

# LECTURE #15

WATERSHED MODEL CALIBRATION AND VALIDATION: ISSUES AND PROCEDURES









### **ASTM DEFINITIONS**

Model - An assembly of concepts in the form of a mathematical equation that portrays understanding of a natural phenomenon

Verification - Examination of the numerical technique in the computer code to ascertain that it truly represents the conceptual model and that there are no inherent numerical problems with obtaining a solution

Calibration - A test of a model with known input and output information that is used to adjust or estimate factors for which data are not available

Validation -

on - Comparison of model results with numerical data independently derived from experiments or observations of the environment





### **THE MODELING PROCESS**

Phase I

Phase II

Phase III

- Data collection
- Model input preparation
- Parameter evaluation
- Calibration
- Validation
- (Post-audit)
- Analysis of alternatives

Model Testing



### ALL MODELS ARE WRONG,

**BUT**....

## SOME ARE USEFUL !

(Depends on the Model Testing Process)

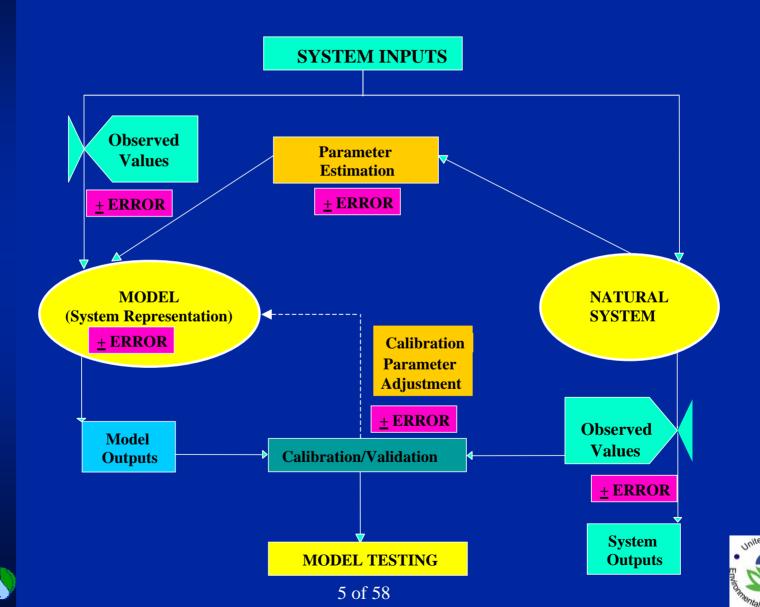


(Source: G.E.P. Box, 1979)

United States

4 of 57

### MODEL VERSUS NATURAL SYSTEM: INPUTS, OUTPUTS, AND ERRORS



AQUA TERRA CONSULTANTS

### **MODEL VALIDATION COMPARISONS**

- Point-to-Point Paired Data Performance
- Time and/or Space Integrated Paired Data Performance
- Frequency Domain Performance





## **CALIBRATION ISSUES**

### 'Basic Truths' in modeling Natural Systems

- Models are approximations of reality; they can not precisely represent natural systems
- There is no single, accepted statistic or test that determines whether or not a model is valid
- Both graphical comparisons and statistical tests are required in model calibration and validation
- Models cannot be expected to be more accurate than the errors (confidence intervals) in the input and observed data





## CALIBRATION/VALIDATION COMPARISONS

### "Weight-of-Evidence" Approach

- Annual and monthly runoff volume (inches)
- Mean runoff volume for simulation period (inches)
- Daily flow timeseries (cfs)
  - observed and simulated daily flow
  - scatter plots
- Flow frequency (flow duration) curves (cfs)
- Storm hydrographs, hourly or less, (cfs)





## CALIBRATION/VALIDATION COMPARISONS

### Water Balance Components

- Precipitation
- Total Runoff (sum of following components)
  - Overland flow
  - Interflow
  - Baseflow
- Total Actual Evapotranspiration (ET) (sum of following components)
  - Interception ET
  - Upper Zone ET
  - Lower Zone ET
  - Baseflow ET
  - Active Groundwater ET
  - **Deep Groundwater Recharge/Losses**





•

## CALIBRATION/VALIDATION COMPARISONS

### **Graphical/Statistical Procedures & Tests**

#### **Graphical Comparisons:**

- Timeseries plots of observed and simulated values for fluxes (e.g., flow) or state variables (e.g., stage, sediment concentration, biomass concentration)
- Observed and simulated scatter plots, with 45° linear regression line displayed, for fluxes or state variables
- Cumulative frequency distributions of observed and simulated fluxes or state variable (e.g., flow duration curves)

#### **Statistical Tests:**

- Error statistics, e.g., mean error, absolute mean error, relative error, relative bias, standard error of estimate, etc.
- Correlation tests, e.g., correlation coefficient, coefficient of modelfit efficiency, etc.
- Cumulative Distribution tests, e.g., Kolmogorov-Smirnov (KS) test



## **ROUGH CALIBRATION/VALIDATION TARGETS**

	% Difference Between Simulated and Recorded Values					
	VERY GOOD	GOOD	FAIR			
Hydrology/Flow	< 10	10 - 15	15 - 25			
Sediment	< 20	20 - 30	30 - 45			
Water Temperature	< 7	8 - 12	13 - 18			
Water Quality/Nutrients	< 15	15 - 25	25 - 35			
Pesticides/Toxics	< 20	20 - 30	30 - 40			

**CAVEATS: 1.) Relevant to monthly and annual values; storm peaks may differ more.** 

- 2.) Quality and detail of input and calibration data.
- 3.) Purpose of model application.4.) Availability of alternative assessment procedures.
- 5.) Resource availability (i.e. time, money, personnel).





Source: Donigian, 2000

11 of 58

## R and R<sup>2</sup> VALUE RANGE FOR MODEL PERFORMANCE

#### Criteria

R	← 0.75 ───	0.80	- 0.85		- 0.90 -		0.95
R <sup>2</sup>	<b>→</b> 0.6		0.7 –		<b>0.8</b> -		0.9 →
Daily Flows	Poor	Fair		Good		Very Good	
<b>Monthly Flows</b>	Poor		Fair		Good		Very Good





HIERARCHY OF WATERSHED MODEL CALIBRATION ( a la HSPF)

- Hydrology / Hydraulics
- Water Temperature
- Sediment Loadings and Instream Sediment Fate / Transport
- Nonpoint Loadings
- Instream Water Quality Processes





## HYDROLOGIC (PWATER) CALIBRATION

• Annual Water Balance -

Runoff = Prec. - Actual ET - Deep Perc. - D Storage

Key Parameters:

Repre. Precipitation (MFACT) LZSN LZETP INFILT DEEPFR

• Groundwater (Baseflow) Volume and Recession -

Runoff = Surface Runoff + Interflow + Baseflow

Key Parameters:

INFILT AGWRC/KVARY DEEPFR BASETP/AGWETP

• Surface Runoff + Interflow (Hydrograph Shape) -

Key Parameters:



UZSN INTFW IRC LSUR, NSUR, SLSUR



## **SEDIMENT CALIBRATION**

#### LAND SEDIMENT LOADING CALIBRATION

- Estimate 'target' sediment loading rates by land use
- Calibrate model sediment loading rates to observed data and/or target rates

#### INSTREAM CALIBRATION

- Estimate initial parameter values for both cohesive (silt, clay) and non-cohesive (sand) sediment fractions
- Perform sediment mass balance to determine land surface versus stream channel contributions
- Make calibration run and output TAU values (max and min daily) calculated by subroutine SHEAR
- Adjust TAUCS and TAUCD to affect scour and deposition of cohesive sediments at appropriate times
- Examine/evaluate sediment load simulation for both mass outflow and composition compared to available data
- Adjust M to improve calibration of cohesive sediments for storms with good flow simulation
- Adjust non-cohesive (sand) parameters based on bed and load composition compared to available data
- Re-do calibration run and output analyses





## WATER QUALITY CALIBRATION

- Estimate all model parameters, including land use specific accumulation and depletion/removal rates, washoff rates, and subsurface concentrations
- Tabulate, analyze, and compare simulated nonpoint loadings with expected range of nonpoint loadings from each land use and adjust loading parameters when necessary
- Calibrate instream water temperature
- Compare simulated and observed instream concentrations at each of the calibration stations
- Analyze the results of comparisons in steps 3, 4, and 5 to determine appropriate instream and/or nonpoint parameter adjustments

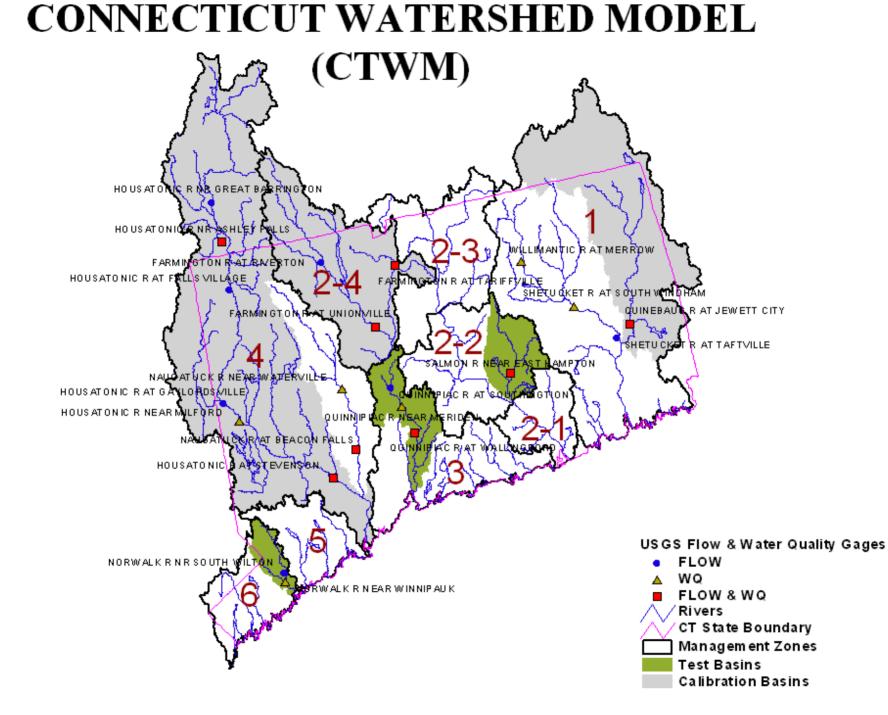


## HSPF CALIBRATION / VALIDATION EXAMPLES

- Connecticut Watershed Model
- Unnamed Northeast Watershed







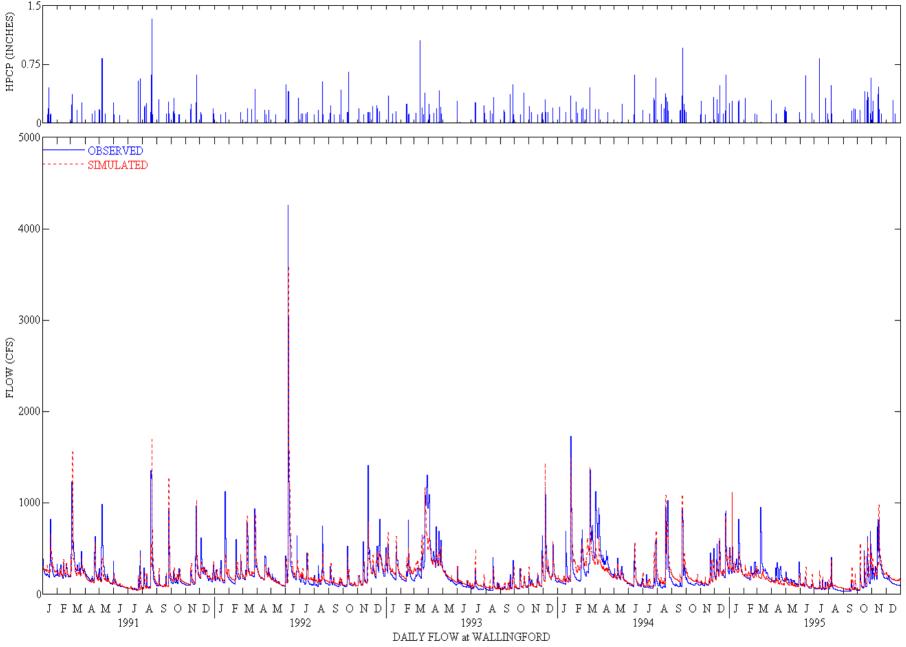
## SUMMARY OF CTWM HYDROLOGIC CALIBRATION/VALIDATION ANNUAL FLOW AND CORRELATION COEFICIENTS

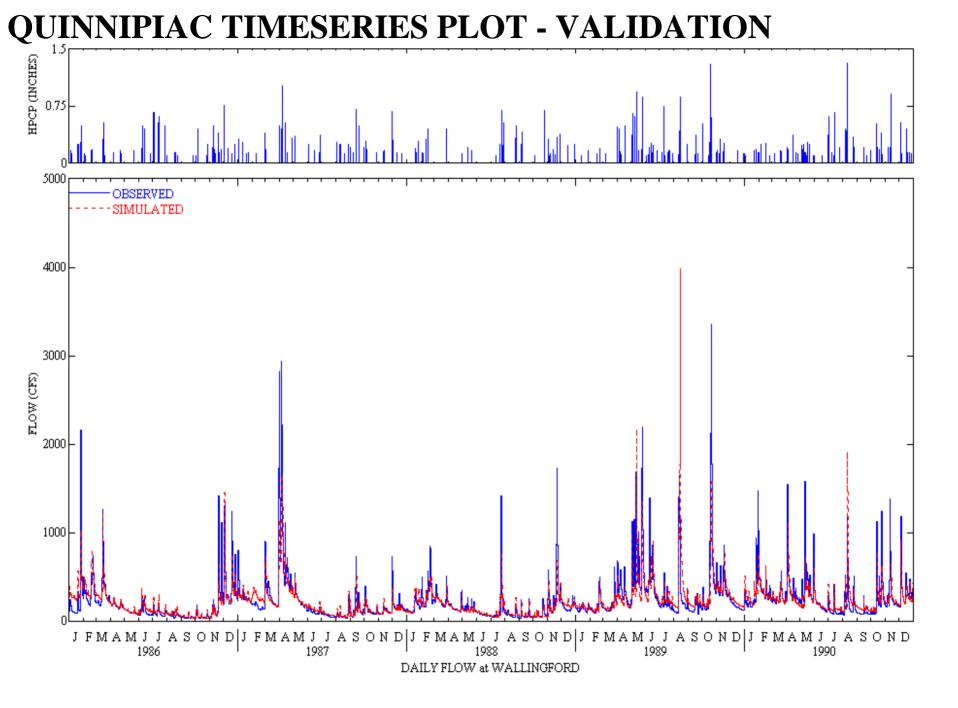
	Calibration Period (1991-1995)				Validation Period (1986-1990)					
	Mean	Mean				Mean	Mean			
	Observed	Simulated				Observed	Simulated			
	Annual	Annual	Percent	R	R	Annual	Annual	Percent	R	R
	Flow	Flow	Difference	Average	Average	Flow	Flow	Difference	Average	Average
Station Name	(inches)	(inches)	(Sim-Obs)	Daily	Monthly	(inches)	(inches)	(Sim-Obs)	Daily	Monthly
Test Watershed										
Gages										
Salmon River nr East										
Hampton	23.6	24.4	3.3	0.83	0.92	26.3	25.8	-1.9	0.79	0.92
Quinnipiac River at										
Wallingford	26.3	26.4	0.4	0.82	0.94	29.0	28.3	-2.5	0.71	0.91
Norwalk River at										
South Wilton	21.4	21.7	1.4	0.84	0.93	25.9	25.2	-2.8	0.75	0.91
Major Basin Gages										
Quinebaug River at										
Jewett City	22.0	22.6	0.0	0 02	0.02	27.2	24.7	10.1	0.96	0.95
,	23.8	23.6	-0.8	0.82	0.93	21.2	24.1	-10.1	0.86	0.95
Farmington River at Tariffville	26.2	26.0	0.0	0.95	0.02	26.2	20.4	10.0	0.07	0.04
	26.2	26.0	-0.8	0.85	0.92	26.2	29.1	10.0	0.87	0.94
Housatonic River at	24 7	21.0	0.6	0.00	0.00	24.6	21 5		0.97	0.06
Stevenson	31.7	31.9	0.6	0.88	0.98	34.6	31.5	-9.8	0.87	0.96



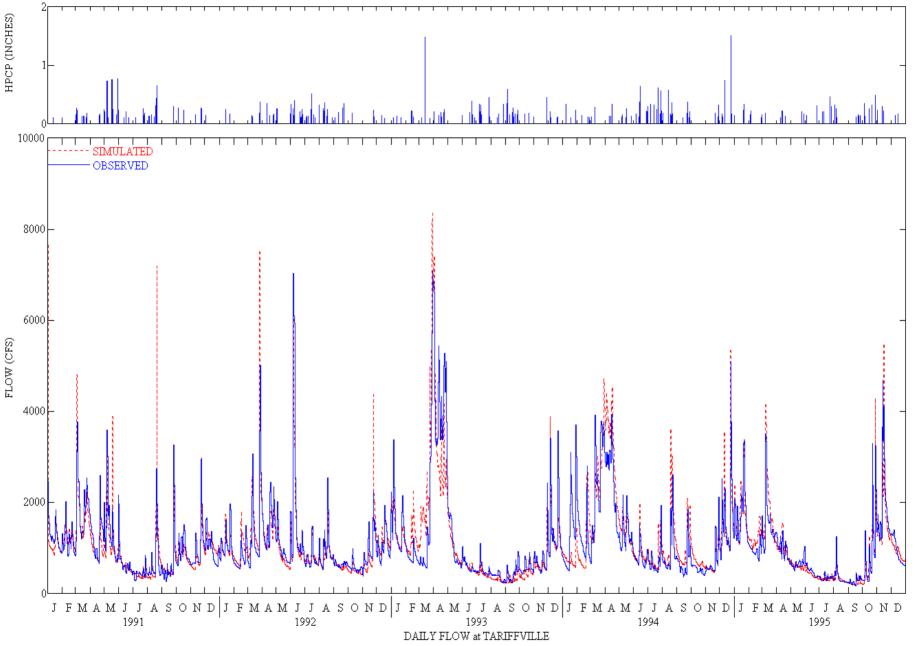


**QUINNIPIAC TIMESERIES PLOT - CALIBRATION** 

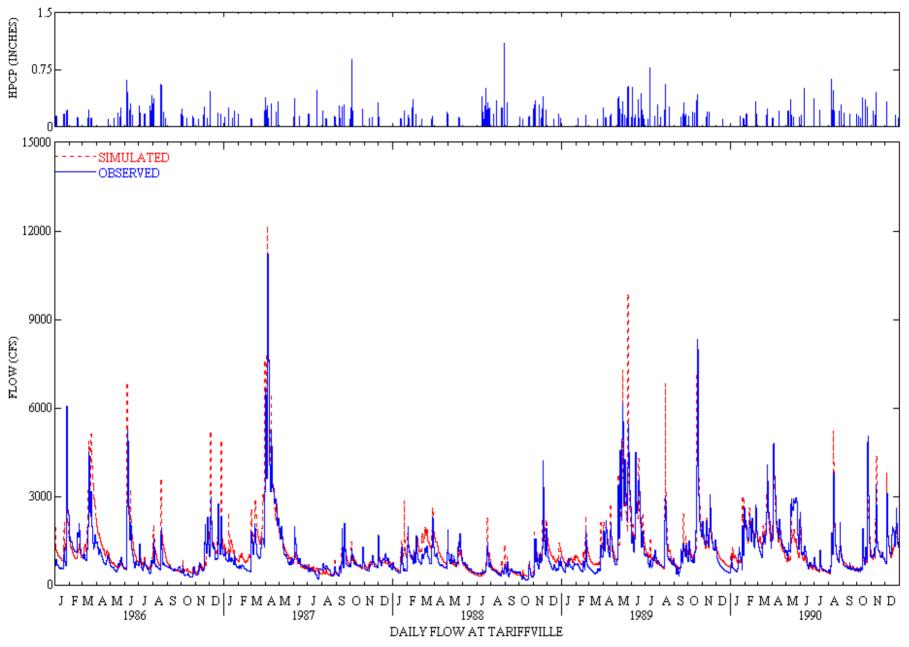




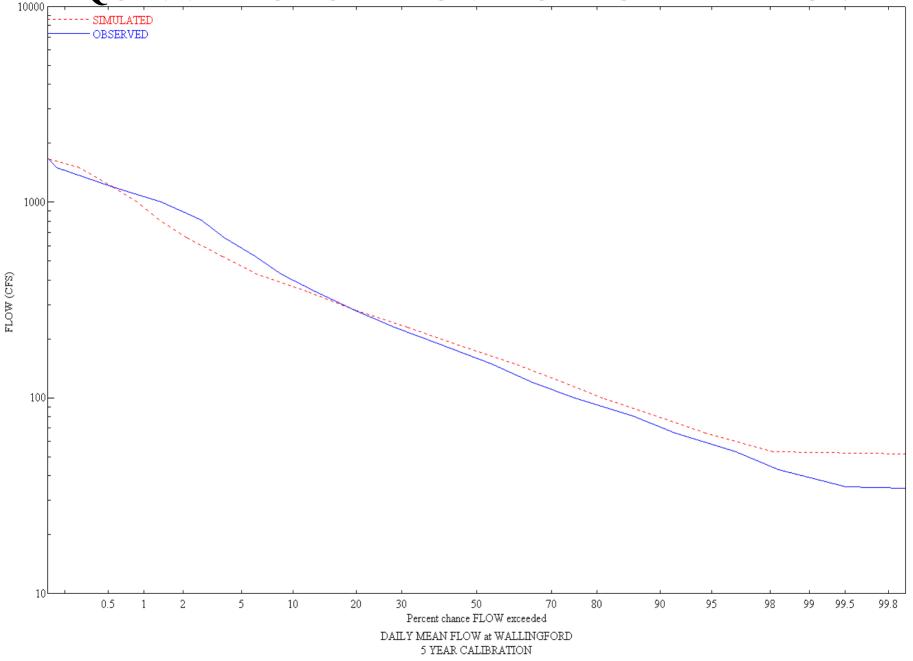
### **FARMINGTON TIMESERIES PLOT - CALIBRATION**



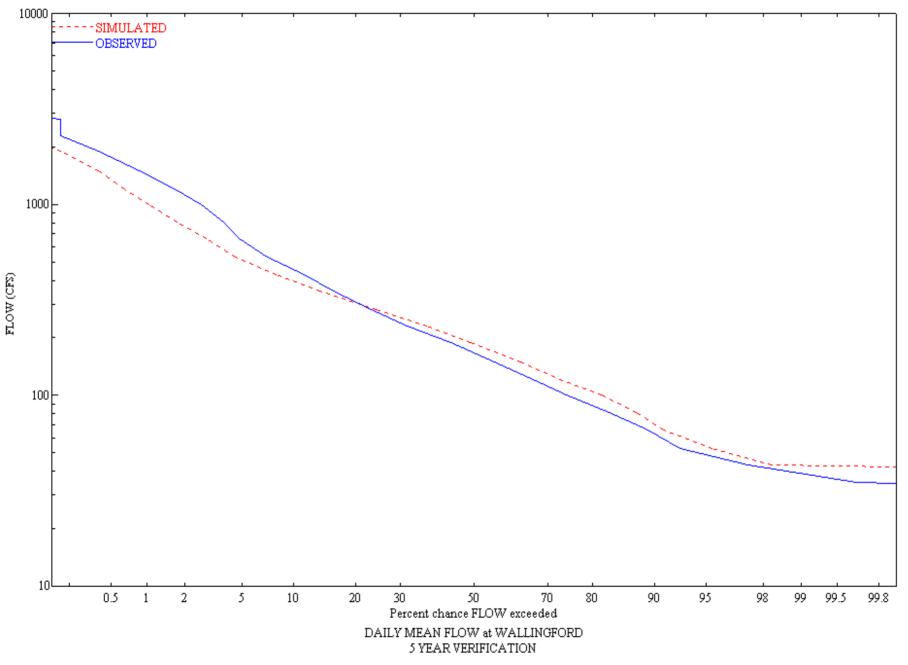
### **FARMINGTON TIMESERIES PLOT - VALIDATION**



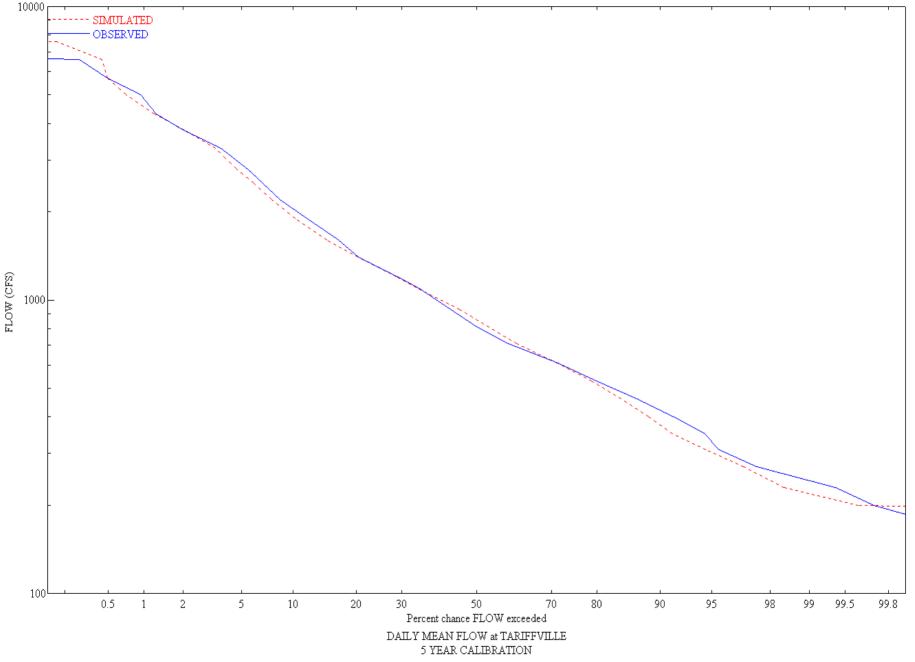
### **QUINNIPIAC DURATION PLOT – CALIBRATION**



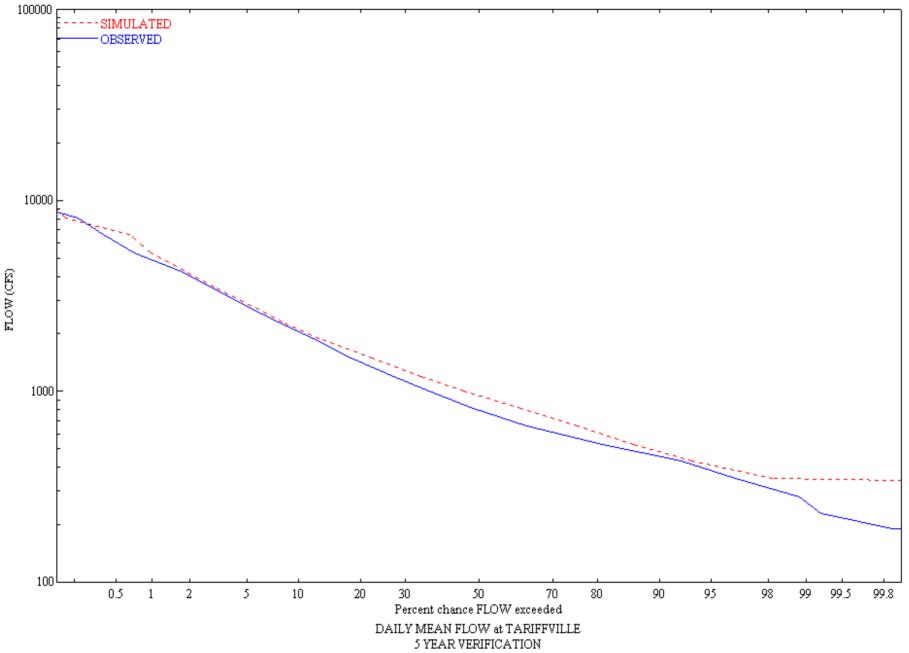
### **QUINNIPIAC DURATION PLOT - VALIDATION**



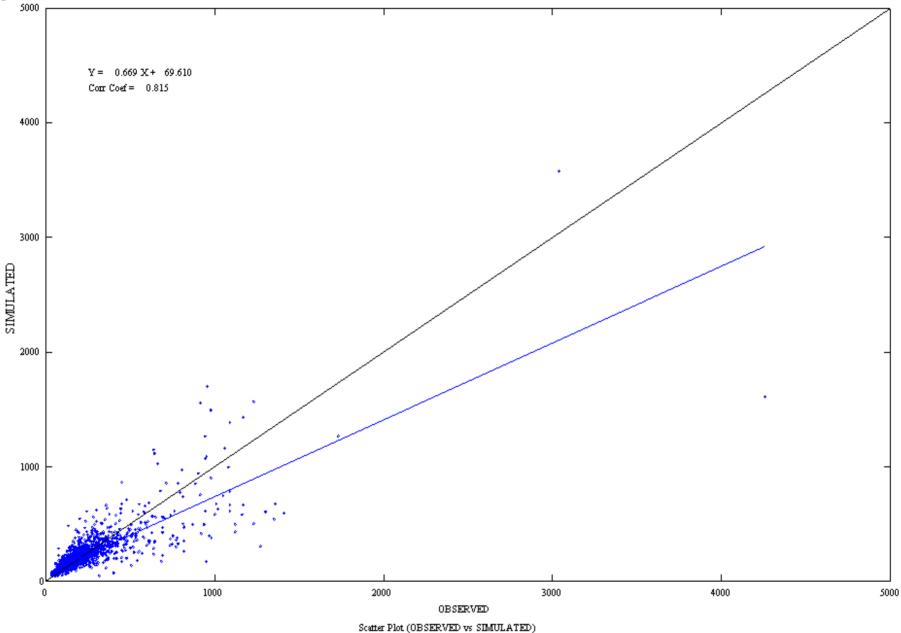
### **FARMINGTON DURATION PLOT - CALIBRATION**



### **FARMINGTON DURATION PLOT - VALIDATION**

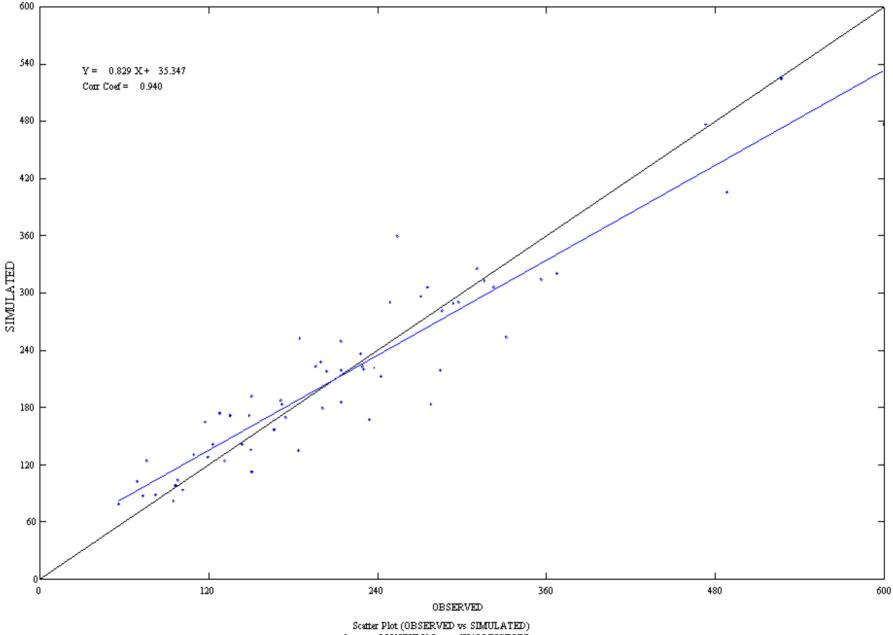


### **QUINNIPIAC SCATTER PLOT - DAILY**



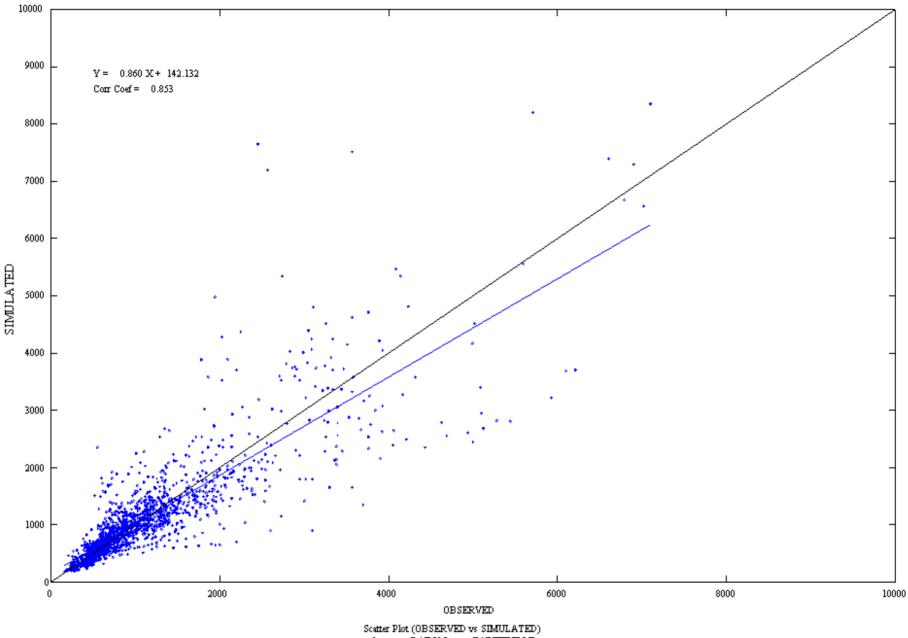
for mean DAILY flow at WALLINGFORD

### **QUINNIPIAC SCATTER PLOT – MONTHLY**



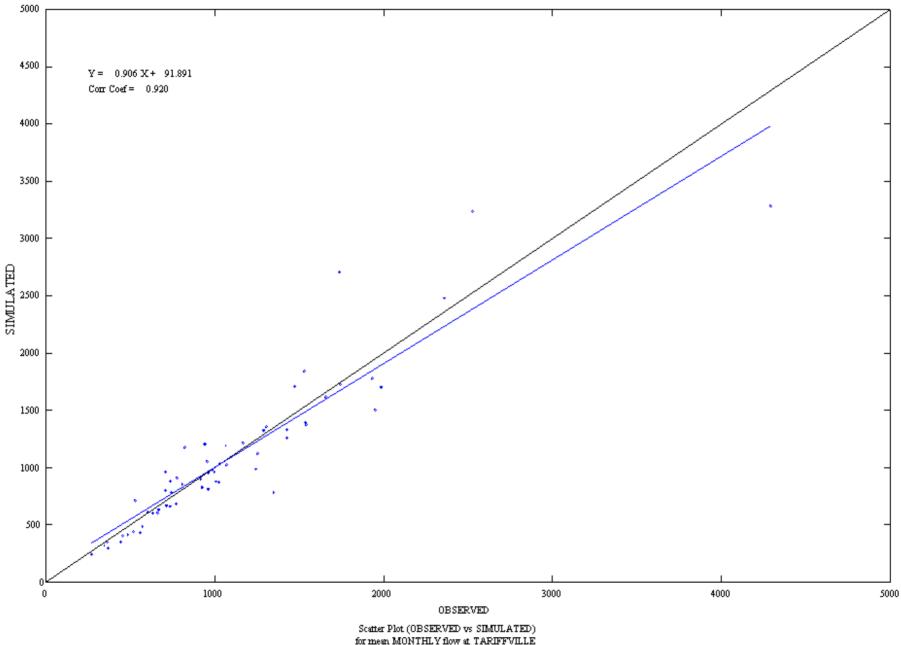
for mean MONTHLY flow at WALLINGFORD

### **FARMINGTON SCATTER PLOT - DAILY**



for mean DAILY flow at TARIFFVILLE

### **FARMINGTON SCATTER PLOT – MONTHLY**





### **LOADING RATES**

#### Frink's Export Coefficients (lb/ac/yr):

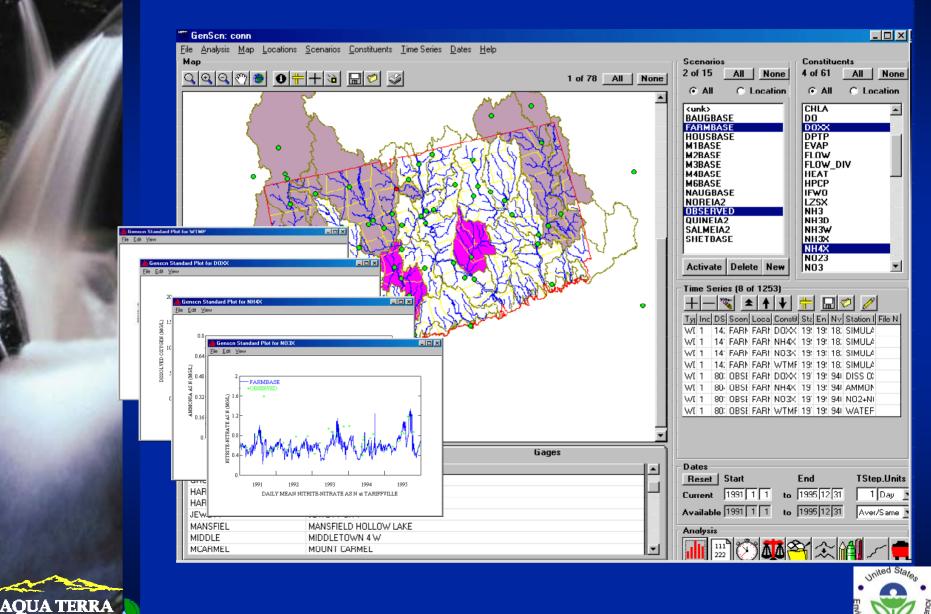
	Total Nitrogen	<b>Total Phosphorus</b>
Urban	$12.0\pm2.3$	$1.5\pm0.20$
Agriculture	$6.8 \pm 2.0$	$0.5\pm0.13$
Forest	$\textbf{2.1} \pm \textbf{0.4}$	$0.1\pm0.03$

#### CTWM Loading Rates (lb/ac/yr): Mean (range)

	Total Nitrogen	Total Phosphorus
Urban - pervious	8.5 (5.6 - 15.7)	0.26 (0.20 - 0.41)
Urban - impervious	<b>4.9</b> ( <b>3.7 - 6.6</b> )	0.32 (0.18 - 0.36)
Agriculture	<b>5.9</b> ( <b>3.4</b> - <b>11.6</b> )	0.30 (0.23 - 0.44)
Forest	2.4 (1.4 - 4.3)	0.04 (0.03 - 0.08)
Wetlands	2.2 (1.4 - 3.5)	0.03 (0.02 - 0.05)

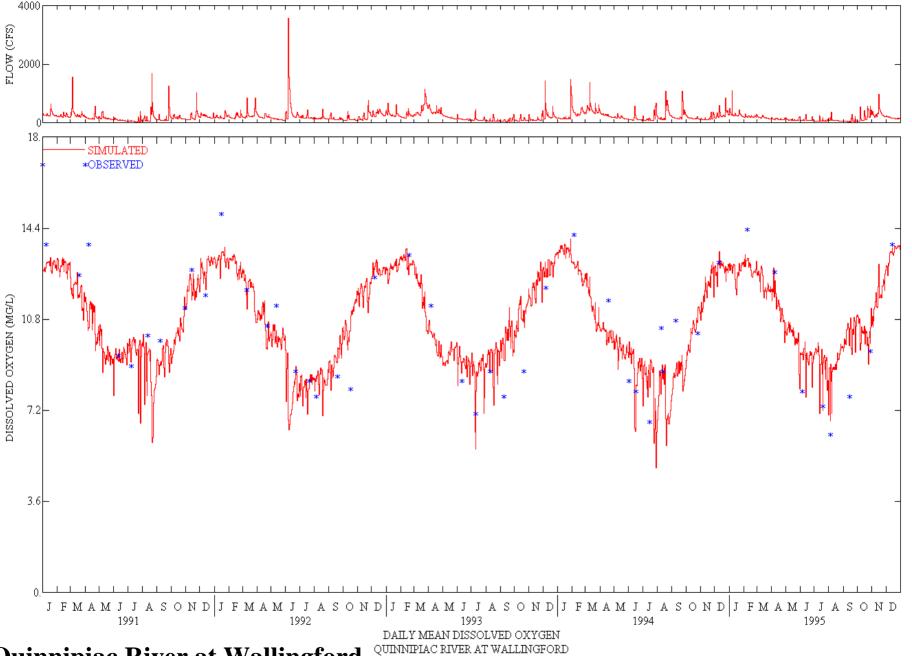


### **GENSCN WITH MULTIPLE WQ PLOTS**



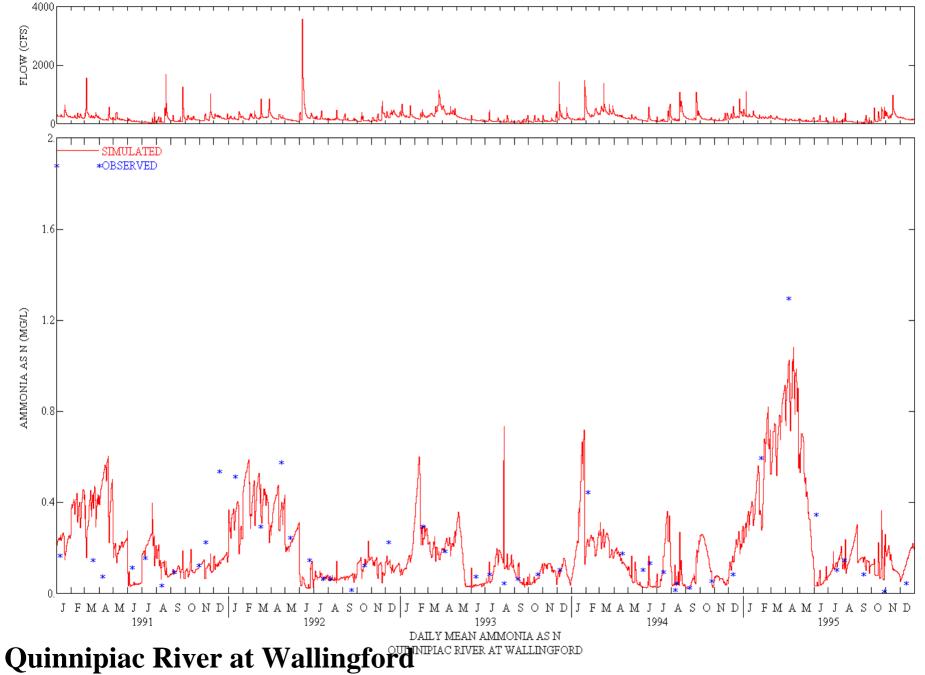
CONSULTANTS

OBSERVED AND SIMULATED DAILY DISSOLVED OXYGEN CONC AT WALLINGFORD C

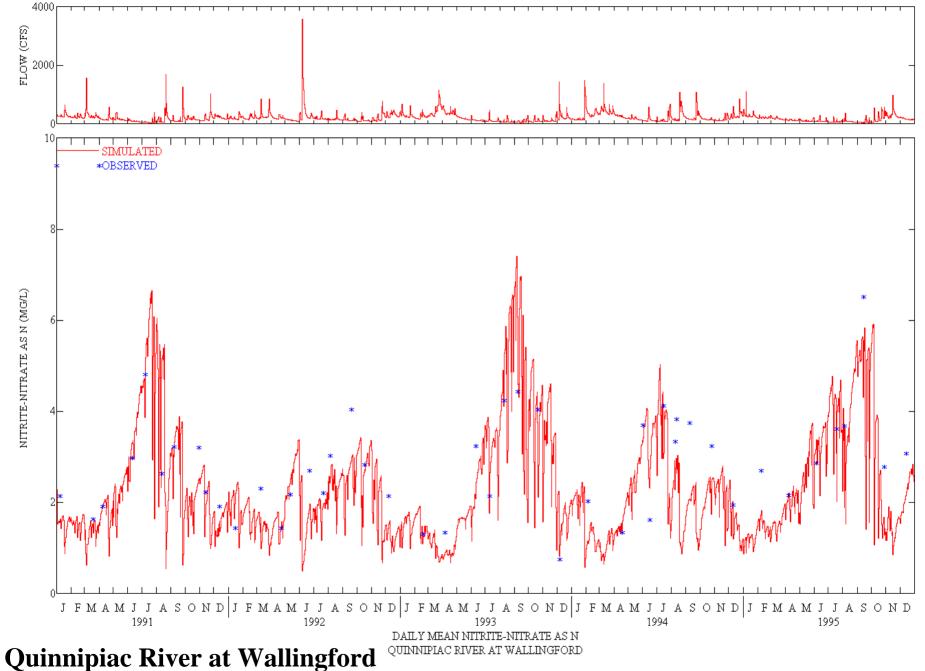


**Quinnipiac River at Wallingford** 

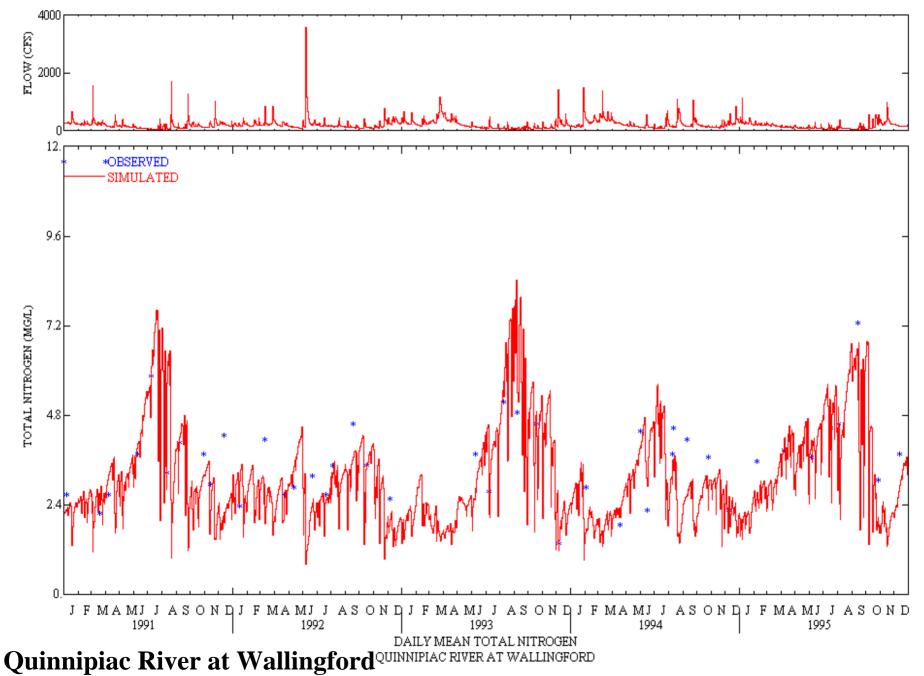
**OBSERVED AND SIMULATED DAILY AMMONIA AS N CONC AT WALLINGFORD CT** 



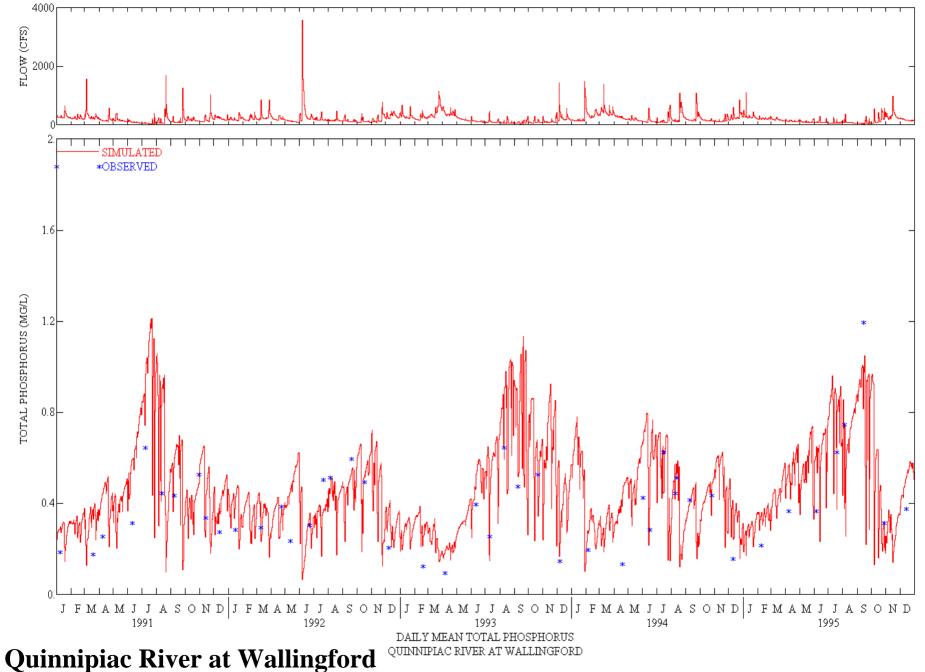
**OBSERVED AND SIMULATED DAILY NITRITE-NITRATE AS N CONC AT WALLINGFORD CT** 



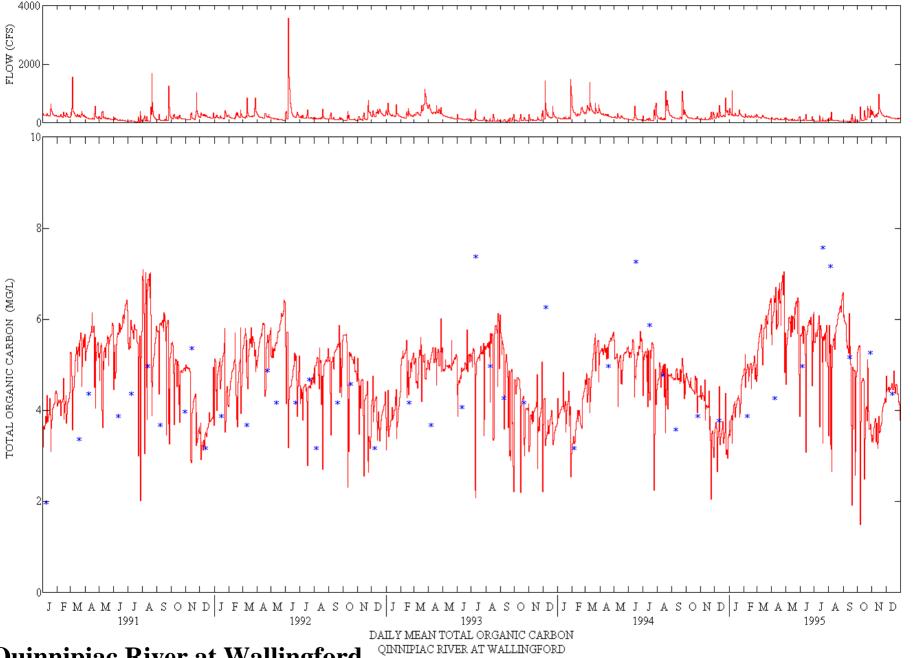
**OBSERVED AND SIMULATED DAILY TOTAL NITROGEN CONC AT WALLINGFORD CT** 



**OBSERVED AND SIMULATED DAILY TOTAL PHOSPHORUS CONC AT WALLINGFORD CT** 

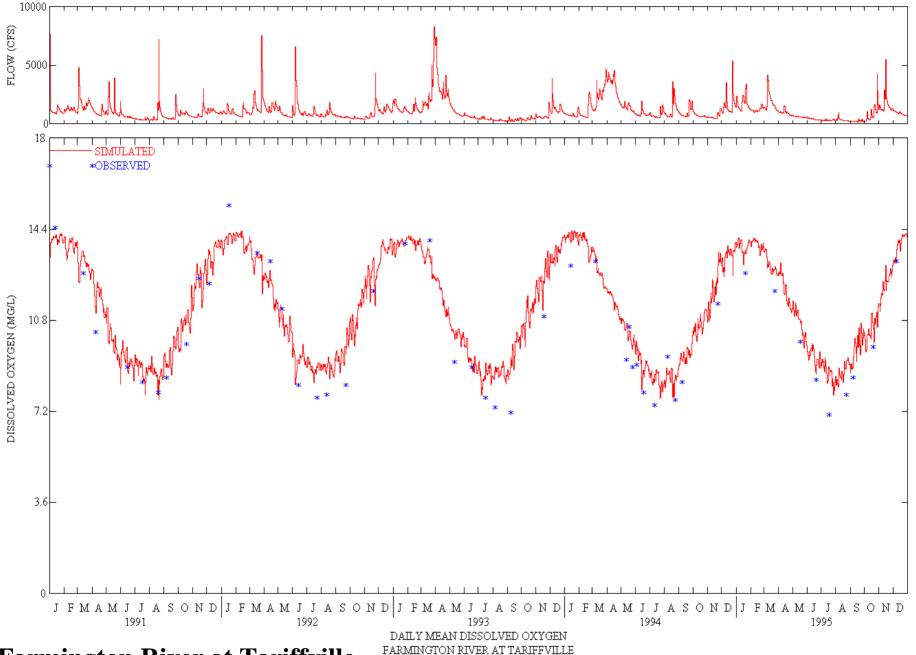


**OBSERVED AND SIMULATED DAILY TOTAL ORGANIC CARBON CONC AT WALLINGFORD** 



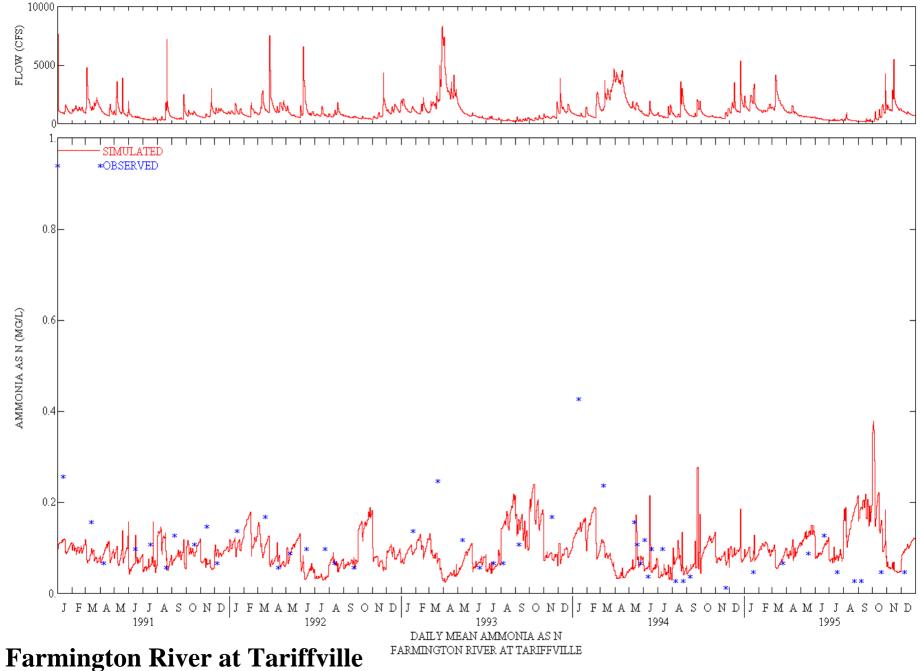
**Quinnipiac River at Wallingford** 

**OBSERVED AND SIMULATED DAILY DISSOLVED OXYGEN CONC AT TARIFFVILLE CT** 

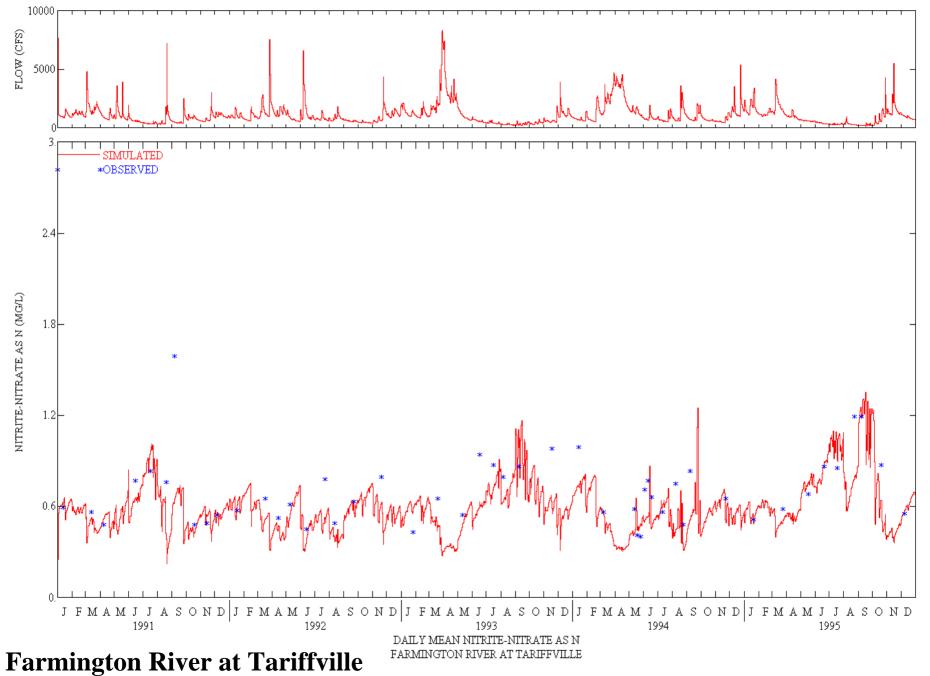


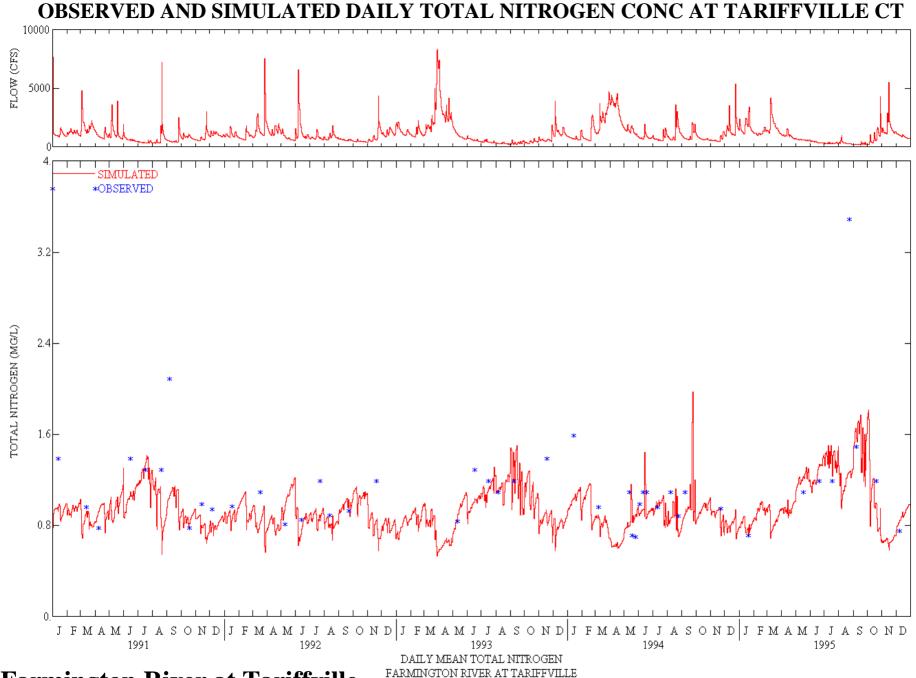
#### **Farmington River at Tariffville**

**OBSERVED AND SIMULATED DAILY AMMONIA AS N CONC AT TARIFFVILLE CT** 

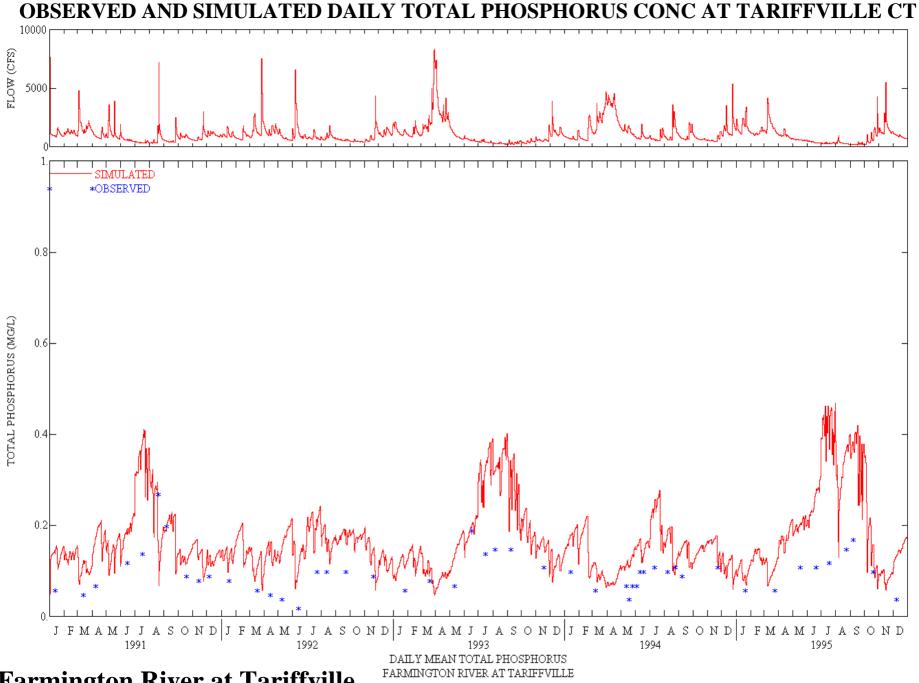


**OBSERVED AND SIMULATED DAILY NITRITE-NITRATE AS N CONC AT TARIFFVILLE CT** 

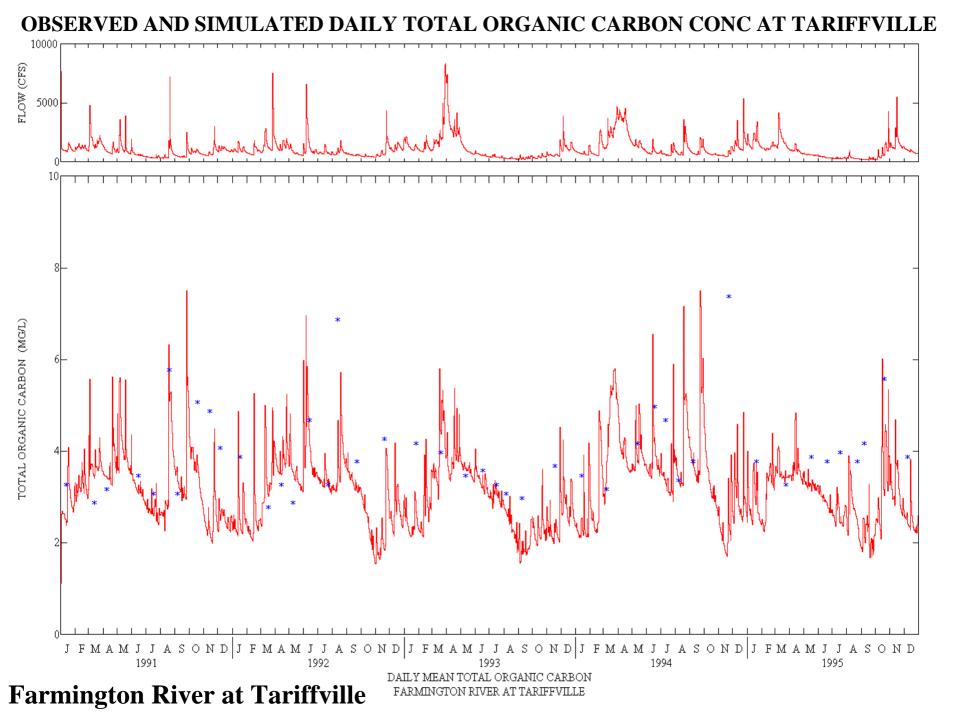




**Farmington River at Tariffville** 



**Farmington River at Tariffville** 



# AVERAGE ANNUAL SIMULATED AND OBSERVED CONCENTRATIONS (mg/L)

		on River nr Hampton	East	-	nipiac Rive Vallingford		Norwalk	River at Wi	nnipauk	Quineba	ug River at City	Jewett		ington Rive Tariffville	er at		atonic Rive Stevenson	er at
Constituent	Observed	Simulated	Ratio * (sample size)	Observed	Simulated	Ratio * (sample size)	Observed	Simulated	Ratio * (sample size)	Observed	Simulated	Ratio * (sample size)		Simulated	Ratio * (sample size)	Observed	Simulated	Ratio * (sample size)
Dissolved Oxygen	10.9	10.5	0.96 (48)	10.4	10.3	0.99 (46)	11.6	10.4	0.90 (97)	10.4	10.3	0.99 (43)		10.8	1.06 (49)	9.5	9.5	1.01 (41)
Ammonia as N	0.03	0.02	0.82 (43)	0.19	0.18	0.92 (46)	0.04	0.04	1.18 (80)	0.08	0.06	0.73 (42)		0.09	0.82 (48)	0.06	0.06	1.10 (33)
Nitrite-Nitrate as N	0.22	0.27	1.21 (46)	2.82	2.45	0.87 (46)	0.39	0.40	1.03 (93)	0.44	0.37	0.84 (42)		0.59	0.83 (49)	0.36	0.41	1.15 (40)
Organic Nitrogen	0.31	0.25	0.80 (30)	0.50	0.60	1.20 (44)	0.33	0.28	0.86 (70)	0.45	0.39	0.86 (40)		0.28	0.90 (45)	0.33	0.28	0.84 (38)
Total Nitrogen	0.53	0.51	0.97 (30)	3.64	3.29	0.90 (44)	0.73	0.69	0.94 (70)	0.96	0.80	0.83 (40)		0.97	0.85 (45)	0.77	0.75	0.97 (38)
Orthophosphate as P	0.01	0.01	0.91 (48)	0.32	0.36	1.10 (46)	0.02	0.02	0.93 (94)	0.02	0.04	1.67 (43)	0.07	0.13	1.90 (49)	0.01	0.02	1.49 (32)
Organic Phosphorus	0.02	0.02	1.30 (48)	0.07	0.11	1.62 (46)	0.02	0.03	1.18 (94)	0.03	0.04	1.23 (43)		0.05	1.59 (49)	0.02	0.03	1.19 (33)
Total Phosphorus	0.02	0.03	1.35 (48)	0.39	0.47	1.19 (46)	0.04	0.05	1.10 (94)	0.06	0.08	1.44 (43)		0.18	1.82 (49)	0.03	0.05	1.47 (40)
Total Organic Carbon	3.9	2.8	0.71 (45)	4.5	4.8	1.06 (44)	4.0	3.2	0.81 (28)	5.6	4.9	0.86 (41)		3.3	0.84 (45)	3.8	2.9	1.06 (49)

AVERAGE AND RANGE OF SIMULATED / OBSERVED CONCENTRATION RATIOS FOR ALL GAGES

Constituent	Average	Range
Dissolved Oxygen	0.99	0.90 - 1.06
Ammonia as N	0.93	0.73 - 1.18
Nitrite-Nitrate as N	0.99	0.83 - 1.21
Organic Nitrogen	0.91	0.80 - 1.20
Total Nitrogen	0.91	0.83 - 0.97
Orthophosphate as P	1.33	0.91 - 1.90
Organic Phosphorus	1.35	1.18 - 1.62
Total Phosphorus	1.4	1.10 - 1.82
Total Organic Carbon	0.89	0.71 - 1.06



### **UNNAMED NORTHEAST WATERSHED**

Western Massachusetts

# • 2 gages:

- Upper watershed about 50 sq mi
- Watershed outlet, about 300 sq mi

# 70% forest, 13% urban, 11% agri, 6% wetlands





# ANNUAL SIMULATED AND OBSERVED RUNOFF (inches)

	Unnamed Watershed						
	Precipitation	Simulated Flow	<b>Observed Flow</b>	Percent Error			
1990	58.9	35.1	35.6	-1.4%			
1991	47.0	23.3	22.8	2.1%			
1992	45.7	23.7	20.1	15.2%			
1993	47.6	27.6	26.0	5.8%			
1994	46.3	25.9	25.5	1.5%			
1995	44.0	20.7	21.0	-1.4%			
1996	62.0	39.4	41.5	-5.3%			
1997	42.2	21.4	23.2	-8.4%			
1998	42.2	22	23.9	-8.6%			
1999	46.9	21.6	24.8	-14.8%			
Total	482.7	260.7	264.4	-1.4%			
Average	48.3	26.1	26.4	-1.4%			
49 of 58							



# ANNUAL FLOW STATISTICS FROM HSPEXP

	Upstream	Tributary	Watershed Outlet		
	Simulated	Observed	Simulated	Observed	
Average runoff, in inches	27.12	26.23	26.07	26.44	
Total of highest 10% flows, in inches	10.88	10.72	8.56	8.94	
Total of lowest 50% flows, in inches	4.22	4.19	5.09	5.13	
Evapotranspiration, in inches	23.77	25.55 <sup>1</sup>	23.41	26.09 <sup>1</sup>	
Total storm volume, in inches <sup>2</sup>	47.07	51.91	38.72	42.36	
Average of storm peaks, in cfs <sup>2</sup>	710.84	791.88	2310.38	2287.19	
	Calculated	Criteria	Calculate d	Criteria	
Error in total volume, %	3.40	10.00	-1.40	10.00	
Error in 10% highest flows, %	1.50	15.00	-4.20	15.00	
Error in 50% lowest flows, %	0.60	10.00	-0.60	10.00	
Error in storm peaks, %	-10.20	15.00	1.00	15.00	

1 – PET (estimated by multiplying observed pan evaporation data by 0.73)

2 – Based on 31 storms occurring between 1990 and 1999

**OUA TERR** 

CONSULTANTS





# DAILY AND MONTHLY AVERAGE FLOW STATISTICS

Unnamed Watershed

I DERIK

CONSULTANTS

	Da	ily	Monthly				
	Simulated	Observed	Simulated	Observed			
Count	3652	3652	120	120			
Mean, cfs	539.85	547.65	540.46	547.56			
Geometric Mean, cfs	376.61	380.86	424.39	428.44			
<b>Correlation Coefficient (R)</b>	0.8	36	0.93				
Coefficient of Determination (R <sup>2</sup> )	0.7	74	0.8	0.87			
Mean Error, cfs	-7.	80	-7.10				
Mean Absolute Error, cfs	152	.97	101.22				
RMS Error, cfs	284	.09	140.26				
Model Fit Efficiency (1.0 is perfect)	0.7	73	0.87				



### AVERAGE OBSERVED MONTHLY RUNOFF RESIDUALS

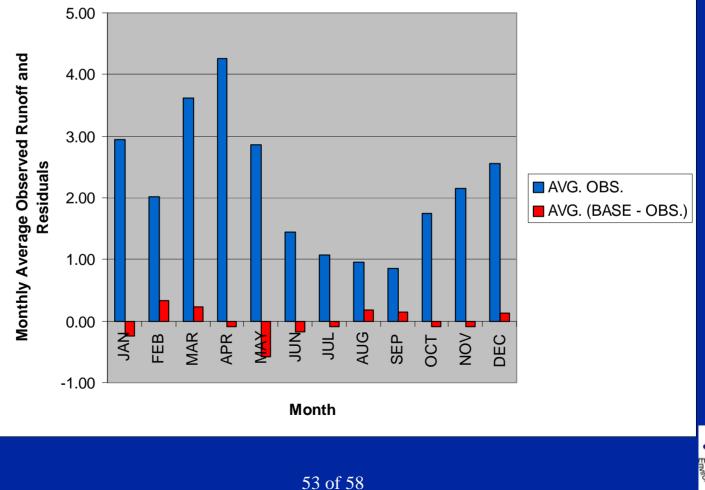
	Unnamed Watershed							
Month	Average Observed (in.)	Average Simulated (in.)	Average Residual (Simulated - Observed)	Percent Error				
JAN	2.94	2.71	-0.24	-8.09%				
FEB	2.01	2.34	0.33	16.46%				
MAR	3.61	3.85	0.23	6.42%				
APR	4.25	4.16	-0.09	-2.07%				
MAY	2.86	2.28	-0.58	-20.19%				
JUN	1.44	1.26	-0.18	-12.55%				
JUL	1.07	0.97	-0.10	-9.03%				
AUG	0.95	1.13	0.18	18.66%				
SEP	0.85	0.98	0.14	16.39%				
ОСТ	1.75	1.66	-0.08	-4.80%				
NOV	2.15	2.05	-0.09	-4.38%				
DEC	2.56	2.70	0.13	5.03%				
Totals	26.46	26.08	-0.35	-1.32%				





# **OBSERVED RUNOFF AND RESIDUALS (inches)**

Unnamed Watershed Yearly Average Observed Runoff and Residuals PRELIMINARY FINAL CALIBRATION



I HERRY

CONSULTANT

# AVERAGE ANNUAL EXPECTED AND SIMULATED WATER BALANCE (inches)

	Expected Ranges	Simulated
Precipitation	43 - 53	48
Total Runoff	23 - 27	24
Total ET	20 - 23	23
Deep Recharge	1 - 4	1





### **SIMULATED WATER BALANCE COMPONENTS BY LAND USE (inches)**

	Forest	Agriculture	Urban Pervious	Wetland	Urban Impervious
Precipitation	48.6	48.4	48.5	48.5	48.3
Total Runoff	22.6	25.8	26.5	21.3	42.8
Surface Runoff	1.0	4.6	4.6	0.3	42.7
Interflow	7.9	8.8	8.8	4.8	0.0
Baseflow	13.6	12.3	13.1	16.2	0.0
Total ET	24.6	22.1	21.2	24.2	5.5
Interception/Retention ET	9.6	6.1	6.3	4.6	5.5
Upper Zone ET	7.8	6.5	9.2	11.1	0.0
Lower Zone ET	6.6	9.2	5.3	4.6	0.0
Active GW ET	0.0	0.0	0.0	2.9	0.0
Baseflow ET	0.6	0.3	0.3	1.0	0.0
Deep Recharge	1.4	0.5	0.8	3.0	0.0



#### EXAMPLE WATERSHED MODEL CALIBRATION "WEIGHT-OF- EVIDENCE" SUMMARY

	Upper Gage	Outlet	Calib Perf.
Entire Period, %ME	0.6	1.6	VG
Annual Volume, %ME	+6/-5	+17/-9	VG
Monthly Volume, %ME	+15/-11	+21/-14	G
R2, Daily	0.76	0.81	G/VG
Monthly	0.9	0.9	VG
Flow-duration	G/VG	G/VG	G/VG
Water Balance	VG	VG	VG
Storm Events:			
Daily Peak, % Error	-7	-3	G
Storm Volumes, % ME	-1	-0.3	VG
10% High Flows, %ME	+2	+3	VG
			United S

**THERE** 

CONSULTANTS



#### **CLOSURE**

Watershed models can be valuable tools for TMDL development when applied and used appropriately, with adequate data, and in recognition of model limitations



#### **QUOTE FOR MODEL USERS**

With poor assumptions, a man can make more mistakes with a computer in a milli-second, than he could in a lifetime of common sense.



