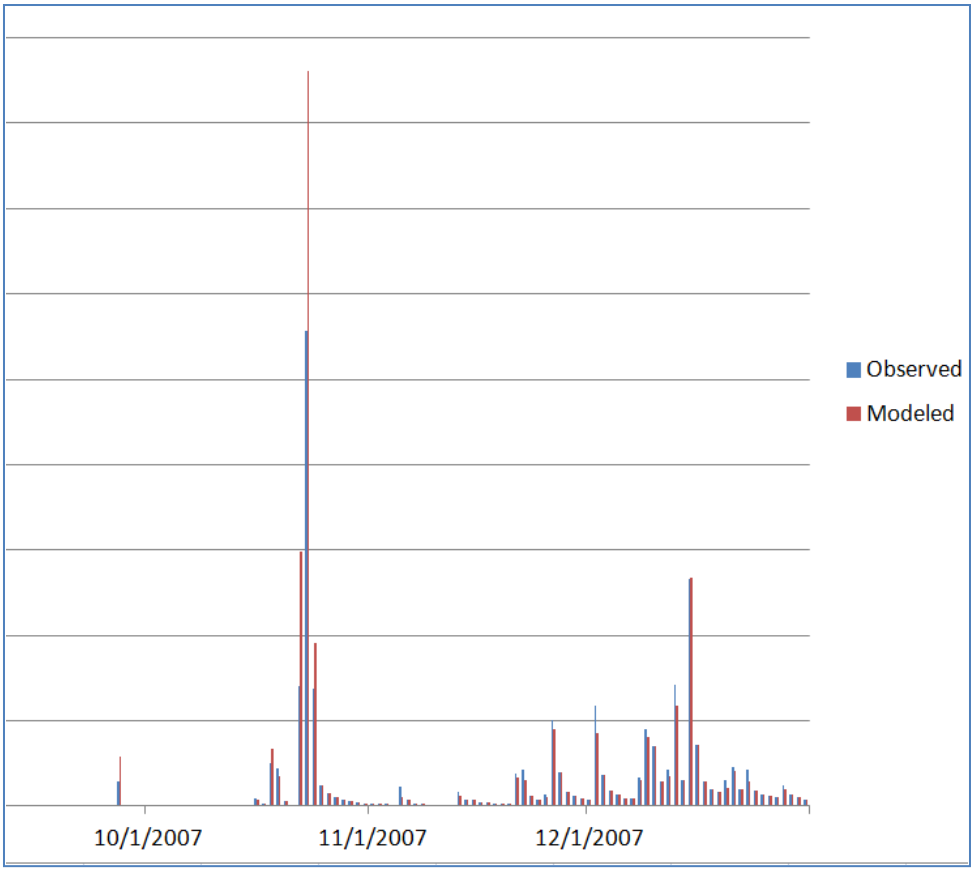

Figure 2.1 Observed and Modeled Daily Streamflows Showing Peak Mismatch Outlier

2.3 Rainfall Calibration

Upon investigation of the total flows plot, it was discovered that there is one day, October 23, 2007, for which the rainfall data is 3.76 inches resulting in a modeled flow of 430 cfs compared to the observed flow of 278 cfs on this day (see ROW 335 of COLUMN E). This one outlier is large enough that it is the reason the NS improves with lower GAF_{NPS}; the lower GAF_{NPS} reduces this high flow and reduces the impact of this point as an outlier. It is likely that the rainfall actually occurring on this day was somewhat less than the 3.76 inches recorded in Louisville. Indeed, looking at another rainfall station in the watershed, near Crestwood, the recorded rainfall on this day was 2.25 inches. After making this rainfall data substitution for this one day, the R^2 becomes 0.88 and the NS becomes 0.88 as well.

2.4 Review of Calibration Results

By making an adjustment to the rainfall, our total annual volume match will now be slightly affected (under the rainfall error hypothesis it was only matching due to offsetting errors). As a result, we will make one more calibration iteration starting back at the base flow. The separated runoff is now 2,404 acre feet while the predicted runoff is 2,091 acre feet, a fair amount high. Therefore we need to increase the base flow slightly. Taking the filter parameter back down to 0.90, we get again an excellent match: the separated runoff is 2,099 acre feet and the predicted runoff is 2,091 acre feet. The total annual volumes also have an excellent match of 7,323 observed acre feet (i.e. CELL M36) compared to 7, 322

predicted acre feet (i.e. CELL N36) The R^2 coefficient is now 0.89 (i.e. CELL M28) and the NS coefficient is now 0.89 (i.e. CELL N32). These represent very good calibration statistics.

The user should also check to see if there is any bias in the model results (i.e. CELL N31). A positive bias would indicate that the model is over-predicting the observed results and the a negative bias would indicate the model is under-predicting the observed results. In our case, the model bias is +1%, which is significantly less than 10% which represents an excellent calibration.

Finally, as one last check, the user should always plot the total observed hydrograph time series against the predicted hydrograph time series (see Figure 2.2) just to confirm the two series are fairly consistent. In our case, the two time series are fairly good. In some cases we are over-predicting the daily peaks and in other cases we are under-predicting the peaks. Our baseflow values are very consistent.

2.5 Model Validation

To make sure that the model is not "over-calibrated" to the peculiar data set (i.e. rainfall and streamflow from 2007), the final calibrated parameters should be validated with another data set (e.g. rainfall and streamflow for 2008). Application of the 2008 data to our calibrated model yields a total volume difference between the observed and predicted of less than two percent.

The result for 2008 hydrology validation is a total annual volume difference of less than two percent, (i.e. CELL M36 and CELL N36) and an R^2 of 0.94 (i.e. CELL N28) and a NS coefficient of 0.84 (i.e. CELL N32). The model bias (i.e. CELL N31) is only 1.7% percent. A review of the observed and predicted time series (i.e. Figure 2.3) also shows a good match. Together, these results indicate a very good model calibration. Thus we can have some assurance that are model parameters are reflective of the actual conditions in the watershed.

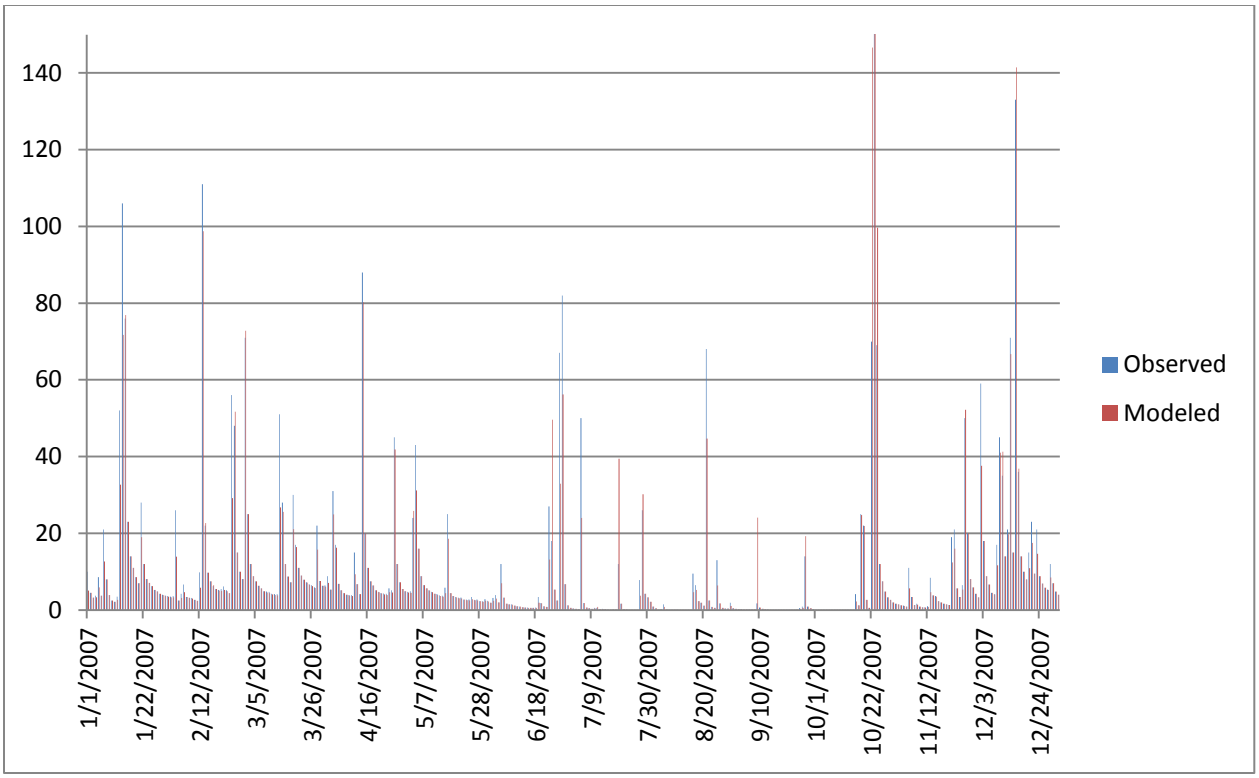

Figure 2.2 Comparison of Observed and Predicted Daily Streamflows for 2007 Calibration Period

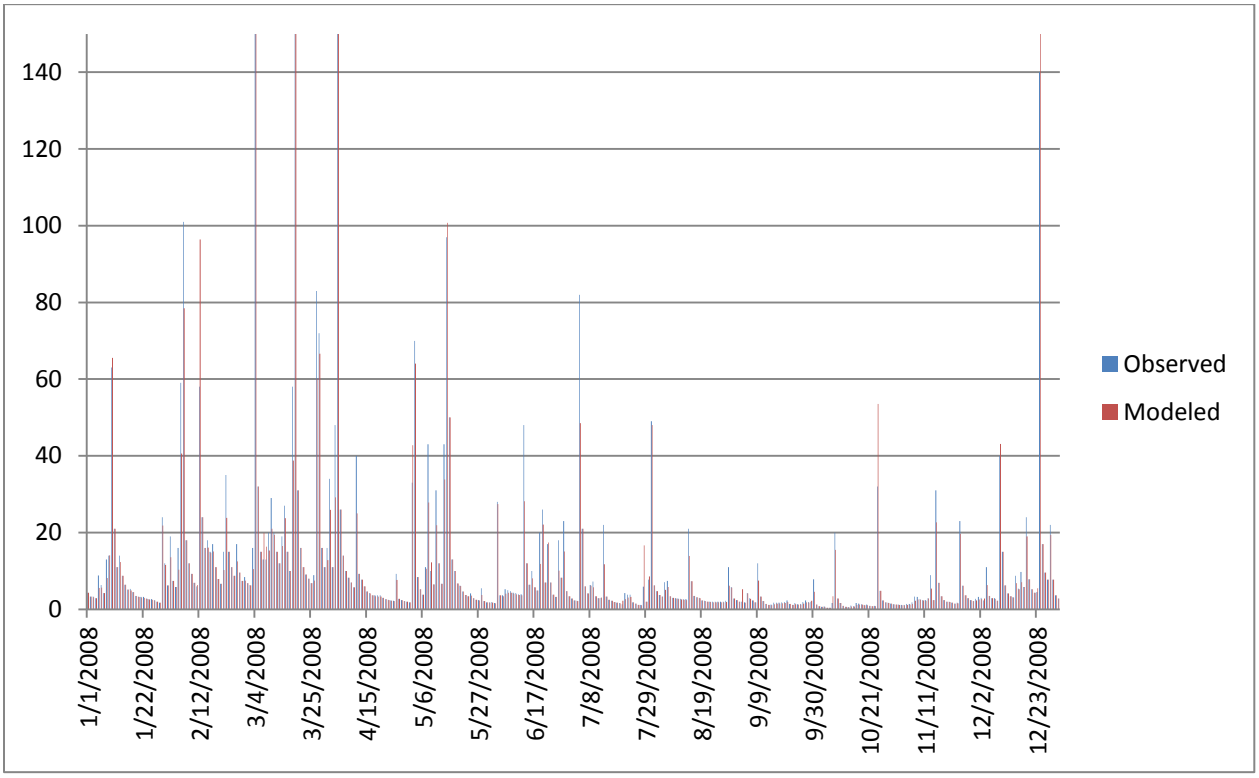

Figure 2.3 Comparison of Observed and Predicted Daily Streamflows for 2008 Validation Period

3.0 Water Quality Calibration

Calibration of the water quality parameters will typically occur in three basic steps: baseflow calibration, point source calibration, and non-point source flow calibration. Before the calibration process is begun, the data for each of these menus must first be entered into the program.

3.1 Water Quality Data

Before the hydrologic calibration process is begun, the user must first input the basic water quality data along with the assumed values of the basic model parameters.

The water quality data include:

1) Observed nitrogen data. (COLUMN N, COLUMN O)

2) Observed phosphorus data. (COLUMN N, COLUMN Q)

3) Observed or assumed total nitrogen and total phosphorus concentrations for all permitted facilities entered either as:

a) daily observed values.

b) monthly observed values.

c) annual observed or assumed values.

4) The α and β parameters for total nitrogen and total phosphorus for each CSO/SSO expressed as:

$TN\ (mg/L) = \alpha_{TN} + \beta_{TN}\,(Q\ mgd)$

$TP\ (mg/L) = \alpha_{TP} + \beta_{TP}\,(Q\ mgd)$

where α = the y-intercept of the relationship between CSO/SSO discharge and the associated discharge concentration of total nitrogen or total phosphorus, and β = the slope of the relationship between CSO/SSO discharge (mgd) and the associated discharge concentration (mg/L) of total nitrogen and total phosphorus. When data are not available to develop such a relationship β can be assumed to be zero and α can be obtained using the values in Table 3.1.

Default CSO concentrations are usually input for each major point source (e.g. CELLS AI31, AI32, AJ31, AJ32) while SSOs concentrations are usually input in the general SSO menu (i.e. CELLS AD31, AD32, AE31, AE32). However, if a facility does not have any CSOs, then one may use the individual menus for SSOs, as long as the appropriate SSO total nitrogen and total phosphorus concentrations are entered in the concentration cells (e.g. CELLS AI31, AI32, AJ31, AJ32).

5) The assumed total nitrogen and total phosphorus concentrations for septic systems (CELLS AD28, AE28). When actual data are not available, the user can use the values in Table 3.1.

6) The assumed average, maximum and minimum event mean concentrations of total nitrogen and total phosphorus associated with the different landuses associated with non-MS4 and MS4 areas. Prior to

calibration, the user should simply use the provided literature values for the event mean concentrations. These can be adjusted later if needed during model calibration.

3.2 Water Quality Parameters

The user should make sure the basic calibration parameters are set equal to their default values before beginning model calibration (see Table 3.1).

Table 3.1 Default Values for Water Quality Parameters

Symbol	Parameter	Default Value	Units	Spreadsheet Cell
	Groundwater			
α_{gN}	Total Nitrogen Y Intercept or Maximum	1.0	mg/L	K28
β_{gN}	Total Nitrogen Slope or Minimum	-.001	mg/L	K33
α_{gP}	Total Phosphorus Y Intercept or Maximum	0.1	mg/L	L28
β_{gP}	Total Phosphorus Slope or Minimum	-.001	mg/L	L33
	Erosion			
αe_N	Total Nitrogen Y Intercept		mg/L	P26
βe_N	Total Nitrogen Slope		mg/L	P27
αe_P	Total Phosphorus Y Intercept		mg/L	R26
βe_P	Total Phosphorus Slope		mg/L	R27
	Non-Point Source Parameters			
GAF_{NPSN}	Global Adjustment Factor for Nitrogen	1	mg/L	U26
GAF_{NPSP}	Global Adjustment Factor for Phosphorus	1	mg/L	V26
	Point Source Parameters			
GAF_{PSN}	Global Adjustment Factor for Nitrogen	1	mg/L	Y32
GAF_{PSP}	Global Adjustment Factor for Phosphorus	1	mg/L	Z32
	Sanitary Sewer/Combined Sewer Overflows			
α_{TN}	Total Nitrogen Y Intercept	40	mg/L	AD31
α_{TP}	Total Phosphorus Y Intercept	8	mg/L	AE31
	Sanitary Sewer/Combined Sewer Overflows			
β_{TN}	Total Nitrogen Slope	-.001	mg/L	AD32
β_{TP}	Total Phosphorus Slope	-.001	mg/L	AE32
	Septic Systems			
GAF_{SSN}	Global Adjustment Factor for Nitrogen	.1263	mg/L	AD28
GAF_{SSP}	Global Adjustment Factor for Phosphorus	.1287	mg/L	AE28
	Air Deposition			
WAD_N	Wet Total Nitrogen Loading Rate	1.4	tn/sqmi/yr	U29
DAD_N	Dry Total Nitrogen Loading Rate	2.0	tn/sqmi/yr	U31
WAY_P	Wet Total Phosphorus Loading Rate	0.04	tn/sqmi/yr	V29
DAD_P	Dry Total Phosphorus Loading Rate	0.08	tn/sqmi/yr	V31

3.3 Water Quality Calibration Steps

In the absence of significant point sources, the main nutrient sources will be groundwater and non-point source runoff. There may still be some influence from household septic systems. Atmospheric

deposition is a source built into the model which can be adjusted but it is recommended to usually leave these values at the literature default values. Erosion sources can be included in the model but this is recommended only as needed for a last step in the calibration process.

The KYNM relies on two different sets of tools for calibrating the water quality parameters of the model: 1) visual tools (scatter plots), and 2) statistics tools (CELLS P28 - P32 and R28 - R32). While adjusting the model parameters the user will try to get the points on the scatter plots to align along the 45 degree diagonal of the scatter plot so that the observed and predicted values will best match (see Figure 3.1). As with hydrology calibration, the KYNM provides five quantitative statistics to also guide the calibration process. The statistics used to measure water quality calibration in the KYNM are mostly different than those used to measure hydrology calibration due to differences in the nature of the data (e.g. the typically much smaller data set for water quality). These five statistics are defined below. The primary two statistics with associated quantitative targets in the KYNM are the relative error and the percent bias. General guidance on desired ranges for these two statistics are provided in Table 3.2 below.

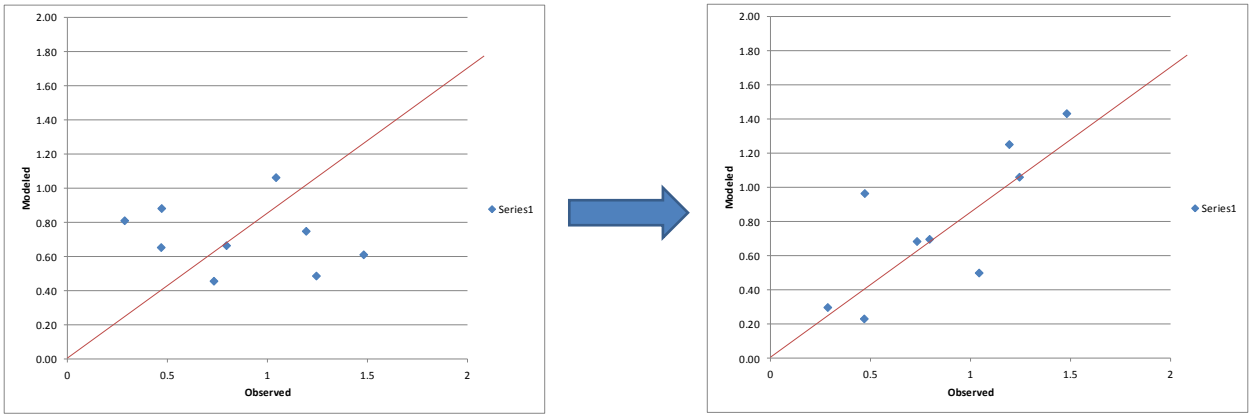

Figure 3.1 Illustration of Calibration Process

Table 3.2 Water Quality Calibration Targets

Statistic	Model CELL	Acceptable Range	Very Good
Relative Error	P32, R32	< .45	< .30
Percent Bias	P31, R31	< 25%	< 10%

Equation 3.1 Relative Error

Relative error (RE) is the absolute error divided by the magnitude of the exact value. It is often used to compare approximations of numbers of widely differing size. There are two features of relative error that should be kept in mind. Firstly, relative error is undefined when the true value is zero as it appears in the denominator. Secondly, relative error only makes sense when measured on a ratio scale, (i.e. a scale which has a true meaningful zero), otherwise it would be sensitive to the measurement units (e.g. Celsius and Kelvin). A relative error of zero is ideal.

$$RE = \left[\frac{\sum_{i=1}^{N} |Y_{obs,i} - Y_{sim,i}|}{\sum_{i=1}^{N} (Y_{obs,i})}\right]$$

Where $Y_{obs,i}$ is the *i*th observation for the constituent being evaluated, $Y_{sim,i}$ is the *i*th simulation for the constituent, N is the total number of observations.

Percent Bias

See Equation 1.2

Equation 3.2 Mean Signed Error

Mean signed error (MSE) is a sample statistic that summarizes how well an estimator matches the quantity that it is supposed to estimate. It is one of a number of statistics that can be used to assess an estimation procedure. This error summarizes performance in ways that takes into account the direction of over- or under- prediction. MSE varies from a large positive to a large negative value. Note that a MSE of zero is ideal however a low MSE does not indicate a better model simulation performance since the MSE is a measure that places emphasis on the direction of error, and so there is a possibility that the average of the positive deviations are cancelling out that of the negative deviations.

$$MSE = \frac{\sum_{i=1}^{N} (Y_{obs,i} - Y_{sim,i})}{N}$$

Where $Y_{obs,i}$ is the *i*th observation for the constituent being evaluated, $Y_{sim,i}$ is the *i*th simulation for the constituent, N is the total number of observations.

Equation 3.3 Mean Absolute Error

Mean absolute error (MAE) is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. As the name suggests, the mean absolute error is an average of the absolute errors. Unlike the MSE, this error summarizes performance in ways that disregard the direction of over- or under- prediction; a measure that does not place emphasis on the direction of error. MAE varies from the optimal value of 0, which indicates zero residual variation and therefore perfect model simulation, to a large positive value. The lower the MAE the better the model simulation performance.

$$MAE = \frac{\sum_{i=1}^{N} |Y_{obs,i} - Y_{sim,i}|}{N}$$

Where $Y_{obs,i}$ is the *i*th observation for the constituent being evaluated, $Y_{sim,i}$ is the *i*th simulation for the constituent, N is the total number of observations.

Equation 3.4 Root Mean Square Error (RMSE)

Root mean square error (RSME) represents the sample standard deviation of the differences between predicted values and observed values. It serves to aggregate the magnitudes of the errors in predictions for various times into a single measure of predictive power. RMSE is a good measure of accuracy, but only to compare forecasting errors of different models for a particular variable and not between variables, as it is scale-dependent. RSME varies from the optimal value of 0, which indicates zero residual variation and therefore perfect model simulation, to a large positive value. The lower the RMSE the better the model simulation performance.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \left(Y_{obs,i} - Y_{sim,i}\right)^2}{N}}$$

Where $Y_{obs,i}$ is the *i*th observation for the constituent being evaluated, $Y_{sim,i}$ is the *i*th simulated value for the constituent being evaluated, and *N* is the total number of observations.

3.3.1 Plot the Observed and Predicted Total Nitrogen and Total Phosphorus

The first step in calibrating the water quality parameters of the model is to develop two scatter plots of the observed and predicted total nitrogen and total phosphorus (see Figures 3.2 and 3.3)

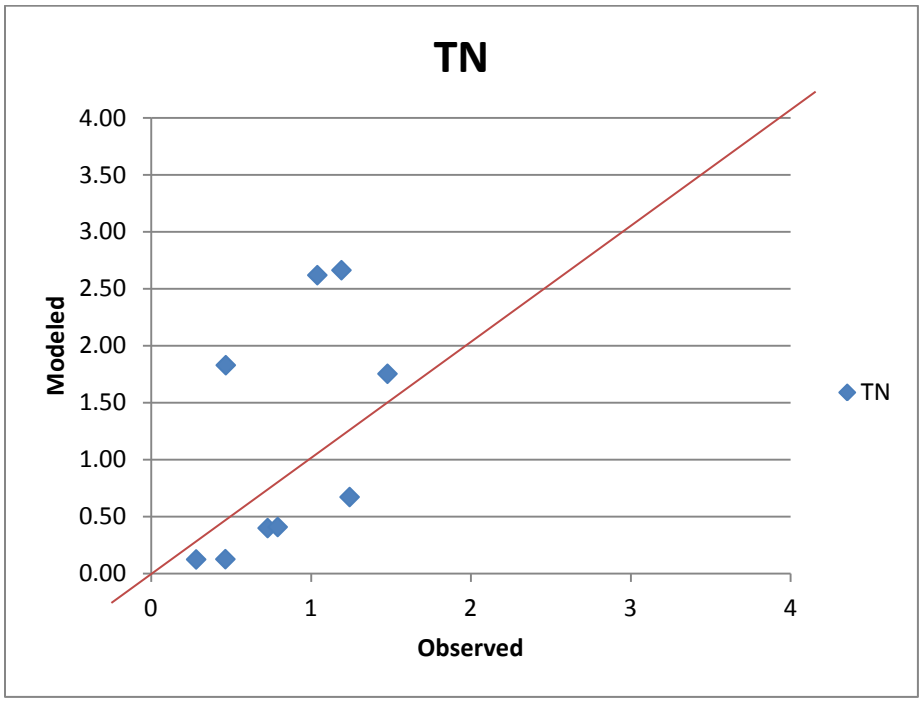

Figure 3.2 Example Scatter Plot for Total Nitrogen

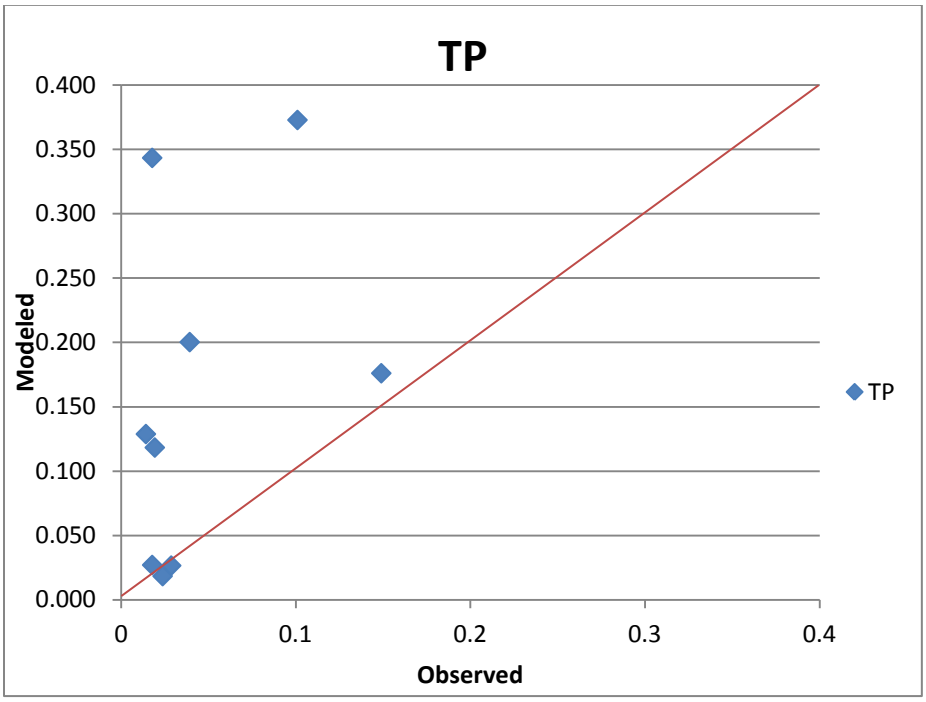

Figure 3.3 Example Scatter Plot for Total Phosphorus

3.3.2 Perform Sensitivity Analysis

The next step in the calibration process is to perform a sensitivity analysis to determine which nutrient sources are most influencing the various points on the scatter plots. Normally, the three most dominant sources will be: 1) point sources, 2) groundwater, and 3) non-point sources. Secondary sources include: 1) air deposition, 2) septic systems, and 3) erosion.

The sensitivity of the model to different pollutant sources can be evaluated by adjusting the global adjustment factors for each source (i.e. point sources, groundwater and non-point sources) and then looking at the scatter plots to evaluate the associated impact. This will then provide insights for use in setting the final values of these parameters.

3.3.3 Adjust the Groundwater Parameters

Step 1: Set the global adjustment factors for the non-point source loads (i.e. GAF_{NPSN} - CELL U2 and GAF_{NPSP} - CELL V2) to zero and then note which points on the scatter plot drop - these will be the data-points that will be associated or influenced by wet-weather events.

Step 2: Now focus on the remaining points. These will be the points that will be mainly influenced by dry event sources: groundwater or point sources. We will normally focus first on the groundwater sources, since theoretically, the point sources should be fairly reliable if they are drawn from DMR data.

The groundwater parameters for both total nitrogen and total phosphorus may be adjusted by either increasing or decreasing the values in model cells (i.e. CELLS K28, K33, L28, L33) in an attempt to get the majority of the points to fall on the 45 degree diagonal line of the scatter plot.

3.3.4 Adjust the Point Source Parameters

Even after adjusting the groundwater parameters, it is likely there will be some points that don't perfectly align on the associated scatter plots. Such deviations may be due to an error with the point source data, septic system loads, flow dependent in-stream loads (or sinks), or air deposition loads. In such cases the user may try to make **slight** adjustments (typically less than 10%) to the values in the cells associated with these parameters. Where monthly or daily nutrient loads are being generated for any point sources, the user may be able to make a direct adjustment to one of these monthly or daily loads in lieu of making a global adjustment for all the values. Unless the user has an adequate data set from which to derive loading coefficients for any in-stream loads, these loads are normally assumed to be zero. Likewise, the air deposition loads are typically assumed to be fixed, but in certain situations the user may elect to make slight adjustments to the default parameter values. The validity of all such adjustments should always be confirmed through the process of model validation.

3.3.5 Adjust the Non-Point Source Parameters

Once you have adjusted the groundwater and point source parameters (and/or secondary parameters) you can now reset the non-point source global adjustment factors (i.e. CELLS U26, V26) to 1 and evaluate how the wet weather data points move on the scatter plot. Ideally they should fall on the 45 degree diagonal on the scatter plots. In the event these do not match (which is more likely), the user can then make adjustment to the non-point source parameters in an attempt to improve this alignment. While

theoretically, the user could adjust the EMCs associated with each land-use, the model has been developed under the assumption that the provided default values should be at least proportionally correct. As a result, it is assumed that the user can simply use the global adjustment factors for both sets of EMCs (i.e. total nitrogen and total phosphorus) to make adjustments to these parameters globally. In this way, the EMCs will be adjusted proportionally. Regardless, the model still allows the user to make individual EMC adjustments if they desire.

In either case, the user should try to either raise or lower these global adjustment factors until the points tend to align along the diagonal of the scatter plot. In making such adjustments it is unlikely all the points will match. At this point, the user should try to make fine adjustments that minimize the total error of the data set. This can be facilitated by using the associated calibration statistics in CELLS (P28-32 and R28-32). In this case, the user should make adjustments so that the calibration statistics meet the acceptable range while trying to approach the optimal range (see Table 3.2). It trying to achieve these objectives, it is possible that all four of the criteria may not be met. In this case , one should focus on the statistics in the order provided (i.e. relative error, percent bias, RSR, and coefficient of determination). That is, the user should first try to get the relative error within the desired range, followed by the percent bias, etc.

4.0 Water Quality Calibration Application to Upper Chenoweth Run Watershed

For the purposes of this tutorial, calibration of total nitrogen (TN) in the Upper Chenoweth Run watershed will be examined first because this segment does not contain any significant point sources. Then after thoroughly dealing with calibration issues regarding groundwater sources and non-point runoff sources in the Upper Chenoweth Run, calibration of TN in the Middle Chenoweth Run watershed will be examined with a special focus on dealing with the major point source facility in that basin.

4.1 Setting Initial Model Parameters

In the absence of significant point sources, the main nutrient sources will be groundwater and non-point source runoff. There may still be some influence from household septic systems. Atmospheric deposition is a source built into the model which can be adjusted but it is recommended to usually leave these values at the literature default values. In-stream sources, such as erosion, can be included in the model but this is recommended only as needed for a last step in the calibration process.

For the purposes of this tutorial, it is assumed that median literature EMC values are inputs to the model for land uses. While the EMC values can be adjusted, a better practice may be to use the global adjustment factor, GAF (CELL U26) to increase or decrease these loadings. Next, initial values for groundwater max is 1 mg/L and for groundwater min. is 0.1 mg/L (CELLs K30 and K33). Groundwater can be modeled using either a linear function or an asymptotic function. Initially the asymptotic function is in use. (Under the linear function the parameters become y-intercept and slope.) The initial scatter plot of modeled versus observed TN concentrations for 2007 is shown below in Figure 4.1.

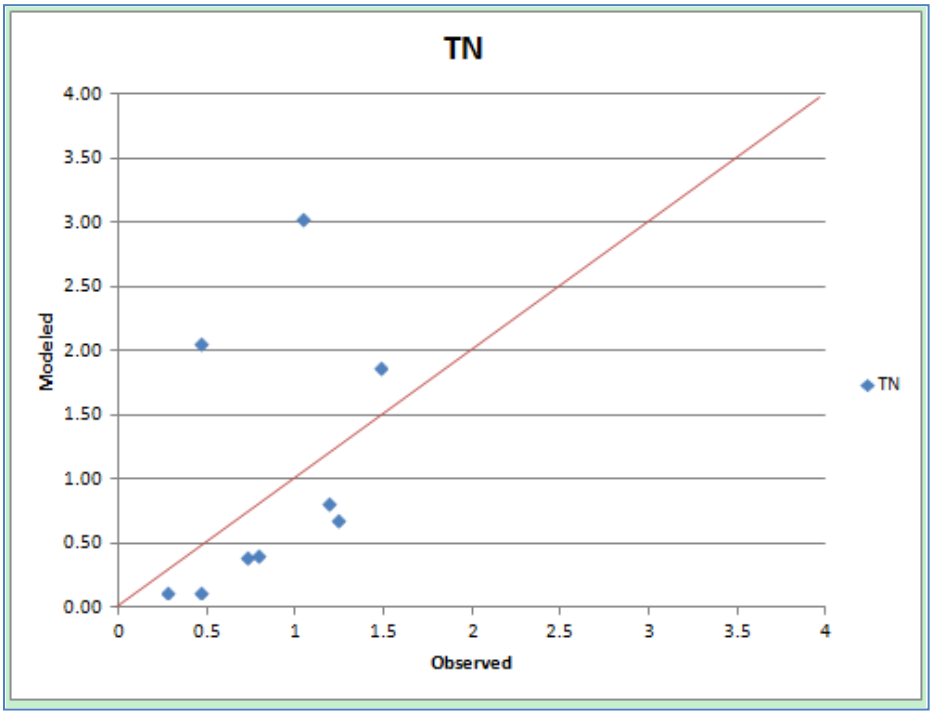

Figure 4.1 Scatter Plot of Total Nitrogen Results

4.2 Setting Non-Point Sources to Zero

The non-point source GAF (CELL U26) can be used to calibrate the water quality, but it can also be used to acquire information about the system by setting it at a range of values. After acquiring the desired information, the GAF can be set to an appropriate value for the calibration. If the modeled concentrations are low, then setting the GAF at multiples of 1can reveal how sensitive the modeled points are to non-point source runoff. If the GAF has been set at 1, 2, 3, 5, etc. with no noticeable rise in any of the modeled points on the scatter plot, then these points are not going to assist in calibrating the non-point source loading in the watershed. Most likely, this is the case when all the observed sampled data were taken on dry days. However, if you find data points that are influenced by the GAF, then those points are an important clue to how to set or calibrate the non-point source loading in the watershed. Another way to investigate the data system is to set the GAF at 0. Figure 4.2 shows the scatter plot result of setting the GAF = 0.

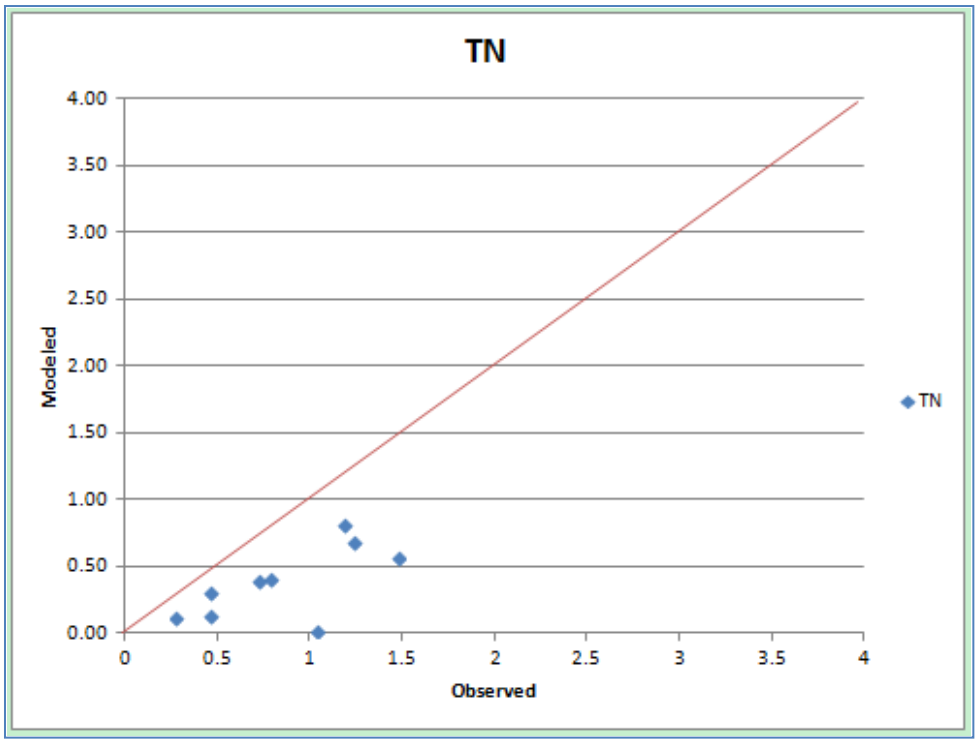

Figure 4.2 Scatter Plot with Non-Point Sources Set at Zero

What is apparent is that the three points above the equivalence line (red line marking y = x) dropped either below the line or even down to zero. This GAF experiment reveals that these three points are the clues to calibrating the non-point source loading. The same experiments can be done with groundwater or point sources.

4.3 Setting the Groundwater Parameters to Zero

Setting the GAF back to 1 and setting the groundwater concentration at 0 results in the plot shown below in Figure 4.3.

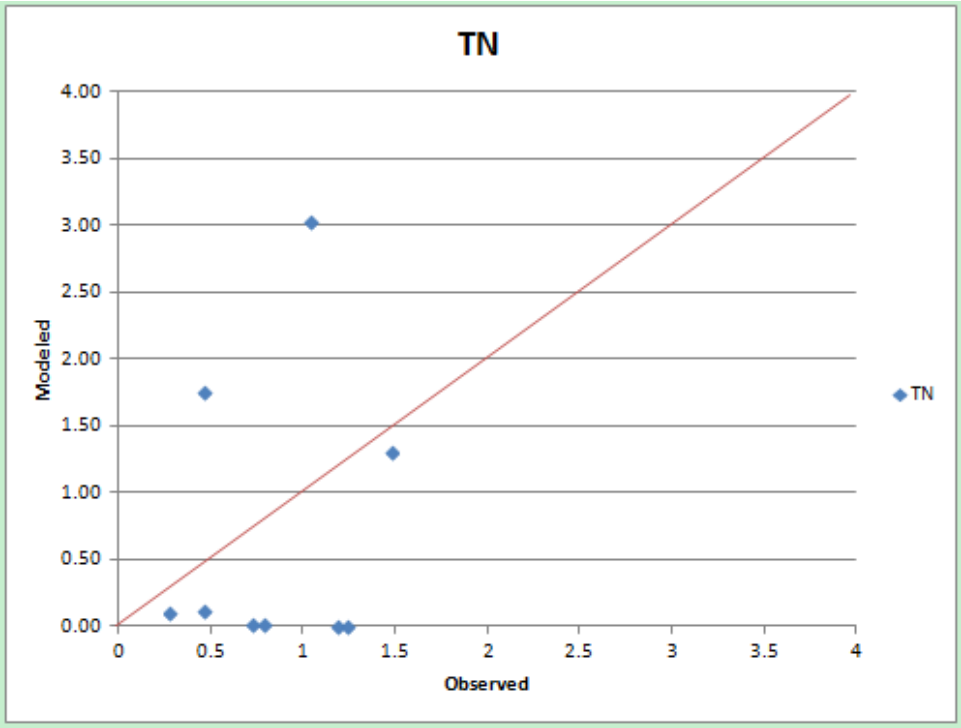

Figure 4.3 Scatter Plot with Groundwater Sources Set at Zero

Several observations can be made at this point. The initial plot shows that the three data points that can give information about non-point source loading suggest that the non-point source loading is too high, as all three points are well above the equivalence line. Even after the groundwater has been set to zero, two of the points are still high above the equivalence line, and the third (which is the highest observed concentration) is nearly on the line. Therefore when calibrating, the GAF will be set somewhere below 1 using trial and error. Next there is a definite influence from groundwater in the watershed since the other points fall to zero when the groundwater is set to 0. However, the initial plot shows that these groundwater influenced points track the equivalence line fairly well, with the only difference being that the modeled points are tracking beneath the line. Therefore the groundwater concentration should be raised so that these points track on top of the equivalence line. First, a GAF = 0.25 brings the non-point source data points down in the line the other data is tracking (see plot below in Figure 4.4).

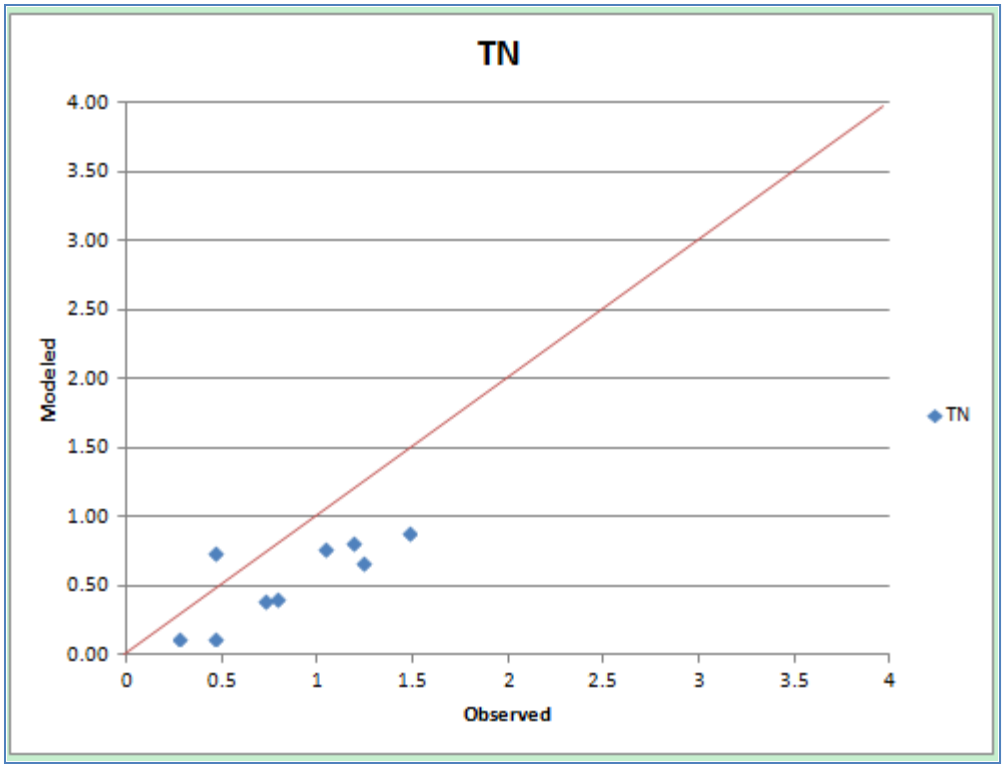

Figure 4.4 Scatter Plot Setting the Non-Point Source GAF = 0.25

Now that there is a fairly good line that the data is forming, tracking the equivalence line, we systematically raise the groundwater concentration to raise this line onto the equivalence line. After some trial and error we have GW max = 1.5 and GW min. = 0.9. In addition to the qualitative metric of visually inspecting the plots, the relative error statistic is of very helpful quantitative importance. The calibration target for nutrient modeling is to meet or be below a relative error of 0.45. The final TN calibration has a relative error of 0.22 which is very good (see CELL P32). Figure 4.5 shows the result of the final calibration parameters.

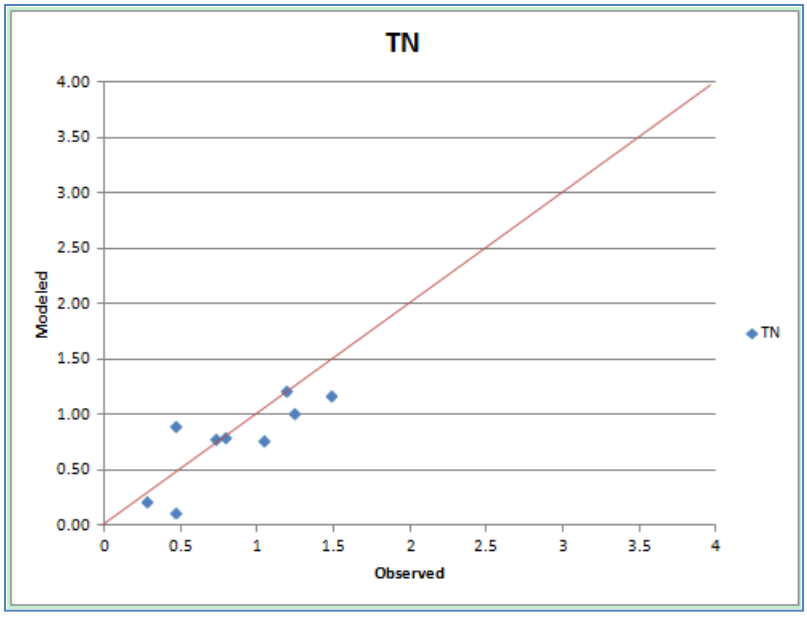

Figure 4.5 Scatter Plot of Final TN Calibration

Validation against 2008 data for these water quality parameters results in a relative error of 0.24, which is very good for a validation statistic. The 2008 validation plot is shown below in Figure 4.6.

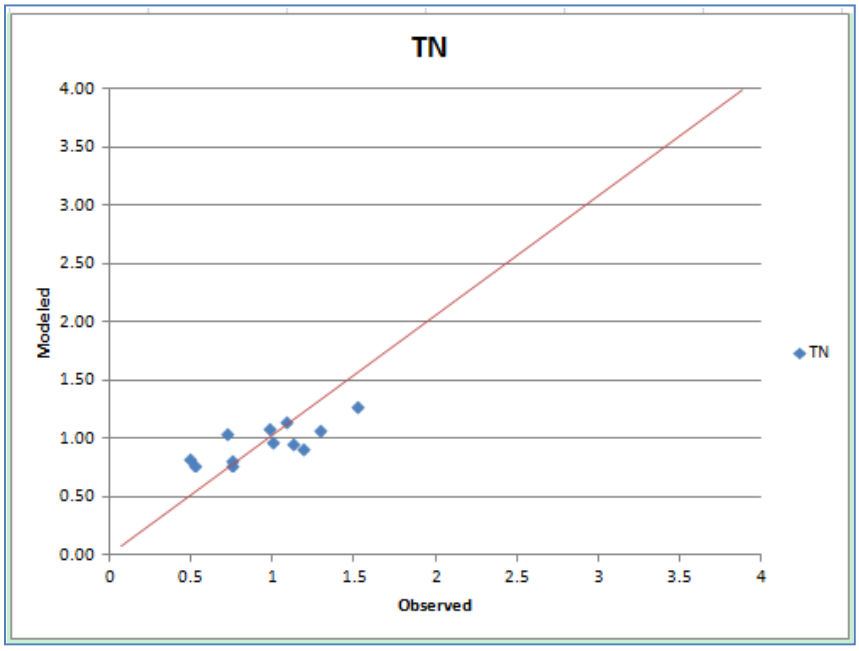

Figure 4.6 Scatter Plot of TN Validation

5.0 Application to Middle Chenoweth Run Watershed

The presence of a large point source such as a wastewater treatment plant can significantly add to the nutrient load and flow of the stream. In the case of the Middle Chenoweth Run watershed, the Jeffersontown WQTC is a major point source.

The process for calibrating the hydrology is relatively the same, the model takes any point source flow input into account automatically. With a filter coefficient and BFI max both set to 0.9, for 2007 there is less than a one percent difference between the separated and predicted runoff, and a less than 1 percent difference between observed and predicted total annual volumes. With K and GAF both set to 1, the R^2 is 0.90 and the NS is 0.85. The 2008 validation has an R^2 of 0.97 and a NS of 0.97.

Initially, land use EMCs are set to median literature values. The land use loading GAF is set to 1. Groundwater is set to a constant concentration of 1.5 mg/L. For the first analysis, point source load is turned "off" by inputting 0 for the point source load GAF (CELL Y32) in order to ascertain what affects the groundwater and runoff sources may be able to exert on the data points. The initial state of the model predictions are shown in the plot below (Figure 5.1).

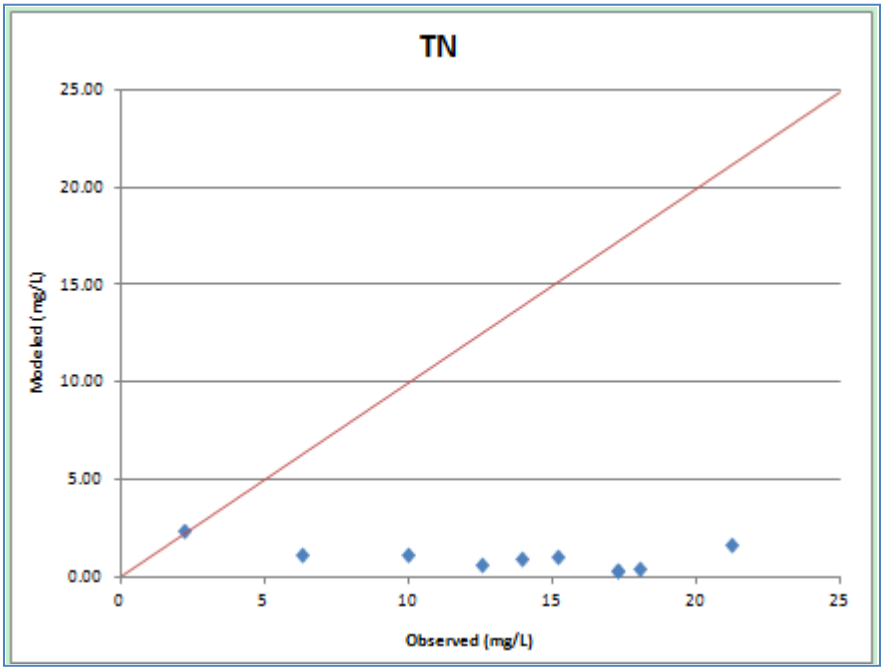

Figure 5.1 Scatter Plot with Point Source GAF = 0

As the scatter plot shows, there are some very high TN concentrations coming out of this basin. Nonpoint source runoff and groundwater sources are only able to produce concentrations ranging from about 0.2 mg/L to about 1.5 mg/L. However, the observed data are mostly above 10 mg/L and a few are near or above 20 mg/L.

Even with a nonpoint source loading GAF of 5, all but one point lies far below the equivalence line. See the scatter plot below (Figure 5.2).

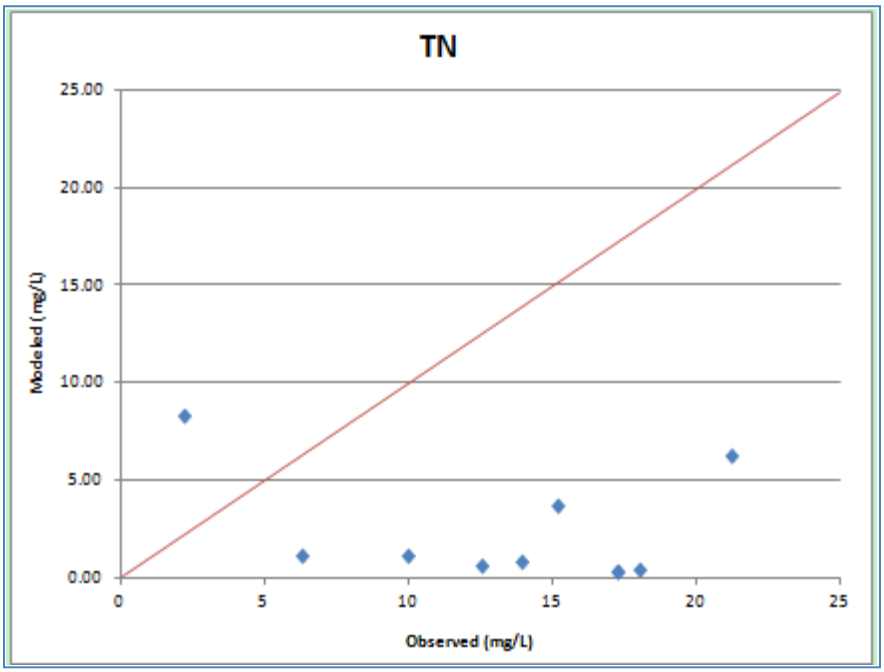

Figure 5.2 Scatter Plot with Non-Point Source GAF = 5 for Sensitivity Analysis

As could be expected, there must exist significant loading from a point source to result in observed concentrations as high as these data. Setting the nonpoint source loading GAF back to 1. We can start turning "on" the point source loading and pull together various pieces of information to determine the loading from the point source. Ideally, all the flows and concentrations for point source inputs could be known from reports. However, this is not always the case. For example, in this study, ammonia data was available but not TN data. At first, a relationship between ammonia and TN was attempted but the attempts did not result in data that seemed to be able to reflect known TN data. Since attempts at a time series of TN data were not successful, an assumed concentration of 10 mg/L was implemented. The scatter plot shown below is the result of this assumed point source concentration (Figure 5.3).

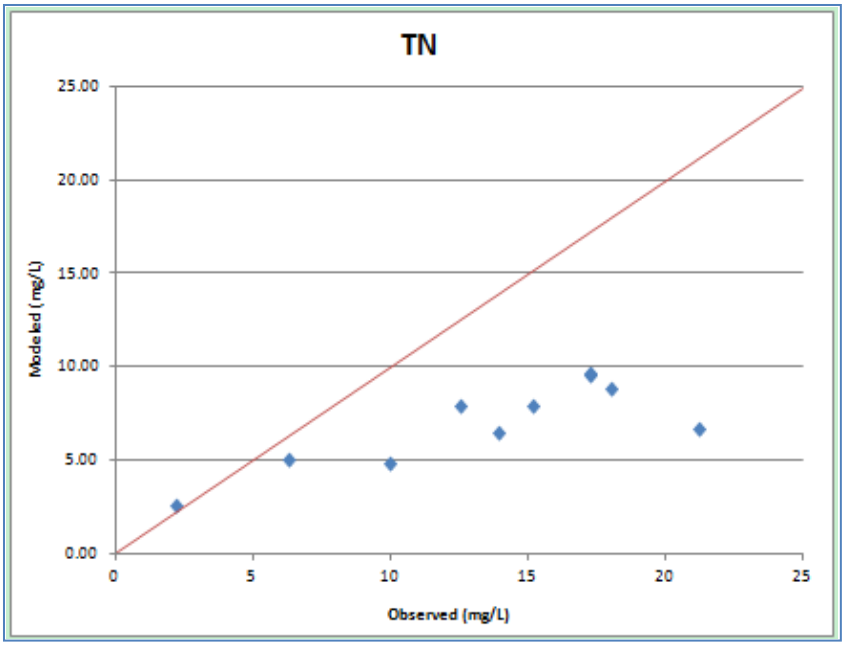

Figure 5.3 Scatter Plot with Point Source Concentration = 10 mg/L

The data points are beginning to look well modeled as they are tracking almost a straight line. However, we can see that there is a definite low bias as the points are still well below the equivalence line. We can use the point source loading GAF to determine if the points can be made to track nicely along the equivalence line. At a GAF of 2, the modeled versus observed concentrations make a extremely good fit. The relative error is as low as 0.174 and the percent bias is an under-prediction of 2.4 percent. The highest observed concentration is still being well under-predicted but all the other points track remarkably well along the equivalence line as shown in the figure below (Figure 5.4).

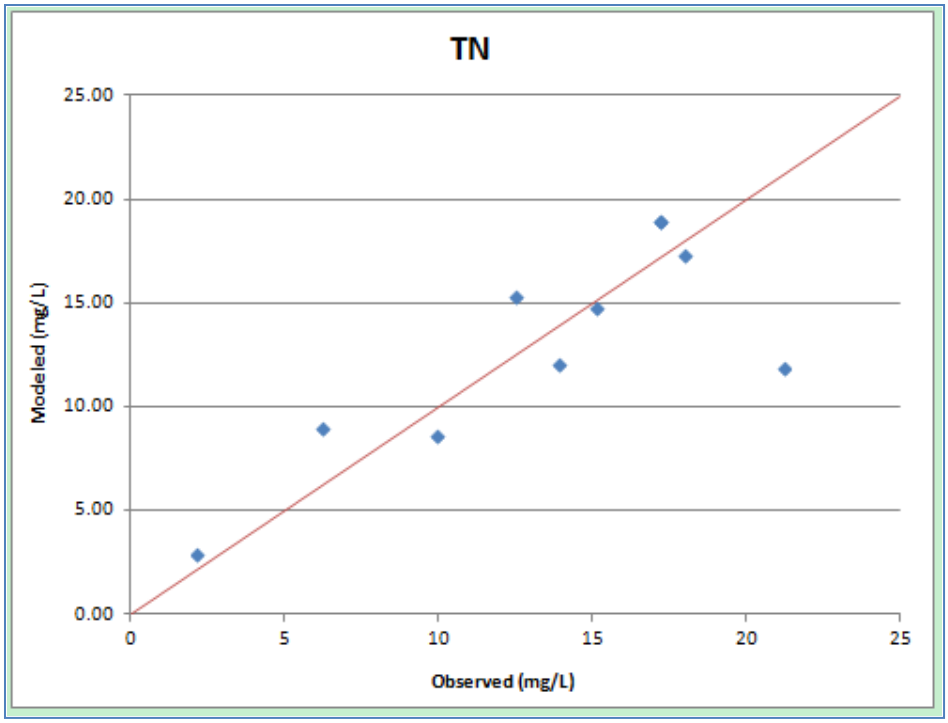

Figure 5.4 Scatter Plot with Point Source Concentration = 20 mg/L

Because the model is doing well at predicting the rise and fall of flow in the stream, this degree of prediction success for the water quality aspect is strong evidence that a constant concentration of 20 mg/L (GAF = 2 at 10 mg/L) is a very good estimate of the point source concentration.

In the case of Middle Chenoweth Run, there is further evidence that the wastewater effluent from this point source is this high. There is water quality data immediately upstream and downstream of the plant discharge. USGS 3298135 is immediately upstream and USGS 3298138 is immediately downstream. The tables below (Tables 5.1 and 5.2) show data for 2007 and 2008. As can be seen in the tables, the concentration downstream of the plant was often right at 20 mg/L (19.6, 21.41, 19.3, 19.1, 22.5, 20.267, 20.04, 20.89, etc) and once even as high as 34.3. The upstream USGS station has flow data which when used in conjunction with the facility effluent flow data allows for an estimate of what percent of the flow downstream of the plant is from the plant effluent. This percent was computed for several of the days where the concentration data downstream of the plant was somewhat lower (15.29, 12.7, 10.26, 15.5, etc.) and it shows substantial dilution occurring in the mixing of the effluent with the flow present in the stream. Therefore the evidence of the model prediction ability over so many data points taken with the hard data evidence justifies a use of 20 mg/L as the concentration of the point source effluent.

Table 5.1 2007 Water Quality Data Upstream and Downstream of the WQTC

2007			
	TN	TN	Percent of stream flow from WWTP
Date	USGS 3298135	USGS 3298138	
10/16/07	0.4684	19.6	
09/20/07	0.4658	21.41	
09/06/07	0.283	19.3	
08/21/07	1.48	no data	
08/14/07	0.493	17.2	
08/01/07	0.73	15.29	84.1
07/17/07	1.041	19.1	
06/25/07	1.192	12.7	53.1
06/11/07	0.793	no data	
05/23/07	1.243	no data	
10/23/07	no data	10.26	10.7

Table 5.2 2008 Water Quality Data Upstream and Downstream of the WQTC

2008			
	TN	TN	Percent of stream flow from WWTP
Date	USGS 3298135	USGS 3298138	
10/23/08	0.744	22.5	
10/16/08	0.749	20.267	
10/09/08	0.716	no data	
10/02/08	0.514	34.3	
09/23/08	0.482	15.5	73.6
08/19/08	0.991	18.9	72.2
07/31/08	0.975	no data	
07/22/08	1.285	20.04	
07/16/08	1.115	17.59	
06/23/08	1.518	14.16	49.0
06/10/08	1.076	16.29	53.8
04/30/08	1.18	18.46	73.3
01/29/08	2.274	20.89	

Validation on 2008 data produced very good statistics as well. A slight under-prediction bias of less than 2 percent and a relative error statistic of 0.195. The scatter plot below (Figure 5.5) shows the qualitative tracking of the data points in the main lying just slightly below the equivalence line. One could further refine the calibration, possibly justifying a slightly higher point source concentration, but the calibration performance measure at this point is well below the target relative error of 0.45.

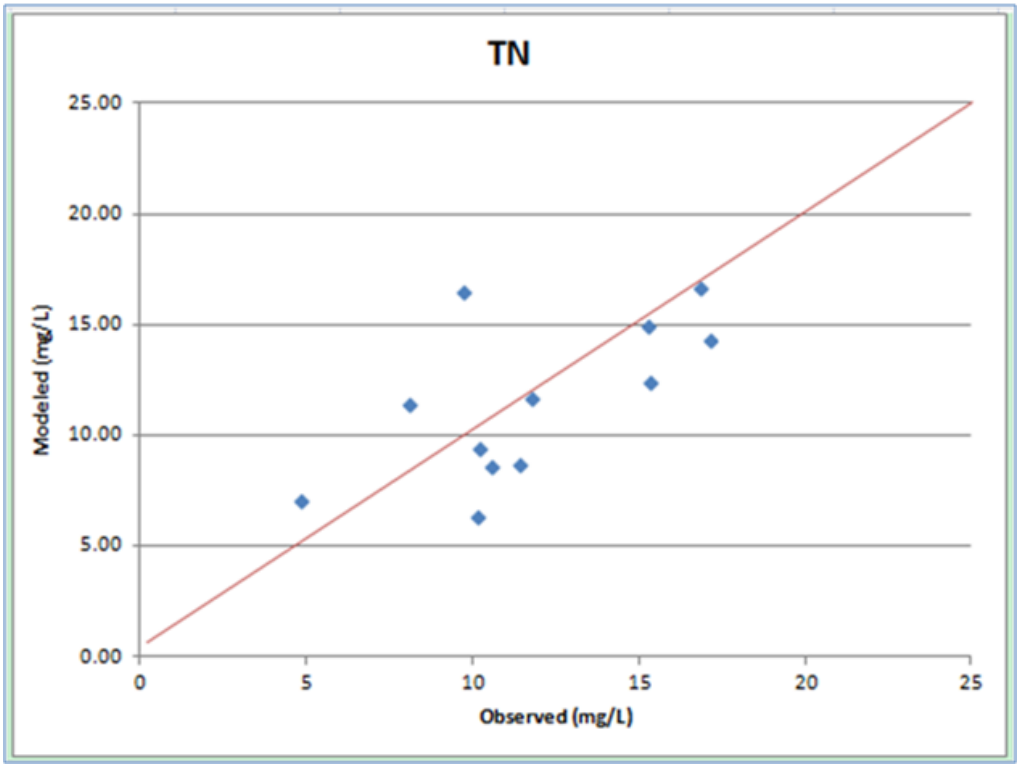

Figure 5.5 Scatter Plot of TN Validation